How "sensory smog" is putting fish at risk *p. 1124*

Indigenous knowledge helps protect wildlife *p. 1135*

Squeezing precision *pp. 1137 & 1163*

Science

$15
21 JUNE 2019
sciencemag.org

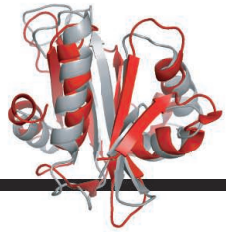
AAAS

RUMINANT GENOMES

Exploring the evolutionary success of a herbivore lineage *pp. 1130 & 1150*

CONTENTS

21 JUNE 2019 • VOLUME 364 • ISSUE 6446



NEWS

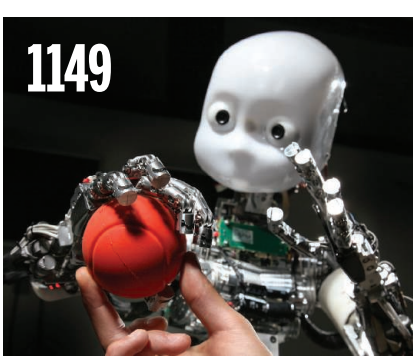
IN BRIEF

IN DEPTH

FEATURES

INSIGHTS

PERSPECTIVES



ON THE COVER

A female steenbok (*Raphicerus campestris*) in Mapungubwe National Park, Limpopo Province, South Africa. Steenboks are ruminants, one of the most widespread lineages of herbivorous mammals across the globe. Three papers in this issue explore ruminant genomes to identify specific traits that underlie their ability to diversify and thrive evolutionarily. See pages 1130 and 1150. *Photo: Neil Aldridge/Minden Pictures*

PHOTOS (FROM TOP): S. KUTZ; BILLARD AND KRAGIC

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. 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): $1971; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (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; 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.

Science AAAS

Editor-in-Chief Jeremy Berg, jberg@aaas.org

Executive Editor Monica M. Bradford

EDITOR, INSIGHTS Lisa D. Chong **EDITORS, RESEARCH** Valda Vinson, Jake S. Yeston **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, Paula A. Kiberstis, Marc S. Lavine (Canada), Steve Mao, Ian S. Osborne (UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Priscilla N. Kelly, Tage S. Rai, Seth Thomas Scanlon (UK), Keith T. Smith (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, 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, Alana Warnke, Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Jeremy Dow, Alexander Kief, Ronmel Navas, Hilary Stewart (UK), 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, Kelly Servick, Robert F. Service, Erik Stokstad (Cambridge, UK), Paul Voosen, Meredith Wadman **INTERN** Alex Fox **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Ann Gibbons, Mara Hvistendahl, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Shanghai), Elisabeth Pain (Careers), Charles Piller, Michael Price, Tania Rabesandratana (Barcelona), Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Donisha Adams, Rachel Bernstein (Editor), Katie Langin **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 **SENIOR DESIGNER** Chrystal Smith **DESIGNER** Christina Aycock **PHOTO EDITOR** Emily Petersen **GRAPHICS EDITOR** Nirja Desai **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel **SCIENTIFIC ILLUSTRATOR** Alice Kitterman **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary

CEO and Executive Publisher Rush D. Holt

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 **QUALITY TECHNICAL MANAGER** Marcus Spiegler **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Antoinette Hodal, Lori Murphy **CONTENT SPECIALIST** Jacob Hedrick **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **DIGITAL AD-OPS SPECIALIST** Patrick Gerrits **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 **MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Tori Velasquez, Mike Romano, Ashley Hylton **SENIOR DESIGNER** Kim Huynh **TRADE SHOW AND MEETINGS ASSOCIATE** Andrew Clamp

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

DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **ASSOCIATE DIRECTOR, PRODUCT DEVELOPMENT AND SPJ** Hannah Heckner **SR. PRODUCT ASSOCIATE** Robert Koepke **DIGITAL PRODUCT STRATEGIST** Michael Hardesty **SPJ ASSOCIATE** Samantha Bruno Fuller

DIRECTOR, INSTITUTIONAL LICENSING Iquo Edim **ASSOCIATE DIRECTOR, RESEARCH & DEVELOPMENT** Elisabeth Leonard **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGERS** Marco Castellan, Christos Skoutas **MANAGER, SYSTEMS AND OPERATIONS** Brian Holiahn **MANAGER, AGENT RELATIONS & CUSTOMER SUCCESS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **FULFILLMENT COORDINATOR** Melody Stringer

DIRECTOR, GLOBAL SALES Tracy Holmes **ASSOCIATE DIRECTOR, ADVERTISING SALES, US** Laurie Faraday **US EAST COAST AND MID WEST SALES** Glen Cox **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 **PROJECT 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

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.

Published by AAAS

EDITORIAL

DNA patents revisited

A bipartisan proposal to modify restraints in patent law was debated before the United States Senate Judiciary Committee this month. Immediately following the third hearing last week, leading scientists called for a study of the relevant issues by the U.S. National Academies of Sciences, Engineering, and Medicine, prior to the draft legislation moving forward (www.patenteligibility.com). This study should be initiated soon, with a short timeline, to ensure that appropriate information and analysis can inform the legislative processes that will substantially affect how the benefits of science and innovation are delivered to society.

Patents balance providing incentives to take the financial risks necessary to convert an invention into useful products with the benefits of sharing information to drive other useful inventions. Since the advent of DNA sequencing and gene identification methods, patenting human genes has been controversial. A notable example involves patents for the genes *BRCA1* and *BRCA2*, variations in which modulate risks for breast and ovarian cancer. These patents supported increased costs and hence, limited accessibility, for diagnostic tests for cancer patients and their families. In 2013, the U.S. Supreme Court ruled that these patents were invalid (*Association for Molecular Pathology* v. *Myriad Genetics*). The Myriad decision shifted the landscape not only for gene patents but also for patents on other "products of nature." The current bill is directed toward clarifying the concepts and language underlying the issues involved in patenting products of nature. It is essential that these issues be examined thoroughly and thoughtfully prior to enshrining language in law that could easily have unintended consequences.

"...patenting human genes has been controversial."

To avoid overly restricting research, there are certain exclusions from patentability, namely for abstract ideas, natural laws, and products of nature. However, for products of nature, exceptions have been made for identifying and isolating substances whose properties differ substantially from the source from which they were obtained. For more than a century, natural products, such as adrenaline, have been deemed patentable in their purified forms because of the enhancement of their activities through purification. Natural products have been quite important as drugs or in driving drug discovery, such as the immunosuppressive compounds cyclosporine and tacrolimus (FK506), which enabled solid organ transplantation.

Isolation of natural substances also enables "composition of matter" patents. These are valuable because they cover the material regardless of its use or the processes in which they are components. The Myriad patents involved the composition of matter of the *BRCA1/2* genes. The Supreme Court ruled that these substances were not eligible as exceptions from the product of nature exclusion because isolation of the *BRCA* genes from their chromosomal environment did not transform their character sufficiently. Moreover, the item of value was information in the gene sequences rather than the materials themselves. However, the language in the Court opinions was subject to different interpretations. Some have interpreted them as applying only to the patenting of genes, whereas others suggested that they change the scope of natural products that can be patented.

Because of the great importance of these somewhat esoteric issues for facilitating the generation of new diagnostics and drugs and for advancing basic research, it is crucial that appropriate steps be taken to guide modification of existing patent law. At this point, case law based on particular court rulings is controlling the evolution of these patent rulings. However, a more intentional approach implemented through legislation has the potential to address these issues in a more comprehensive manner. This task should be informed by a thorough process that the National Academies can effectuate—bringing together stakeholders including academic scientists, industry leaders, patent experts, and patient advocates. Striking the right balance with language that clearly conveys the intended interpretations should strengthen patent protection, bolster science, and improve societal well-being.

–Jeremy Berg

Jeremy Berg
Editor-in-Chief,
Science *Journals.*
jberg@aaas.org

10.1126/science.aay3990

NEWS

> “Ending the manel begins at the top.”
>
> **Francis Collins,** director of the U.S. National Institutes of Health, saying he will not appear on all-male public panels—“manels”—common at scientific meetings.

IN BRIEF

A woman in Mpondwe, Uganda, washes with chlorinated water as a precaution against Ebola.

INFECTIOUS DISEASE

As Ebola spreads, no emergency call

In a controversial decision, the World Health Organization (WHO) in Geneva, Switzerland, decided last week not to declare Africa’s continuing Ebola outbreak, which has killed more than 1400 people, a Public Health Emergency of International Concern (PHEIC), even as the disease moved from the Democratic Republic of the Congo (DRC) into a neighboring country. Two people died in Uganda last week after crossing the border from the DRC, which has battled the outbreak since August 2018. Some public health experts argued that a declaration would focus global attention on what is now the largest outbreak of Ebola since the one that ravaged West Africa 5 years ago; a PHEIC would allow WHO to share information about the disease’s spread without the affected countries’ consent and make temporary recommendations that member states must follow. But the outbreak has not satisfied all PHEIC criteria, said Preben Aavitsland, acting chair of a WHO expert committee. He said a declaration could be counterproductive because countries are already sharing information and following WHO’s advice, and a PHEIC could lead to restrictions on travel and trade that might hamper public health efforts in the region. Even so, he and the critics agree that the international community must provide more funding and resources to control the outbreak.

Edited by **Jeffrey Brainard**

Trump targets advisory panels

SCIENCE POLICY | U.S. President Donald Trump thinks the U.S. government has too many expert committees giving strategic advice to federal research agencies. That’s the gist of a 14 June executive order that gives each of those agencies, including NASA, the National Science Foundation, and the Department of Energy’s Office of Science, until 30 September to wipe out at least one-third of its top-level advisory panels. The order exempts panels at the National Institutes of Health and other agencies that review grant proposals or help ensure the safety and efficacy of drugs, food, and other consumer products. It also excludes—for now—panels created by law rather than agency edict, although the order invites agencies to propose a hit list of those statutory committees in their upcoming budget requests. A list of hundreds of these committees is available at http://bit.ly/2INdueZ.

NIH workers report harassment

#METOO | More than one in five employees at the U.S. National Institutes of Health in Bethesda, Maryland, said they were sexually harassed at work recently, according to interim results from a survey released last week, to which nearly 16,000 NIH employees and agency contractors responded. Eighteen percent of the respondents to the survey, conducted between January and March, said that in the prior 12 months, they had experienced gender harassment such as sexist put-downs, 10.3% reported receiving unwanted sexual attention, and 0.3% said they were subject to sexual coercion. In all, 21.6% said they had coped with one or more of these three forms of sexual harassment.

China regulates genetic material

GENETICS | China last week announced national regulations for study and transport of human biological material containing DNA. The government intends the policies to encourage R&D of drugs and medical treatments. But the rules

PHOTO: RONALD KABUUBI/AP PHOTO

Published by AAAS

CLIMATE SCIENCE

Abnormal heat causes early Greenland melt

The melting season in Greenland abruptly accelerated last week, as Steffen Olsen, a climate scientist at the Danish Meteorological Institute in Copenhagen, and a colleague (above) discovered when they set off on dog sleds to retrieve research equipment from a fjord on Greenland's northwest coast. An unusual spell of heat, with temperatures briefly more than 20°C above normal, left meltwater on top of ice and set more than 40% of Greenland's entire ice sheet melting at its surface, an extent seen only at the peak of summer last year. Human-driven climate change has warmed the Arctic more than any other region, and researchers fear this summer could rival records set in 2012, when a July heat wave drove surface melt across nearly all of Greenland's ice sheet.

PHOTOS: (TOP TO BOTTOM) STEFFEN OLSEN/DANISH METEOROLOGICAL INSTITUTE; JOSHUA STEVENS/U.S. GEOLOGICAL SURVEY/EUROPEAN SPACE AGENCY/NASA

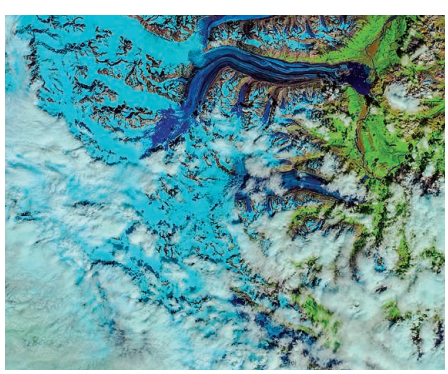
are also meant to protect patient privacy and control of the material: One of the rules requires foreign scientists to have Chinese collaborators to access and use the resources and data. That reflects concern by the government that foreign researchers have collected and removed biological materials from China for study without permission. The rules cover the collection, storage, and use of organs, tissues, and cells as well as the underlying genetic data. They also specify fines of up to 5 million Chinese yuan (about $724,000) for violations.

New York ends vaccine exemption

PUBLIC HEALTH | New York state last week eliminated an exemption that had allowed parents to refuse to vaccinate their school-attending children because of "genuine and sincere religious beliefs." State lawmakers in Albany acted despite vocal protests by scores of vaccination opponents at the Capitol. Governor Andrew Cuomo immediately signed the bill ending the exemption, calling it a response to a public health crisis. The state has been the epicenter of a U.S. measles epidemic this year, logging more than 850 of the 1044 cases nationwide as of 13 June. Most of New York state's cases have occurred in ultra-Orthodox Jewish communities. New York joins five other states that bar religious exemptions: Arizona, California, Maine, Mississippi, and West Virginia. New York requires vaccination of children who attend any public or private daycare, preschool, or school.

Panel: Keep Landsat images free

EARTH SCIENCE | A federal advisory panel last week called on the U.S. Department of the Interior (DOI) to maintain free and public access to widely used imagery from its Landsat satellite program. DOI could decide nevertheless to charge fees for access, but is not expected to. Begun in 1972, Landsat has produced the world's longest-running series of satellite images, used by academics and industry to document global change. In 2017, DOI asked the National Geospatial Advisory Committee to evaluate the possibility of charging for Landsat data, which the government did until 2008. Such a move would be unwise, the panel concluded, driving users to free alternatives and generating little money.



This false-color photo taken by Landsat-8 captured a developing "snow swamp" in Canada's Yukon territory.

U.K. integrity watchdog gears up

RESEARCH ETHICS | The United Kingdom's main funding agency, UK Research and Innovation (UKRI), is creating a watchdog to oversee research integrity. Last year, a Parliament committee called for national oversight, citing concerns that universities were not fully complying with requirements to report on their internal investigations of research misconduct. UKRI is establishing an independent entity to make sure institutions follow "appropriate processes to investigate misconduct," according to the agency's annual strategic

IN OTHER NEWS

SPACE STATION China announced nine scientific experiments it will run on its new space station, scheduled for completion in 2022. Projects were selected in an international competition and involve scientists from 17 countries, but not the United States. Topics include effects of space flight on tumor formation in human organoid tissue.

FETAL TISSUE The Democratic-controlled U.S. House of Representatives amended a proposed 2020 spending bill to block a key part of President Donald Trump's administration's new policy that restricts U.S.-funded research with fetal tissue donated after elective abortions. The amendment prohibits funding ethics advisory boards to review new fetal tissue projects.

NEW WEED KILLERS Bayer AG in Leverkusen, Germany, announced it will spend €5 billion over the next decade on research to develop weed-killing alternatives to its controversial, widely used herbicide glyphosate. Countries in Europe have been moving toward banning the product, and Bayer faces multimillion-dollar U.S. jury verdicts from lawsuits alleging that it caused illnesses.

AGENCIES MOVED The U.S. Department of Agriculture has picked the Kansas City, Missouri, region to host two research units that many scientists say should be kept in Washington, D.C. "[Agriculture] Secretary [Sonny] Perdue is well on his way to dismantling one of the best agricultural economics research institutions in the world," says Ron Wasserstein of the American Statistical Association in Alexandria, Virginia, about one of the units, the Economic Research Service. The other is the National Institute of Food and Agriculture.

DOCUMENTARY QUESTIONED Cell biologists voiced qualms about being shown in a documentary touting stem cell therapies after discovering the project was funded by for-profit companies under investigation for offering the treatments, the *San Francisco Chronicle* reported. Makers of *The Healthcare Revolution* removed Jeanne Loring, a professor emeritus at Scripps Research in San Diego, California, from the film at her request. No stem cell therapy sold at for-profit clinics has been approved by the U.S. Food and Drug Administration.

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

plan released last week. UKRI provided no additional details but says it will issue an update this summer. The strategic plan also says UKRI will create an ethics statement, evaluate training of Ph.D. students in research integrity, and fund research on how academic incentives could be adjusted to support research integrity.

Fecal transplant risks arise

CLINICAL RESEARCH | After a patient died from an infection acquired during a fecal transplant, the U.S. Food and Drug Administration (FDA) last week issued an alert to clinicians about the procedure's risks and suspended an undisclosed number of clinical trials. Inserting a healthy person's stool into a patient's colon has shown dramatic results in clearing intestinal infections caused by the bacterium *Clostridium difficile*; the method is also under investigation to treat other illnesses. But in two cases described in the agency's 13 June warning, a donor's stool was not screened for other dangerous bacteria and contained drug-resistant *Escherichia coli*. Two immunocompromised recipients developed infections, and one died. FDA is now requiring clinical researchers to establish procedures to screen donors and their stools before resuming the trials.

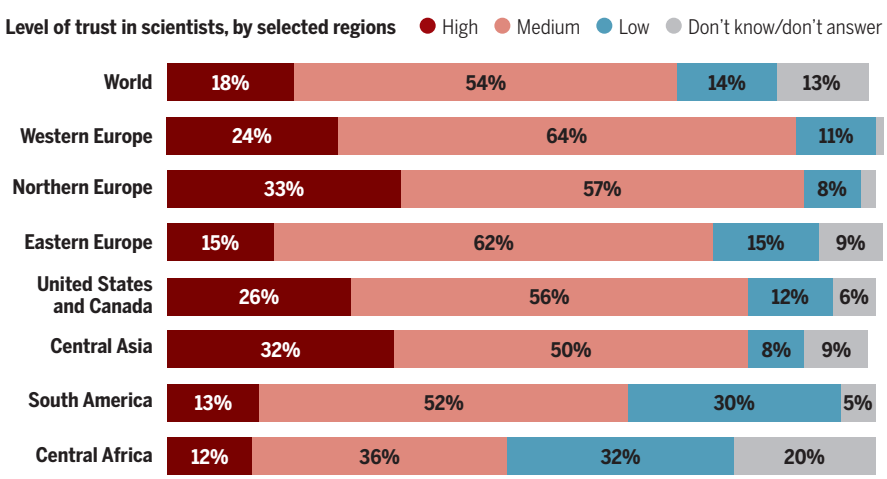
Dutch college seeks only women

DIVERSITY | The Eindhoven University of Technology (TUE) in the Netherlands said this week it is taking radical action to hire more female academics by opening job vacancies to women only. The engineering institution says it will not allow men to apply for permanent academic jobs for the first 6 months of the recruitment process under a new fellowship program. Dutch and EU laws allow policies to recruit underrepresented groups, TUE says. But only 29% of TUE's assistant professors are women; about 15% are at the associate and full professor level. The university wants women to make up 50% of its assistant and associate professor positions and 35% of its full professors.

PUBLIC OPINION

Global survey finds strong support for scientists

Nearly three-quarters of people worldwide solidly trust scientists: That's one of the main findings of the Wellcome Global Monitor, a new survey that asked 140,000-plus people in more than 140 countries how they think and feel about health and science. Other polls have asked similar questions, but this one, conducted by Gallup World Poll on behalf of London-based biomedical charity the Wellcome Trust, claims to be the first to study on a global scale how attitudes vary by nationality, gender, income, and education. An index based on five questions found that 54% of people trust scientists at a medium level and 18% at a high level, whereas 14% have a low level of trust. But regional differences are striking: People in Uzbekistan say they trust scientists the most, residents of Gabon the least. On the benefits of science, more than one-third of people in southern Africa and Latin America say science helps "very few" people in their country. The survey also explored attitudes toward vaccines and found people in France are the most skeptical (see p. 1122). The survey reveals a global gender gap in self-assessments of scientific knowledge: Globally, 49% of men say they know "some" or "a lot" about science—a full 11 percentage points more than women.



Percentages are rounded and may not add up to 100%.

CREDITS: (GRAPHIC) A. CUADRA/*SCIENCE*; (DATA) HTTPS://WELLCOME.AC.UK/WHAT-WE-DO/OUR-WORK/WELLCOME-GLOBAL-MONITOR

Francis Collins, director of the National Institutes of Health, says he's pleased with bold recommendations to help the agency curb sexual harassment.

PHOTO: STEPHEN VOSS

SCIENCE POLICY

NIH prepares to toughen harassment rules

Advisory group urges mandatory reporting of #MeTooSTEM investigations to agency

By **Jocelyn Kaiser**

Last year, on the heels of several high-profile cases of sexual harassment in science, the National Institutes of Health (NIH) came under fire for standing by while another major U.S. research funding agency, the National Science Foundation (NSF), quickly tightened its policies. But now, NIH officials appear ready to act—and even to go beyond NSF's rules. NIH leaders said last week that they plan to move forward with four recommendations from a group advising the agency, including hard-hitting reporting requirements aimed at making sure NIH doesn't unknowingly fund researchers found guilty of sexual harassment.

NIH Director Francis Collins said he welcomed the advice from the 14 outside scientists and seven NIH staff members who had been meeting since February: "I'm happy the recommendations are quite bold," he said after the group presented to his Advisory Committee to the Director at NIH in Bethesda, Maryland, on 13 June. Collins later announced NIH would "be pursuing mechanisms and approaches to implement these interim recommendations, where feasible."

The group's report is drawing praise from the #MeTooSTEM movement and others. "This is a very dramatic shift. I'm jumping up and down about the progress," says Kathy Hudson, formerly the senior policy official at NIH and now a consultant in Washington, D.C. But she and others hope the group's final recommendations, due in December, will go even further.

NSF last fall began to require that institutions report when a principal investigator (PI) has been found guilty of sexual harassment or when institutions take actions such as putting a scientist on leave as they investigate allegations of such misconduct. The goal was to avoid repeating past cases in which the agency learned belatedly, through media reports, that scientists it was funding had been found to have harassed women. But NIH held off on new policies and instead appointed the working group to explore possible changes.

Topping the list of its proposed changes is that sexual harassment be treated "as seriously as research misconduct," a recommendation that comes from a 2018 National Academies of Sciences, Engineering, and Medicine (NASEM) report (*Science*, 15 June 2018, p. 1159). That would mean establishing rules parallel to those now requiring institutions to inform the Department of Health and Human Services about research misconduct cases. Specifically, the group says, NIH should require that institutions tell the agency about any finding or investigation of alleged misconduct—including sexual harassment—within 1 week after an investigation begins or findings are issued.

The advisers also recommended that PIs and co-PIs themselves "attest" on grant applications and progress reports that they have not violated and will not violate their institution's code of conduct. This second reporting step goes beyond NSF's guidelines. The group proposed specific questions asking whether the applicant has been found guilty of or been involved in a legal settlement concerning sexual harassment or other misconduct within the past 7 years. The idea is to encourage good behavior as well as to potentially prevent some harassers from continuing to get NIH funding. Francis Cuss, a former Bristol-Myers Squibb executive in New York City who co-chaired the working group, said at the meeting that responses would go to NIH staff, who would consider whether to deny or revoke funding or impose other penalties on a case-to-case basis.

"Right now NIH has no way to know" about harassment allegations, says biologist Carol Greider of Johns Hopkins University in Baltimore, Maryland, part of the advisory group. "To know it's happening … is a large step in the right direction."

NIH should also take steps to restore the careers of sexual harassment survivors, the

group says, for example by encouraging them to apply to programs that help researchers restart their careers after dropping out for reasons such as having a baby. Finally, NIH should give trainees more control over their financial support, perhaps by giving more awards directly to individual trainees, rather than to institutions or mentors.

The time limit on reporting past misdeeds was set so that a reformed harasser would not be "branded for life," explained working group co-chair and NIH Associate Director for Science Policy Carrie Wolinetz at the meeting. The group will examine the 7-year limit before it issues a final report in December, added co-chair Kristina Johnson, chancellor of the State University of New York system. It will also explore whether the reporting policies should apply to staff who are not PI or co-PI on a grant.

The recommendations "while important, will only catch a tiny minority of harassers," predicts Lilia Cortina, a psychologist at the University of Michigan in Ann Arbor who helped write the NASEM report. She hopes NIH will also have institutions survey labs it funds to assess whether sexual harassment is happening, as it has recently done for its intramural researchers (see p. 1114).

NIH officials said previously that adopting a reporting policy like NSF's could require a lengthy, formal rulemaking process. Greider thinks that bureaucratic obstacle can be surmounted. "I do hope the working group goes further," she says. "We will have until December to work on a lot of this."

One prominent #MeTooSTEM activist was pleased with the recommendations. "There are a lot of good things," including that NIH will be more directly involved in sanctioning harassers, says BethAnn McLaughlin of Vanderbilt University in Nashville, who has sharply criticized NIH for its lack of action.

McLaughlin also finds new data NIH shared at the meeting last week to be "hugely inspiring." The numbers show that disciplinary actions involving sexual harassment are up, suggesting NIH is taking the problem seriously. So far this year, NIH says it has reviewed 31 cases of possible harassment (more than in all of 2018), removed five scientists from grants and stopped using 19 peer reviewers.

Congress is also moving to reconcile sexual harassment policies across federal science agencies. This week, the science committee in the U.S. House of Representatives was expected to approve a measure intended to push the largest federal research agencies, including NIH, toward a "uniform set of policies" on sexual harassment. ■

With reporting by Meredith Wadman and Jeffrey Mervis.

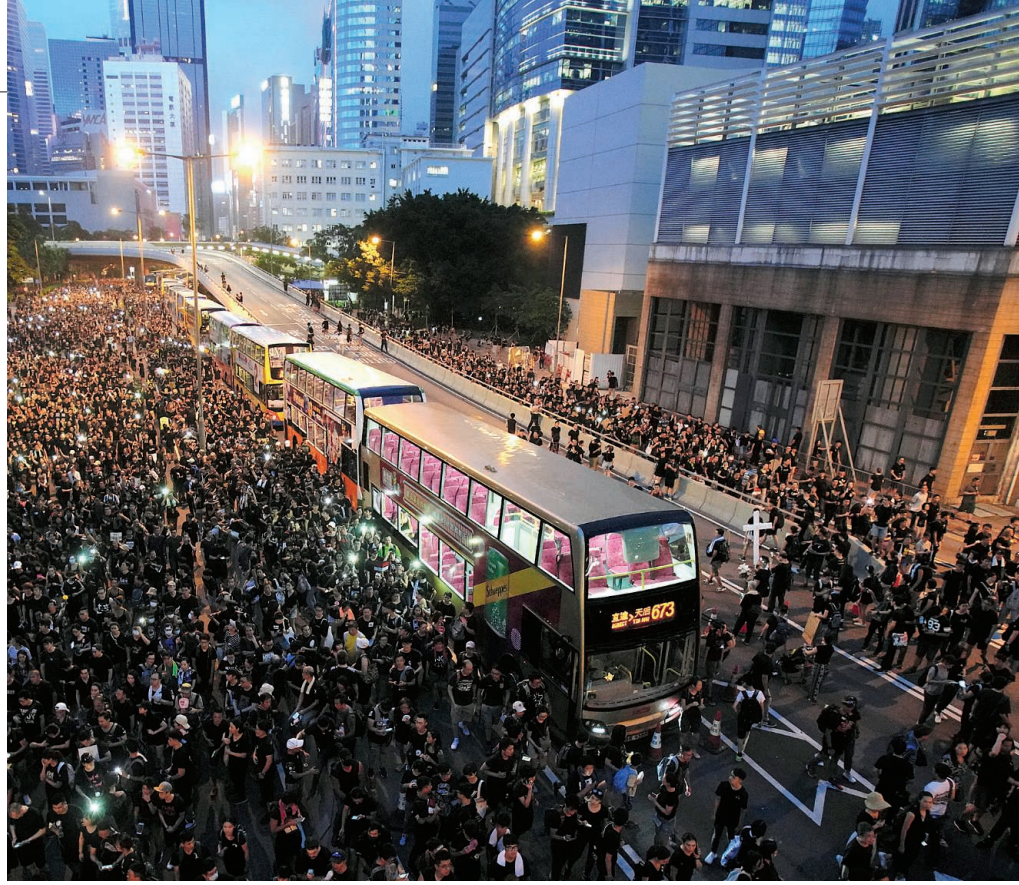


PHOTO: *THE ASAHI SHIMBUN*/GETTY IMAGES

CHINA

Hong Kong scientists protest, but also forge mainland ties

Cross-border research collaborations expand, even as conflict continues over city's semiautonomous status

By **Dennis Normile**

After a series of massive protests by Hong Kong's residents, including many academics, the leaders of the semiautonomous Chinese city last week shelved controversial legislation that would have allowed people there to be extradited to mainland China. But even as that battle to preserve independence continues, Hong Kong's researchers are forging closer ties with the mainland.

Those links will be strengthened this year, with several new cross-border funding programs set to make their first awards. And although many researchers welcome the new opportunities for funding and collaboration, some worry they could give Beijing greater influence over Hong Kong's research agenda.

The tension arises from Hong Kong's special political status. In 1997, China regained control of the former U.K. colony under a "one country, two systems" policy that gives Hong Kong's 7.4 million residents a greater say in their economic and political affairs. Academic efforts have thrived under the arrangement. The city now hosts nearly 30,000 researchers, creating a per capita ratio triple that found on China's mainland, according to United Nations statistics. Hong Kong's research spending has risen from just 0.4% of its gross domestic product in 1998 to 0.8% in 2017. Several of the city's universities are among the top 50 in the world, according to this year's *Times Higher Education* rankings.

Research ties with mainland China have grown since the handover. In 1998, 16.5% of all scientific papers produced in Hong Kong involved collaborations with the mainland; by 2017, the share had jumped to 53.2%, information scientists Ma Qian and Li Wenlan of Tianjin University in China reported in September 2018 in *Scientometrics*.

Funding ties are also deepening. Hong Kong researchers have long been able to win grants from the Chinese government, but the money had to be spent within the mainland. Last year, however, the government dropped

Two million people, a quarter of Hong Kong, China's residents, joined protests against an extradition bill.

that requirement at the request of prominent Hong Kong scientists. In the first grants under the new policy, China's Ministry of Science and Technology (MOST) in May 2018 awarded 22 million Chinese yuan ($3.2 million) to 22 Hong Kong research groups.

Two new cross-border programs will announce their first grants later this year. One is backed by MOST and Hong Kong's Innovation and Technology Bureau, and the other by Hong Kong's Research Grants Council and China's National Natural Science Foundation. The latter will focus on six research areas, including medicine and materials science.

Even as Hong Kong has strengthened scientific ties with the mainland, questions about the durability of the city's special status have grown. The current protests began soon after the extradition bill was unveiled in April. Opponents said it would enable mainland authorities to seize political opponents on flimsy charges. And more than 1700 academics from around the world voiced support for their Hong Kong colleagues by signing an online petition warning that the bill was "jeopardizing the rule of law and human rights in Hong Kong." On 15 June, Hong Kong officials "indefinitely" postponed action on the bill.

The bill was "definitely a concern for academics," because it could have had "a chilling effect on people working on so-called 'sensitive' topics," says philosopher Timothy O'Leary of the University of New South Wales in Sydney, Australia, who taught at the University of Hong Kong (HKU) for 17 years. Some scientists also worried the law would hamper recruitment efforts.

Authorities on both sides argue cross-border collaborations advance science and help Hong Kong become an innovation hub. But such schemes are also "a very useful mechanism" for integrating Hong Kong into China, notes one senior Hong Kong scientist, who asked to remain anonymous because of the issue's sensitivity. Even so, he says it would be a leap from tighter integration "to Hong Kong institutions losing their autonomy."

Other researchers are even more sanguine. HKU microbiologist Yuen Kwok-Yung doubts "such additional funding will erode academic freedom in [Hong Kong] as long as … the independent judiciary and free press are still being protected." And O'Leary says the recent protests show that Hongkongers "will not easily acquiesce to an encroachment on their civil liberties." But he urges the city's universities to follow the protesters' lead "and continue to insist on the nonnegotiable importance of academic freedom." ■

IMAGES: ALBERT HSIAO AND BRIAN HURT/UNIVERSITY OF CALIFORNIA SAN DIEGO AIDA LABORATORY

CLINICAL RESEARCH

Medicine contends with how to use artificial intelligence

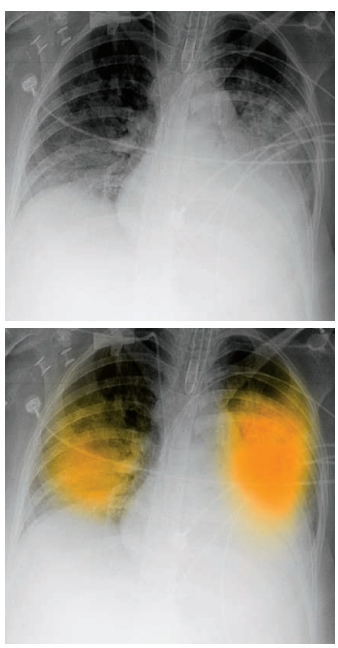
Barriers include lack of reproducibility across hospitals and populations

By **Jennifer Couzin-Frankel**

Artificial intelligence (AI) is poised to upend the practice of medicine, boosting the efficiency and accuracy of diagnosis in specialties that rely on images, such as radiology and pathology. But as the technology gallops ahead, researchers and physicians are grappling with its potential downsides. "Just working with the technology, I see lots of ways it can fail," says Albert Hsiao, a radiologist at the University of California, San Diego, who has developed algorithms for reading cardiac images and improving their quality. One major concern: Most AI software is designed and tested in one hospital, and it risks faltering when transferred to another.

U.S. government scientists, regulators, and doctors last month published a road map describing how to convert research-based AI into improved medical imaging for patients. Among other things, the commentary in the *Journal of the American College of Radiology* urged more collaboration across disciplines in building and testing AI algorithms, and intensive validation of algorithms before they reach patients. For now, Hsiao says, "I would want a human physician no matter what," even if a machine hums alongside.

Most AI in medicine is used in research, but regulators have already approved some algorithms for radiologists. Physicians are also developing their own—which they're permitted to use without regulatory approval as long as companies aren't marketing the new technology. Studies are testing algorithms to read x-rays, detect brain bleeds, pinpoint tumors, and more.



In a chest x-ray (top), artificial intelligence marks a likely infection (bottom).

The algorithms learn as scientists feed them hundreds or thousands of images—of mammograms, for example—training the technology to recognize patterns faster and more accurately than a human could. "If I'm doing an MRI of a moving heart, I can have the computer predict where the heart's going to be in the next fraction of a second and get a better picture instead of a blurry" one, says Krishna Kandarpa, a cardiovascular and interventional radiologist at the National Institute of Biomedical Imaging and Bioengineering in Bethesda, Maryland.

Or AI might analyze computerized tomography head scans of suspected strokes, label those more likely to harbor a brain bleed, and put them on top of the pile for the radiologist to examine. An algorithm could help spot breast tumors in mammograms that a radiologist's eyes risk missing.

But Eric Oermann, a neurosurgeon at Mount Sinai Hospital in New York City, has explored one downside of the algorithms: The signals they recognize can have less to do with disease than with other patient characteristics, the brand of MRI machine, or even how a scanner is angled. With colleagues, Oermann developed a mathematical model for detecting patterns on chest x-rays consistent with pneumonia and trained it with images from patients at Mount Sinai. The hospital has a busy intensive care unit with many elderly people, who are often admitted with pneumonia; 34% of the Mount Sinai x-rays came from infected patients.

When the algorithm was tested on a different batch of Mount Sinai x-rays it performed admirably, accurately detecting pneumonia 93% of the time. But

Oermann also tested it on tens of thousands of patient images from two other sites: the National Institutes of Health Clinical Center in Bethesda and the Indiana Network for Patient Care. With x-rays from those locations—where pneumonia rates just squeaked past 1%—the success rate fell, ranging from 73% to 80%, the team reported last year in *PLOS Medicine*. "It didn't work as well because the patients at the other hospitals were different," Oermann says.

At Mount Sinai, many of the infected patients were too sick to get out of bed, and so doctors used a portable chest x-ray machine. Portable x-ray images look very different from those created when a patient is standing up. Because of what it learned from Mount Sinai's x-rays, the algorithm began to associate a portable x-ray with illness. It also anticipated a high rate of pneumonia, boosting misdiagnoses.

> ***"Just working with the technology, I see lots of ways it can fail."***
>
> **Albert Hsiao**, University of California, San Diego

Few multisite studies like Oermann's have been published, and last month's road map deemed this worrying. This year, a South Korean team reported in the *Korean Journal of Radiology* an analysis of 516 studies of AI algorithms designed to interpret medical images. The authors found that just 6% of the studies tested their algorithm at more than one hospital. "It's very concerning," says Elaine Nsoesie, a computational epidemiologist at Boston University. Even the brand of scanner matters, as the pixel pattern can vary, disrupting how AI assesses the image.

One way to avoid this pitfall, Nsoesie says, is to test an algorithm using data from several hospitals. Researchers are starting to do this, she says, "but less than you would think." A rare example is an algorithm first trained and tested on data at Stanford University's Lucile Packard Children's Hospital in Palo Alto, California, and Children's Hospital Colorado in Aurora. It's now undergoing testing in a clinical trial on scans from nine sites. The software measures skeletal maturity in hand x-rays, which orthopedists use to guide treatment for growth disorders in children and teenagers.

In pathology, another field that relies on images, Jeroen van der Laak, a computer scientist at Radboud University Medical Center in the Netherlands, tried a new way to encourage researchers to test their algorithms across hospitals: a competition. In 2015, van der Laak gathered and digitized 400 lymph node slides from breast cancer patients at two Dutch centers. Then, he invited all comers to train their algorithm on 270 of those slides and test it on the remaining 130, to see whether it could do better than pathologists hunting for tiny cancers. Twenty-three teams submitted 32 algorithms.

The results, published in 2017 in *JAMA*, showed that 10 matched or exceeded a panel of 11 pathologists. The top-performing algorithm, from a group at Harvard University and the Massachusetts Institute of Technology in Cambridge, matched a pathologist who took an entire weekend to go over 130 slides.

"To actually see that you could be as good as a pathologist [was] a shock," van der Laak says. He and others say that to achieve that kind of accuracy, AI algorithms should train on data that are diverse not only in their hospital of origin, but also in race and geography, because disease can manifest differently across populations.

The U.S. Food and Drug Administration (FDA) continues to weigh how to assess algorithms for patient care. The agency considers "locked" AI software, which is unchanging, as a medical device. But earlier this year, it announced it was developing a framework for regulating more cutting-edge AI software that continues to learn over time. Still, there "are major questions everyone is struggling with" around regulation, says Hugo Aerts, who directs the computational imaging and bioinformatics laboratory at the Dana-Farber Cancer Institute in Boston. What if developers update an algorithm that works in 96% of cases to achieve a 99% success rate; do they have to go through the regulatory process again? What if an approved algorithm is applied to a patient population it wasn't originally tested on?

FDA has already issued some approvals. One algorithm, created by Hsiao, measures heart size and blood flow in a cardiac MRI. Hsiao was frustrated that analyzing the data by hand took several hours, so he returned to his roots as a computer science major and wrote his own software. He subsequently formed a company, Arterys, based in San Francisco, California, and won FDA approval in about 6 months, he says. Hsiao is now working on algorithms that help identify pneumonia infections in the lungs.

But, he says, the doctor, not the machine, is still the boss, and entitled to override the technology. "If I think it's not pneumonia," Hsiao says, "it's not." ■

OCEAN SCIENCE

Fight for the Arctic Ocean is a boon for science

Armed with seafloor maps, Russia, Denmark, and Canada seek control over the North Pole

By **Richard Kemeny**

A competition for the North Pole heated up last month, as Canada became the third country to claim—based on extensive scientific data—that it should have sovereignty over a large swath of the Arctic Ocean, including the pole.

Canada's bid, submitted to the United Nations's Commission on the Limits of the Continental Shelf (CLCS) on 23 May, joins competing claims from Russia and Denmark. Like theirs, it is motivated by the prospect of mineral riches: the large oil reserves believed to lie under the Arctic Ocean, which will become more accessible as the polar ice retreats. And all three claims, along with dozens of similar claims in other oceans, rest on extensive seafloor mapping, which has proved to be a boon to science, whatever the outcome for individual countries. The race to obtain control over parts of the sea floor has "dramatically changed our understanding of the oceans," says marine geophysicist Larry Mayer of the University of New Hampshire in Durham.

Coastal nations have sovereign rights over an exclusive economic zone (EEZ), extending by definition 200 nautical miles (370 kilometers) out from their coastline. But the 1982 United Nations Convention on the Law of the Sea opened up the possibility of expanding that zone if a country can convince CLCS that its continental shelf extends beyond the EEZ's limits.

Most of the 84 submissions so far were driven by the prospect of oil and gas, although advances in deep-sea mining technology have added new reasons to apply. Brazil, for example, filed an application in December 2018 that included the Rio Grande Rise, a deep-ocean mountain range 1500 kilometers southeast of Rio De Janeiro that's covered in cobalt-rich ferromanganese crusts.

Canadian and U.S. Coast Guard ships worked together to map the Arctic sea floor for continental shelf claims.

To make a claim, a country has to submit detailed data on the shape of the sea floor and on its sediment, which is thicker on the shelf than in the deep ocean. The data come from sonar, which reveals seafloor topography, and seismic profiling, which uses low-frequency booms to probe the sediment. Canada's bid also enlisted ships to conduct high-resolution gravimetry—measurements of gravity anomalies that reveal seafloor structure. Elevated gravity readings are found over higher-density mantle rocks found in oceanic crust, and lower readings over lighter, continental structures. And the bid used analyses of 800 kilograms of rock samples dredged up from the sea floor, whose composition can distinguish continental from ocean crust.

The studies don't come cheap; Canada's 17 Arctic expeditions alone cost more than CA$117 million. But the work by the three countries vying for the Arctic—and that of dozens of others elsewhere in the world—has been a bonanza for oceanography. In the Arctic alone, the mapping has revealed several sunken mountains, previously missed or undetected by older sonar methods. Hundreds of pockmarks found on the Chukchi Cap, a submarine plateau extending out from Alaska, suggest that bursts of previously frozen methane have erupted from the seabed, a phenomenon that could accelerate climate change. And gaps discovered across submarine ridges allow currents to flow from basin to basin, with "important ramifications on the distribution of heat in the Arctic and on overall modeling of climate and ice melting," Mayer says.

CLCS, composed of 21 scientists in fields such as geology and hydrography who are elected by member states, has accepted 24 of the 28 claims it has finished evaluating, some partially or with caveats; in several cases, it has asked for follow-up submissions with more data. Australia was the first country to succeed, adding 2.5 million square kilometers to its territory in 2008. New Zealand gained undersea territory six times larger than its terrestrial area. But CLCS only judges the merit of each individual scientific claim; it has no authority to decide boundaries when claims overlap. To do that, countries have to turn to diplomatic channels once the science is settled.

The three claims on the North Pole revolve around the Lomonosov Ridge, an underwater mountain system that runs from Ellesmere Island in Canada's Qikiqtaaluk region to the New Siberian Islands of Russia, passing the North Pole. Both countries claim the ridge is geologically connected to their continent, whereas Denmark says it is also tied to Greenland, a Danish territory. As the ridge is thought to be continental crust, the territorial extensions could be extensive. (U.S. scientists should finish mapping in the Arctic in about 2 years, says Mayer, who is involved in that effort, but as one of the few countries that hasn't ratified the Law of the Sea convention, the United States can't file an official submission.)

Tensions flared when Russia planted a titanium flag on the sea floor beneath the North Pole in 2007, after CLCS rejected its first claim, saying more data were needed. The Canadian foreign minister at the time likened the move to the land grabs of early European colonizers. Not that the North Pole has any material value: "The oil potential there is zip," says geologist Henry Dick of the Woods Hole Oceanographic Institution in Massachusetts. "The real fight is over the Amerasian Basin," Dick says (see map, below) where large amounts of oil are thought to be locked up.

It will take years, perhaps decades, for CLCS to rule on the overlapping Arctic claims. Whoever wins the scientific contest still faces a diplomatic struggle.

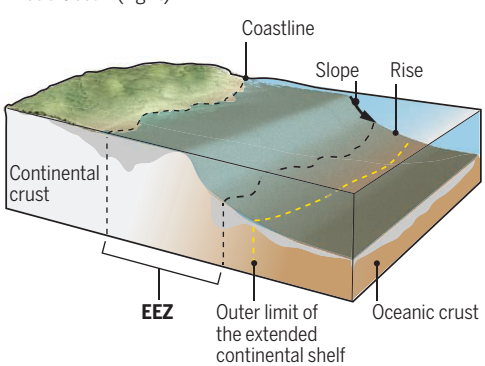
Denmark, Russia, and Canada have expressed their desire to settle the situation peacefully. "Russia actually has played nice on this and stopped at the North Pole," rather than extending its claim along the length of the ridge, says Philip Steinberg, a political geographer at Durham University in the United Kingdom. Denmark had no such qualms. and put in a claim up to the edge of Russia's EEZ, "even though there's no way in hell they'll get that," when it comes to the diplomatic discussions, Steinberg says.

One solution would be to use the equidistance principle, by drawing a median line between the coastlines, as has been done when proposed marine territories overlapped in the past; doing so would mean the North Pole falls to Denmark. There's also a proposal to make the pole international, like Antarctica, as a sign of peace, says Oran Young, a political scientist at the University of California, Santa Barbara. "It seems a very sensible idea." ■

Richard Kemeny is a science journalist in São Paulo, Brazil.

Who owns the North Pole?

Countries can claim the sea floor beyond the 200-nautical-mile (370-kilometer) exclusive economic zone (EEZ) if data show it to be an extension of the continental shelf (below). Russia, Denmark, and Canada have submitted overlapping claims in the Arctic Ocean (right).



CREDITS: (PHOTO) DVIDSHUB/FLICKR/CC BY; (GRAPHICS) N. DESAI/*SCIENCE*

A woman gets vaccinated at a Paris hospital.

PHOTO: PHANIE/ALAMY STOCK PHOTO

SOCIOLOGY OF SCIENCE

France is wary of science and vaccines, global survey finds

Researchers attribute suspicion to institutional scandals

By **Tania Rabesandratana**

Quelle surprise. France, cradle of the Concorde, the face transplant, and the first isolation of HIV, is more wary of vaccines and the economic value of science than more than 140 other countries, according to a global survey of public attitudes toward science and health released this week. But French scientists say the skepticism is familiar and doesn't affect their work; some suggest it instead reflects a deep-seated mistrust of institutions. "We think there's a problem of trust in government, in particular in health authorities," says Pierre Verger, an epidemiologist who studies vaccine hesitancy at the French biomedical research institute INSERM in Marseille.

When asked whether vaccines are safe, one-third of the 1000 French respondents to the survey disagreed—far more than in other nations. (In the United States, 11% disagreed.) The mistrust didn't vary much across age, gender, or education, according to the survey, conducted by the Gallup World Poll for the Wellcome Trust, a biomedical charity based in London.

Verger's research points to high-profile health scandals as an explanation. In 2009, for example, years after other countries, French regulators withdrew approval for Mediator, an amphetamine-based diabetes drug linked to hundreds of deaths, amid concerns of undue industry influence. The French have also been unhappy with heavy-handed public health campaigns, such as a costly mass vaccination for swine flu that same year. "France is one of a few countries where [such controversies] have been so frequent," Verger says.

The survey also reveals French pessimism about the economic value of science. Some 55% say they see science and technology as a threat to local jobs in the next 5 years. Although France is the only country scoring above 50% on this question, a similar gloominess spans other parts of Europe, whereas most regions in Africa and Asia are optimistic about science boosting job prospects. The survey suggests France's sluggish economy and relatively high unemployment as plausible causes. That fear about the future is understandable, says Catherine Pélachaud, an artificial intelligence (AI) researcher at Sorbonne University in Paris. "We see factories closing down and it can be very hard for less-qualified people to adapt."

It might be unfair to label France as averse to science based on a single survey. There is plenty of enthusiasm for specific technologies, according to a 2017 OpinionWay survey of 1059 people in France commissioned by Quattrocento, a scientific business incubator in Paris. It found that more than three-quarters of respondents were hopeful about research in transport and renewable energies. At the same time, about two-thirds were worried about nuclear energy research and studies of genetically modified (GM) foods. The French government, a global leader in producing nuclear energy and exporting its technology, has largely ignored public dissent in that arena, but it has taken up public suspicions about GM foods by, for example, banning the cultivation of GM maize.

Some French scientists are unsurprised by the survey results, and point out that opinions don't always correlate to behavior. Of the French parents surveyed, 91% say their children are vaccinated, in line with the global average of 92%. "It's the French paradox: We have doubts about many things; we grumble. But thankfully, vaccine coverage remains high," says Olivier Schwartz, scientific director of the Pasteur Institute in Paris, which depends on public donations to carry out its work, including vaccine research. "I don't perceive a hostile climate," Schwartz adds. "On the contrary, I feel that [people in France have] a thirst for knowledge."

The survey data back this up: Some 71% of respondents in France say they know "some" or "a lot" about science—placing France in the top 10 globally. And 46% say they sought scientific information in the past month—compared with a median of 30% worldwide. Schwartz attributes the French vaccine skepticism to a lack of information, and says researchers and institutions need to fill that gap in a "simple and rigorous way."

But Brice Laurent, a sociologist at the Center for the Sociology of Innovation in Paris, warns against this "deficit model" of communication, which implies that ignorance breeds skepticism and that people would embrace technologies if only they knew enough science. Studies show that more informed people are often more skeptical, he notes.

"It would be annoying if decision-makers saw this [survey] and thought, 'French people just don't get it, so we'll repeat the message,'" Laurent cautions. He believes skepticism is ingrained in the culture of France, a place where all high schoolers are taught philosophy. "You could look at these stats and say: The French are critical thinkers; they are interested and ask questions," he says.

Pélachaud agrees that a vigilant public can help push scientists to consider societal impacts. "When I started 30 years ago, we didn't ask ourselves ethical questions" in the AI field, says Pélachaud, who works on animated chatbots capable of nonverbal communication. "The French may be grumblers, but their critical thinking is important." ■

Just say '*non*'

When asked whether vaccines are safe, a survey finds French people disagree the most.

COUNTRY	DISAGREE
France	33%
Gabon	26%
Togo	25%
Russia	24%
Switzerland	22%
Armenia	21%
Austria	21%
Belgium	21%
Iceland	21%
Burkina Faso	20%
Haiti	20%

STRUCTURAL BIOLOGY

Mutant power resolves protein shapes

Clever calculations on mutated proteins reveal spatial arrangement of amino acids

By **Robert F. Service**

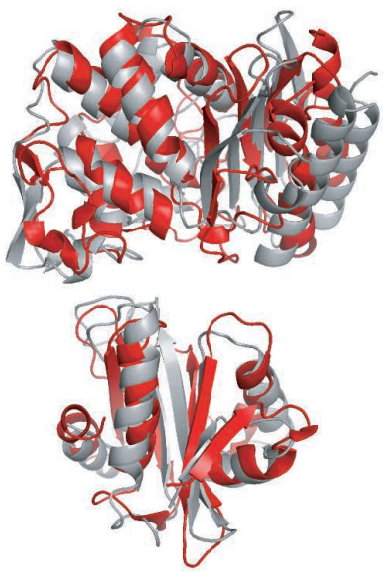
Mapping the atomic structure of proteins is crucial to understanding how they behave, but it's painstaking work that typically requires dedicated, expensive facilities with supercooled, powerful magnets or stadium-size synchrotrons. Now, two research teams independently report this week that they've hit on a way to use genetic and biochemical techniques to do the job, potentially opening structural biology to many more labs. And unlike traditional methods, which visualize proteins in crystals or solution, the approach can also reveal proteins' natural shape inside cells as they do their work, which could uncover how mutations that disrupt protein function lead to disease.

"It's fantastic," says Douglas Fowler, a genome scientist at the University of Washington in Seattle. Fowler notes the new approach doesn't offer the full atomic map that standard approaches do, but the general shape it provides for a protein nonetheless offers "extremely valuable" clues to its function. Plus, he says, "It could have a really big impact" on efforts to determine structures of proteins that are membrane-bound or part of large complexes, both of which are difficult to study with standard methods.

Today's most common approach for determining a protein's structure requires coaxing millions of copies of a protein into an ordered crystal, blasting it with x-rays, and tracking their ricochets to reveal the identity and position of each atom. Alternative approaches, including nuclear magnetic resonance spectroscopy and cryoelectron microscopy, also require large amounts of a protein and can take months. Researchers have unraveled the structures of only about 150,000 of the millions of proteins thought to exist—and the process has taken decades.

Some have tried to speed things up by predicting the most likely shape of a protein simply from its sequence of amino acids and probable interactions between atoms. The accuracy of such computational efforts typically lags behind experimental methods. One recent trick to improve matters has been to compare the same protein in multiple species to find pairs of amino acids that have evolved together even though they are far apart on the protein's linear sequence. That's a strong indication the two are close together and interact within the folded 3D molecule (*Science*, 22 July 2016, p. 338). But this approach works only if researchers can identify proteins that are shared by many organisms but different enough across them to identify multiple pairs of amino acids evolving in tandem, says Debora Marks, a systems biologist at Harvard Medical School in Boston.

Now, separate teams led by Marks and Ben Lehner, a geneticist at the Barcelona Institute of Science and Technology in Spain, have bypassed the need for evolution's help.



Genetic tests have revealed protein structures (red) with nearly the accuracy of x-ray structures (gray).

They independently hit on the idea of finding interacting amino acids within a protein by systematically mutating each amino acid and tracking how the changes alter the protein's function, such as an ability to bind to another molecule. Both groups built on work on a bacterial protein fragment called GB1 by a team led by Ren Sun, a systems biologist at the University of California, Los Angeles. In 2014, Sun's team reported creating more than half a million copies of the *GB1* gene, each with one or two of its 56 amino acids changed. For the so-called single mutants, the researchers systematically swapped every amino acid for one of the 19 other options. In the double mutants, they changed pairs of amino acids, working through nearly all possible combinations. After growing bacteria with these mutant genes and isolating the proteins, Sun's team determined the importance of GB1's amino acids by seeing which mutants bound most tightly to its natural target, human immunoglobulin G antibodies.

Marks's and Lehner's groups realized they could combine the binding data of the single and double mutants to determine which amino acids interact most strongly and are therefore likely sit next to each other in the protein's 3D structure. "Sometimes we see mutations that combine to have a much more dramatic effect," Fowler says. By tracking dozens of such occurrences and feeding the results into a structure prediction program, the teams computed the shape of GB1's main backbone to within a few angstroms of the resolution of the already known experimental x-ray structure.

The teams, who report their success this week in *Nature Genetics*, also showed that their technique worked with other small proteins and an RNA with analogous available data. Fowler notes the same approach may be more difficult for proteins with hundreds or thousands of amino acids, because the number of mutant proteins that must be made increases exponentially as the proteins grow. But Marks is optimistic: Early indications suggest the technique can solve structures with only a fraction of all possible mutants. The approach could even work using measures of stability for proteins that lack known binding partners, Lehner's team says.

The method has other strengths. In March, Lehner and his team posted a preprint on the bioRxiv server describing its use to learn how an RNA-binding protein called TDP-43 may cause the neurodegenerative disease amyotrophic lateral sclerosis (ALS). Insoluble aggregates of TDP-43 have been seen in neurons in the autopsied brains of many ALS patients, but the deposits may not cause the disease. So Lehner and his colleagues made 50,000 mutants of TDP-43 and tracked their toxicity in yeast cells. They found that mutant forms that aggregated were less toxic than other versions of the protein. "This is the exact opposite of what we expected," Lehner says, cautioning that they need to confirm this result in mammalian cells. Either way, he says, it shows that scanning mutations by the thousands may offer new insights into both proteins themselves and how their structures affect health in living cells. ■

IMAGES: M. STIFFLER *ET AL.*/BIORXIV. HTTPS://DOI.ORG/10.1101/667790

Published by AAAS

LOST AT SEA

A growing sensory smog threatens the ability of fish to communicate, navigate, and survive

By **Elizabeth Preston**; *Illustrations by* **Valerie Altounian**

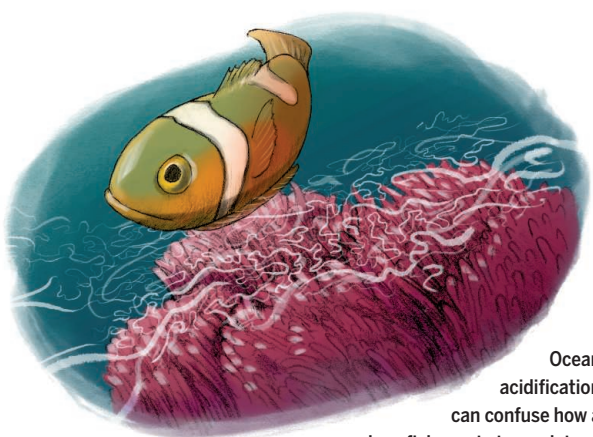
Try, for a moment, to be a fish. As you swim through dim waters, you see shapes moving past and watch for threats. You hear other animals calling or producing rasps and crackles by scraping together rigid body parts. The water is a tapestry of smells that reveals predators and potential mates, food, and the route home.

Now, imagine that nothing makes sense—the tapestry has unraveled. Smells still reach you, but their meanings are muddled. You listen for calls from your kin, but all you hear is the roar of a passing boat. You can't tell whether that looming shadow is a friend or foe.

When many people think of threats to the world's fish, overfishing or vanishing reefs might leap to mind. Increasingly, however, scientists also worry about a subtler danger: how human activities might interfere with the senses fish use to perceive the world. Noise from ships and construction, murkier waters caused by pollution, and rising ocean acidification from the buildup of atmospheric carbon dioxide (CO_2) are all possible culprits. In laboratories and in the wild, scientists study exactly how those factors might affect a fish's ability to communicate, navigate, and survive.

The studies face both logistical and conceptual challenges. Observing the behavior of fish in the vast sea is nearly impossible, but a laboratory aquarium is a far cry from their natural environment. And we can't know exactly what seeing, smelling, or hearing as a fish is like. But by drawing on tools as elaborate as simulated underwater environments and as simple as bits of thread tethering baby fish to stream bottoms, researchers are gaining a better understanding of how fish use their senses—as well as the consequences of disrupting them.

Driving the work is the concern that humans are creating a pervasive threat to the fish stocks that provide food and livelihoods for millions of people: a kind of sensory smog. Combating causes such as pollution and acidification is a staggering challenge, the scientists admit. But the stakes are high. "The knock-on implications are huge," says Jennifer Kelley, a behavioral ecologist at the University of Western Australia in Perth. When fish with compromised senses settle in the wrong homes or fail to recognize predators, the results could ripple outward "to how individuals interact and how communities operate, and the whole ecology of the system altering."



Ocean acidification can confuse how a clownfish reacts to predators.

AN 11-DAY-OLD CLOWNFISH, pale orange and about as long as a grain of rice, searches for a place to settle down on a reef. Its keen sense of smell helps it both navigate to a safe home and steer away from the mouths of predators.

In the wild, clownfish inhabit living coral reefs. But in behavioral ecologist Danielle Dixson's laboratory at the University of Delaware in Lewes, the habitats beckoning the fish are made mostly of wires. Dixson will use the experimental setup to study how ocean acidification could alter how fish perceive and respond to their world.

The laboratory—a converted garage with black bags taped over the windows to block bright light—holds two bays of fish tanks. A network of densely draped cables connects sensors in the aquariums to a black box on the wall, which researchers call "the brain." It helps them monitor and control water temperatures and acidity levels in each tank.

In some tanks, Dixson will keep clownfish and other species in seawater with the acidity levels found now in the ocean. In other tanks, the water will be more acidic, to mimic the ocean chemistry that's forecast for later this century if humans do not curb CO_2 emissions. In both cases, conditions fluctuate during the day, as on a real reef. The researchers will look at how pH levels affect the way fish behave, interact, and respond to olfactory cues.

Dixson is building on research showing that acidification can jumble a fish's response to smells. For example, she and colleagues reported in *Ecology Letters* in 2010 that young clownfish raised in more acidic waters were strongly attracted to the smell of a predator, which the fish would normally avoid. In some predatory fish, acidification had the opposite effect. Dixson and coauthors found that after 5 days in water with high levels of CO_2, sharks called smooth dogfish avoided the smell of their prey. Other researchers found a similar

result in the brown dottyback, a reef-dwelling predator.

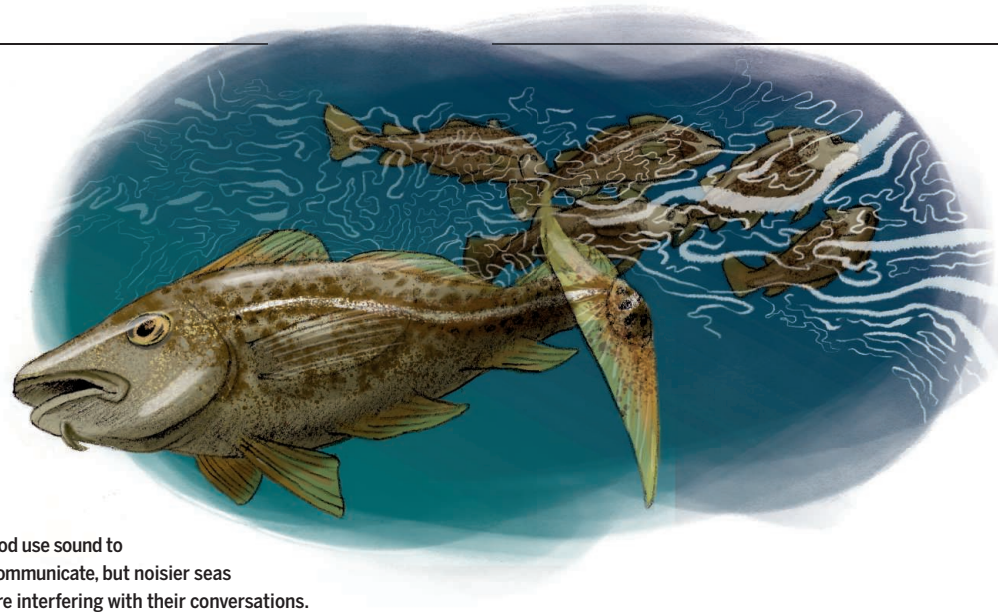
Acidification also seems to cause other behavioral changes. In boldness tests—in which researchers approach fish with a blunt object to make them retreat—fish treated with acidic water come back out of hiding sooner. Those fish are like angry teenagers, Dixson says. "They're really aggressive, but they really don't know what they're doing."

The problem isn't that acidified water damages fish noses, Dixson says. Instead, the issue is apparently in the brain. Multiple mechanisms may be at work. One strong possibility is that acidic conditions interfere with brain cell receptors that respond to γ-aminobutyric acid (GABA), a neurotransmitter. All vertebrates share GABA receptors, which inhibit neuron activity. They are "like the brake on electrical circuits in the brain," says biologist Göran Nilsson of the University of Oslo who, along with Dixson and others, first highlighted the potential connection between GABA receptors, fish behavior, and acidification in a 2012 *Nature Climate Change* paper.

Normally, when GABA binds to a receptor, it opens a channel that lets negatively charged chloride ions flow into the neuron. But that flow reverses in fish living in acidified water, studies suggest. That's because, physiologically, the fish cope with the more acidic environment by accumulating bicarbonate ions, which are basic, and by shedding chloride ions. When the channel is opened, chloride ions flow out of the neuron instead of into it, and the neuronal "brake actually becomes an accelerator," Nilsson says. That hypothesis has drawn support from studies in which researchers dose fish with gabazine, a drug that suppresses GABA receptors. After a quick dip in gabazine-laced water, formerly confused fish act normal again.

Acidification similarly affects fish vision and hearing, perhaps also by scrambling GABA receptors. In one study, researchers treated young damselfish with high-CO_2 water and then placed a plastic bag holding a predatory fish in their tank—so the damselfish could see the predator, but not smell it. Normally, that apparition would make the damselfish lie low. Instead, fish from acidified water ignored or swam close to the bag. In another study, researchers used an underwater speaker to play young clownfish a recording of a coral reef teeming with predators. Clownfish raised in high-CO_2 water didn't flee from the sound, as they normally would.

Other researchers have found that otoliths, the little chunks of calcium carbonate in the inner ears of vertebrates, are larger than normal in some (but not all) of the fish species they raised in acidified waters. Changes to otolith size could affect fish hearing and orientation, researchers say.



Cod use sound to communicate, but noisier seas are interfering with their conversations.

ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

Different fish species are vulnerable to different degrees of ocean acidification. In one study, researchers found that conditions mimicking an atmospheric CO_2 level of 700 parts per million (ppm) addled about half of clownfish; all the fish were affected at 850 ppm. At current emission rates, fish populations could experience those levels well before the end of the century, leaving little time to adapt. "It doesn't seem like there's a lot of wiggle room for evolution or natural selection to take place," Dixson says.

IT'S EARLY SPRING in the Atlantic Ocean, and an adult cod is journeying back to his preferred spawning grounds to find a mate. He grunts while he swims. He and the other migrating cod add their grunts to a marine cacophony: the barks, mutters, clucks, and chirps of other fish; the singing of whales; the steady munching of sea urchins below. Light rain above tinkles musically, reaching a crescendo when heavier storms pass. The noises reverberate in the cod's belly, where his balloonlike swim bladder acts as an extension of his ears.

What cod or other migrating fish are saying as they travel is not clear. They may be calling to stragglers, synchronizing migration or spawning, or establishing dominance. But clearly, the fish increasingly must compete with human noise sources such as recreational boats, commercial ships, pile driving, sonar, and deep-sea mining.

Marine ecologist Jenni Stanley at the National Oceanic and Atmospheric Administration's Northeast Fisheries Science Center and the Woods Hole Oceanographic Institution in Massachusetts has measured the impact of the added noise. She and colleagues recorded ambient underwater sounds at spawning sites for cod and haddock in Massachusetts Bay. Unlike the grunting cod, haddock produce an insistent knocking sound that can go on for hours. At times with less competing noise from vessels, the fish calls could carry more than 20 meters, the researchers reported in 2017 in *Scientific Reports*. But at noisier times, the calls carried only a meter or so before being drowned out.

Researchers don't know whether such interference has affected overall populations of cod and haddock. But scientists worry about subtle harms that could ultimately take a toll. "I think the critical issue … is the impact on the soundscape," says biologist Arthur Popper of the University of Maryland in College Park. At least 800 fish species make sounds. But "even those that don't communicate with sound are still listening to their environment," Popper says. They use sound to identify predators, for example, or suitable habitat.

A noisier world could therefore have "potentially dire consequences" for fish, warned the authors of a 2018 meta-analysis, published in *Global Change Biology*, which examined 42 pertinent studies. Biologist Kieran Cox of the University of Victoria in Canada and co-authors found that noise can hurt fishes' ability to find food, increase their risk of being eaten, and reduce their reproductive success.

Just how well many fish can hear in the first place is uncertain. "What we don't know is so gigantic," Popper says. Studying fish behavior in the ocean over long periods is challenging. But laboratory studies have their own limitations, Popper points out: Aquariums can change how sound waves travel, and confined fish may not behave as they would in the wild.

Scientists do know that fish are vulnerable to loud events. Popper's laboratory, for

example, used a large contraption to mimic the sounds created by a pile being hammered into the sea floor. The researchers found that as the pounding sound passes through a fish's body, the swim bladder can knock into other tissues hard enough to cause serious injuries. In the wild, fish deaths have been observed near areas of pile driving. Fish in the open ocean can often flee far from loud noises, Popper notes, but those in more constrained freshwater lakes and rivers likely can't escape the din.

THE BELLY OF A MALE three-spined stickleback, normally dull in color but stained red-orange for the breeding season, blushes even more deeply as he swims toward a female. He faces her and repeats a zigzagging dance, darting left and right. If he's lucky, the female follows him to his nest in the sand and lays her eggs.

"They're quite cute when they perform their courtship behavior!" says ecologist Ulrika Candolin of the University of Helsinki. Sticklebacks are widespread in the ocean and in lakes; the population she studies lives in the slightly salty Baltic Sea. In recent years, she says, nitrogen and phosphorus pollution has fertilized algal blooms so severe that, on warm days, they turn the Baltic's water green.

In waters thick with algae, both laboratory and field studies have shown that romance-minded sticklebacks "have more difficulties detecting each other," Candolin says. When a pair does connect, the male spends longer than usual performing his display. And the female spends more time inspecting the male, as though unsure of what she's seeing.

Such visual interference can harm sexual selection, Candolin and colleagues found. Stickleback mating becomes more random in murky water, with females having a greater chance of selecting a less fit male, they reported in a 2016 *Ecology* paper. The consequences include having fewer surviving offspring. (Candolin notes that smell might also be a factor in poor mate choices—the algal blooms may obscure scent information that female sticklebacks would normally use to judge males.)

So far, Candolin notes, the sloppier mating caused by murky waters isn't hurting stickleback populations in the Baltic. Their numbers are actually growing; a decline in predators, also due to ecosystem changes, might be a factor. But Candolin's work has helped highlight how visual pollution might take a toll in the future. "Females are wasting their time and energy" when they select suboptimal mates, says Grant Brown, an ecologist at Concordia University in Montreal, Canada, who says Candolin's work on mate choice is "very nicely done."

Another risk of not clearly seeing your partner is that you will mate with the wrong species. Africa's Lake Victoria, for example, hosts hundreds of species of fish called cichlids. Females choose mates on the basis of their distinctive bright colors. But the number of cichlid species has declined, researchers found in the 1990s. Meanwhile, runoff from deforestation and agriculture has fertilized and clouded the lake waters, causing females to struggle to recognize males from their own species. They seem to end up mating with other species, resulting in hybrid offspring that sport duller hues. Not only could hybridization take a toll on species diversity, Brown says, but the mismatched mates may produce offspring that can't compete or reproduce, harming populations.

Murky waters can complicate stickleback mating.

Polluted water may obscure other crucial signals. Brown, for example, has found in the lab that chemical pollutants appear to mask important molecules that fish use to sniff out danger. Fish release those molecules in their urine when disturbed; the odors warn other fish that a predator may be nearby. But high levels of nitrogenous pollutants, such as fertilizer runoff in streams, can overpower the warning scents. Brown says detecting the signal becomes "like trying to hear someone in a very crowded room with lots and lots of background noise."

Pollution can also chemically alter a different warning signal, he found: molecules that leak from fish skin when damaged, for example, by a predator's attack. Freshwater acidification, caused not by rising CO_2 levels, but by airborne sulfate and nitrate pollution, may render those molecules unrecognizable. To better understand the possible consequences, Brown and co-author Chris Elvidge, now of Carleton University in Ottawa, used meter-long threads to tether baby Atlantic salmon to stream bottoms. Normally, the young fish would respond to the smell of other fishes' alarm cues by dropping to shelter in the gravel. But when the researchers returned 6 hours later, the fish in more acidic streams, which were less able to detect those cues, were likelier to have vanished, apparently because they didn't duck from danger. Sometimes a full-bellied trout had taken the place of the young salmon on the end of the thread.

RESEARCHERS ARE NOT only documenting how sensory smog harms fish, but also searching for ways to protect them from it. Some marine reserves, for example, already ban certain activities, such as construction or exploration for oil. But as researchers learn more about aquatic soundscapes, they can imagine other ways of protecting fish from unwanted sound, such as by barring loud motors or industrial activities from key fish spawning grounds during certain seasons.

Larger global issues, such as ocean acidification and polluted runoff, could be harder to tackle. Even here, however, "There are rays of hope," Brown says. For instance, he found that fish can associate a given smell not only with risk, but with a certain habitat or time of day. If a predator hunts only at dawn or dusk, the prey fish may learn to ignore the smell of that predator at midday. In a world where some of their sensory cues are masked or changed, fish may find new cues and learn new rules to live by.

Fish may also be able to cope with some interference because they have multiple ways of sensing the world. For example, fish have lateral lines, a set of organs on the outside of the body that sense pressure and currents and help fish orient themselves. Some species can also sense magnetic fields. Sharks and their relatives can detect electricity. "In some cases, if one sense is blocked or unusable, then fish will just switch to another one," Kelley says. "That makes them very resilient in that context."

The rapid pace of environmental change, however, is testing the sensory resilience of fish as never before. The outcome of rising sensory smog "could be not as bad as we're anticipating," Dixson says. "Or it could be 1000 times worse." ■

Elizabeth Preston is a journalist in Arlington, Massachusetts.

ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

ECOLOGY

Do tiny fish rule the reefs?

Covert fish larvae may serve as crucial cuisine in coral reef ecosystems

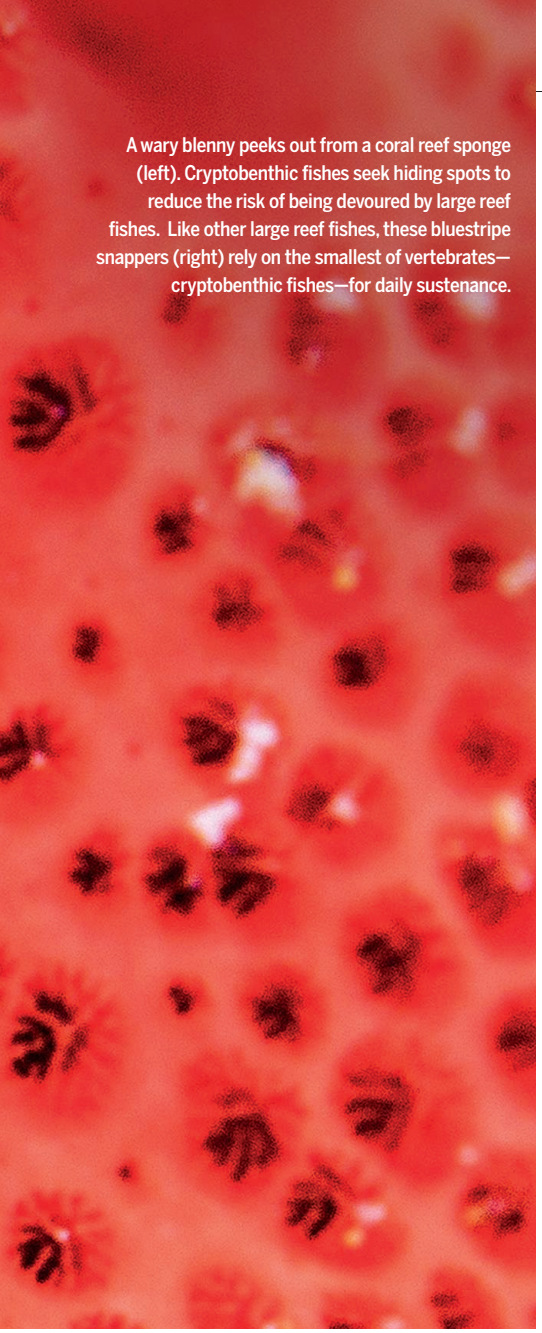
By **Cynthia Riginos**[1] *and* **Jeffrey M. Leis**[2,3]

Coral reef fishes are famous for their fantastic colors and forms. Easily overlooked are the cryptic and diminutive (cryptobenthic) bottom-dwelling fishes that also call coral reefs home. By linking empirical data and ecosystem modeling, Brandl *et al.* (*1*), on page 1189 of this issue, propose that pelagic (open-water) larvae of cryptobenthic fishes and their small juveniles that recently settled on reefs constitute a key food source for other reef residents. Such a scenario could help explain why coral reefs in nutrient-poor waters teem with life.

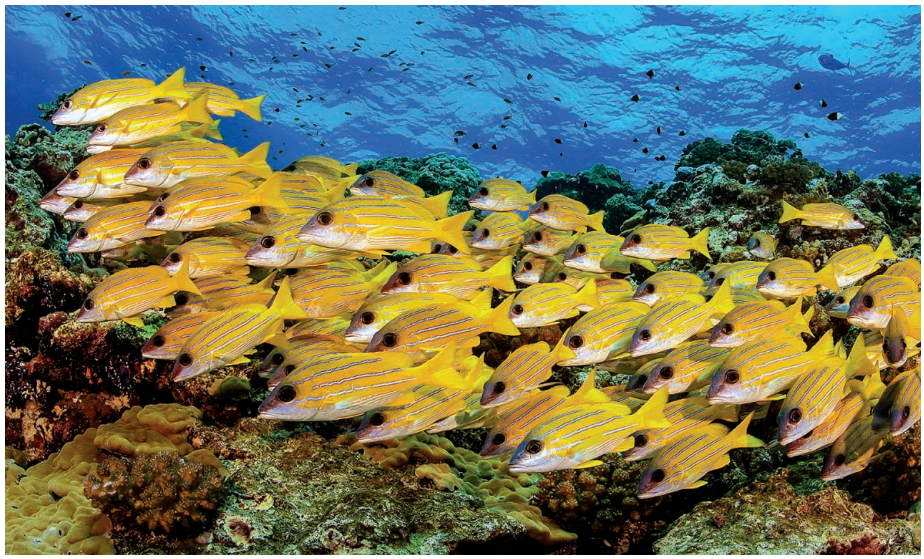
Small fishes live fast and die young (*2*). Although each fish might only spawn a few eggs at a time, Brandl *et al.* argue that the sheer number of spawning fishes at any time could provide a near-constant supply of pelagic larvae (see the figure). Pelagic larvae from cryptobenthic reef fishes (CRFs) tend to remain near reefs and lagoons (*3*, *4*) and thus are ready prey for reef residents. By contrast, large fishes live long and mature slowly. Although they produce more eggs per spawn than small fishes (*5*), large fishes are less numerous and tend to spawn seasonally, so the total number of pelagic larvae produced is smaller and these lar-

[1]*School of Biological Sciences, The University of Queensland, St Lucia, Australia.* [2]*Institute of Marine and Antarctic Studies, University of Tasmania, Hobart, Australia.* [3]*Ichthyology, Australian Museum Research Institute, Sydney, Australia. Email: c.riginos@uq.edu.au*

PHOTO: TANE SINCLAIR-TAYLOR

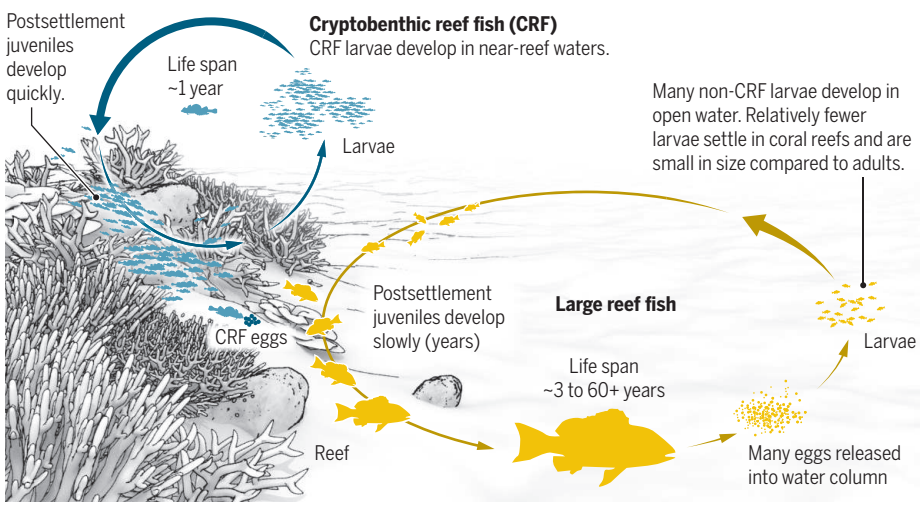


A wary blenny peeks out from a coral reef sponge (left). Cryptobenthic fishes seek hiding spots to reduce the risk of being devoured by large reef fishes. Like other large reef fishes, these bluestripe snappers (right) rely on the smallest of vertebrates—cryptobenthic fishes—for daily sustenance.



Distinct life histories shape coral reefs

CRFs produce few eggs, but because adults are abundant, pelagic larvae and postsettlement juveniles are plentiful and might constitute a generous food source. Large reef fishes each spawn many eggs, but there are fewer adult fishes. Thus, the amount of larvae is lower and provides less of a food resource.



vae develop offshore more often relative to CRFs. The net effect of these contrasting life histories, Brandl *et al.* suggest, is that larvae of CRFs greatly outnumber other larval fishes in near-reef waters and thus feed the "wall of mouths" of reef-dwelling planktivores. The abundant CRF juveniles and short-lived small adults also provide nutrient inputs to the reef system. This new demographic model suggests that the contributions of CRFs to nutrient cycling in coral reef ecosystems are more important than previously appreciated (*6*).

The notion that substantial proportions of nutrients are recycled within a reef system by way of CRFs, if confirmed, represents an important advance toward understanding how coral reefs can be so productive. Coral reefs are surrounded by oceanic "deserts" of low-nutrient waters, and thus their high productivity and biodiversity have long perplexed scientists. If CRFs are instrumental for local nutrient retention, then they would constitute a crucial guild of species within reef ecosystems. In turn, if CRFs fulfill a key ecological role, then the ecological and evolutionary dynamics that ensure their local persistence—especially the extent to which young-adult fishes return to their parents' reef location or disperse to other neighborhoods—are highly consequential to the health of coral reef ecosystems.

Although the idea that CRFs perform critical ecosystem services is compelling, definitive demonstration of such a seminal role is challenging. In particular, the pelagic environment in which larvae develop remains poorly understood; thus, much of the information about fish life histories and species interactions necessary to reconstruct demographic dynamics is unknown. Because of these many uncertainties, Brandl *et al.* were understandably forced to estimate most life-history parameters and make simplifying assumptions.

For example, few key life-history parameters—such as fecundity, mortality, settlement size, and growth rate—are well understood at the species level, which required the authors to estimate such parameters at the family or order level. Such inferences based on higher taxonomic levels might obscure CRF characteristics. Brandl *et al.* define CRF families as having at least 10% of species smaller than 5 cm, but in only four of the 17 known CRF families are >50% of species known to be smaller than 5 cm (*2*). Thus, most trait estimates were derived from species that do not match the CRF criterion set by Brandl *et al.* For instance, cardinalfishes (Apogonidae) are often neither

small (80% of species are >5 cm) nor cryptic, with many species conspicuously schooling in the water column in daytime. Determining life-history attributes for more species would enable future updates of demographic ecosystem models based on species-level characteristics, thereby increasing precision and confidence in results.

A second simplification reflects the paucity of information regarding near-reef plankton composition. Because of biases in sampling methods (*7*) and challenges in identifying animal larvae, studies typically focus on selected taxonomic groups. For example, invertebrates are the most numerous constituents of zooplankton in reef waters, yet typical inventories focused on fishes do not report invertebrate abundances (*3*, *4*). To aggregate larval fish data across disparate studies, Brandl *et al.* focused on familial proportions of reef fishes rather than abundances, which obscures important aspects of nutrient cycling, especially invertebrate contributions to reef residents' nutrition. Comprehensive plankton-sampling regimens combined with species-level identifications of larvae would yield improved estimates of nutrient exchange in reefs and also show how larval retention differs among species.

Also unknown is the extent to which CRF larvae and juveniles are consumed by other fishes and invertebrates. It is possible that conventional approaches for examining the contents of fish guts underestimate CRF consumption (*1*). DNA sequencing data show promise for evaluating the full complexity of reef fish demographics and nutrient cycling (*8*, *9*).

Against the backdrop of pervasive coral reef bleaching lies an urgent need to accurately understand and quantify coral reef biodiversity. The study of Brandl *et al.* underscores our ignorance regarding the lives of larvae and their participation in coral reef ecosystem dynamics. Collecting life-history information at a species level can seem mundane, but such fundamental investigations with standardized sampling methods are crucial to the design of biologically realistic models. ■

REFERENCES AND NOTES

1. S. J. Brandl *et al.*, *Science* **364**, 1189 (2019).
2. S. J. Brandl *et al.*, *Biol. Rev. Camb. Philos. Soc.* **206**, 227 (2018).
3. J. M. Leis, *Bull. Mar. Sci.* **53**, 362 (1993).
4. J. M. Leis *et al.*, *Coral Reefs* **22**, 271 (2003).
5. K. Kasimatis, C. Riginos, *Coral Reefs* **35**, 387 (2015).
6. J. E. Allgeier, K. E. Speare, D. E. Burkepile, *Ecosphere* **9**, e02216 (2018).
7. G. S. Santos, M. Brito-Lolaia, R. Schwamborn, *J. Exp. Mar. Biol. Ecol.* **491**, 27 (2017).
8. J. M. Casey *et al.*, *Methods Ecol. Evol.* 10.1111/2041-210X.13206 (2019).
9. M. Leray *et al.*, *Mol. Ecol.* 10.1111/mec.15090 (2019).

10.1126/science.aax8961

EVOLUTION

Ruminants: Evolutionary past and future impact

The genomes of ruminant animals promise advances for agriculture, conservation, and biomedicine

By **Dai Fei Elmer Ker**[1,2] *and* **Yunzhi Peter Yang**[3,4,5]

Ruminants are mammals of considerable agricultural, conservational, and biomedical importance. The approximately 200 extant species of this clade include traditional livestock such as cattle and sheep, endangered species such as Père David's deer or milu, and species of biomedical interest, such as antlered deer (*1*–*8*). Ruminants are extremely diverse but share in common a multichambered stomach specialized for digesting tough plant fibers through microbial-aided fermentation (*1*, *4*). Collectively, ruminants are highly successful, having colonized multiple terrestrial environments, including the unforgiving conditions of the Arctic tundra (*3*, *4*). On pages 1152, 1153, and 1154 of this issue, Chen *et al.* (*1*), Wang *et al.* (*2*), and Lin *et al.* (*3*), respectively, sequenced the genomes of multiple ruminant species, offering resources and analyses for understanding their evolution and diversity. The findings provide vital insights into genetic adaptations that are responsible for their biological success, as well as how they have been affected by human activity.

Chen *et al.* assembled the genomes of 44 ruminant species from all six Ruminantia families (Tragulidae, Antilocapridae, Giraffidae, Cervidae, Moschidae, and Bovidae), This substantially adds to the library of previously sequenced ruminant genomes, including cattle, yak, goat, sheep, Masai giraffe, okapi, and red deer. With their dataset, Chen *et al.* analyzed Ruminantia evolutionary divergences. Their methodology used the distantly related killer whale as an outgroup, the well-annotated goat genome as a reference for orthologous syntenic sequences (evolutionarily related genes colocalized on the same chromosome), and the fossil record as a means to calibrate molecular rates of evolution, producing a whole-genome phylogenetic tree for 51 ruminant species. The evolutionary insights gained include identifying a closer (sister-group) relationship between Antilocapridae and Giraffidae. Previous mitochondrial DNA–based phylogenies placed Antilocapridae as an outgroup relative to other pecorans. Additionally, they confirmed Moschidae as a sister group of Bovidae and better refined Bovidae subfamily relations and species positions (*1*). These findings provide strong evidence for adjudicating the positions of ambiguous taxa.

Chen *et al.*'s dataset also provides a window to the past. Using methods to infer ruminant demographics during the Pleistocene (~2.58 million to 12 thousand years ago), they found that major population declines coincided with increased human (effective) population size. This implies that humans may have contributed to ruminant declines during the late Pleistocene and holds important conservation lessons. When analyzed with nonruminant genomes and ruminant transcriptomic (gene expression) data, genome evolution could be deciphered. Such analyses identified fast-evolving genes, positively selected genes, and newly evolved genes that may have important roles in ruminant trait evolution.

Wang *et al.* studied the evolution of ruminant headgear (antlers, horns, ossicones, and pronghorns), which has implications for regenerative biology and potentially for cancer biology. These cranial appendages serve diverse functions that include mate selection, feeding, and sensing. Ruminant headgear share similarities in that they comprise a bony core covered by integument (skin and/or keratinized sheath) and are located on the frontal cranial bones. However, these headgear are unique in that four of the ruminant families—Cervidae, Bovidae, Giraffidae, and Antilocapridae—each possess a different type—antlers, horns, ossicones, or pronghorns, respectively (see the figure). In particular, antlers are a research focus for tissue engineering and regenerative medicine because they are deciduous bony structures that typically regenerate rapidly. Indeed,

[1]*Institute for Tissue Engineering and Regenerative Medicine, The Chinese University of Hong Kong, New Territories, Hong Kong SAR.* [2]*School of Biomedical Sciences, Faculty of Medicine, The Chinese University of Hong Kong, New Territories, Hong Kong SAR.* [3]*Department of Orthopaedic Surgery, Stanford University, Stanford, CA 94305, USA.* [4]*Department of Materials Science and Engineering, Stanford University, Stanford, CA 94305, USA.* [5]*Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. Email: ypyang@stanford.edu*

antlers can grow as much as 2.5 cm per day and produce 10 kg of osseous (bone) tissue within a few months (*4*, *6*, *7*). Although this phenomenon was noted in Aristotle's *History of Animals* more than 2000 years ago (*9*), the underlying molecular mechanism(s) are only just beginning to be understood.

By leveraging the data of Chen *et al.*, Wang *et al.* found that the most parsimonious explanation for headgear evolution is that this trait evolved once and was independently lost in Tragulidae and Moschidae. Comparative analysis of 68 goat, 162 sheep, 20 roe deer, and 20 sika deer transcriptomes sourced from 16 tissues identified 201 common headgear-specific genes as well as 624 horn-specific and 761 antler-specific genes. Additional analysis identified highly conserved gene regulatory elements implicated in stem cell pluripotency and the transforming growth factor–β (TGF-β) signaling pathway, indicating that a single mechanism—neural crest–mediated bone differentiation—may be involved. This provides further support that ruminant headgear may have originated once. Further bioinformatics analysis revealed that oncogenesis- and neurogenesis-associated pathways may mediate rapid antler regeneration. In particular, high expression of tumor suppressors, including promyelocytic leukemia protein (PML), was identified as being under positive selection. This may explain the low rate of cancer in cervids, while also promoting rapid antler growth.

Lin *et al.* examined molecular adaptations that enable reindeer to thrive in harsh Arctic and subarctic conditions, including severe cold, limited food availability, and prolonged periods of light or darkness. Reindeer have developed several distinct features, including specialized fat metabolism to limit heat loss, little-to-no circadian rhythm (daily cycle) to adapt to extreme light periodicity, and efficient vitamin D metabolism to facilitate rapid antler growth in the presence of low light. In addition, reindeer are unique within Cervidae in two aspects: They are the only species in which females possess antlers and the only fully domesticated species.

Using comparative genomics analyses with ruminant and nonruminant species, Lin *et al.* identified mutations that facilitate the Arctic adaptation and domestication of reindeer. These include mutations in genes and regulatory sequences that presumably enhanced vitamin D metabolism, modified fat production and transport, facilitated female antler growth in the presence of low androgen levels, disrupted circadian rhythm, and enhanced docility. Therefore, the study of Lin *et al.* reveals a genetic basis for the evolutionary success of reindeer.

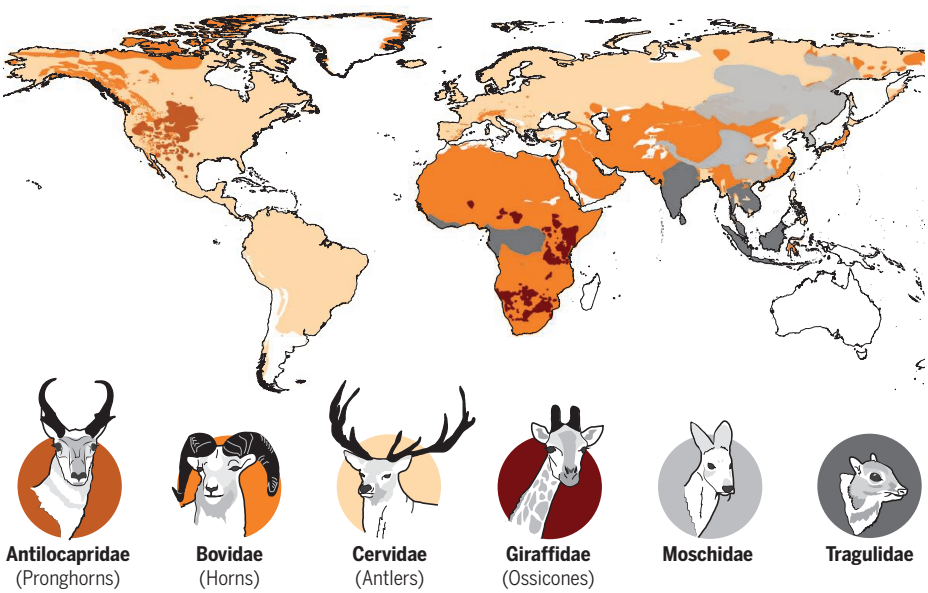
The genomic resources and evolutionary insights derived from these studies are likely to contribute to ruminant agriculture and conservation. These contributions may include improving agricultural economics and food security; reducing disease transmission; and minimizing human environmental impact, for example, by identifying and monitoring endangered species (*10*, *11*).

Additionally, studies of deer antlers offer attractive approaches for tissue engineering and regenerative medicine. For instance, deer antlers have inspired a commercially promising prosthesis for amputees (*8*). This prosthesis mimics the soft-to-hard (skin-to-bone) tissue interface of deer antlers, allowing secure intraosseous and transcutaneous attachment while minimizing infection. During antler regeneration, open wounds form on top of a deer's pedicles (the cranial structures from which antlers regenerate), which is quickly followed by scab formation, velvet skin transformation (which is crucial to antler formation), and antler morphogenesis. The ability of cervids to rapidly generate large quantities of innervated bone with low tumor and infection incidence suggests that studying antlerogenesis may offer insights for addressing large skeletal defects, achieving infection control, regenerating nerves, and possibly even limiting cancer growth (*4*–*7*).

Typically, regenerative biology is studied using small organisms such as hydras (*12*), lizards (*13*), naked mole-rats (*14*), and salamanders (*15*). Studying large mammals such as deer is logistically and ethically challenging (*4*). Although such challenges can be avoided with in vitro approaches to identify differentially expressed genes of interest (*6*), the lack of genomic resources and their accompanying insights presents tremendous research barriers. Thus, these three studies (*1*–*3*) are extremely valuable. Future work that validates the function(s) of identified genes will likely have major impacts on society. ■

Geographic distribution of endangered ruminants

Ruminants are highly successful, having colonized multiple terrestrial environments, but they remain threatened or endangered. This diverse group comprises six families, four of which have different types of headgear. The map shows distributions for threatened or endangered ruminant species included in the International Union for Conservation of Nature Red List of Threatened Species (*16*).



GRAPHIC: N. CARY/*SCIENCE*

REFERENCES AND NOTES

1. L. Chen *et al.*, *Science* **364**, eaav6202 (2019).
2. Y. Wang *et al.*, *Science* **364**, eaav6335 (2019).
3. Z. Lin *et al.*, *Science* **364**, eaav6312 (2019).
4. R. J. Goss, *Deer Antlers: Regeneration, Function, and Evolution* (Academic Press, 1983).
5. U. Kierdorf, H. Kierdorf, *Gerontology* **57**, 53 (2011).
6. D. F. E. Ker *et al.*, *Stem Cell Res. Ther.* **9**, 292 (2018).
7. C. Li, W. Chu, *Front. Biosci.* **21**, 455 (2016).
8. C. J. Pendegrass *et al.*, *J. Anat.* **209**, 59 (2006).
9. Aristotle, *History of Animals* (350 BCE).
10. M. Gilbert *et al.*, *Sci. Data* **5**, 180227 (2018).
11. L. W. Traill *et al.*, *Raffles Bull. Zool.* **S25**, 111 (2012).
12. J. A. Chapman *et al.*, *Nature* **464**, 592 (2010).
13. T. P. Lozito, R. S. Tuan, *Development* **143**, 2946 (2016).
14. E. B. Kim *et al.*, *Nature* **479**, 223 (2011).
15. S. Nowoshilow *et al.*, *Nature* **554**, 50 (2018).
16. The IUCN Red List of Threatened Species, version 2019-1; www.iucnredlist.org.

ACKNOWLEDGMENTS

We thank colleagues at Stanford University (C. Kim, E. Lui, and A. Stahl) for helpful discussions and feedback. We are supported by The Chinese University of Hong Kong Faculty Innovation Award, Innovation Technology Fund (ITS/090/18), and NIH grants R01AR057837, R01AR074458, R01AR072613, and U01AR069395.

10.1126/science.aax5182

CANCER

Hiding in plain sight

A common biomarker of pancreatic disease has a functional role in pathogenesis

By **Christopher J. Halbrook**[1] and **Howard C. Crawford**[1,2,3]

Pancreatic cancer remains one of the deadliest human malignancies. This is in part due to the lack of a reliable method of early detection, because late-stage disease is largely refractory to treatment. Biomarkers for early disease detection have remained elusive. However, the glycan carbohydrate antigen 19-9 (CA19-9), which is produced by pancreatic cancer cells, is increased in the serum of most patients (*1*). It is clinically useful as a biomarker of tumor burden during treatment, rather than for early detection, because serum CA19-9 is also increased in other diseases, including inflammation of the pancreas (pancreatitis), a risk factor for the development of pancreatic cancer. To date, there has been little understanding of CA19-9 function in pancreatic pathophysiology. On page 1156 of this issue, Engle *et al.* (*2*) report that CA19-9 drives the development of pancreatitis and accelerates pancreatic tumor progression.

CA19-9, also known as sialyl-Lewis^a, is a glycan moiety that modifies proteins, lipids, and other glycans. In humans, the expression of CA19-9 is dependent on Lewis (Le) blood type (*3*), a human blood group system based on the expression of two fucosyltransferase (FUT) enzymes, FUT2 and FUT3. Individuals with the blood type Le(a-b-) do not generate CA19-9 because they lack FUT3; mice also do not express this enzyme and therefore cannot produce CA19-9. Functionally, CA19-9 can promote cancer cell metastasis by binding to the adhesion protein E-selectin that is expressed on the surface of endothelial and mesothelial cells (*4*). This binding presumably enhances migration of metastatic cells through these cell layers. Additionally, CA19-9 expression in colon cancer cells replaces the structurally related disialyl Lewis^a expressed by normal colon epithelia (*5*). Disialyl Lewis^a produced by normal epithelium serves an immune inhibitory role by binding sialic acid–binding immunoglobulin-like lectins (siglecs) on resident macrophages, suppressing the expression of proinflammatory cyclooxygenase-2 (*6*). Accordingly, it is thought that the replacement of disialyl Lewis^a by CA19-9 during carcinogenesis relieves this immune suppression through its failure to engage siglecs, thereby promoting tumor progression. Apart from these generic functions, a pathophysiological function induced by CA19-9 modification of a specific protein has not been reported until now.

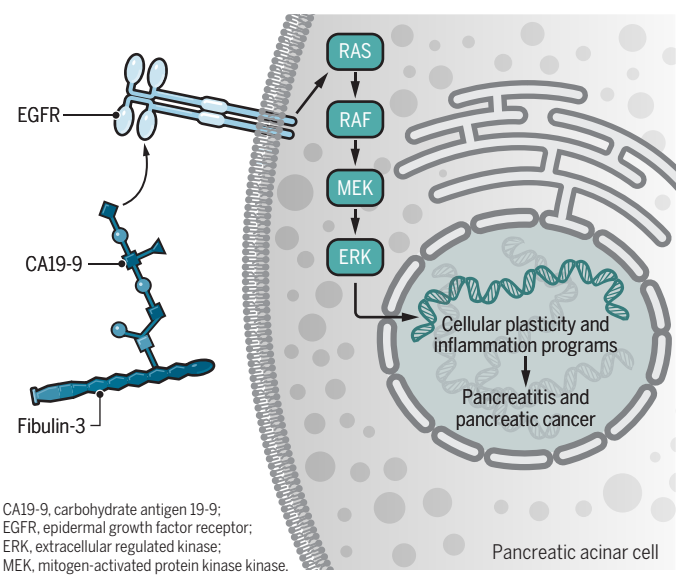
Pancreatic cancer cells nearly universally contain activating mutations in the *KRAS* oncogene. Expression of epidermal growth factor receptor (EGFR) is required for induction of *Kras*-driven tumorigenesis in the pancreas of mice (*7*, *8*) by increasing total RAS activity above a minimal threshold that is required for transformation. Concomitantly, the EGFR–RAS–mitogen-activated protein kinase (MAPK) signaling axis in pancreatic epithelium increases the production of proinflammatory cytokines (*9*). These cytokines recruit immune cells, which are key drivers of pancreatitis and pancreatic tumor development and maintenance (*10*–*12*). The numerous reactive inflammatory and stromal cells are in turn a rich source of EGFR ligands (*13*, *14*), invoking the involvement of a vicious cycle of chronic EGFR signaling and inflammation in benign pancreatic disease. In later stages of pancreatic cancer, EGFR is largely dispensable (*7*), and inhibiting it in pancreatic cancer patients provides minimal benefit (*15*).

Using genetic engineering, Engle *et al.* created a mouse model that expresses both FUT3 and the galactosyltransferase β3GALT5, allowing the production of CA19-9. They find that CA19-9 synthesis promotes modification of the extracellular matrix protein fibulin-3, producing CA19-9:fibulin-3. This activates pancreatic EGFR signaling, inducing a pathology that is similar to pancreatitis, which is reversible either by shutting down the expression of the CA19-9 synthesis machinery or by targeting CA19-9 itself (see the figure). Thus, inhibiting this newly described function of CA19-9 presents an attractive potential target for the treatment of early-stage pancreatic cancer and pancreatitis patients.

The results of Engle *et al.* suggest that the canonical EGFR ligands produced by the pancreatic epithelium or the inflammatory and stromal cells (*7*, *12*) are minimally involved in maintaining the disease state. One potential explanation for this is that CA19-9:fibulin-3 engages and activates EGFR with distinctive kinetics compared to other ligands. Alternatively, CA19-9:fibulin-3 is necessary but not sufficient to maintain the phenotype without collaborating with other CA19-9–modified molecules. This latter possibility seems more likely because, as shown by Engle

Signaling underlying pancreatic disease

Posttranslational modifications by CA19-9 can confer new functions to proteins. CA19-9:fibulin-3 can act as an EGFR ligand, thereby activating the downstream RAS-RAF-MEK-ERK signaling cascade.



CA19-9, carbohydrate antigen 19-9; EGFR, epidermal growth factor receptor; ERK, extracellular regulated kinase; MEK, mitogen-activated protein kinase kinase.

GRAPHIC: A. KITTERMAN/*SCIENCE*

[1]*Department of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, MI, USA.* [2]*Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA.* [3]*The Rogel Cancer Center, University of Michigan, Ann Arbor, MI, USA. Email: howcraw@med.umich.edu*

et al., blocking CA19-9 is more effective at reversing the pancreatitis phenotype than inhibiting EGFR itself, which causes severe pancreatic atrophy (wasting).

Although the discovery of a new function of CA19-9 is exciting, Engle *et al.* have also potentially revitalized a role for CA19-9 in the early detection of pancreatic cancer. A major issue for many putative early detection markers, including CA19-9, has been the limited ability to distinguish between chronic pancreatitis and pancreatic cancer. With their transgenic mouse model and the organoids (cells grown as three-dimensional cultures) derived from them, they have generated an optimal system with which to identify CA19-9–modified proteins that can distinguish the metaplastic epithelia found in pancreatitis from the cancerous epithelia of adenocarcinoma. Finding specific cancer-associated CA19-9–modified proteins in the circulation could greatly enhance the specificity of a biomarker compared to either CA19-9 or the unmodified protein alone.

To date, there are still no effective therapies or early detection methods for pancreatic cancer. This is arguably because of a lack of understanding of the biology underlying the disease, a void illustrated by the study of Engle *et al.*, which ascribes a critical function to the most commonly used biomarker of the disease, first identified over 30 years ago (*1*). The translational potential of the findings of Engle *et al.* also reinforce the value of studying the events that precede and promote the formation of pancreatic cancer. These observations lend much needed insight into the understanding of pancreatic cancer and can inform new treatment strategies for the advanced stages in which it is usually diagnosed in patients. ■

REFERENCES AND NOTES

1. M. Herlyn, H. F. Sears, Z. Steplewski, H. Koprowski, *J. Clin. Immunol.* **2**, 135 (1982).
2. D. D. Engle *et al.*, *Science* **364**, 1156 (2019).
3. M. A. Tempero *et al.*, *Cancer Res.* **47**, 5501 (1987).
4. F. Gebauer *et al.*, *Gut* **62**, 741 (2013).
5. K. Miyazaki *et al.*, *Cancer Res.* **64**, 4498 (2004).
6. K. Miyazaki *et al.*, *J. Immunol.* **188**, 4690 (2012).
7. C. M. Ardito *et al.*, *Cancer Cell* **22**, 304 (2012).
8. C. Navas *et al.*, *Cancer Cell* **22**, 318 (2012).
9. C. J. Halbrook *et al.*, *Cell. Mol. Gastroenterol. Hepatol.* **3**, 99 (2017).
10. C. Guerra *et al.*, *Cancer Cell* **11**, 291 (2007).
11. G. Y. Liou *et al.*, *Cancer Discov.* **5**, 52 (2015).
12. F. McAllister *et al.*, *Cancer Cell* **25**, 621 (2014).
13. Y. Zhang *et al.*, *eLife* **6**, e27388 (2017).
14. K. C. Ray *et al.*, *Oncogene* **33**, 823 (2014).
15. M. J. Moore *et al.*; National Cancer Institute of Canada Clinical Trials Group, *J. Clin. Oncol.* **25**, 1960 (2007).

ACKNOWLEDGMENTS

C.J.H. is supported by a MICHR PTSP grant (UL1TR000433) and F32CA228328. H.C.C. is supported by the Sky Foundation and by Cancer Center Support Grant (P30CA046592) and U01 CA224145.

10.1126/science.aax9341

CANCER

The gut microbiota and colon cancer

Microbiome data should be incorporated into the prevention, diagnosis, and treatment of colon cancer

By **Wendy S. Garrett**

The human microbiota is the collection of microorganisms—bacteria, archaea, viruses, fungi, protozoa, and helminths—that populate the human body. They are emerging as an important feature of human health and disease. Currently, access to the genomic data of human cells and of microbiota (microbiomes) is more affordable and accessible than ever before. A major challenge is to unravel how we integrate microbiome data into precision medicine approaches for the prevention, diagnosis, and treatment of diseases such as cancer. The gastrointestinal (GI) tract is densely populated with microorganisms. Colorectal cancer (CRC) is the third most prevalent cancer worldwide. It is increasing in individuals less than 50 years old and is associated with specific dietary factors and eating patterns that affect the gut microbiota. Therefore, CRC seems ripe for microbiome-based prevention, diagnostics, and therapeutics.

Over the past 10 years, several species of bacteria have garnered attention for their associations with, and potential roles in, colorectal carcinogenesis (see the figure). Some, such as *Fusobacterium nucleatum*, initially emerged from sequencing-based studies of human CRC samples (*1*, *2*), and mechanistic data from preclinical models have since demonstrated a panoply of roles for *F. nucleatum* in several aspects of CRC biology. In vitro, *F. nucleatum* promotes CRC cell proliferation; in mice, the presence of *F. nucleatum* in patient-derived CRC xenografts (where patient CRC samples are implanted in mice) increases tumor growth rates (*3*). A possible mechanism to explain these findings is that the *F. nucleatum* adhesin, FadA, binds to E-cadherin on the CRC cell surface and activates oncogenic Wnt/β-catenin signaling (*4*). *F. nucleatum* can also alter the function of tumor-infiltrating lymphocytes and natural killer (NK) cells by binding to the inhibitory immune receptor TIGIT (T cell immunoreceptor with Ig and ITIM domains) through another adhesin, Fap2 (*5*). Fap2 also binds a disaccharide sugar motif [galactose *N*-acetyl-D-galactosamine (Gal-GalNAc)], which is expressed at high levels on the surface of many tumor cell types, as well as other cell types, and facilitates *F. nucleatum* binding to CRC cells and enrichment in CRC (*5*).

The mechanisms by which *F. nucleatum* affects carcinogenesis also extend to facilitating resistance to chemotherapy. In preclinical models of CRC with high tumor loads of *F. nucleatum*, tumors are more resistant to the commonly used CRC chemotherapeutic, oxaliplatin. *F. nucleatum* activates autophagy (a cellular recycling process that influences cell survival) through Toll-like receptor 4 (TLR4) expressed on CRC cells, rendering these tumors more resistant to oxaliplatin-induced cell death (*5*).

Other bacteria, such as enterotoxigenic *Bacteroides fragilis* (ETBF), are long-studied human GI pathogens that cause diarrhea and GI inflammation. ETBF also potentiates colorectal carcinogenesis in mice, and recently, ETBF was detected in biofilms coating human CRCs and precancerous colonic lesions called adenomas (*6*). Similarly, *Escherichia coli* expressing the genomic island polyketide synthase (pks^+) enhance tumorigenesis in preclinical CRC models and are enriched in human CRC tissues (*7*). pks^+ *E. coli* produce the genotoxin colibactin, which alkylates DNA, resulting in DNA adducts (a form of potentially mutagenic DNA damage) in colonic epithelial cells. The recent biochemical resolution of this process illustrates how host-microbe interactions might lead to cancer (*8*).

There is still much to learn from this triad of species, including whether they interact sequentially, in tandem, or at all to influence colorectal carcinogenesis. Although continuing to study and translate knowledge of *F. nucleatum*, ETBF, and pks^+ *E. coli* for human health is important,

Department of Immunology and Infectious Diseases and Department of Genetics and Complex Diseases, Harvard T. H. Chan School of Public Health, Boston, MA 02115, USA; Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA; and Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA 02215, USA. Email: wgarrett@hsph.harvard.edu

these are not the only microorganisms that are relevant for CRC. Therefore, a valuable resource needed for the advancement of CRC-microbiota studies is an atlas to map not only what microorganisms are present in the tumor microenvironment but also where they are within and on tumors, how they interact with one another and the host, and how their arrivals and departures over time shape the evolving tumor.

The questions that such a cancer-microbiome atlas can address are wide-ranging. For example, what are the consequences, if any, of the presence of a given microbiota for how cancer cells function, the metabolic wiring of tumors, the oncogenic programs that drive growth and metastasis, and whether and how a cancer will respond to therapy? A nascent area of special interest is understanding how the gut microbiota in a patient may influence why thousands of individuals who have been, and continue to be, closely studied over several decades prior to, and at presentation with, colorectal adenomas and CRCs. Such longitudinal studies are ideal for understanding how the microbiota may be associated with CRC risk or influenced by other modifiable or protective factors. Broad screening registries for CRC such as the United Kingdom's National Health Service Bowel Cancer Screening Programme, in which millions of fecal occult blood test (FOBT) kits have been collected from individuals aged 60 to 74 biennially, also provide rich opportunities to study the stool microbiome and its correlation with CRC over time.

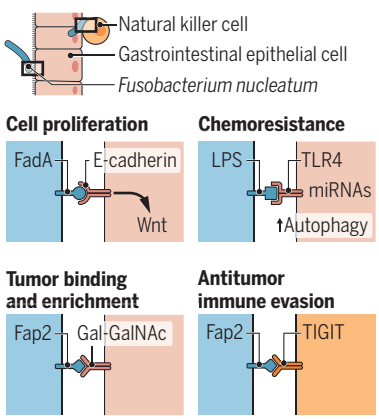
Evaluating the associations that emerge from population-based studies in model systems remains an essential step in moving microbiota discovery toward CRC diagnostics and therapeutics. Preclinical models in-

peutic targets as well as allow for patient stratification and prognostication.

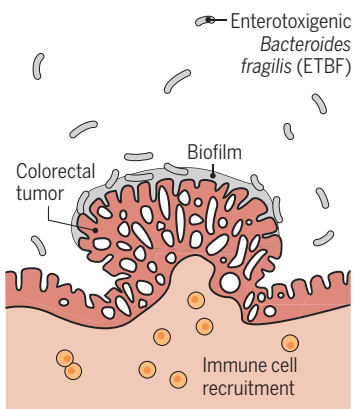
Candidate microbes and metabolites gleaned from longitudinal prospective cohort and cross-sectional studies of colorectal adenomas and CRCs can also be evaluated in gnotobiotic animal models with tractable genetics, such as fruit flies, zebrafish, and mice. Gnotobiotic models are entirely devoid of endogenous microbes or harbor small, well-defined microbial communities. They are a useful system for unraveling the multiplicity of influences that microbes and their metabolites have in cancer development and cancer therapeutic efficacy and toxicity. Recently, testing healthy human stool samples and cocktails of human-derived microbial isolates in gnotobiotic mouse CRC models led to the observation that groups of bacteria can shape interferon-γ–expressing (IFNγ$^+$)

Microorganisms may drive colorectal cancer

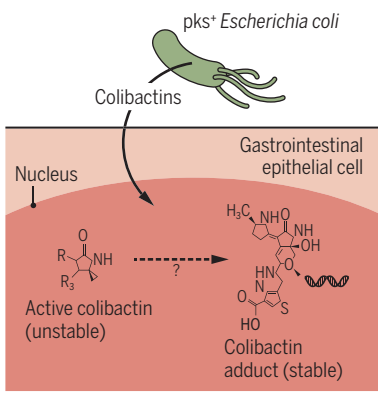
Three examples of the possible mechanisms by which bacteria might influence colorectal cancer. It remains unclear whether they have a causative role in colorectal carcinogenesis and the potential mechanisms involved.



F. nucleatum expresses adhesins and lipopolysaccharide (LPS), which can have multiple influences on cell behavior.



ETBF coats tumors and recruits other bacteria to create a biofilm. ETBF also recruits immune cells and promotes inflammation.



Polyketide synthase–expressing (pks$^+$) *E. coli* may influence carcinogenesis through the generation of potentially mutagenic DNA adducts.

Gal-GalNAc, galactose *N*-acetyl-D-galactosamine; miRNA, microRNA; TIGIT, T cell immunoreceptor with Ig and ITIM domains; TLR4, Toll-like receptor 4.

immunotherapies work or fail in patients.

Given that geography, diet, environmental exposures, and lifestyle choices all influence the composition of an individual's microbiota, it is important that such metadata are collected and studied in global CRC cohorts. This is crucial to begin to make sense of a range of findings including conflicting CRC stool and tumor microbiome meta-analyses (*5*) and variation in chemotherapy toxicities (side effects) observed in CRC patients from different continents.

Patient populations provide a window of opportunity to study cancer and the microbiota. However, prospective, longitudinal cohorts of healthy populations are also of critical importance to understand factors that not only increase susceptibility for CRC but also decrease CRC risk. The Nurses' Health Study and Health Professional Follow-up Study have enrolled hundreds of

cluding organoids (tissue cultures that form three-dimensional organlike structures) and animal models are resolving whether and how microorganisms are disease drivers and key influencers. Organoids can now be generated from patient tumor tissue, can be genetically manipulated to mirror the acquisition of critical genetic mutations observed in patient tumors, and can be transplanted back into animal models to study tumor growth and spread in vivo (*9*, *10*). When cocultured with microbes or microbial metabolites, organoid systems, especially when nonepithelial host cell components such as stromal cells and immune cells are included (*11*), provide a powerful, reductionist system to directly interrogate the roles of microbes and their products. Studies from model systems facilitate the identification of specific microbial metabolites and host-microbial signaling pathways that may provide therapeutic targets as well as allow for patient stratification and prognostication.

Candidate microbes and metabolites gleaned from longitudinal prospective cohort and cross-sectional studies of colorectal adenomas and CRCs can also be evaluated in gnotobiotic animal models with tractable genetics, such as fruit flies, zebrafish, and mice. Gnotobiotic models are entirely devoid of endogenous microbes or harbor small, well-defined microbial communities. They are a useful system for unraveling the multiplicity of influences that microbes and their metabolites have in cancer development and cancer therapeutic efficacy and toxicity. Recently, testing healthy human stool samples and cocktails of human-derived microbial isolates in gnotobiotic mouse CRC models led to the observation that groups of bacteria can shape interferon-γ–expressing (IFNγ$^+$) CD8$^+$ T cell populations to promote antitumor immunity (*12*). Similarly, gnotobiotic mouse tumor models are being used to evaluate whether a cause of nonresponsiveness to a cancer treatment, such as immunotherapy, results from microbiota (*13*, *14*). From these and other studies, single microorganisms, consortia of microbes, and microbial metabolites have been identified that are being developed to improve cancer treatment efficacy.

Interventions targeting the microbiota and cancer frequently mirror the spectrum of preventive and therapeutic strategies for infectious diseases, including vaccines, antibiotics (as well as antivirulence drugs), fecal microbiota transplants (FMTs), and phage-based therapeutics. A recent proof-of-principle study demonstrated that the antibiotic metronidazole could slow the growth of *F. nucleatum*–positive human

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

tumor samples in patient-derived xenograft mouse models (*3*). However, it is unclear whether antibiotics could be used in CRC prevention, given the years to decades over which cancerous lesions in the colon develop and the consequent potential of developing antibiotic resistance. Highly selective antibiotics or antivirulence approaches might represent a more viable strategy that is less disruptive for human microbial ecology. FMT, the transfer of fecal material from healthy individuals to patients, is increasingly being used to combat colitis caused by antibiotic-resistant *Clostridium difficile* infection. More recently, FMT has been considered for the treatment of a number of diseases, including obesity and inflammatory bowel disease, with several clinical trials under way. FMT has also been piloted in a small number of cancer patients who developed severe colitis associated with the use of immunotherapy (*15*).

Beyond FMT, carefully curated cocktails of microbes are being tested for *C. difficile* colitis and for CRC. The goal of such a therapeutic is to use microbial consortia to "push out" or exclude a disease-associated microorganism from a patient's gut or tumor. Vaccines also hold tremendous potential for cancer prevention, with cancer-associated microbes and tumor neoantigens used to elicit antitumor immunity. For many CRC-potentiating microbes, their toxins, adhesins, and outer membrane proteins are attractive vaccine targets. Additionally, exciting opportunities also exist for using microbiota profiling information not only in CRC prevention, diagnostics, and therapeutics but also for treatments targeting GI microorganisms to decrease the toxicities of CRC therapeutics. ■

REFERENCES AND NOTES

1. M. Castellarin *et al.*, *Genome Res.* **22**, 299 (2012).
2. A. D. Kostic *et al.*, *Genome Res.* **22**, 292 (2012).
3. S. Bullman *et al.*, *Science* **358**, 1443 (2017).
4. M. R. Rubinstein *et al.*, *EMBO Rep.* **20**, e47638 (2019).
5. C. A. Brennan, W. S. Garrett, *Nat. Rev. Microbiol.* **17**, 156 (2019).
6. C. M. Dejea *et al.*, *Science* **359**, 592 (2018).
7. J. C. Arthur *et al.*, *Science* **338**, 120 (2012).
8. Y. J. R. Wilson *et al.*, *Science* **363**, eaar7785 (2019).
9. J. Drost *et al.*, *Science* **358**, 234 (2017).
10. K. P. O'Rourke *et al.*, *Nat. Biotechnol.* **35**, 577 (2017).
11. J. T. Neal *et al.*, *Cell* **175**, 1972 (2018).
12. T. Tanoue *et al.*, *Nature* **105**, 44 (2019).
13. V. Gopalakrishnan *et al.*, *Science* **359**, 97 (2018).
14. V. Matson *et al.*, *Science* **359**, 104 (2018).
15. Y. Wang *et al.*, *Nat. Med.* **24**, 1804 (2018).

ACKNOWLEDGMENTS

W.S.G. thanks L. Ricci for help with the figure. W.S.G. is a member of the scientific advisory boards of Evelo Biosciences, Kintai Therapeutics, Leap Therapeutics, and uBiome and has consulted for BiomX, Merck, and Pfizer Consumer Health in the past 2 years. W.S.G. is a coprincipal investigator with M. Meyerson of a Cancer Research UK Grand Challenge Grant.

10.1126/science.aaw2367

PHOTO: M. TOMASELLI

Inuit harvester Julia Ogina, of the Kitikmeot Inuit Association (left), shares her knowledge on muskox health with wildlife veterinarian Matilde Tomaselli (right), in Cambridge Bay, Nunavut, Canada.

BRIDGE: INDIGENOUS AND SCIENTIFIC KNOWLEDGE

"Two-eyed seeing" supports wildlife health

Bridging Indigenous and scientific knowledge improves wildlife surveillance and fosters reconciliation

By **Susan Kutz**[1] *and* **Matilde Tomaselli**[1,2]

The cry "Don't shoot the leaders!" is central to the traditional knowledge of Indigenous peoples across the Canadian North. For countless generations, northern Indigenous peoples have witnessed the annual caribou migrations, understanding their mechanisms and patterns. They know that if the caribou leading the migration are removed, the rest of the herd do not know where to migrate and will not return to the traditional harvesting grounds. A recent scientific study on the migratory behavior of hoofed animals also concludes that they learn from their conspecifics where and when to migrate (*1*). Experiential-based knowledge such as that of the northern Indigenous peoples, acquired through practice and over generations, has been central to human adaptation and survival for millennia. Combining this knowledge with scientific knowledge will help to achieve better-informed and more timely and effective decision-making on wildlife health and conservation.

The urgency of ethically documenting and effectively using local knowledge for wildlife health surveillance is underscored by the current rate of climate change in the Arctic and elsewhere (*2*), where disease emergence and wildlife population declines are threatening species survival (*3*) and the livelihoods, food safety, and food security for Indigenous peoples (*4*). Many of the principles that apply to local knowledge held by Indigenous peoples also apply to local knowledge held by nonindigenous people around the world.

VALUING INDIGENOUS KNOWLEDGE

Accounts related by Indigenous knowledge holders, and even local idioms, provide tremendous insights into wildlife biology, health, and disease ecology. People living in close intimacy with their environment have a comprehensive understanding of the dynamic nature of the systems in which they live, including indicators of ecosystem health and their natural variability. Systematically documenting Indigenous knowledge can result in early detection of ecological changes (*5*). For example, observations by subsistence hunters of thinner

[1]*Department of Ecosystem and Public Health, Faculty of Veterinary Medicine, University of Calgary, Calgary, Alberta, Canada.* [2]*Polar Knowledge Canada, Government of Canada, Cambridge Bay, Nunavut, Canada. Email: skutz@ucalgary.ca; matilde.tomaselli@ucalgary.ca*

animals can predict an impending population decline long before detection by conventional population surveys.

Whether it be overt events such as disease outbreaks or more insidious population changes, early detection is essential to respond effectively to conserve wildlife and to protect human and domestic animal health, food security, and livelihoods (*6*). Conservation organizations are increasingly recognizing that respecting Indigenous knowledge is essential for equitable development and environmental sustainability (*7*).

Despite these advances, application of Indigenous knowledge for wildlife health assessment and conservation is in its infancy, and there remains uncertainty, and at times skepticism, among decision-makers on how to combine local perspectives with scientific insights for evidence-based decision-making around wildlife. Fundamental barriers include the "need to meet scientific credibility" to produce reliable accounts and the difficulty to "translate user perceptions into categories/criteria that rely on numbers" (*8*). Addressing these barriers does not mean that Indigenous knowledge must be coopted into the scientific knowledge system or to be validated with science; rather, it requires engagement in meaningful dialogues that enable different knowledges to interface.

THE POWER OF "TWO-EYED SEEING"

Recognizing the legitimacy of multiple ways of knowing is the first step in bridging knowledge types. This includes mutual respect for different philosophical theories and assumptions underlying knowledge systems, acknowledging limitations, and capitalizing on strengths (*9*, *10*).

Scientific knowledge is widely regarded as the gold standard because of its focus on measurable and objective criteria. However, scientific surveillance of wildlife health is often limited in space and time, only documents certain types of information, and is difficult to interpret when data are fragmented. These limitations are especially evident in remote settings where fieldwork is expensive, logistically difficult, and often restricted to specific months of the year.

Indigenous knowledge, a holistic body of knowledge acquired through experience at broad spatial and temporal scales, cumulatively captures information across multiple interdependent elements (*10*). Indigenous knowledge holders generally cannot talk about caribou populations, behavior, or health without discussing the environment, wolves, insects, and caribou-human interactions. In many ways, Indigenous knowledge provides an ideal example of the "One Health" approach, in which animals and the complexities of their environment, including the human element, are considered simultaneously. Unfortunately, the legacy of colonial practices, together with rapid socioecological changes and an aging population of knowledge holders, has led to the breakdown of intergenerational transmission of Indigenous knowledge, putting the continuity of this rich knowledge system at risk.

Bridging multiple knowledge systems requires drawing on natural and social sciences' methodologies and constant consideration for the value systems of all knowledge holders, a process that is based on ongoing iteration and feedback (see fig. S1). The Mi'kmaq principle of "Etuaptmumk" or "two-eyed seeing" captures the concept of bringing different knowledge systems together to increase our collective breadth and depth of understanding: "learning to see from one eye with the strengths of Indigenous knowledges…and from the other eye with the strengths of Western knowledges…and learning to use both these eyes together, for the benefit of all" (*11*).

Participatory epidemiology provides a pragmatic framework for applying this principle to understand wildlife health. Participatory epidemiology was developed, and is still largely applied, in low-income countries to enhance livestock disease surveillance. The two-eyed seeing principle is achieved through triangulation, or corroboration of information by using multiple methods and sources, including knowledge from resource users and conventional veterinary methods. This approach was essential for the global eradication of the infectious viral disease rinderpest: Pastoralist knowledge identified the last foci of rinderpest in remote areas of East Africa where scientific methods had failed, guiding targeted control (*12*).

In Arctic Canada, participatory epidemiology was used to understand health in a declining muskox population (*5*). Knowledge of Inuit harvesters identified and characterized changes in population demographics, condition, emerging disease syndromes, and mortality. Together, scientific and Inuit knowledge generated a greater understanding of mechanisms that drive muskox populations. The utility of participatory epidemiology applied to wildlife health extends beyond the Arctic and Indigenous peoples—particularly to developing countries, where data on wildlife are often scarce, and strong interactions between people, wildlife, and livestock create hotspots for emerging zoonoses.

Indigenous knowledge of bowhead whale behavior provides new insights that are critical for management.

PHOTO: PAUL NICKLEN/NATIONAL GEOGRAPHIC IMAGE COLLECTION

OVERCOMING BARRIERS

Bridging Indigenous and scientific knowledge holds considerable promise in wildlife health ecology and surveillance; however, challenges remain. As the combined use of Indigenous and scientific knowledge becomes increasingly encouraged and mandated, there is risk of implementing approaches that produce power imbalances

or include Indigenous knowledge only superficially. To ethically bridge knowledge systems, meaningful engagement with knowledge holders is necessary at all stages, from project design to data interpretation and project evaluation (*10*, *13*).

To move from anecdote to the data that will be used by decision-makers, documentation of local knowledge must follow rigorous qualitative research design, including repeatable and transparent methods ("need to meet scientific credibility") (*5*, *9*). Identification of expert knowledge holders, thematic saturation and triangulation (qualitative research methods used to avoid missing and to corroborate information, respectively), and validation and cointerpretation of results will ensure accurate accounts, interpreted and applied in context, and will serve as an internal peer review process (*5*).

Through participatory appraisal exercises (such as mapping, timelines, and proportional piling), qualitative information from resource users can be translated into spatial, temporal, and numerical representations ("into categories/criteria that rely on numbers") (*5*). These types of data enable inclusion of local knowledge into familiar formats used for management decisions. As with scientific data, quantifying local observations allows us to understand variability around observations and act with knowledge of that uncertainty. However, qualitative and quantitative data must be documented and interpreted together to avoid misinterpreting information and loss of deeper insights.

Growing efforts to incorporate local knowledge into quantitative models [such as Bayesian Belief Networks (*14*)] can provide tools that further facilitate inclusion of this knowledge into decision-making. However, sophisticated mathematical models can pose technical barriers to participation, producing power imbalances (*15*). When applying these tools, it is paramount to ensure that Indigenous knowledge holders actively participate at all stages and share control of the process.

Inevitably, at times, Indigenous and scientific knowledges will conflict. Partners then need to coevaluate the strengths and limitations of the respective studies and jointly reflect on disparate findings. Such a contention arose in 1977 between scientists and indigenous whalers around the number of bowhead whales, which resulted in a hunting ban (*10*). This conflict was resolved only when the knowledge of Iñupiat whalers on bowhead whale behavior was considered, leading scientists to realize that they had severely underestimated the stock size based on faulty assumptions. This example is a reminder that such instances of conflict are opportunities for the greatest insights into ecological processes to be revealed, transformative knowledge generated, and improved dialogue among partners achieved, enabling the development of shared solutions.

"Don't shoot the leaders" echoes as a call for action to enable today's elders to pass their knowledge to the youth who will be tomorrow's leaders. The creation of an inclusive platform for knowledge exchange around wildlife health, a theme so central to many Indigenous cultures, improves wildlife conservation and public health protection, promotes intergenerational learning and retention of Indigenous knowledge within contemporary contexts, and contributes to reconciliation with Indigenous peoples. Respecting Indigenous knowledge and working collaboratively with knowledge holders will build capacity and empower Indigenous communities to set their own agendas as advocated by the United Nations (*7*). The latest report by the Intergovernmental Panel on Climate Change (IPCC) highlights the rate at which climate change is affecting the Arctic and other regions of the world (*2*), emphasizing the urgency with which we all need to embrace the concept of Etuaptmumk, to cogenerate knowledge and to learn from each other "for the benefit of all." ■

REFERENCES AND NOTES

1. B. R. Jesmer *et al.*, *Science* **361**, 1023 (2018).
2. IPCC, *Global Warming of 1.5°C, Summary for Policymakers* (IPCC, 2018).
3. R. A. Kock *et al.*, *Sci. Adv.* **4**, eaao2314 (2018).
4. S. Kutz *et al.*, *Can. Vet. J.* **56**, 560 (2015).
5. M. Tomaselli *et al.*, *Biol. Conserv.* **217**, 337 (2018).
6. OIE–World Organization for Animal Health, *Manual for Terrestrial Animal Health Surveillance* (OIE, 2014).
7. United Nations, *Declaration on the Rights of Indigenous Peoples* (United Nations, 2008); www.un.org/esa/socdev/unpfii/documents/DRIPS_en.pdf.
8. International Union for Conservation of Nature (IUCN), *Integrating Traditional Knowledge into Red List Assessments, WCC Workshop no. 9744* (IUCN, 2016); www.iucn.org/sites/dev/files/content/documents/tk_workshop_summary_draft_1.pdf.
9. K. Moon, D. Blackman, *Conserv. Biol.* **28**, 1167 (2014).
10. N. C. Ban *et al.*, *Nat. Ecol. Evol.* **2**, 1680 (2018).
11. S. K. Denny, L. M. Fanning, *Int. Indig. Policy J.* **7**, 1 (2016).
12. J. C. Mariner *et al.*, *Science* **337**, 1309 (2012).
13. J. V. Lavery, *Science* **361**, 554 (2018).
14. C. S. Mantyka-Pringle *et al.*, *Environ. Int.* **102**, 125 (2017).
15. M. Barber, S. Jackson, *Hum. Ecol.* **43**, 119 (2015).

ACKNOWLEDGMENTS

We thank the Indigenous knowledge holders we work with. Ideas were formed through collaborations with residents and Indigenous organizations Cambridge Bay, Kugluktuk, Ulukhaktok , and Sahtu in northern Canada and through discussions with S. Checkley, C. Gerlach, B. Elkin, C. Ribble, and the Kutz lab. We thank A. Dobson, B. Moorman, R. Barry, B. Buntain, J. Adamczewski, J. Pellissey, and anonymous reviewers for comments.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1135/suppl/DC1

10.1126/science.aau6170

METROLOGY

Squeezing out higher precision

Well-timed kicks to an ion's momentum enable better position measurements

By **Monika Schleier-Smith**

Making measurements at a precision limited only by quantum uncertainty is crucial for applications such as the detection of gravitational waves (*1*) or the search for dark matter (*2*). In this quest, physicists can reduce uncertainty in measurements of one variable at the expense of another (*3*, *4*). Ultimately, however, the precision is often limited by the classical noise of the measurement apparatus. In principle, a solution to this problem is to amplify the signal relative to the classical noise. Readers familiar with cassette tapes may recognize this strategy, which is much like turning up the volume of music in a car so that it can be heard above the engine noise, only to have the tape hiss obscure the finer features of the music. On page 1163 of this issue, Burd *et al.* (*5*) report extremely sensitive measurements of the position of an individual ion, achieved by both suppressing quantum noise and amplifying signal.

Trapped ions are exquisitely well-controlled quantum systems, used as ultraprecise clocks (*6*) and as building blocks for quantum computers (*7*, *8*). Electric fields can be used to suspend a single ion in ultrahigh vacuum, and lasers are then used to bring it to rest in its lowest energy state. Even in this stable resting state, the ion has zero-point energy associated with quantum fluctuations in its position and momentum. In the case of the low-mass magnesium atom studied by Burd *et al.*, the motion created by the zero-point energy is 70 times the size of the ion in its trapped ground state. Does this mean that precise measurements of its position are futile?

The limit imposed by the Heisenberg Uncertainty Principle makes knowing both an object's position and its momentum at

Department of Physics, Stanford University, Stanford, CA 94305, USA. Email: schleier@stanford.edu

the same time impossible. However, we can know the position alone to arbitrary precision. When the imprecision becomes smaller than the quantum uncertainty of the resting state, the result is called a squeezed state. Squeezing the uncertainty in position requires hiding all information about the object's momentum, which in turn requires pumping in energy. The energy must, however, be added carefully to avoid just heating the ion up and increasing the uncertainties in both momentum and position.

The ideal way to pump in energy, it turns out, is known to every child who has sat on a swing and pumped her legs. Imagine that the child starts out sitting as still as she can, like an ion at rest in its trap. Her position nonetheless might be ever so slightly perturbed from equilibrium, depending on the direction of the latest gust of wind. If she now starts to pump her legs, she can amplify that slight impulse. A distant observer who was unable to discern the size or direction of the initial perturbation will nonetheless easily see the amplified motion with each pump of the child's legs. Burd *et al.* achieve the same effect by modulating the electric fields that trap the ion, relaxing and tightening its confinement in the same rhythm as the child extending and compressing her legs.

You may recall from your own childhood that the timing of this rhythm is crucial, to avoid inadvertently slowing yourself down. Specifically, depending on exactly when you extend and contract your legs, you can amplify either your initial momentum or your initial displacement, but not both, as the other will be attenuated. Hence, Burd *et al.*'s amplification of the initial quantum fluctuations in momentum at the same time attenuates the quantum fluctuations in position, which places the ion into a squeezed state (see the figure). At a later time during the experiment, the same pumping trick amplifies a perturbation of the ion's motion, which is the signal that the experimenters wish to precisely measure. In amplifying the signal, they also amplify the quantum noise back to its original unsqueezed level. As a result, even with a measurement apparatus that can barely resolve the quantum noise of the initial state, the complete protocol can detect much smaller displacements, down to the level of the squeezed uncertainty.

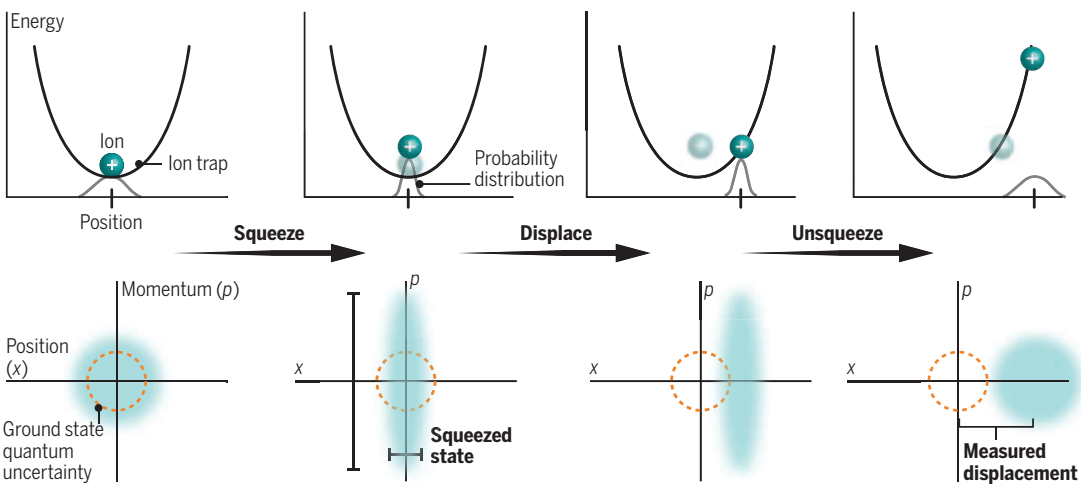
This principle of amplified quantum sensing has been used to enhance optical interferometers (*9*) and has been demonstrated in atomic clocks and sensors whose signals are stored in internal states of an ensemble of many atoms (*10*–*12*). Its application to detecting tiny displacements of a single particle, however, required surmounting different technical challenges. The authors showed amplification of the quantum motion without adding any appreciable extra noise. In other words, they pumped energy into the system without heating it. They detected a displacement of less than 1 nm, which is smaller than the initial quantum uncertainty in the ion's position by a factor of 7.

The study by Burd *et al.* is an important milestone in advancing toward ultimate quantum limits in precision measurement and opens the door to several applications. One prospect is to make ions even better at what they already do best, which is precision spectroscopy. For example, by comparing the absorption spectra of different ionic species, it is possible to address fundamental questions such as whether the constants of nature are really constant. One method of measuring these spectra is notable for being applicable to a wide range of different ions and relies on detecting the tiny momentum kick imparted when an ion absorbs a photon (*13*). Quantum squeezing and amplification can make this method of photon recoil spectroscopy even more precise.

Extended to groups of multiple ions, noiseless amplification can enhance interactions that allow for transmitting quantum information between the ions (*14*) and may thus ultimately find applications in quantum computing. Another exciting prospect is to precisely detect the motion of increasingly massive objects that can be cooled to the quantum regime (*15*). A grand challenge is to detect minute gravitational forces in a regime where quantum mechanics simultaneously plays a role. Every bit of extra precision that can be squeezed from such systems will be essential to achieving that goal. ■

Improving sensitivity to small displacements

Carefully manipulating the quantum fluctuations in an ion's motion improves the measurement precision.



The trapped ion is in its lowest energy state with an uncertainty in position and momentum (blue circle) set by quantum mechanics.

The energy of the ion is increased by modulating the electric field, which increases uncertainty in momentum but decreases it in position below the quantum limit.

The ion is displaced while in the squeezed state.

Amplifying the displacement returns the original quantum uncertainty, and both are now larger than the measurement error.

GRAPHIC: N. DESAI/*SCIENCE*

REFERENCES AND NOTES

1. J. Aasi *et al.*, *Nat. Photonics* **7**, 613 (2013).
2. M. Malnou *et al.*, *Phys. Rev. X* **9**, 021023 (2019).
3. C. M. Caves, *Phys. Rev. D* **23**, 1693 (1981).
4. L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, P. Treutlein, *Rev. Mod. Phys.* **90**, 035005 (2018).
5. S. C. Burd *et al.*, *Science* **364**, 1163 (2019).
6. T. Rosenband *et al.*, *Science* **319**, 1808 (2008).
7. R. Blatt, D. Wineland, *Nature* **453**, 1008 (2008).
8. C. Monroe, J. Kim, *Science* **339**, 1164 (2013).
9. B. Yurke, S. L. McCall, J. R. Klauder, *Phys. Rev. A* **33**, 4033 (1986).
10. E. Davis, G. Bentsen, M. Schleier-Smith, *Phys. Rev. Lett.* **116**, 053601 (2016).
11. O. Hosten, R. Krishnakumar, N. J. Engelsen, M. A. Kasevich, *Science* **352**, 1552 (2016).
12. D. Linnemann *et al.*, *Phys. Rev. Lett.* **117**, 013001 (2016).
13. Y. Wan *et al.*, *Nat. Commun.* **5**, 3096 (2014).
14. W. Ge *et al.*, *Phys. Rev. Lett.* **122**, 030501 (2019).
15. J. Chan *et al.*, *Nature* **478**, 89 (2011).

10.1126/science.aax0143

POLICY FORUM

RESEARCH FUNDING

Government-funded research increasingly fuels innovation

Nearly a third of U.S. patents rely directly on federal research

By **L. Fleming[1], H. Greene[2], G. Li[1], M. Marx[3], D. Yao[4]**

Innovation increasingly relies on scientific knowledge (*1, 2*). Research to generate that knowledge has historically been funded by both industry and government. Although industry and government research spending was relatively equal in 1980 in the United States, by 2010 their shares had shifted to 60% and 30%, respectively (*3*). Yet, despite this increase in industrial spending, firms appear to be pursuing—or at least publishing—less basic science (*4*). If corporations are doing less basic research, then where do they find the ideas to fuel their innovation? Here, we detail individual bibliometric linkages across tens of millions of documents and quantify the broad sweep and impact of U.S. federally supported research on patented innovation over most of the past century. We illustrate how patentees, both U.S. and non-U.S., and corporations in particular, increasingly depend upon federally supported research as a source of scientific knowledge. Although multiple mechanisms interact and contribute to the trend, federal research increasingly appears to fuel the innovation that ultimately leads to jobs, industrial competitiveness, and entrepreneurial success.

Though research has established the rise of "open innovation" communities, research consortia, and markets for ideas as sources of innovative ideas (*5*), the role of government funding for research has not been ignored. Recent research—at the micro level of individual science papers and patents—has quantified how 10% of grants awarded by the U.S. National Institutes of Health (NIH) generate patents (*2*). Macro-level and nonquantitative histories have also described the broad sweep of government impact (*6*). Here, we consider every U.S. patent assigned to U.S. inventors since 1926, and show that the proportion of patents relying on federally funded science has outstripped the overall increase in patenting, plateauing since a high of 30.0% in 2011 (see the first figure). Despite this plateau, the absolute number of patents that rely on federally supported research has almost doubled recently, from 22,647 in 2008 to 45,220 in 2017. We count reliance when a patent is owned by the government, acknowledges government support, or directly cites a patent or science paper that acknowledges support.

Corporations, identified as the owner of a patent from its "assignee" field, account for most of the increased reliance on government-supported research since the 1970s (see the second figure). Corporations have increased their own acknowledgment of support and their citation of government-supported science papers and patents [although alternatives cannot be entirely ruled out, see supplementary materials (SM) for evidence that the trends are robust to more versus less stringent definitions of reliance and also do not result from a change in reporting requirements or in response to requirements, an increase in the citable stock of federal research, or a mechanical process due to an increase in the number of citations per patent].

The sharp increase following World War II (WWII) illustrates the invention behind the vision articulated by Vannevar Bush in 1945, of how federal investment in science research could train the technical workforce and support innovation, resulting in "…new products, new industries, and more jobs…new and improved weapons" (*7*). The almost monotonic increase in reliance has been partially enabled by the steady though slower rise in inflation-adjusted federal research dollars over the time period, even as federal science investment as a percentage of gross domestic product has decreased (*8*).

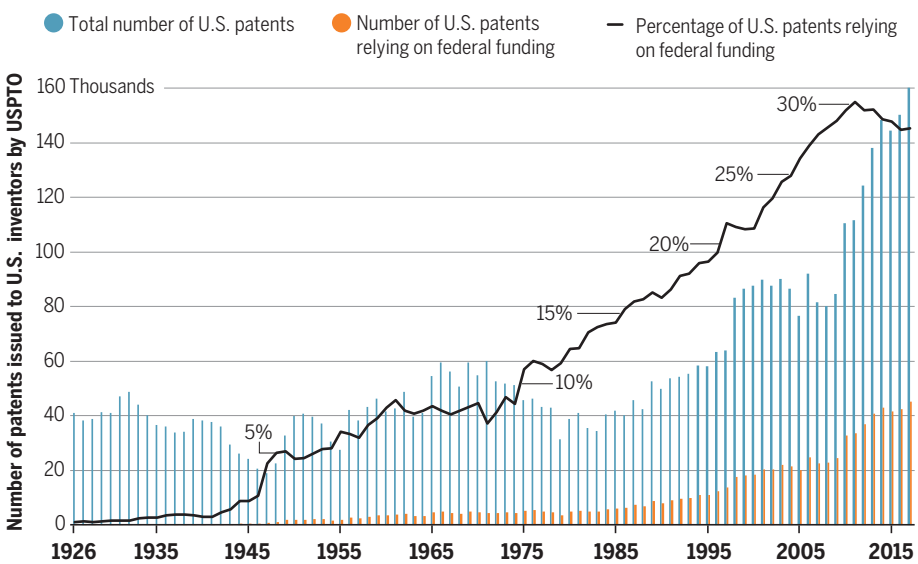
Prior to the 1970s, much of the government-supported patenting in the United States occurred inside the government. Universities, often supported by government research grants, began patenting more in the 1980s, motivated in part by the Bayh-Dole Act, which sought to spur commercialization through assignment of property rights to universities. Lone inventors constituted a large proportion of reliance upon government research in the early part of the last century, and their reliance has remained steady.

Although some portion of the overall increase shown in the first figure results from a change in the composition of patented technologies, growing investment in

[1]College of Engineering, University of California, Berkeley, CA, USA. [2]School of Law, University of Connecticut, Hartford, CT, USA. [3]Questrom School of Business, Boston University, Boston, MA, USA. [4]Harvard Business School, Harvard University, Cambridge, MA, USA. Email: lfleming@berkeley.edu

Patentees increasingly depend upon federally supported research

Total granted U.S. patents by U.S. inventors (blue bars), and subtotal that rely on federal research (orange bars), and proportion of patents (black line = orange bars/blue bars) that rely on federally supported research.



GRAPHIC: N. CARY/SCIENCE

health research, and an accompanying rise in patents supported by the Department of Health and Human Services (HHS) (which includes the NIH), all technologies have relied increasingly upon government support, albeit to varying degrees. Federal support for U.S. invention gets routed through a variety of agencies. In 2017, for example, the Department of Defense (including the armed services) supported 6.2% of the total of U.S. inventions, HHS 5.4%, Department of Energy 3.9%, National Science Foundation 2.9%, NASA 1.0%, and Department of Agriculture 0.50% (others 5.5% and unspecified 2.5%; see SM for illustrations of R&D investment and patenting by agency and by technology class).

Large, small, and micro firms all draw upon government research, as observed in patent-renewal data. Startups also depend heavily on government research, as 34.6% of the 121,765 patents assigned to venture-backed companies from 1976 to 2016 cited federally supported research; by comparison, for all corporate patents during the same time period, only 21.7% rely on federally supported research. The higher reliance of startups on government-supported research may be explained by resource constraints, which inhibit internal R&D expenditure in young ventures and encourage identification and external licensing of promising technologies. It also illustrates the intended consequence of the Bayh-Dole Act and an important path of commercialization for federal science and technology research.

Foreign inventors' reliance on federal research in their U.S. patenting has steadily increased since WWII (see the third figure), though the proportion of total foreign patenting in the U.S. system that relies on federal research still lags behind that of U.S. inventors (in 2017, 28.2% of U.S. patents by U.S. inventors, and 12.4% of the U.S. patents by foreign inventors, relied on federally supported research). China has recently begun to rely more heavily on U.S. federal research; however, its reliance is still small relative to most (in order of reliance in 2017): Japan, Germany, Korea, England, France, China, Taiwan, and India. The greatest recent increase has come from the category of "other" nations, perhaps reflecting easier access via the web and other electronic libraries to the patent and scientific literature.

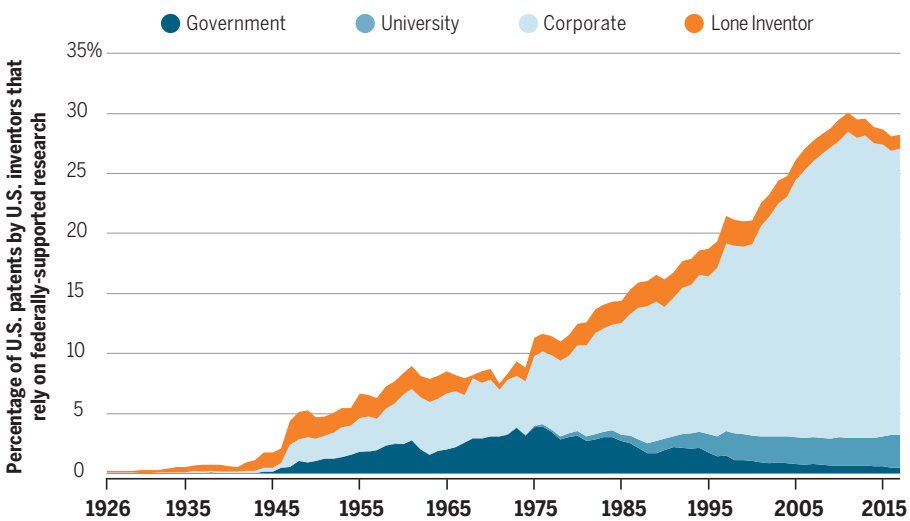
Many factors make it more difficult for foreign inventors to gain access to U.S. science—including language differences, the cost of international travel, fewer social ties between scientists and engineers, problems of security clearances and access to sensitive technologies, and surely less mobility from (U.S.) university positions into private firms in foreign countries. Moreover, countries vary in their degree of local specialization and may find U.S. research less helpful, if working in areas not supported by U.S. research. Although other nations' inventors have increased their reliance on federal research, the rate of increase appears less than that of U.S. inventors, even as the world's scientific and technical knowledge has become digitized and thus more easily searched and used.

Corporate patents that rely on federal research are consistently more important than those that do not, as measured by the number of prior art citations from subsequent patents (such citations establish the departure point of a new patent). Comparing population averages for all U.S. corporate patents granted in 2010, patents that rely on federal research receive 6.33 citations on average in the 5 years following the grant versus 4.42 for those that do not. Additional panels in 2000, 1990, and 1980 show similar but not always significant effects, possibly because of thinner data or changes in corporate patenting strategy (see SM). These population averages aggregate a great variety of different technologies, industry-specific patenting strategies, and number of inventors on the patent, however, which motivates a comparison of similar types of patents. Matching and pairing corporate patents that cite federally supported patents to similar corporate patents that did not (see SM), the former receive on average 3.39 more citations in the 5 years after issuance. The citation increase arises from both the firm's subsequent patents (1.53 additional self-citations on average) and citations in other firms' patents as well (1.86 additional citations from other organizations on average), indicating that both the inventing firm and its competitors find these technological trajectories more fertile.

The same pattern holds for corporate patents that cite federally supported science publications relative to those that cite science publications that are not government supported, with a matched and paired mean difference of 3.57 citations, and a mean difference in self- and nonself-citations of 1.17 and 2.41, respectively. Prior research has estimated that an additional citation is associated with an increase in the value of the patent of up to 3% (*9*).

Reliance upon federal support has been dominated by corporations

Stacked area chart of percentage of total U.S. patents by U.S. inventors that are owned by government or rely on federally supported research papers or patents, by type of patent owner.



Citation increases also benefit the larger economy as citations that occur from outside the inventing firm can be interpreted as knowledge spillovers and positive externalities from the inventing firm to the larger economy (*10*).

Consistent with the citation measure and again comparing matched patents, patents that rely on federal research are slightly more likely (<2%) to pay renewal fees (required payments that keep the patent in force) after 4 years. They are also more likely, on average, to introduce words that are new to the patent corpus, indicating foundational patents with greater novelty (*11*).

Although the bulk of reliance historically can be traced through citation of prior patent art, U.S. inventors have recently shifted their citation patterns to favor more im-

GRAPHIC: N. CARY/*SCIENCE*

mediate citation of government-supported science papers. The average percentage of prior art citations within each patent that are supported by federal research fell from a high of 15.8% in 2003 to 9.9% in 2017. By contrast, citation of scientific papers that are supported by federal research reached a historical high of 10.7% in 2017, up from 6.9% in 2004 (see SM).

If federal science research is so important to corporations, why don't they undertake all (or at least more of) their research internally? Putting aside the question of the optimal rate and types of corporate taxation, why should the government subsidize private firms through federally funded and managed research investment? Much research has argued and demonstrated that corporations probably underinvest in research (*12*), owing to uncertainty and the difficulty of commercializing research by the firm doing the research (as opposed to that firm's competitors). Furthermore, from a societal viewpoint, the discoveries from research are best distributed widely throughout an economy, rather than being kept in a single firm, thus reducing duplicative efforts. Research aimed at fundamental understanding thus constitutes a public good, typically precedes commercialization, and can be applied by many corporations, across many industries, for many years following publication. Although these results might be interpreted as evidence of short-sighted lack of research investment on the part of corporations, they can also be interpreted as evidence of the effectiveness of public investment in science and technology research in spurring innovation.

Our bibliometrics enable explicit linkage of government support and output but remain vulnerable to many limitations in methodology, interpretation, and causal attribution. Patents are only one—and an admittedly imperfect—measure of innovation, particularly prior to the availability of comprehensive data in the mid-1970s. Moreover, the use of patents varies across industries and over time (though the matched analysis above controls for technology, number of inventors, and time). Our focus does not extend, for example, to innovations protected as trade secrets or through copyright. Other factors that complicate causal interpretation include the concurrent (though slower) rise in federal spending on science over the period, changes in compliance requirements (*13*), changes in corporate patenting and citation strategies, and improvements in database and search technology that enabled increasingly accurate acknowledgment of government support (see SM for discussion).

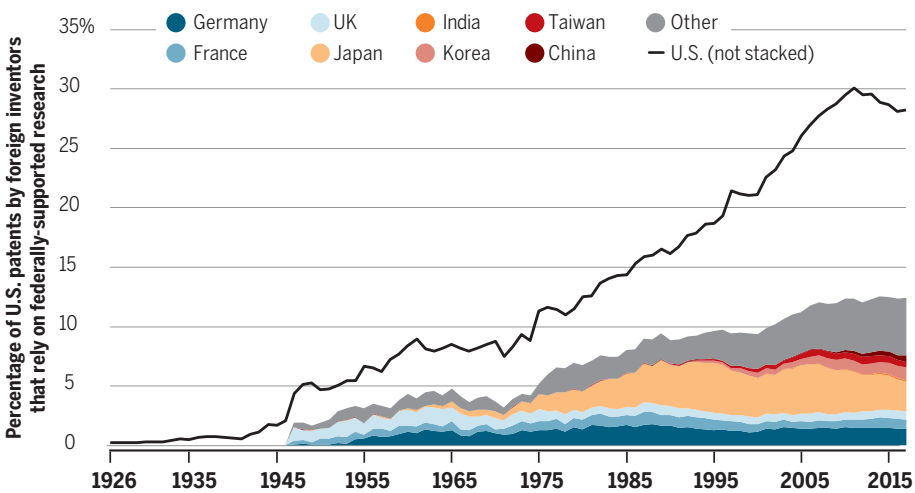
Furthermore, this study does not compare observed output to the counterfactual: patenting and publishing that would have (not) occurred in the absence of federally supported research. Hence, one cannot estimate the extent to which federal support of R&D has crowded out private investment, though there is increasing evidence that thoughtful applications of such support can encourage additional and complementary private investment (*2*).

These issues notwithstanding, these data provide a quantitative estimate of the changing impact of federal research on U.S. patenting. They show that almost one-third of U.S. invention—and the more important part as measured by future citations, renewals, and novelty—relies on federal research investment. Our measures conservatively use only direct linkages, ignoring dubious or missing acknowledgments of support as well as knowledge flow from second- and later-generation citations. These analyses also ignore the nonpecuniary (and possibly more substantial) benefits such as better weapons, improved health and medical practices, technical workforce training, directly and indirectly supported jobs, and entrepreneurial ideas that spark startups (Google's Page Rank patent acknowledges government support, for example).

If the sponsoring nation's firms can more effectively take advantage of that nation's research—as appears to be the case to date in the United States—then firms can exploit a fundamental competitive advantage in international competition. Federal research increases the likelihood that new industries will locate in that country and provide jobs and productivity spillovers to older industries. Policy-makers should consider these newly observable benefits of federal research when they formulate tax policy and science research budgets.

Foreign reliance on federal research has steadily increased

Proportion of U.S. patents by non-U.S. inventors that rely on federally supported research, stacked by country (for comparison, the black line above is not stacked, and illustrates the proportion of U.S. patents by U.S. inventors that rely on federally-supported research, from the first and second figures).



GRAPHIC: N. CARY/*SCIENCE*

REFERENCES AND NOTES

1. M. Ahmadpoor, B. F. Jones, *Science* **357**, 583 (2017).
2. D. Li, P. Azoulay, B. N. Sampat, *Science* **356**, 78 (2017).
3. National Science Board, Science & Engineering Indicators 2018: Figure 4-4, "U.S. Total R&D Expenditures, by Source of Funds: 1953-2015"; https://nsf.gov/statistics/2018/nsb20181/report.
4. A. Arora, S. Belenzon, A. Patacconi, *Nat. Index* 6 (2017).
5. E. von Hippel, *Democratizing Innovation* (MIT Press, 2005).
6. M. Mazzucato, *The Entrepreneurial State* (Perseus Books, 2013).
7. V. Bush, *Science: The Endless Frontier* (U.S. Government Printing Office, Washington, D.C., 1945).
8. https://www.aaas.org/programs/r-d-budget-and-policy/historical-trends-federal-rd.
9. B. Hall *et al.*, *Rand J. Econ.* **36**, 16 (2005).
10. A. B. Jaffe *et al.*, *Q. J. Econ.* **108**, 577 (1993).
11. B. Balsmeier *et al.*, *J. Econ. Manage. Strategy* **27**, 535 (2018).
12. K. Arrow, *Rev. Econ. Stud.* **29**, 155 (1962).
13. A. K. Rai, B. N. Sampat, *Nat. Biotechnol.* **30**, 953 (2012).
14. E. Garfield, *J. Pat. Off. Soc.* **39**, 583 (1957).
15. L. Fleming *et al.*, Replication Data for: Government-funded research increasingly fuels innovation. Harvard Dataverse, V4 (2019); https://doi.org/10.7910/DVN/DKESRC.

ACKNOWLEDGMENTS

The authors acknowledge support from NSF 1536022, the Coleman Fung Institute for Engineering Leadership, the University of Connecticut School of Law, and the Harvard Business School. We thank M. Ewens for sharing data and B. Balsmeier and D. Gross for helpful comments. All opinions and errors remain with the authors. The authors declare no competing interests. Data and code are posted on Harvard Dataverse (*15*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1139/suppl/DC1

10.1126/science.aaw2373

People can become more empathic with time and practice, finds Zaki.

BOOKS *et al.*

SOCIAL SCIENCE

Exercising empathy

Compassion can be cultivated, argues a psychologist

By **Sara Konrath**

Google searches for "empathy" have been increasing over the past 15 years, and it's the latest buzzword at companies as diverse as Facebook and Ford. Joining several recent books on the topic [such as (*1-3*)] is *The War for Kindness* by Stanford psychologist Jamil Zaki, which sets out to help people increase their empathy in sustainable ways.

Zaki opens the book by revealing that many people believe that empathy is a fixed trait with genetic underpinnings. He calls this the "Roddenberry hypothesis," named after Gene Roddenberry, the creator of *Star Trek*, a show that featured characters at the extremes of the empathy scale.

Zaki is deeply opposed to this idea. His research finds that people can become more empathic with practice, and when people believe empathy is changeable, they spend more time and effort helping others (*4*). This is important in a time of both declining empathy (*5*) and rising political polarization.

The War for Kindness
Building Empathy in a Fractured World
Jamil Zaki
Crown, 2019. 268 pp.

Zaki is a compelling writer, and even an android could not help but respond to his prose. His chapters, which treat readers to a broad tour of the most interesting recent empathy research, read like a series of long essays. This is a clever approach because, as he notes, one effective empathy-building technique is storytelling. Zaki opens the book with the story of his parents' acrimonious divorce to show how powerful empathy can be when used as a tool to maintain relationships. By embracing each of their perspectives, he writes, he was able to remain connected to both parents even as they were increasingly disconnected from each other.

The chapter "Caring too much" starts with another personal example: the story of the birth of Zaki's daughter Alma, who suffered a stroke during a difficult birth, and the compassionate staff who cared for her during the early days of her life. He then artfully transitions into how neonatal nurses and others in similar professions can sometimes be negatively affected by others' suffering. This is a very important discussion, especially in light of recent debates on the potential negative implications of empathy (*2*, *3*).

Yet after sharing the potential harms associated with too much empathy, Zaki confusingly shifts to the opposite conclusion at the end of the same chapter, arguing that empathy helps to protect people from negative effects of stress. This betrays a common problem with the term "empathy," which is its many definitions (*6*). If empathy is defined as personal distress or emotion sharing, there is indeed research that shows negative implications (*7*). Yet most empathy scholars would define it as more akin to compassion. Using this definition, several studies find that empathy actually serves as a buffer for burnout (*8, 9*).

The book includes extensive footnotes, a detailed appendix on the definitions of empathy, and a second appendix that attempts to evaluate the evidence discussed in the book in an accessible way. Scientists will likely appreciate these sections, but they may not be satisfied by them, and lay readers will probably ignore them. Still, I understand why Zaki included them: to appease persnickety academics such as me. Academic audiences, however, would ultimately be better served by more scholarly tomes [such as (*10, 11*)].

The War for Kindness may inspire readers to learn more, but Zaki's goals go beyond sharing the science of empathy with the masses. He hopes to inspire people to actually practice more kindness in their lives, writing in the book's closing lines, "Building a new sort of empathy takes effort and sacrifice, for people who might not repay it.... We each have a choice, and the sum of our choices will create the future. What are you going to do?"

Although he does not offer clear action points, the stories told throughout the book reveal many ways to increase empathy. Ultimately, however, there is no quick fix. Building empathy takes time and effort (*12*). Take heart, though; it gets easier with practice. ■

The reviewer is director of the Interdisciplinary Program on Empathy and Altruism Research, Indiana University Lilly Family School of Philanthropy, Indianapolis, IN 46202, USA. Email: skonrath@iupui.edu

REFERENCES AND NOTES

1. H. Riess, L. Neporent, *The Empathy Effect: Seven Neuroscience-Based Keys for Transforming the Way We Live, Love, Work, and Connect Across Differences* (Sounds True, 2018).
2. P. Bloom, *Against Empathy: The Case for Rational Compassion* (Ecco, 2016).
3. F. Breithaupt, *The Dark Sides of Empathy* (Cornell Univ. Press, 2017).
4. K. Schmann *et al.*, *J. Person. Soc. Psych.* **107**, 3 (2014).
5. S. H. Konrath *et al.*, *Pers. Soc. Psychol. Rev.* **15**, 2 (2011).
6. S. Konrath, *Greater Good Magazine*, 24 January (2017).
7. O. Klimecki, T. Singer, in *Pathological Altruism* (Oxford Univ. Press, 2012), pp. 368–383.
8. E. Gleichgerrcht, J. Decety, *PLOS ONE* **8**, 4 (2013).
9. T. D. Shanafelt *et al.*, *J. Gen. Intern. Med.* **20**, 7 (2005).
10. M. H. Davis, *Empathy: A Social Psychological Approach* (Westview Press, 1995).
11. C. D. Batson, *Altruism in Humans* (Oxford Univ. Press, 2011).
12. C. D. Cameron *et al.*, *J. Exp. Psych.* **148**, 962 (2019).

10.1126/science.aax2199

PHOTO: GEOVIEN SO/SOPA IMAGES/LIGHTROCKET VIA GETTY IMAGES

PALEONTOLOGY

The dinosaur as social capital

A new history reveals how the wealthy elite helped shape modern natural history museums in America

By **Ilja Nieuwland**

Assembling the Dinosaur
Fossil Hunters, Tycoons, and the Making of a Spectacle
Lukas Rieppel
Harvard University Press, 2019. 335 pp.

The first 6 months of 1905 saw two prominent, celebrity-studded, and copiously publicized unveilings of full-size dinosaur exhibits. In February, the American Museum of Natural History (AMNH) in New York unveiled its *Brontosaurus* mount; in April, Andrew Carnegie's donation of a cast of *Diplodocus carnegii* (his "namesake," as he liked to call it) was presented to the British Museum in London. As Lukas Rieppel explains in his book, *Assembling the Dinosaur*, neither the timing nor the place nor the form of these events rested on coincidence.

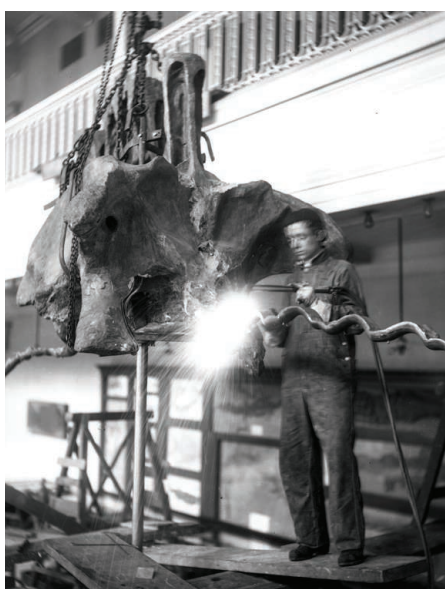
In the past decade, the "Long Gilded Age" of American natural history museums—roughly the last third of the 19th century and the first decade of the 20th—has been subjected to an unprecedented number of treatments (*1*–*3*). Rieppel's book is a welcome addition to this body of literature, not in the least because of his approach, which seeks to integrate dinosaurs' status "as an object of technical knowledge with social, cultural, and economic history."

The result is a book about dinosaurs that features surprisingly few of them (and fewer fossil hunters still). If that sounds like criticism, it is not. Rieppel's focus is on the people who used dinosaur mounts to convert monetary capital into social capital and, in so doing, built the natural history museums that define the museum landscape today.

At the turn of the 20th century, a time in which big things mattered, dinosaurs were a natural point of attraction for wealthy plutocrats. And because natural history museums could be instrumentalized to translate them into social capital and ideological metaphor, this created an almost natural bond between such institutions and wealthy benefactors looking for recognition.

Rieppel frames America's big natural history museums as tools for the dissemination of ideas about capitalism and society as much as they are tools for disseminating ideas about natural history. He shows how various financial and administrative mechanisms from capitalist ventures (particularly industry) shaped the operations of natural history museums and led to a strict regulation of both their internal and external relations.

The main focus is on the AMNH, arguably the United States's most important natural history museum, and its main figurehead, the paleontologist Henry Fairfield Osborn. In a sense, this book is a careful deconstruction of Osborn's efforts to instrumentalize natural history for ideological ends—ideologies that usually served the agenda of the AMNH's backers. A good example of this would be his advocacy of "aristogenesis," a concept that holds that evolution proceeds in a predetermined path that favors the "best" genes—an idea obviously attractive to patricians at the top of New York's social order.



The armature for a *Diplodocus* is welded at the Carnegie Museum of Natural History, ca. 1904.

PHOTO: CARNEGIE MUSEUM OF NATURAL HISTORY

The same events were taking place in other new institutions, including Chicago's Field Museum and Pittsburgh's Carnegie Museum, both of which feature prominently in Rieppel's narrative. He describes how wealthy philanthropists such as Marshall Field and Andrew Carnegie shaped these institutions in much the same form as they had done with their commercial ventures: by applying a strict social hierarchy and rigorous bookkeeping throughout their respective organizations.

With the exception of Osborn, Rieppel may be somewhat guilty of underestimating intermediary figures such as William Holland and Frederick Skiff, who actually converted their proprietors' ideas into practice. These museum directors were not just delegates of their wealthy proprietors; they were their social equals and representatives, and although they rarely misunderstood the museum's role in generating social capital for their bosses, they also had agendas of their own. This could make things difficult when the museum's interests failed to coincide with the interests of its patrons. Even Holland, for example—no stranger to sycophancy—occasionally issued stern reprimands to Carnegie when he felt his boss was losing sight of the museum's interests.

Unlike the previous generation of collectors, the pressure was on new museum curators to prove themselves as the defenders of "pure" scientific interests. But, as Rieppel emphasizes, they eventually came to serve the same sociocommercial agenda because "pure" science was a highly valued ideological commodity whose aura could be made to fit more openly commercial agendas, too.

Assembling the Dinosaur is a solid entry into the growing body of literature on Gilded Age American paleontology, but it is particularly valuable for its contribution to enhancing our understanding of how science and its representation during that period were influenced by, and in turn affected, society as a whole. By incorporating cultural, economic, and scientific developments, Rieppel shines new light on the history of both American paleontology and museum exhibition practice. ■

The reviewer is at the Huygens Institute for the History of the Netherlands, the Royal Netherlands Academy of Arts and Sciences, Oudezijds Achterburgwal 185, 1012 DK Amsterdam, Netherlands. Email: ilja.nieuwland@huygens.knaw.nl

REFERENCES AND NOTES

1. P. D. Brinkman, *The Second Jurassic Dinosaur Rush: Museums and Paleontology in America at the Turn of the Twentieth Century* (Univ. Chicago Press, 2010).
2. R. J. Gangewere, *Palace of Culture: Andrew Carnegie's Museums and Library in Pittsburgh* (Univ. Pittsburgh Press, 2011).
3. I. Nieuwland, *American Dinosaur Abroad: A Cultural History of Carnegie's Plaster Diplodocus* (Univ. Pittsburgh Press, 2019).

10.1126/science.aax5532

Only three known Yangtze giant softshell turtles remain.

PHOTO: JOEL SARTORE/NATIONAL GEOGRAPHIC

LETTERS

Edited by **Jennifer Sills**

Invest in amphibians and reptiles

The Yangtze giant softshell turtle (*Rafetus swinhoei*) is listed in Appendix I of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (*1*) and categorized as Critically Endangered by the International Union for Conservation of Nature (*2*). Threats to *Rafetus swinhoei* have long been known, including habitat loss caused by humans' excessive hunting for food and trade, resource competition, climate change, disease, and genetic stochasticity (*3*–*5*). Despite our understanding of the risks, on 14 April, the only known female *Rafetus swinhoei* died in China, leaving one captive male in China and two known wild males in Vietnam. Unless another female is found in the wild, the species is now functionally extinct (*6*, *7*).

More investment in conservation may have been able to save this species. However, insufficient funding still hinders amphibian and reptile conservation, and public awareness of wildlife protection remains weak. China must further expand its nature reserves and implement scientific management and protection policies to save other endangered amphibians and reptiles. In addition, developed countries and international organizations should help developing countries that lack the capital, management, and technology required to protect reptiles and amphibians. Long-term international collaborations could prevent other endangered amphibians and reptiles from following in the path of *Rafetus swinhoei*.

Huinian Liu[1,2], Xin Li[1,2]*, Chang Zhang[1,2]
[1]College of Environmental Science and Engineering, Hunan University, Changsha 410082, China. [2]Key Laboratory of Environmental Biology and Pollution Control (Hunan University), Ministry of Education, Changsha 410082, China.
*Corresponding author. Email: hgxlixin@hnu.edu.cn

REFERENCES AND NOTES

1. C. J. Xia *et al.*, *Pakistan J. Zool.* **45**, 5 (2013).
2. IUCN, Red List of Threatened Species, Yangtze Giant Softshell Turtle (2018); www.iucnredlist.org/species/39621/97401328.
3. J. Chen *et al.*, *Mod. Agricult. Tech.* **11**, 224 (2017) [in Chinese].
4. Y. Xi, *Green China* **15**, 68 (2013) [in Chinese].
5. T. N. Engstrom, H. B. Shaffer, W. P. Mccord, *Biol. Conserv.* **104**, 173 (2002).
6. "The only known female Yangtze Giant Softshell Turtle dies in China, and there are only three left in the world" (2019); www.nbd.com.cn/articles/2019-04-14/1320976.html [in Chinese].
7. "The world's fourth surviving *Rafetus swinhoei* has been discovered in Vietnam" (2018); http://news.sina.com.cn/o/2018-04-12/doc-ifyzeyqc2180200.shtml [in Chinese].

10.1126/science.aax7259

Brazil's biodiversity researchers need help

Between 2006 and 2014, Brazil experienced a boom in its national science and technology budget (*1*) that fomented local research groups and enhanced international collaboration, resulting in the rise of Brazil in scientific rankings (*2*). During that period, Brazil engaged in environmental conservation (*3*), earning recognition as a global biodiversity leader (*4*). However, recently elected president Jair Bolsonaro gutted the budget for science and higher education, causing funding agencies to run out of cash by July 2019 (*5*) and universities to struggle to pay maintenance bills (*6*). Environmental regulations are collapsing, benefiting major landowners and commodity companies that demand easier environmental licensing, permits for harmful pesticides, and elimination of protected areas (*7*). The advances in science and conservation from the past decade will soon be reversed, which has caused international concern (*8*). As Brazilian researchers struggle to adjust to the national budget and biodiversity crises, the global biodiversity community can help by recognizing the challenges Brazil's scientific community now faces.

Many international scientific societies are dedicated to fostering research through scientific meetings, research grants, and scientific publications. Researchers from low-income and lower-middle-income countries are often offered grants [e.g., (*9*)], discounts for membership and meetings [e.g., (*10*)], and reduced publication fees (*11*). Brazilian scientists have not been eligible for these incentives since 2006 (*12*), which made sense before the recent backslide in policy. However, in Brazil's new antiscience climate, the eligibility requirements exclude an important share of qualified researchers now working under precarious conditions. We urge the international scientific community to help ameliorate the crisis by making Brazilian researchers eligible for incentives.

Maria Tereza Chiarioni Thomé[1]* and Célio Fernando Baptista Haddad[2]
[1]Postgraduate Program in Biological Diversity and Conservation in the Tropics, Institute of Biological and Health Sciences, Federal University of Alagoas, Maceió, AL, Brazil. [2]Department of Zoology, Institute of Biosciences, São Paulo State University, Rio Claro, SP, Brazil.*Corresponding author.
Email: mtcthome@gmail.com

REFERENCES AND NOTES

1. Coordenação de Indicadores e Informação (COIND), Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), "Brasil: Dispêndio nacional em ciência e tecnologia (C&T), em valores correntes, por atividade, 2000-2016" (2018); www.mctic.gov.br/mctic/opencms/indicadores/detalhe/recursos_aplicados/indicadores_consolidados/2.1.1.html.
2. D. Cross, S. Thomson, A. Sinclair, "Research in Brazil: A report for CAPES by Clarivate Analytics" (Clarivate Analytics, 2018); www.capes.gov.br/images/stories/download/diversos/17012018-CAPES-InCitesReport-Final.pdf.
3. R. Mittermeier *et al.*, *Nat. Conserv.* **8**, 197 (2010).
4. R. A Mittermeier *et al.*, *Megadiversity: Earth's Biologically Wealthiest Nations* (Cemex, Mexico, 1997).
5. C. Angelo, *Nature* **568**, 155 (2019).
6. Associated Press, "Brazil plans to slash funding of universities by 30 percent," *The New York Times* (2019).
7. D. Abessa, A. Famá, L. Buruaem, *Nat. Ecol. Evol.* **2019**, 1 (2019).
8. H. Araújo, "Save the Amazon from Bolsonaro," *The New York Times* (2019).
9. The Darwin Initiative (www.gov.uk/government/groups/the-darwin-initiative).
10. Registration Fees, 56th Annual Meeting of the Association for Tropical Biology and Conservation (www.atbc2019.org/registration-fees).

11. Wiley, Waivers and Discounts (https://authorservices.wiley.com/open-research/open-access/for-authors/waivers-and-discounts.html).
12. The World Bank, Data, "Historical classification by income" (http://databank.worldbank.org/data/download/site-content/OGHIST.xls), "Country analytical history" tab, row 38.

10.1126/science.aax9478

China's ineffective plastic solution to haze

The Beijing-Tianjin-Hebei (BTH) region of China, one of the most densely populated regions in the world (*1*, *2*), suffers from severe atmospheric haze (*2*, *3*). In an effort to minimize fine dust blown by the wind [called eolian dust (*4*)], a precursor to haze (*2*), local governments have adopted a strategy of covering all bare land and soil in the region with plastic gauze [e.g., (*5*)]. However, the plastic has no effect on haze and causes soil pollution, increasing ecological risk.

Statistically, persistent haze in the BTH region occurs during low wind speeds (less than 2 m/s), allowing an accumulation of fine particles [e.g., particulate matter with a diameter of less than 2.5 μm ($PM_{2.5}$)] in the stagnant air (*3*). Although covering land with gauze may be useful to prevent localized dust generation during high winds, the gauze has no direct effect in the fight against haze (*6*–*8*). However, using gauze to cover land introduces a substantial amount of plastic into the soil, which can affect soil properties. Meanwhile, plastic decomposition may be a source of the microplastics in water and coastal sediments of the region (*9*). The plastic cover also hinders plant recolonization and harms natural habitats, including the habitat of migratory birds, which use the coast of Bohai Bay as an important staging site on the East Asian-Australasian Flyway (*10*).

It is clear that the vast amount of plastic deposited in the environment does not help ameliorate haze, but it is causing ecological problems. Policy-makers must stop treating one environmental issue by creating another one. Instead of plastic cover, the bare soil and land should be rewilded with ecological restoration (*11*).

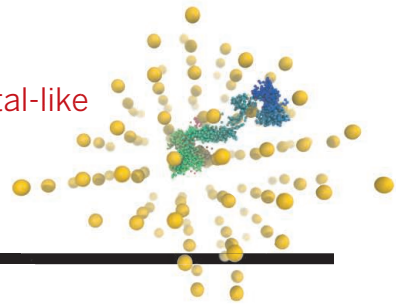
Xueqiang Lu[1,2,3]*, Rolf D. Vogt[3,4], Haixiao Li[1,2,3], Suqin Han[5], Xunqiang Mo[6], Yufen Zhang[1], Sami Ullah[7], Chen Chen[8], Xiaoxin Han[1,2,3], Hongyuan Li[1,2,3], Hans M. Seip[4]
[1]College of Environmental Science and Engineering, Nankai University, Tianjin 300350, China. [2]Tianjin Key Laboratory of Environmental Technology for Complex Trans-Media Pollution, Tianjin 300350, China. [3]Tianjin International Joint Research Center for Environmental Biogeochemical Technology, Tianjin 300350, China. [4]Department of Chemistry, University of Oslo, 0315 Oslo, Norway. [5]Tianjin Institute of Meteorological Science, Tianjin 300074, China. [6]College of Geographic and Environmental Sciences, Tianjin Normal University, Tianjin 300387, China. [7]School of Geography, Earth, and Environmental Sciences, University of Birmingham, Birmingham B15 2TT, UK. [8]Tianjin Huanke Future Ecological Technology Company, Tianjin 300191, China.
*Corresponding author. Email: luxq@nankai.edu.cn

REFERENCES AND NOTES

1. M. Tan *et al.*, *Appl. Geogr.* **90**, 239 (2018).
2. H. Xu *et al.*, *Sci. Total Environ.* **658**, 280 (2019).
3. H. Fu, J. Chen, *Sci. Total Environ.* **578**, 121 (2017).
4. T. M. Zobeck, R. S. Van Pelt, *J. Hazard. Mater.* **132**, 26 (2006).
5. Beijing Municipality Government, Beijing International, "Integrated measures and collaborative authorities to promote the refined management of fine dust" (2019); www.beijing.gov.cn/ywdt/zwzt/ltbwz/tj/t1583679.htm [in Chinese].
6. J. F. Kok, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1016 (2011).
7. R. Kimura, M. Shinoda, *Geomorphology* **114**, 319 (2010).
8. J. H. Seinfeld, S. N. Pandis, Eds., *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (John Wiley & Sons, ed. 3, 2016).
9. A. A. de Souza Machado *et al.*, *Environ. Sci. Technol.* **52**, 9656 (2018).
10. S. Xia *et al.*, *Biol. Conserv.* **210**, 72 (2017).
11. A. Perino *et al.*, *Science* **364**, 351 (2019).

10.1126/science.aax5674

RESEARCH

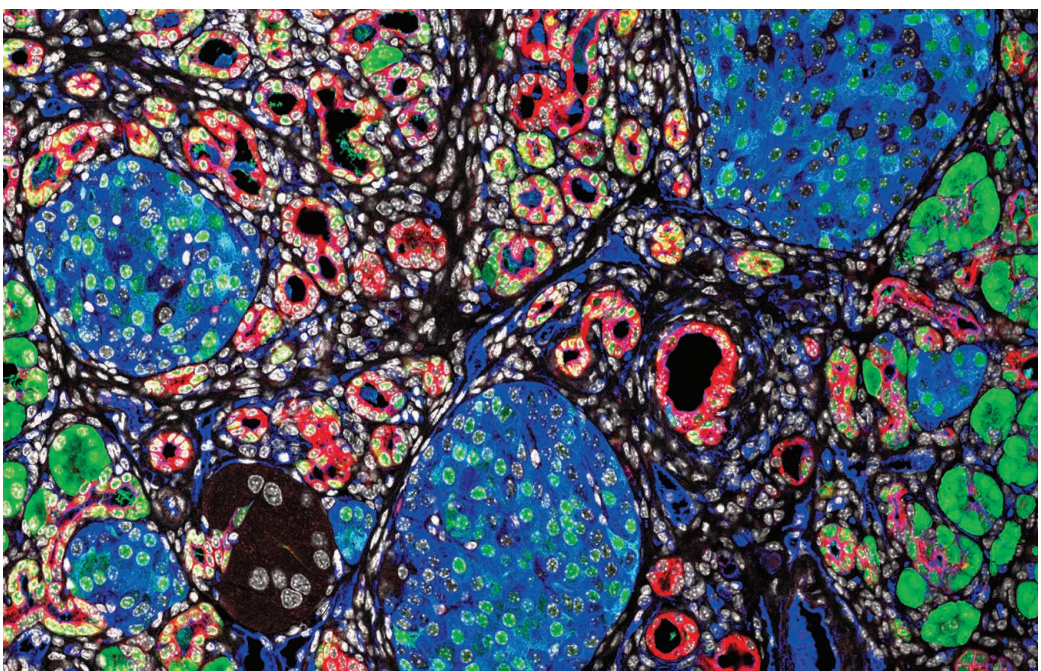


Colloids with metal-like particle mobility

Girard et al., *p. 1174*

IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**



CANCER

Sweet bystander becomes a villain

Patients with pancreatic cancer often have elevated blood levels of CA19-9, a carbohydrate antigen present on many proteins. CA19-9 is thus commonly used as a biomarker for diagnosing and monitoring disease progression. In a study of mice, Engle *et al.* found that CA19-9 may be more than an innocent bystander that marks the presence of pancreatic disease; it may play a causal role in disease (see the Perspective by Halbrook and Crawford). Transgenic mice expressing the human enzymes that add CA19-9 to proteins developed severe pancreatitis that could be reversed by treatment with CA19-9 antibodies. When the transgenic mice also harbored a *Kras* oncogene, they went on to develop pancreatic cancer. These unexpected observations suggest new avenues for the treatment of pancreatic disease. —PAK

Science, this issue p. 1156; see also p. 1132

An immunofluorescence image of a pancreas in a mouse model of pancreatic disease

PHYSICS

Improving precision with quantum amplification

Quantum mechanically, an object can be described by a pair of noncommuting observables, typically by its position and momentum. The precision to which these observables can be measured is limited by unavoidable quantum fluctuations. However, the method of "squeezing" allows the fluctuations to be manipulated, while preserving the Heisenberg uncertainty relation. This allows improved measurement precision for one observable at the expense of increased fluctuations in the other. Burd *et al.* now show that an additional displacement of a trapped atom results in amplification of the squeezing and a further improvement in the precision with which the displacement can be determined (see the Perspective by Schleier-Smith). This technique should be useful for a number of applications in metrology. —ISO

Science, this issue p. 1163; see also p. 1137

RADIOTRACER CHEMISTRY

A metal-free route to PET probes

Positron emission tomography (PET) is a widely used imaging technique for medical diagnostics and pharmaceutical development. As the name implies, it requires tracers that emit positrons, typically through labeling with fluorine or carbon radioisotopes. W. Chen *et al.* devised a versatile technique to incorporate radioactive fluoride into aromatic rings. The metal-free photochemical method directly substitutes aryl carbon-hydrogen bonds with [^{18}F]fluoride and so is particularly well suited to late-stage transformation of complex molecules into tracers. —JSY

Science, this issue p. 1170

CORAL REEFS

Little fish make a big contribution

Coral reefs represent one of the most biodiverse and rich ecosystems. Such richness conjures up images of coral heads and large colorful reef fishes. Brandl *et al.* show, however, that one of the most striking and important parts of the reef ecosystem is almost never seen

CREDITS (TOP TO BOTTOM): GIRARD *ET AL.*; DANNIELLE ENGLE, COLD SPRING HARBOR LABORATORY

Published by AAAS

(see the Perspective by Riginos and Leis). Small cryptobenthic fish, like blennies, make up nearly 40% of reef fish biodiversity. Furthermore, the majority of cryptobenthic fish larvae settle locally, rather than being widely dispersed, and have rapid turnover rates. Such high diversity and densities could thus provide the biomass base for larger, better-known reef fish. —SNV

Science, this issue p. 1189; see also p. 1128

BIOCATALYSIS

Light teaches (co)enzymes new tricks

Light is widely used in organic synthesis to excite electrons in a substrate or catalyst, opening up reactive pathways to a desired product. Biology uses light sparingly in this way, but coenzymes such as flavin can be driven to excited states by light. Biegasiewicz *et al.* investigated this reactivity and found a suite of flavoenzymes that catalyze asymmetric radical cyclization when exposed to light. "Ene"-reductases, when reduced and illuminated, converted starting materials containing an α-chloroamide and an alkene into five-, six-, seven-, or eight-membered lactams. Different enzymes furnished different stereochemistry in the products, likely because of changes in active-site pocket geometry. —MAF

Science, this issue p. 1166

STRUCTURAL BIOLOGY

Keeping the gate open

Cystic fibrosis is a progressive disease that affects lung function and is often fatal. It is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR). One class of mutants impairs ion conductance, and the drug ivacaftor acts by increasing the ion flux. Liu *et al.* describe high-resolution structures of CFTR bound to ivacaftor and to an investigational drug GLPG1837 that also potentiates ion flow. The two drugs bind at the same site in the transmembrane region. This site coincides with a hinge involved in channel gating, suggesting that the drugs may stabilize the open conformation of the channel. —VV

Science, this issue p. 1184

ISLET TRANSPLANTATION

Revascularizing eye-lets

Pancreatic islet transplantation is a potentially promising therapy for type 1 diabetes, but poor revascularization hinders islet long-term viability. Figueiredo *et al.* studied the role of protein tyrosine phosphatase 1B (PTP1B) in regulating islet vascularization. Islets from PTP1B$^{-/-}$ mice retained more endothelial cells during in vitro culture and restored normoglycemia when transplanted into the anterior chamber of the eye in diabetic mice. Future work is needed to determine whether targeting PTP1B can improve islet transplantation in human patients. —CC

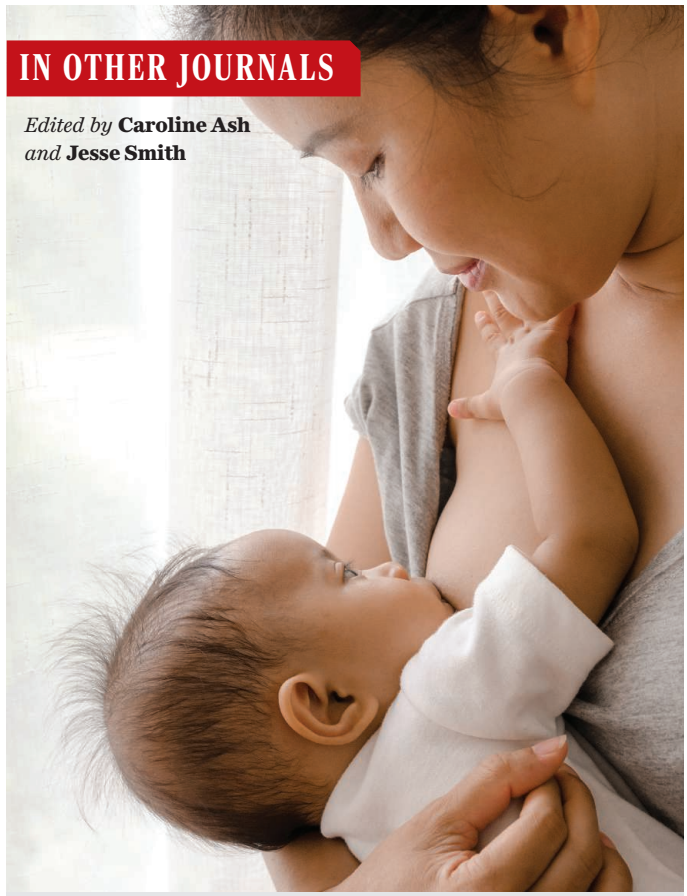
Sci. Transl. Med. **11**, eaar6294 (2019).

NEURODEVELOPMENT

Rescued by a dietary supplement

Patients with Rett-like syndrome show impaired motor and cognitive development. Soto *et al.* studied this disorder in a 5-year-old patient with a mutation in an *N*-methyl-D-aspartate (NMDA) receptor subunit. The electrophysiological and morphological defects in neurons expressing this mutant were improved by the NMDA receptor agonist D-serine. D-Serine itself can be toxic, but it can be converted to its stereoisomer L-serine, a natural component of various foods. Supplementing the patient's diet with L-serine powder increased her D-serine levels and improved her motor and cognitive performance after about 1 year of treatment. —LKF

Sci. Signal. **12**, eaaw0936 (2019).



IN OTHER JOURNALS

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

IMMUNOLOGY

BeATing back obesity

Children who are breastfed appear to be at a lower risk for obesity later in life. One explanation is that breast milk may somehow prevent the premature transition from beige adipose tissue (BeAT) to white adipose tissue during childhood and adolescence. Yu *et al.* report that alkylglycerol ether lipids found in breast milk maintain the BeAT of infant mice. Adipose tissue macrophages metabolize these lipids into platelet-activating factor, which then induces macrophage release of interleukin-6, signal transducer and activator of transcription 3 (STAT3) signaling in adipocytes, and the development of BeAT. Although usually inactivated in adulthood, this pathway resurrects in obese adipose tissue, indicating a degree of metabolic flexibility that could be exploited for therapeutic applications. —STS

J. Clin. Invest. **129**, 2485 (2019).

Lipids in breast milk delay the transition from beige to white adipose tissue and protect against obesity.

INFORMATION STORAGE

Storing information in molbytes

Present technologies for storing information have been developed over many decades, and each has its own weaknesses and strengths. In this work, Cafferty *et al.* present a possible alternative storage approach that uses mixtures of small organic molecules ("molbytes," oligopeptides in the present

PHOTO: LIFEBRARY/SHUTTERSTOCK.COM

PSYCHOLOGY

Social media use and mental health

There has been public concern that increasing social media use among adolescents may cause reduced well-being. Orben *et al.* analyzed social media use among more than 12,000 teenagers in the United Kingdom over the course of 8 years to determine whether increased social media use predicted reduced life satisfaction over time. They found virtually no effect of social media use on life satisfaction either between individuals or in the same individual across time. These results suggest that concern over the use of social media and its relationship to mental health may be unwarranted. —TSR

Proc. Natl. Acad. Sci. U.S.A. **116**, 10226 (2019).

Adolescents using social media on their smartphones

study), distinguishable by mass spectrometry, to encode, write, store, and read any information in a binary manner. Each new message depends entirely on simple physical manipulations with the molecules, so no additional synthesis is required. The proposed strategy is in its early development stage but demonstrates that the chemistry of small molecules can potentially offer alternative approaches for secure long-term, zero-energy storage of information. —YS

ACS Cent. Sci. **5**, 911 (2019).

AGING

Senescent beta cells

Removal of senescent cells, which accumulate during aging, may have clinical utility in managing chronic diseases. One chronic disease commonly associated with aging is type 2 diabetes. Diabetes develops when pancreatic beta cells age and senesce, which then contributes to failure of glucose homeostasis. Aguayo-Mazzucato *et al.* investigated senescence of the insulin-producing beta cells in mouse models of diabetes. Removal of senescent cells with drug treatment or by specific depletion of cells expressing the marker $p16^{Ink4a}$ in transgenic mice helped restore beta-cell function. Similar properties were observed in cultured human beta cells—results that hint of translational possibilities. —LBR

Cell Metab. 10.1016/j.cmet.2019.05.006 (2019).

NEUROSCIENCE

Imagination as a dialogue

The hippocampus constructs scene imagery to facilitate recollected or imagined mental representations. However, input from the ventromedial prefrontal cortex (vmPFC) is also needed for scene construction. How do these regions interact when we imagine a scene? Barry *et al.* addressed this question using magnetoencephalography. Participants imagined novel scenes or single objects after being given a cue. The direction of information flow during scene imagination mirrored that observed during episodic memory retrieval, in which vmPFC drives hippocampal activity. These results indicate that the vmPFC selects the elements for a scene, whereas the hippocampus is necessary to construct the scene imagery. The vmPFC strongly modulates the construction of spatially coherent, contextually appropriate scene imagery. Episodic memory and imagination thus use similar regional dialogue in the brain. —PRS

J. Neurosci. **39**, 4375 (2019).

CELL BIOLOGY

QC keeps on keeping on

Several protein quality control processes have been identified that monitor and remove misfolded or damaged proteins in the endoplasmic reticulum. Schmidt *et al.* asked whether additional pathways monitor proteins further along the secretory pathway in the Golgi complex and the endosome. They identified a degradation pathway in budding yeast cells associated with endosome and Golgi, which they named EGAD. EGAD extracts membrane proteins from the Golgi and endosomes, allowing them to become substrates for cytosolic proteasomal degradation. An important cellular substrate of this pathway is a protein called Orm2, which is involved in down-regulating sphingolipid biosynthesis. The selective degradation of Orm2 facilitates the homeostatic regulation of sphingolipid biosynthesis. —SMH

EMBO J. 10.15252/embj.2018101433 (2019).

NANOMATERIALS

Enhancing nanodiamond sensors

The fluorescence and spin properties of nitrogen vacancy (NV) centers in nanodiamonds could make them useful sensors when linked to biomolecules. Although nanodiamonds with diameters of about 5 nanometers can be extracted from the products of closed-chamber detonation of explosives like trinitrotoluene, the number of spin-triplet (negatively charged) NV centers can be too low to ensure that each nanodiamond is active. Terada *et al.* show that irradiation of detonation nanodiamonds with a sufficient fluence of 2 mega–electron volt electrons can increase the number of negatively charged NV centers by a factor of 4. The irradiation process aggregated the nanodiamonds, but they were redispersed with boiling oxidizing acid. —PDS

ACS Nano 10.1021/acsnano.8b09383 (2019).

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

ROBOTICS

Hand it to you

Our ability to grab, hold, and manipulate objects involves our dexterous hands, our sense of touch, and feedback from our eyes and muscles that allows us to maintain a controlled grip. Billard and Kragic review the progress made in robotics to emulate these functions. Systems have developed from simple, pinching grippers operating in a fully defined environment, to robots that can identify, select, and manipulate objects from a random collection. Further developments are emerging from advances in computer vision, computer processing capabilities, and tactile materials that give feedback to the robot. —MSL

Science, this issue p. 1149

RUMINANT GENOMICS

Phylogeny and characteristics of ruminants

Ruminants are a diverse group of mammals that includes families containing well-known taxa such as deer, cows, and goats. However, their evolutionary relationships have been contentious, as have the origins of their distinctive digestive systems and headgear, including antlers and horns (see the Perspective by Ker and Yang). To understand the relationships among ruminants, L. Chen *et al.* sequenced 44 species representing 6 families and performed a phylogenetic analysis. From this analysis, they were able to resolve the phylogeny of many genera and document incomplete lineage sorting among major clades. Interestingly, they found evidence for large population reductions among many taxa starting at approximately 100,000 years ago, coinciding with the migration of humans out of Africa. Examining the bony appendages on the head—the so-called headgear—Wang *et al.* describe specific evolutionary changes in the ruminants and identify selection on cancer-related genes that may function in antler development in deer. Finally, Lin *et al.* take a close look at the reindeer genome and identify the genetic basis of adaptations that allow reindeer to survive in the harsh conditions of the Arctic. —LMZ

Science, this issue p. 1152, p. 1153, p. 1154; see also p. 1130

MOLECULAR MACHINES

Flexible domains in a well-oiled machine

Motors convert one form of energy into another. For biological motors, adenosine triphosphate (ATP) serves as chemical energy and its hydrolysis is coupled to conformational changes that exert mechanical force. ATP synthases reverse this process in a multistep process: first converting an electrochemical gradient to rotational kinetic energy, and then coupling rotation to formation of high-energy phosphodiester bonds. Murphy *et al.* investigated these energy changes in the dimeric mitochondrial F_1-F_o ATP synthase from *Polytomella* sp., a unicellular alga. They solved high-resolution cryo–electron microscopy structures of the ATP synthase complex, extracting 13 rotational substates. This collection of structures revealed that the rotation of the F_o ring and central stalk is coupled with partial rotations of the F_1 head. This flexibility may enable the head to better couple continuous rotation with discrete ATP synthesis events. —MAF

Science, this issue p. 1155

FISHERIES

A small, interconnected world

Countries manage their fisheries as if they were a local resource. To some degree, this may reflect reality, but marine fish, perhaps more than any other vertebrate group, are connected across large distances through ocean currents. Ramesh *et al.* model how these currents distribute the fish larvae of more than 700 species. They used network analysis to assess the degree to which populations found in one part of the world may have come from another. It seems that global fish populations represent a small-world network where connections across populations are tight and particular hubs of productivity are widely important. Such connectivity has wide-ranging implications for conservation, management, and food supplies globally. —SNV

Science, this issue p. 1192

MUCOSAL IMMUNITY

Context shapes anticommensal immunity

The gut bacterium *Akkermansia muciniphila* is associated with protection from obesity, enhanced wound healing, and augmented antitumor responses. Ansaldo *et al.* found that this microbe induces antigen-specific immunoglobulin G1 (IgG1) antibodies generated by B cells with $CD4^+$ T cell help. This is in contrast to most anticommensal responses, which involve the T cell–independent production of IgA antibodies. In a gnotobiotic setting in which all components of the microbiome are defined, *A. muciniphila*–specific T cells expanded only when *A. muciniphila* was present. The T cells primarily displayed a phenotype associated with B cell help. However, in mice with a conventional gut microbiota, other proinflammatory *A. muciniphila*–specific T cell populations also expanded. Thus, anti–*A. muciniphila* immunity is context dependent, which may explain the variable immune responses to this microbe reported in patients. —STS

Science, this issue p. 1179

CONSERVATION

Integrating knowledge systems for wildlife health

Knowledge held by local or indigenous populations is often crucial for supporting wildlife health and conservation efforts. In a Perspective, Kutz and Tomaselli highlight efforts to integrate these different forms of knowledge in a mutually respectful manner. For example, use of a participatory epidemiology approach that integrates scientific and indigenous knowledge was key to eradicating rinderpest in its last foci in East Africa. A similar approach led to a greater understanding of the mechanisms driving muskox population declines in the Canadian Arctic. Thus, going forward, rigorous research design that is based on ongoing iteration and feedback could enable inclusion of qualitative local indigenous knowledge in models and decision-making. —JFU

Science, this issue p. 1135

COLLOIDAL MATERIALS

Mobile particles in colloidal crystals

The crystallization of nanoparticles can be controlled by functionalizing them with DNA strands that direct assembly through hybridization. The design rules for interactions between pairs of particles resemble those for ionic compounds. Inspired by molecular dynamics simulations, Girard *et al.* show that larger particles (~10 nanometers in diameter) that have mutual repulsive interactions can form a stable lattice only if much smaller conjugate particles (~1.5 nanometers in diameter) are present. These smaller particles are mobile and diffuse through the lattice, so the bonding interaction resembles the classical picture of electrons in metals. —PDS

Science, this issue p. 1174

CANCER

The microbiota in colorectal cancer

Several types of bacteria have been associated with colorectal cancer, but do these bacteria have a role in tumorigenesis? Data remain unclear, but evidence suggests that some types of bacteria can influence tumor progression and responses to therapy. In a Perspective, Garrett discusses how to advance microbiota research to potentially improve prevention, diagnostics, and therapy of colorectal cancer. —GKA

Science, this issue p. 1133

INFLAMMATION

A colitis circuit

Cytokines are known to play a critical role in maintaining gut homeostasis, but their specific cellular sources are less well understood. Bernshtein *et al.* used a murine model of inflammatory bowel disease in which macrophages specifically lacked expression of the interleukin-10 receptor (IL-10R). The mice developed symptoms of spontaneous colitis similar to those observed in children with IL-10R mutations. The macrophage-derived cytokine IL-23 was critical for inducing the pathology. IL-23 triggered accumulation of IL-22–producing T helper 17 cells that, in turn, promoted production of chemokines by colonic epithelial cells and destructive neutrophil recruitment. —CNF

Sci. Immunol. **4**, eaau6571 (2019).

HYDROLOGY

Modeling groundwater depletion

Understanding the full effects of human activities on groundwater resources has been a persistent challenge to hydrologic models. Such models often seek to inform best practices for resource management. One important gap in our understanding is how groundwater depletion affects surface water behavior. Condon and Maxwell simulated the impact of 100 years of groundwater declines across the continental United States. Their integrated model illuminates the nature of connections between groundwater pumping and changes in natural watersheds, including the widespread impacts of pumping on evapotranspiration rates and streamflow. —KVH

Sci. Adv. 10.1126/sciadv.aav4574 (2019).

REVIEW SUMMARY

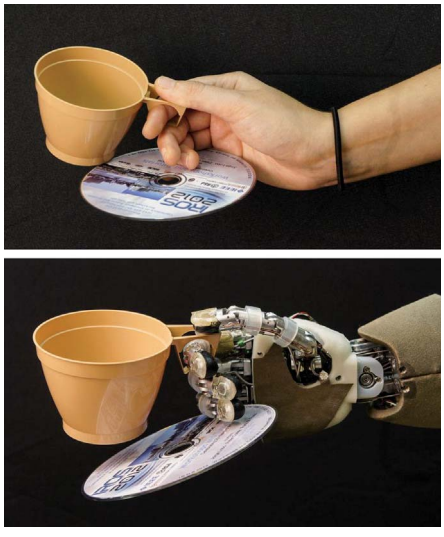
ROBOTICS

Trends and challenges in robot manipulation

Aude Billard* **and Danica Kragic**

BACKGROUND: Humans have a fantastic ability to manipulate objects of various shapes, sizes, and materials and can control the objects' position in confined spaces with the advanced dexterity capabilities of our hands. Building machines inspired by human hands, with the functionality to autonomously pick up and manipulate objects, has always been an essential component of robotics. The first robot manipulators date back to the 1960s and are some of the first robotic devices ever constructed. In these early days, robotic manipulation consisted of carefully prescribed movement sequences that a robot would execute with no ability to adapt to a changing environment. As time passed, robots gradually gained the ability to automatically generate movement sequences, drawing on artificial intelligence and automated reasoning. Robots would stack boxes according to size, weight, and so forth, extending beyond geometric reasoning. This task also required robots to handle errors and uncertainty in sensing at run time, given that the slightest imprecision in the position and orientation of stacked boxes might cause the entire tower to topple. Methods from control theory also became instrumental for enabling robots to comply with the environment's natural uncertainty by empowering them to adapt exerted forces upon contact. The ability to stably vary forces upon contact expanded robots' manipulation repertoire to more-complex tasks, such as inserting pegs in holes or hammering. However, none of these actions truly demonstrated fine or in-hand manipulation capabilities, and they were commonly performed using simple two-fingered grippers. To enable multipurpose fine manipulation, roboticists focused their efforts on designing humanlike hands capable of using tools. Wielding a tool in-hand became a problem of its own, and a variety of advanced algorithms were developed to facilitate stable holding of objects and provide optimality guarantees. Because optimality was difficult to achieve in a stochastic environment, from the 1990s onward researchers aimed to increase the robustness of object manipulation at all levels. These efforts initiated the design of sensors and hardware for improved control of hand–object contacts. Studies that followed were focused on robust perception for coping with object occlusion and noisy measurements, as well as on adaptive control approaches to infer an object's physical properties, so as to handle objects whose properties are unknown or change as a result of manipulation.

ADVANCES: Roboticists are still working to develop robots capable of sorting and packaging objects, chopping vegetables, and folding clothes in unstructured and dynamic environments. Robots used for modern manufacturing have accomplished some of these tasks in structured settings that still require fences between the robots and human operators to ensure safety. Ideally, robots should be able to work side by side with humans, offering their strength to carry heavy loads while presenting no danger. Over the past decade, robots have gained new levels of dexterity. This enhancement is due to breakthroughs in mechanics with sensors for perceiving touch along a robot's body and new mechanics for soft actuation to offer natural compliance. Most notably, this development leverages the immense progress in machine learning to encapsulate models of uncertainty and support further advances in adaptive and robust control. Learning to manipulate in real-world settings is costly in terms of both time and hardware. To further elaborate on data-driven methods but avoid generating examples with real, physical systems, many researchers use simulation environments. Still, grasping and dexterous manipulation require a level of reality that existing simulators are not yet able to deliver—for example, in the case of modeling contacts for soft and deformable objects. Two roads are hence pursued: The first draws inspiration from the way humans acquire interaction skills and prompts robots to learn skills from observing humans performing complex manipulation. This allows robots to acquire manipulation capabilities in only a few trials. However, generalizing the acquired knowledge to apply to actions that differ from those previously demonstrated remains difficult. The second road constructs databases of real object manipulation, with the goal to better inform the simulators and generate examples that are as realistic as possible. Yet achieving realistic simulation of friction, material deformation, and other physical properties may not be possible anytime soon, and real experimental evaluation will be unavoidable for learning to manipulate highly deformable objects.



Holding two objects in one hand requires dexterity. Whereas a human can grab multiple objects at the same time (top), a robot (bottom) cannot yet achieve such dexterity. In this example, a human has placed the objects in the robot's hand.

PHOTOS: LEARNING ALGORITHMS AND SYSTEMS LABORATORY, EPFL

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aat8414

OUTLOOK: Despite many years of software and hardware development, achieving dexterous manipulation capabilities in robots remains an open problem—albeit an interesting one, given that it necessitates improved understanding of human grasping and manipulation techniques. We build robots to automate tasks but also to provide tools for humans to easily perform repetitive and dangerous tasks while avoiding harm. Achieving robust and flexible collaboration between humans and robots is hence the next major challenge. Fences that currently separate humans from robots will gradually disappear, and robots will start manipulating objects jointly with humans. To achieve this objective, robots must become smooth and trustable partners that interpret humans' intentions and respond accordingly. Furthermore, robots must acquire a better understanding of how humans interact and must attain real-time adaptation capabilities. There is also a need to develop robots that are safe by design, with an emphasis on soft and lightweight structures as well as control and planning methodologies based on multisensory feedback. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: aude.billard@epfl.ch
Cite this article as A. Billard, D. Kragic, *Science* **364**, eaat8414 (2019). DOI: 10.1126/science.aat8414

REVIEW

ROBOTICS

Trends and challenges in robot manipulation

Aude Billard[1]* and Danica Kragic[2]

Dexterous manipulation is one of the primary goals in robotics. Robots with this capability could sort and package objects, chop vegetables, and fold clothes. As robots come to work side by side with humans, they must also become human-aware. Over the past decade, research has made strides toward these goals. Progress has come from advances in visual and haptic perception and in mechanics in the form of soft actuators that offer a natural compliance. Most notably, immense progress in machine learning has been leveraged to encapsulate models of uncertainty and to support improvements in adaptive and robust control. Open questions remain in terms of how to enable robots to deal with the most unpredictable agent of all, the human.

Have you ever found yourself busy foraging in your bag in search of a set of keys? If so, you may recall that it took only a few seconds to find them among the disparate contents of the bag. For certain, you did not reflect on your abilities and may have carried on this display of unique dexterity through a swift in-hand manipulation, taking out the correct key and inserting it into the lock even though the corridor lights had gone out. All day long, our fingers grasp, move, and transform objects and interact with objects in various media such as air, water, and oil. We do not spend time thinking about what our hands and fingers are doing or how the continuous integration of various sensory modalities—such as vision, touch, proprioception, and hearing—help us outperform any other biological system in the breadth of the interaction tasks we can execute. Largely overlooked, and perhaps most fascinating, is the ease with which we perform these interactions, resulting in a belief that they are also easy to accomplish in artificial systems such as robots.

Manipulating objects is such a ubiquitous activity that we forget how difficult it was to acquire this competence as a child. Children are born with simple grasp reflexes. It takes them 3 years to develop an individuated control of each finger and another 6 years to display an adult-equivalent ability for making smooth contact and for planning sequences of manipulation skills (*1*). Even for humans, some dexterous activities may pose a challenge. For example, tying shoes may be done in various ways, and there may be several valid models of how to execute such an activity. In addition, we can visually demonstrate how to do something and what the expected result may be, but we cannot easily communicate the magnitude of the applied forces and torques or the size of the friction coefficient necessary to satisfy stability conditions. Still, we find ways of achieving manipulation goals through training and exploration even if the end result is not always optimally performed. We may also adapt as circumstances dictate (e.g., tying shoes with an excess of free shoelace or when the ends are quite short), forcing us to deviate from our normal methods. Thus, the context in which interactions are performed affects various parameters of the execution.

Although robotics has made vast progress in mechanical design, perception, and robust control targeted to grasping and handling objects, robotic manipulation is still a poor proxy for human dexterity. To date, no robots can easily handwash dishes, button a shirt, or peel a potato.

What can robots do today?

Robots are skilled at picking up and manipulating objects in repetitive and familiar settings such as industrial assembly setups. In such settings, the geometry, material properties, and weight of the objects are commonly known. Robots can handle some variation in routine movements in terms of adapting to small differences in the object properties, but the whole process is typically optimized to a limited set of expected variations. In early factory settings, robot arms followed predetermined trajectories and assumed that objects would always appear at the same place. Today, robots can adapt their trajectory to retrieve objects at different locations, making it possible for objects to be placed by humans or simply dropped on a conveyer belt instead of being deposited at exact positions by other machines. The classical assembly lines in which robots were bolted into the floor and placed one after another, typical for the automobile industry, can now be made more flexible. Objects moving on conveyers can be detected fairly easily by cameras and picked up if fully visible. However, detection of transparent objects or objects partially hidden (e.g., when stacked on top of one another) remains difficult.

With the need to frequently change the type of goods produced, the robotics industry strives for multipurpose object grasping and handling solutions. One step toward this objective is to provide robots with a choice of grippers varying in size and strength and to enable robots with tool-changing mechanisms so that they can select the correct tool. To determine which tool to use for a given task, a robot must have knowledge of an object's properties, such as shape, weight, material, and so forth. This information is readily available in factories where all objects are known. However, this requirement presents a limitation for robots in other settings, where the set of objects to be manipulated may not be known beforehand.

What can robots not do today?

Although robots are adept at handling rigid objects, they still struggle with flexible materials—such as fruits and vegetables or clothing items—that differ in size, weight, and surface properties. Manipulations that produce a deformation (e.g., inserting, cutting, or bending) are particularly difficult, as accurate models of the deformations are needed. Industrial grippers often use pneumatic vacuum pumps to pick up objects by sucking. This technique is unbeatable when it comes to grasping an object but is much less useful for object manipulation (e.g., reorienting the object and placing it in a confined space). One step to address this challenge is to provide robots with more dexterous hands. Yet creating hands as dexterous as human hands is difficult, owing to a lack of sensors and actuators equivalent in size, precision, and efficiency to our skin and muscles.

Improvements in robots' dexterity are not limited to the engineering of more-capable hands. Advanced software programs are required to analyze in real time the large flux of visual, tactile, and force information and to relate these different senses to recognize objects and model their transformations. Additionally, robots need advanced cognitive capabilities to predict where, how, and why to manipulate objects. The rest of this Review describes why overcoming these challenges is difficult and where the field of robotics stands today.

Why is designing robotic hands difficult?

Although research on robot hands has been ongoing for more than five decades (*2*–*4*), the most common hand used in many applications to date is still a parallel jaw gripper, usually without any extra sensing. Picking up objects with a gripper devoid of sensing is akin to grasping with the tip of your thumb and index finger when both are numb! This tool may suffice for simple pick-and-place actions, but not for more-complex motions such as shuffling keys. Because the human hand performs intricate movements with

[1]Learning Algorithms and Systems Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. [2]Robotics, Perception and Learning (RPL), EECS, Royal Institute for Technology (KTH), Stockholm, Sweden.
*Corresponding author. Email: aude.billard@epfl.ch

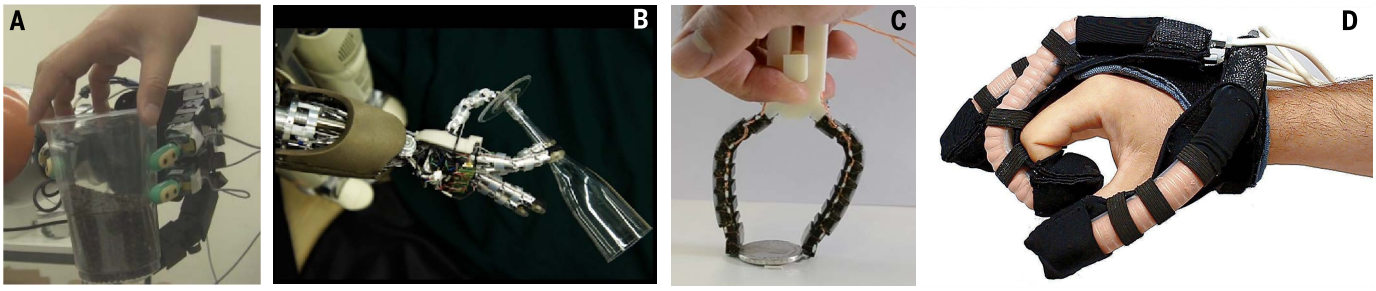


Fig. 1. Soft hands. Robotic hands are traditionally made of hard materials with rigid control of fingers. Recent designs aim to mimic the human hand's natural compliance by using soft actuators, soft materials, and advanced controllers. (**A**) Rigid material and actuators and (**B**) a rigid cover with partially soft cable-driven actuation: Both hands become soft through software intervention, modulating pressure at the fingertips via tactile feedback. [Reproduced from (*17*, *33*)] (**C**) A soft, foldable gripper that can adapt shape and stiffness. [Reproduced from (*11*)] (**D**) Soft actuation and material for a rehabilitation glove that can be worn by a human. [Reproduced from (*12*)]

ease, it is a natural inspiration for robotics. But designing robotic hands with sensors and actuators similar to those of the human hand is difficult for many reasons.

When constructing anthropomorphic robot hands, it is challenging to fit all of the necessary actuators, sensors, and mechanical structure in the limited available space. Another obstacle is to keep the total weight of the hand low so that it satisfies the payload requirements of the arm to which it is attached. Hence, compared with human hands, most anthropomorphic hands and prostheses do not have nearly as many controllable degrees of freedom (*5*, *6*).

The human hand is soft and flexible, with a dexterous thumb whose distinctive range of motion remains difficult to replicate mechanically (*7*), as the intricate combinations of tendons and muscles differ markedly from traditional serial robotic joint design (*8*). Today, robotic hands are still largely composed of rigid plastic and metal components, with electric motors as actuators. This rigidity is partially the cause of the lack of dexterity, as it allows no room for mistakes when executing grasps. Rigid fingers closing on an object may easily move rather than grasp an object if its pose is not perfectly estimated, and applying too much force may crush the object. A growing trend in robotics is the development of soft hands that can conform to an object's shape, absorb unexpected forces at contact, and compensate for load change during manipulation (*9*, *10*).

Softness can be achieved through a change in hardware or software or a combination of both (Fig. 1). Softness from material used to construct hands builds on solutions from 3D manufacturing and materials science. For instance, one can manufacture rigid and flexible materials in a layer-by-layer manner to create foldable fingers that can deploy and retract as needed (*11*). Currently, the low payload and slow speed of these elastomers restricts manipulation to light objects only. As an alternative to generate more power, pneumatic or hydraulic actuation may be used (*12*, *13*).

Human hands are covered with a multipurpose skin that provides the appropriate level of friction and damping. Human skin is a high-frequency and high-resolution sensor that provides precise information on normal and tangential forces, information that is critical for grip adjustment. Human skin can also measure stretch and temperature. By contrast, robot hands typically measure exerted forces through miniature force sensors placed solely at the fingertips (*14*). Force sensors yield very accurate 3D measurements, but they cannot easily reveal the exact location of contact. To move objects once held in the hand or to hold multiple objects at once (Fig. 1), one needs to measure precise contact points, not just at the fingertips but also along the length and side of the fingers and inside the palm. This can be achieved through artificial skins that provide contact measurements all along the limbs. Interest in artificial skins can be traced back to the 1980s (*15*, *16*), but major advances were achieved in the past decade.

At present, we find a variety of affordable commercial products, several of which can be customized to a robot's shape. Touch sensors measure the normal contact force; a few also provide data on tangential forces, torque, temperature, vibrations, or surface properties. Nevertheless, most touch sensors are rigid, and their placement is constrained to fingertips and along limb segments. Touch detection at the joint (knuckle, elbow, knee) is, however, crucial to detect entrapment. It is also useful to guide exploration inside objects (*17*). Such contact can be detected only by soft sensors that bend and extend along the flexion and extension points of the limb (Fig. 2, right) (*18*). Hence, flexible and stretchable skins are of utmost interest to roboticists (*19*). Prototypes exist in laboratories, and we can expect to see their deployment soon, given the current interest in soft electronics (*20*).

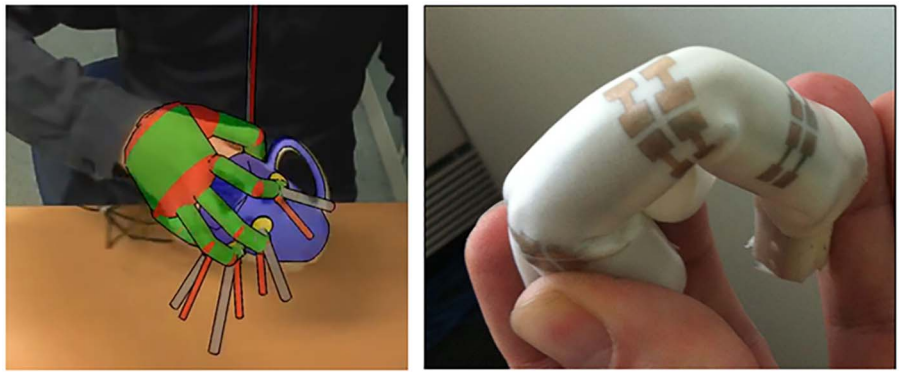


Fig. 2. Sense of touch for robots. (**Left**) Vision can be used to infer contact forces (red). [Reproduced from (*21*)] (**Right**) Stretchable artificial skin measures contact at knuckles, which may be useful for exploring internal parts of objects. [Reproduced from (*18*)]

As an alternative to using skin, it is possible to deduce haptic (contact and force) information from vision. One can, for instance, infer forces from vision through a dynamic model of contacts (*21*) (Fig. 2, left) or use an optical sensor that renders deformation of an object's geometry at a high spatial resolution (*22*). The necessity of estimating the exact position of an object, its local geometry, and other properties such as weight and weight distribution depends strongly on the application. It is the interplay among hand design, material, and internal and external sensing that offers the appropriate redundancy. One additional challenge is the need to measure contact at a very high frequency for accurate and timely detection of slip (*23*). Such high spatial and temporal resolution, together with

the real-time processing of vision data for object tracking, leads to a computational overload with a massive data stream that must be interpreted in real time. This processing is usually carried out by a CPU (central processing unit) located away from the hand. Alternatively, processing may be performed on the hand itself through the use of a dedicated CPU (*24*), but such CPUs have focused solely on processing vision data. Additional studies are needed to develop hardware for processing tactile information in conjunction with vision.

Dexterous robot hands may hence be realized by using research in materials science for the design of soft actuation, enabling contact sensing along the entire surface of the hand, and by using advances in electronics for onboard, real-time processing of multisensory data.

Design beyond anthropomorphism

Although the human hand is fascinating, it does not have to be the ultimate solution for robotics. A human hand design may be desirable for aesthetic reasons; for instance, when designing hand prostheses or humanoid robots. But this same design may be superfluous for many robots. Industrial hands remain a good solution for specific tasks. Rather than try to replicate the positioning of human fingers, these hands have two or three fingers arranged symmetrically around the palm, a design particularly suited for industrial screwing.

Robotics keeps oscillating between anthropomorphic and traditional industrial designs for hands. But the gripping systems of simpler animals may also provide inspiration. For instance, fish suck in their prey. Adding suction at robots' fingertips is useful under water, as this technology cancels the flow generated by the hand (*25*).

Why not create hands that both leverage and go beyond nature? For instance, the human thumb is amazing, but it creates an asymmetry that constrains the orientation of the hand for manipulation. Two thumbs on the same hand, however, would provide a dexterity beyond human capability (Fig. 3).

Desiderata for the next generation of robotic hands

The objects around us have been built and adapted to our hands, which are still rather small and very robust in comparison with contemporary robot hands. Enabling robots to pick up small items such as pens, raisins, screws, and needles is a clear functionality goal. Today, robotic arms and hands are commonly developed separately, and integrating them is an engineering job of its own. Industrial arms have substantial payloads but are commonly designed to be bolted into the floor and are too large to be deployed outside industrial settings. The arms of humanoid robots and robots intended for fine assembly tasks have low payloads, which are typically not sufficient to carry a hand and an object held by the hand. Adding sensing functionality to arms and hands requires cabling that can quickly become complicated. Furthermore, many hands come with no or limited means of measuring contact and forces. Thus, a change in paradigm is needed to move away from developing robotic arms devoid of hands and hands devoid of arms. We must further ensure that hands are developed in a "plug-and-play" manner and can easily be attached and detached through existing tool-switching systems. State-of-the-art force and tactile sensors must become an inherent part of the arm–hand system.

Robot dexterity is as much a by-product of advances in hardware as it is of advances in software. It requires suitable algorithms to rapidly and efficiently process the vast amount of information collected through sensors and actuators. At the same time, it needs algorithms to adequately control the movement of a hand in relation to object, scene, and task properties. We next review advances in perception, control, and learning for manipulation.

Perception for manipulation

As for humans, robot perception for manipulation is multimodal (Fig. 4). Vision is instrumental for recognizing and localizing objects. When associated with a database of existing objects, robot vision can help infer geometric and physical properties of known and even unknown objects (*26*), and this information is important for shaping the aperture of the hand and the forces to be applied. Proprioception—namely, knowledge of where the robot's limbs are located—is needed to guide the arm and hand toward the object, with visual support to continuously track the object. Touch and force measurements become important once contact has occurred and the object is held or explored by the hand. The associated control algorithms are used to guide the grasp and/or to infer the object's physical properties, such as rigidity and mass distribution, that may have been poorly estimated or unknown previously. Sound has also received attention recently as a means to infer an invisible object's content and to monitor changes in content during manipulation (*27*).

As an example, a robot is tasked with fetching a package of milk from a refrigerator. Before the robot holds the package in hand, it may not know how much milk the package contains nor the package's actual weight. Given that the package may be made of cardboard, the robot needs to know the weight in order to apply a suitable grasping force and avoid destroying

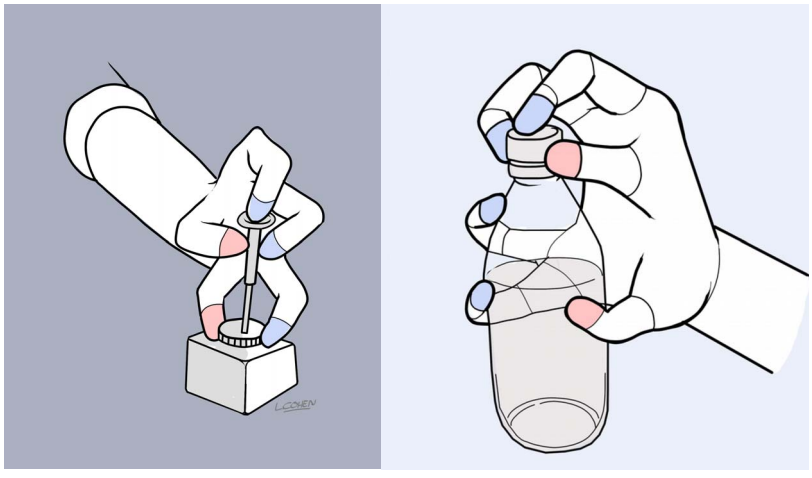


Fig. 3. Designing hands beyond human dexterity. Two thumbs would make it possible to execute screwing and unscrewing motions with one hand rather than two. This capability may be useful for robots and humans via prostheses. [Illustrations: Laura Cohen]

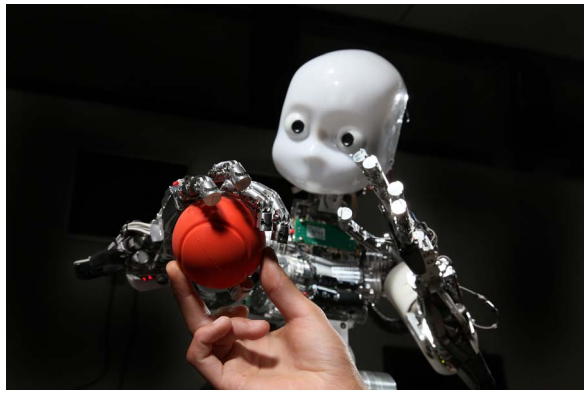


Fig. 4. Manipulation is multimodal. Vision is used before contact, whereas haptics and sound are involved upon contact to estimate an object's physical properties that cannot be directly observed. [Photo: Learning Algorithms and Systems Laboratory, EPFL]

the package. In the case of milk, sound may also provide information about the viscosity when the package is shaken, as milk will sound different from another substance, such as yogurt.

Over the past few years, major efforts have been undertaken to analyze visual information, and progress has been considerable. Nevertheless, robots still struggle to recognize objects that are partially occluded (*28*), particularly when viewed from a moving camera or when an object moves in a robot's hands (*29*). In comparison to developing vision algorithms, much less effort has been devoted to analyzing haptic information, given that solutions for covering entire hands with haptic sensors are still lacking. Today, visual and haptic information are still used primarily in a sequential manner [e.g., with visual information provided in the preparation phase and haptic data provided upon contact (*30*)], and only a few recent works integrate both modalities for recognition, grasping, in-hand adaptation, and shape reconstruction (*31*–*35*). In comparison, humans are proficient at alternating between different senses, from vision to touch and back, and can do so rapidly even if these senses change in processing frequency. By contrast, robots still lack the ability to decide what sensors to use, when to use them, and when to switch between sensors.

Grasping: A stepping stone

Before a robot can manipulate an object in hand, it must be able to grasp its fingers around the object. If grasping is conceptualized only as getting fingers around an object with no additional constraints considered, the challenge of grasping may appear to be solved. However, grasping an object is a far more daunting problem. For decades, researchers have worked to establish the theory of how to form a stable grasp. This became an intricate mathematical exercise aimed at determining the minimal number and optimal positions of the fingertips on the object's surface to ensure stability (*36*).

Although it is valuable, most of this theoretical work relies on assumptions such as a known 3D model of the object, a rigid point contact, and no uncertainty in the process. To incorporate uncertainty originating from imperfect object models and dynamics in the interaction process, we must go beyond modeling a single point contact and pursue substantial advances in the basic theory.

Thus, many of the more recent approaches are data driven (*37*). To avoid computing an optimal grasp each time a robot encounters an object, one can build a database of grasps and employ methodologies for sampling and ranking candidate grasps in real time. This approach deals with uncertainty in perception and provides fast and online generation of grasps for known, familiar, and even unknown objects. Prior knowledge of object properties determines the necessary perceptual processing and associated object representations for generating and ranking grasp candidates. Although this method works well for known and familiar objects, unknown objects necessitate additional heuristics for the discovery of geometric structures (e.g., handles, for which a robot would have a candidate grasp). This challenge is closely related to the classical problems of instance recognition and categorization in computer vision, but the notion that grasping is not an isolated process adds a new dimension.

In addition to being object dependent, grasps are also robot dependent. Moreover, as the number of degrees of freedom of the hand increases, so does the complexity of the control. This is particularly an issue for anthropomorphic hands. One avenue of research to simplify the control draws inspiration from biology and promotes the use of postural synergies (*38*). Synergies form a basis of the subspace of effective human movements in relation to those that are possible by the kinematics of the body. These have been used as a tool for robot hand analysis, control, and design choice (*39*–*42*). Several studies have also demonstrated how underactuated hands can be leveraged to grasp and manipulate objects in unstructured environments and how this work may lead to adaptive hands that are relatively cheap, lightweight, and easy to control compared with fully actuated hands (*43*–*48*). More recent work has optimized hand design to improve manipulation capabilities (*49*, *50*), providing open-source software for such design. Other recent work has suggested that the ability of compliant hands to deform in and with the environment may reduce the cognitive load of manipulation (*51*). Furthermore, this idea can be studied systematically using morphological computation (*52*), in which compliant interactions allow adaptation of behavior to a particular context, without the need for explicit control.

From grasping to manipulation

Grasping is not an end on its own; it is also related to the task a human or robot is executing. For example, one grasps a cup differently depending on whether the goal is to drink from it, fill it with fluid, put it in a dishwasher, or serve it to another person (*53*) (Fig. 5). Similarly, although a knife, fork, or spoon may be held with the same grasp when used to mix soup, this grasp differs from those employed when these utensils are used for eating or cutting. To determine the optimal way to grasp an object, one must understand the purpose of the grasp. Hence, while roboticists aimed to solve the problem of how to grasp an object, they first had to identify the reason for executing the grasp. Today, researchers consider grasping as part of an overall plan for object manipulation.

To determine the correct grasp to use with the correct tool, one must first have the correct tool at their disposal. When in need of a hammer but no hammer is in reach, a human will instead select the first object sturdy enough to act as a hammer. Future efforts to develop robots that can reason in this manner when the most appropriate tool is not available will be critical to facilitate deployment of robots in natural environments. Additionally, robots with this capability will be able to use tools originally designed for human dexterity to perform household tasks without making undesired modifications to our households. How to program such "common sense" tool use is hence an important avenue for research, and some initial work has been conducted in this direction (*54*–*57*).

Manipulations that remain difficult

The previous sections detail the many problems that remain to be solved before robots can perform grasps with a human level of intelligence. This said, robots are already fairly efficient at grasping and releasing certain types of objects. They are also capable of performing a variety of simple manipulation actions such as throwing (*58*), sliding (*59*), poking (*60*), pivoting (*61*), and pushing (*62*). Difficulties arise when these actions must be performed in cluttered environments or require contact-rich interactions (e.g., when an object of interest is placed close to or

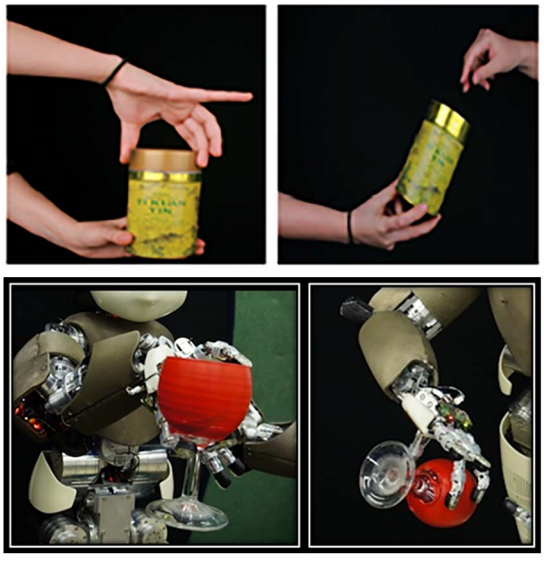


Fig. 5. Grasp functionality. (**Top**) A human grasps an item differently depending on whether the aim is to hold it, open the cap, or hand it to someone else. (**Bottom**) Robots can also be programmed to hold the same glass differently depending on whether they are tasked with handing it to a human or pouring out its contents. [Photos: Learning Algorithms and Systems Laboratory, EPFL]

Fig. 6. Remaining challenges for robot manipulation. Dexterous movement of objects within the hand (**left**), manipulation of deformable objects (e.g., fruits and vegetables) (*65*) (**middle**), and manipulation of objects in collaboration with humans (**right**) present ongoing difficulties. [Photos: Learning Algorithms and Systems Laboratory, EPFL]

is covered by other objects or is located in a confined space such as a shelving unit). It is necessary to plan a feasible path and generate a set of intermediate actions to ensure no damage to the hand or other objects. Today, it is also recognized that perception and control are tightly coupled, and the field of interactive perception (*63*) regards manipulation as a means to perceive and perception as a means to achieve better manipulation.

Manipulation actions that generate changes on the object (cutting, crushing) remain particularly difficult, as they require a model of the deformation and advanced perception to monitor the alterations (*64*). To facilitate adaptation to the changes induced, these actions also necessitate that the forces be applied by the hand (e.g., a reduction of friction when unscrewing a bottle cap, an increase in viscosity when digging into a melon) (Fig. 6) (*65*). Thus, modeling an object's friction and viscosity properties is still an important open problem.

In-hand manipulations in which an object is moved while being held are also particularly complicated. Examples include twirling a pen across fingers or preparing a key to be inserted into a keyhole. These actions comprise an extensive combination of (re)grasping movements and sliding and rotating maneuvers, as well as interactions between two arms and hands, in some cases. When discussing such advanced interactions with objects in robotics, we commonly talk about intrinsic and extrinsic dexterity. The former denotes the ability of the hand to manipulate objects using its available degrees of freedom. Hands with high intrinsic dexterity often mimic the structure of the human hand (*66*). Alternatively, the hand can be simpler, and the end-effector is designed specifically for a particular task (*67*, *68*). Extrinsic dexterity is the ability to compensate for the lack of degrees of freedom by using external support, such as friction, gravity, and contact surfaces (*69*). This functionality also enables dexterous manipulation with simple parallel grippers.

One of the largely underdeveloped areas in robotics is dual-arm or bimanual (*70*) manipulation, as well as the use of the second hand and/or arm to support both intrinsic and extrinsic dexterity (*71*). Some recent work in this area (*72*) proposes integration of object representation, definition of simple movement primitives, and planning to model the problem in an efficient manner. This area will gradually produce more and more contributions in the future, given that most of today's humanoid robots have bimanual capabilities. Furthermore, manipulation does not stop at simply controlling the hand; it requires control of the arm, torso, and ultimately the entire body (*73*). The challenges listed above only increase in scale when one wishes to enable a full humanoid robot to manipulate objects while maintaining its balance (*74*) (Fig. 7). Finally, control of more-complex manipulation skills that require reasoning, such as using an object to retrieve another object, are still in infancy.

Learning for manipulation

Human dexterity is a skill acquired during childhood and further refined throughout life, in activities such as playing a musical instrument or practicing a craft. Similarly, robot dexterity cannot be achieved in the confines of our laboratories. To be able to manipulate the vast array of objects that exist throughout the world, robots must be able to learn continuously, adapt their perception, and control unfamiliar objects.

Learning also addresses some challenges linked to the lack of accurate models of objects and contact dynamics and the increasing complexity of control for robots with large degrees of freedom. Hence, many of the present approaches to dexterous manipulation rely on learning methodologies in place of control-theoretic approaches. For instance, learning can be used to embed representations of stable or suitable grasps (*75*–*78*), which can then be applied to verify stability and generate regrasping motions at run time or to catch a fast-moving object (*79*). Learning is particularly suitable for embedding the dynamic nature of grasping and manipulation, as well as for modeling manipulation of complex nonrigid objects. Learning has been used to model contacts (*80*) and is also beneficial for reducing control dimensions by determining latent space, as required in bimanual dynamics (*65*).

Nevertheless, solving all problems by solely relying on learning is not a viable solution and has certain limitations. First, learning requires data for training, and a common approach is to generate data from trial-and-error experiments. However, this process is tedious and may damage the robot. A growing trend in providing training data is to test the algorithms in simulation first and then refine the learning on a real platform; e.g., for learning dexterous in-hand manipulation (*81*–*83*). Training in simulation depends on having an accurate simulator of the tasks. Alternatively, robots may learn from image data and videos available on the internet (*84*) or from demonstration by a live expert, usually a human. Yet it may not always be possible to find an expert, especially when the tasks are dangerous or require extreme precision. Hence, although learning is important, it cannot be the answer to every problem in robotics.

Fig. 7. Whole-body manipulation. Manipulation of a heavy object by a humanoid robot requires coordination of arms and body to maintain balance. [Reproduced from (*74*)]

Manipulating objects in interaction and collaboration with humans: Reality and challenges

Human–robot collaboration in manufacturing setups has been deemed crucial for the industry (*85*, *86*). Although historically, humans were prohibited from entering the robot's environment (ISO 10218; ANSI/RIA R15.06-1999), it is now accepted that robots can work in close proximity to and in collaboration with humans. However, potentially dangerous scenarios may still occur and need to be addressed. Currently, human–robot collaboration is allowed through the use of manipulators that remain fairly light and are endowed with internal force sensors for detecting unexpected contact or collision with humans. For applications that require maneuvering heavy weight, robots capable of managing the weight may be coupled with an external vision system for monitoring human presence. Yet challenges remain for accurately detecting human presence. At present, the best solution is to combine sensing of proximity and force with external vision-based monitoring. Nonetheless, the 100% fence-based safety paradigm is gone, and industrial standards now target risk minimization and mitigation (ISO/TS 15066).

In addition to facing a world in which objects move and change, robots are now expected to manipulate these objects in collaboration with humans. Interactive and collaborative manipulation adds a new dimension to robot manipulation but presents a wealth of challenges (*87*). For example, when a robot is tasked with handing an object to a person or carrying a large object jointly with someone, the robot must grasp and move the object carefully and with foresight, so that the robot can infer where the human will move and the human does not get injured. As simple as it may seem, the act of a robot handing an object to a human entails several complex questions, which have in turn inspired studies on how to enable a robot to properly perform this task (*88*–*92*). These questions range from how to present an object for optimal human grasping to others related to social factors, such as the role of gaze, social cues, and awareness of user state. There is no agreement on what factors are most important in determining how handovers are carried out between two humans, let alone between a robot and a human. Although most research has focused on a robot handing objects to humans, there have also been studies on robots taking objects from humans (*93*–*95*). Additionally, several efforts have been aimed at enabling a robot to manipulate objects jointly with humans, and the act of carrying objects jointly with humans has been demonstrated using both humanoid robots (*96*–*99*) and mobile manipulators (*100*–*103*). Notable recent efforts have explored human–robot joint manipulation of deformable materials (*104*), helping humans to dress (*105*), and assistive support (*106*).

Hence, for robots to work seamlessly with humans, researchers are striving to equip robots with the tools for better perception of humans and more adaptive control modes. In addition, roboticists seek guarantees in terms of machine performance and the use of common evaluation scenarios and benchmarks.

Outlook

Since the 1960s, substantial progress has been achieved in several areas of robotic manipulation. We have established the basic theory of evaluating stability of a grasp, control algorithms that can adapt to unpredicted situations, and changing dynamics when the appropriate sensor feedback is available to perform state estimation. Lately, the field has also seen advances in data-driven approaches in which even dexterous in-hand manipulation can be accomplished, but only for very specific problems and in highly tailored environments. Achievement of robust, flexible, and adaptive grasping and manipulation of completely unknown objects in media such as water and oil (not solely in air) is expected to result in a major manufacturing revolution, which will affect most of the work that relies on fine manipulation and high dexterity. However, systematic development is ongoing toward several technologies vital for meeting and exceeding human dexterity and fine manipulation capabilities.

First, there is still a need for basic theoretical development. We must seek to understand and model soft point contact and provide stability rules for both point contacts and surface contacts. We also need to develop a better method for modeling objects whose states change markedly after manipulation (e.g., a cucumber after being sliced, an onion after being chopped). A thorough description of manipulation and task goals will be required for planning and generating appropriate intermediate grasping and manipulation actions. This emphasis on theory and planning is also relevant for data-driven approaches, as we need better tools for simulating soft bodies and generating relevant scenarios and examples that include force and torque information.

In addition to the aforementioned modeling and software aspects, we also seek to achieve substantial progress in hardware development and design. One area of particular relevance is that of robot sensing. It will be important to develop skinlike sensors that are well integrated with hand design but do not require excessive cabling or add substantial weight. This sensing functionality should facilitate force and torque measurements, determining shear forces to detect and counteract slippage. To achieve dexterous in-hand manipulation, we also need actuated hands that can be controlled with high frequencies. Such hands must function in different media (air, water, and oil) without being damaged or needing to be covered by special gloves. Overall, we need hands that are light, cheap, robust, and easily integrated with any type of robotic arm.

Finally, an important industrial challenge will be to bring robots in closer proximity to humans and enable safe physical interaction and collaboration. Fences that used to separate humans from robots will disappear gradually. Robots will thus need to be engaged in collaborative tasks to jointly manipulate objects with humans while adapting to unexpected human behavior. Equipping robots with advanced physical interaction capabilities to achieve safe and smooth synchronization of motion between machine and human is still a major hurdle. This objective will require advances in detailed tracking of human fine body movement, as well as a better understanding of how humans collaborate and achieve joint goals through planning and direct physical interaction. Furthermore, there is a demand for robots that are safe by design, putting focus toward soft and lightweight structures as well as control and planning methodologies based on multisensory feedback. Human ways of acting will continue to serve as inspiration for future robot systems, and robots will serve as a tool for better understanding humans.

REFERENCES AND NOTES

1. J. P. Thibaut, L. Toussaint, Developing motor planning over ages. *J. Exp. Child Psychol.* **105**, 116–129 (2010). doi: 10.1016/j.jecp.2009.10.003; pmid: 19919864
2. H. Hanafusa, H. Asada, Stable prehension by a robot hand with elastic fingers. *T. Soc. Instr. Control Eng.* **13**, 370–377 (1977). doi: 10.9746/sicetr1965.13.370
3. A. Bicchi, Hands for dexterous manipulation and robust grasping: A difficult road toward simplicity. *IEEE Trans. Robot. Autom.* **16**, 652–662 (2000). doi: 10.1109/70.897777
4. R. A. Grupen, T. C. Henderson, I. D. McCammon, A survey of general-purpose manipulation. *Int. J. Robot. Res.* **8**, 38–62 (1989). doi: 10.1177/027836498900800103
5. A. M. Okamura, N. Smaby, M. R. Cutkosky, "An overview of dexterous manipulation" in *Proceedings of the 2000 IEEE International Conference on Robotics and Automation* (IEEE, 2000), pp. 255–262.
6. T. Feix, J. Romero, C. H. Ek, H. B. Schmiedmayer, D. Kragic, A metric for comparing the anthropomorphic motion capability of artificial hands. *IEEE Trans. Robot.* **29**, 82–93 (2013). doi: 10.1109/TRO.2012.2217675
7. M. Chalon, M. Grebenstein, T. Wimböck, G. Hirzinger, "The thumb: Guidelines for a robotic design" in *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2010), pp. 5886–5893.
8. L. Y. Chang, Y. Matsuoka, "A kinematic thumb model for the ACT hand" in *Proceedings of the 2006 IEEE International Conference on Robotics and Automation* (IEEE, 2006), pp. 1000–1005.
9. Y. Li, I. Kao, "A review of modeling of soft-contact fingers and stiffness control for dextrous manipulation in robotics" in *Proceedings of IEEE International Conference on Robotics and Automation* (IEEE, 2001), pp. 3055–3060.
10. B. S. Homberg, R. K. Katzschmann, M. R. Dogar, D. Rus, "Haptic identification of objects using a modular soft robotic gripper" in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2015), 1698–1705.
11. A. Firouzeh, J. Paik, Grasp mode and compliance control of an underactuated origami gripper using adjustable stiffness joints. *IEEE/ASME T. Mech.* **22**, 2165–2173 (2017). doi: 10.1109/TMECH.2017.2732827
12. P. Polygerinos, Z. Wang, K. C. Galloway, R. J. Wood, C. J. Walsh, Soft robotic glove for combined assistance and at-home rehabilitation. *Robot. Auton. Syst.* **73**, 135–143 (2015). doi: 10.1016/j.robot.2014.08.014
13. R. Deimel, O. Brock, A novel type of compliant and underactuated robotic hand for dexterous grasping. *Int. J. Robot. Res.* **35**, 161–185 (2016). doi: 10.1177/0278364915592961
14. A. Bicchi, J. K. Salisbury, D. L. Brock, Contact sensing from force measurements. *Int. J. Robot. Res.* **12**, 249–262 (1993). doi: 10.1177/027836499301200304
15. L. D. Harmon, Automated tactile sensing. *Int. J. Robot. Res.* **1**, 3–32 (1982). doi: 10.1177/027836498200100201

16. H. R. Nicholls, M. H. Lee, A survey of robot tactile sensing technology. *Int. J. Robot. Res.* **8**, 3–30 (1989). doi: 10.1177/027836498900800301
17. N. Sommer, A. Billard, Multi-contact haptic exploration and grasping with tactile sensors. *Robot. Auton. Syst.* **85**, 48–61 (2016). doi: 10.1016/j.robot.2016.08.007
18. A. P. Gerratt, N. Sommer, S. Lacour, A. Billard, "Stretchable capacitive tactile skin on humanoid robot fingers - First experiments and results" in *Proceedings of the 2014 IEEE International Conference on Humanoid Robots* (IEEE, 2014), pp. 238–245.
19. M. Amjadi, K. U. Kyung, I. Park, M. Sitti, Stretchable, skin-mountable, and wearable strain sensors and their potential applications: A review. *Adv. Funct. Mater.* **26**, 1678–1698 (2016). doi: 10.1002/adfm.201504755
20. S. Xu *et al.*, Soft microfluidic assemblies of sensors, circuits, and radios for the skin. *Science* **344**, 70–74 (2014). doi: 10.1126/science.1250169; pmid: 24700852
21. T. H. Pham, N. Kyriazis, A. A. Argyros, A. Kheddar, Hand-object contact force estimation from markerless visual tracking. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**, 2883–2896 (2018). doi: 10.1109/TPAMI.2017.2759736; pmid: 29989962
22. W. Yuan, S. Dong, E. H. Adelson, Gelsight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors* **17**, 2762 (2017). doi: 10.3390/s17122762; pmid: 29186053
23. W. Chen, H. Khamis, I. Birznieks, N. F. Lepora, S. J. Redmond, Tactile sensors for friction estimation and incipient slip detection - towards dexterous robotic manipulation: A review. *IEEE Sens. J.* **18**, 9049–9064 (2018). doi: 10.1109/JSEN.2018.2868340
24. D. Bhowmik, K. Appiah, "Embedded vision systems: A review of the literature" in *Applied Reconfigurable Computing. Architectures, Tools, and Applications*, N. Voros *et al.*, Eds. (Lecture Notes in Computer Science Series, Springer, 2018), pp. 204–216.
25. M. R. Cutkosky, "Reach, grasp, and manipulate" in *Living Machines: A Handbook of Research in Biomimetics and Biohybrid Systems*, T. J. Prescott, N. Lepora, P. F. M. J. Verschure, Eds. (Oxford Univ. Press, 2018), chap. 30.
26. N. Vahrenkamp, L. Westkamp, N. Yamanobe, E. E. Aksoy, T. Asfour, "Part-based grasp planning for familiar objects" in *Proceedings of the IEEE/RAS 16th International Conference on Humanoid Robots* (IEEE, 2016), pp. 919–925.
27. S. Clarke, T. Rhodes, C. G. Atkeson, O. Kroemer, "Learning audio feedback for estimating amount and flow of granular material" in *Proceedings of the 2nd Conference on Robot Learning* (Microtome Publishing, 2018), pp 529–550.
28. Y. Guo, M. Bennamoun, F. Sohel, M. Lu, J. Wan, 3D object recognition in cluttered scenes with local surface features: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 2270–2287 (2014). doi: 10.1109/TPAMI.2014.2316828; pmid: 26353066
29. S. Ojha, S. Sakhare, "Image processing techniques for object tracking in video surveillance-A survey" in *2015 IEEE International Conference on Pervasive Computing* (IEEE, 2015), pp 1–6.
30. R. S. Dahiya, G. Metta, G. Cannata, M. Valle, Special issue on robotic sense of touch. *IEEE Trans. Robot.* **27**, 385–388 (2011). doi: 10.1109/TRO.2011.2155830
31. J. Ilonen, J. Bohg, V. Kyrki, "Fusing visual and tactile sensing for 3-D object reconstruction while grasping" in *Proceedings of the 2013 IEEE International Conference on Robotics and Automation* (IEEE, 2013), pp. 3547–3554.
32. H. Dang, P. K. Allen, "Grasp adjustment on novel objects using tactile experience from similar local geometry" in *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2013), pp. 4007–4012.
33. M. Li, Y. Bekiroglu, D. Kragic, A. Billard, "Learning of grasp adaptation through experience and tactile sensing" in *Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2014), pp. 3339–3346.
34. K. Hang *et al.*, Hierarchical fingertip space: A unified framework for grasp planning and in-hand grasp adaptation. *IEEE Trans. Robot.* **32**, 960–972 (2016). doi: 10.1109/TRO.2016.2588879
35. M. M. Zhang, R. Detry, L. Matthies, K. Daniilidis, "Tactile-vision integration for task-compatible fine-part manipulation" in Robotics: Science and Systems (RSS) Workshop on Revisiting Contact - Turning a Problem into a Solution (2017); www.seas.upenn.edu/~zmen/MabelZhang_rss2017ws.pdf.
36. A. Bicchi, V. Kumar, "Robotic grasping and contact: A review" in *Proceedings of the 2000 IEEE International Conference on Robotics and Automation* (IEEE, 2000), pp. 348–353.
37. J. Bohg, A. Morales, T. Asfour, D. Kragic, Data-driven grasp synthesis-A survey. *IEEE Trans. Robot.* **30**, 289–309 (2014). doi: 10.1109/TRO.2013.2289018
38. M. L. Latash, *Fundamentals of Motor Control* (Academic Press, 2012).
39. A. Bicchi, M. Gabiccini, M. Santello, Modelling natural and artificial hands with synergies. *Philos. Trans. R. Soc. B* **366**, 3153–3161 (2011). doi: 10.1098/rstb.2011.0152; pmid: 21969697
40. C. Y. Brown, H. H. Asada, "Inter-finger coordination and postural synergies in robot hands via mechanical implementation of principal components analysis" in *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2007), pp. 2877–2882.
41. M. Ciocarlie, C. Goldfeder, P. Allen, "Dexterous grasping via eigengrasps: A low-dimensional approach to a high-complexity problem" in Robotics: Science and Systems (RSS), Manipulation Workshop - Sensing and Adapting to the Real World (2007); http://projects.csail.mit.edu/manipulation/rss07/paper__dexterous_grasping_via_eigengrasps_a_low_dimensional_approach_to_a_high_complexity_problem__ciocarlie.pdf.
42. M. G. Catalano *et al.*, Adaptive synergies for the design and control of the Pisa/IIT softhand. *Int. J. Robot. Res.* **33**, 768–782 (2014). doi: 10.1177/0278364913518998
43. L. Birglen, T. Laliberté, C. M. Gosselin, *Underactuated Robotic Hands* (Springer, ed. 1, 2007), vol. 40.
44. M. Ciocarlie *et al.*, The velo gripper: A versatile single-actuator design for enveloping, parallel and fingertip grasps. *Int. J. Robot. Res.* **33**, 753–767 (2014). doi: 10.1177/0278364913519148
45. R. R. Ma, W. G. Bircher, A. M. Dollar, "Toward robust, whole-hand caging manipulation with underactuated hands" in *Proceedings of the 2017 IEEE International Conference on Robotics and Automation* (IEEE, 2017), pp. 1336–1342.
46. P. Kyberd, A. Clawson, B. Jones, The use of underactuation in prosthetic grasping. *Mech. Sci.* **2**, 27–32 (2011). doi: 10.5194/ms-2-27-2011
47. C. Meijneke, G. A. Kragten, M. Wisse, Design and performance assessment of an underactuated hand for industrial applications. *Mech. Sci.* **2**, 9–15 (2011). doi: 10.5194/ms-2-9-2011
48. A. M. Dollar, R. D. Howe, The highly adaptive SDM hand: Design and performance evaluation. *Int. J. Robot. Res.* **29**, 585–597 (2010). doi: 10.1177/0278364909360852
49. J. Borràs, A. M. Dollar, Dimensional synthesis of three-fingered robot hands for maximal precision manipulation workspace. *Int. J. Robot. Res.* **34**, 1731–1746 (2015). doi: 10.1177/0278364915608250
50. R. Ma, A. Dollar, Yale openhand project: Optimizing open-source hand designs for ease of fabrication and adoption. *IEEE Robot. Autom. Mag.* **24**, 32–40 (2017). doi: 10.1109/MRA.2016.2639034
51. K. Ghazi-Zahedi, R. Deimel, G. Montúfar, V. Wall, O. Brock, "Morphological computation: the good, the bad, and the ugly" in *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2017), pp. 464–469.
52. R. Pfeifer, J. C. Bongard, *How the Body Shapes the Way We Think: A New View of Intelligence* (MIT Press, 2006).
53. R. de Souza, S. El-Khoury, J. Santos-Victor, A. Billard, Recognizing the grasp intention from human demonstration. *Robot. Auton. Syst.* **74**, 108–121 (2015). doi: 10.1016/j.robot.2015.07.006
54. M. Madry, D. Song, D. Kragic, "From object categories to grasp transfer using probabilistic reasoning" in *Proceedings of the 2012 IEEE International Conference on Robotics and Automation* (IEEE, 2012), pp. 1716–1723.
55. R. Detry, J. Papon, L. Matthies, "Task-oriented grasping with semantic and geometric scene understanding" in *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2017), pp. 3266–3273.
56. M. Kokic, J. A. Stork, J. A. Haustein, D. Kragic, "Affordance detection for task-specific grasping using deep learning" in *Proceedings of the 2017 IEEE International Conference on Humanoid Robots* (IEEE, 2017), pp. 91–98.
57. A. Nguyen, D. Kanoulas, D. G. Caldwell, N. G. Tsagarakis, "Detecting object affordances with convolutional neural networks" in *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2016), pp. 2765–2770.
58. A. Ghadirzadeh, A. Maki, D. Kragic, M. Björkman, "Deep predictive policy training using reinforcement learning" in *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2017), pp. 2351–2358.
59. J. Shi, J. Z. Woodruff, K. M. Lynch, "Dynamic in-hand sliding manipulation" in *Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2015), pp. 870–877.
60. P. Agrawal, A. Nair, P. Abbeel, J. Malik, S. Levine, "Learning to poke by poking: Experiential learning of intuitive physics" in *Proceedings of IEEE Conference on Neural Information Processing Systems* (IEEE, 2016), pp. 5074–5082.
61. R. Antonova, S. Cruciani, C. Smith, D. Kragic, Reinforcement learning for pivoting task. arxiv:1703.00472 [cs.RO] (1 March 2017).
62. J. Z. Woodruff, K. M. Lynch, "Planning and control for dynamic, non-prehensile, and hybrid manipulation tasks" in *Proceedings of the 2017 IEEE International Conference on Robotics and Automation* (IEEE, 2017), pp. 4066–4073.
63. J. Bohg *et al.*, Interactive perception: Leveraging action in perception and perception in action. *IEEE Trans. Robot.* **33**, 1273–1291 (2017). doi: 10.1109/TRO.2017.2721939
64. M. Gemici, A. Saxena, "Learning haptic representation for manipulating Deformable food objects" in *Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2014), pp. 638–645.
65. L. P. Ureche, A. Billard, Constraints extraction from asymmetrical bimanual tasks and their use in coordinated behavior. *Robot. Auton. Syst.* **103**, 222–235 (2018). doi: 10.1016/j.robot.2017.12.011
66. R. Ozawa, K. Tahara, Grasp and dexterous manipulation of multifingered robotic hands: A review from a control view point. *Adv. Robot.* **31**, 1030–1050 (2017). doi: 10.1080/01691864.2017.1365011
67. N. Rahman, L. Carbonari, M. D'Imperio, C. Canali, D. G. Caldwell, F. Cannella, "A dexterous gripper for in-hand manipulation" in *Proceedings of the 2016 IEEE International Conference on Advanced Intelligent Mechatronics* (IEEE, 2016), pp. 377–382.
68. W. G. Bircher, A. M. Dollar, N. Rojas, "A two-fingered robot gripper with large object reorientation range" in *Proceedings of the 2017 IEEE International Conference on Robotics and Automation* (IEEE, 2017), pp. 3453–3460.
69. N. C. Dafle, A. Rodriguez, R. Paolini, B. Tang, S. S. Srinivasa, M. Erdmann, M. T. Mason, I. Lundberg, H. Staab, T. Fuhlbrigge, "Extrinsic dexterity: In-hand manipulation with external forces" in *Proceedings of the 2014 IEEE International Conference on Robotics and Automation* (IEEE, 2014), pp. 1578–1585.
70. N. Vahrenkamp, M. Przybylski, T. Asfour, R. Dillmann, "Bimanual grasp planning" in *Proceedings of the 2011 IEEE/RAS International Conference on Humanoid Robots* (IEEE, 2011), pp. 493–499.
71. C. Smith *et al.*, Dual arm manipulation-A survey. *Robot. Auton. Syst.* **60**, 1340–1353 (2012). doi: 10.1016/j.robot.2012.07.005
72. S. Cruciani, C. Smith, D. Kragic, K. Hang, "Dexterous manipulation graphs" in *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2018), pp. 2040–2047.
73. M. T. Mason, Toward robotic manipulation. *Annu. Rev. Control Rob. Auton. Syst* **1**, 1–28 (2018). doi: 10.1146/annurev-control-060117-104848
74. K. Bouyarmane, K. Chappellet, J. Vaillant, A. Kheddar, Quadratic programming for multirobot and task-space force control. *IEEE Trans. Robot.* **35**, 64–77 (2018). doi: 10.1109/TRO.2018.2876782
75. O. B. Kroemer, R. Detry, J. Piater, J. Peters, Combining active learning and reactive control for robot grasping. *Robot. Auton. Syst.* **58**, 1105–1116 (2010). doi: 10.1016/j.robot.2010.06.001
76. Y. Bekiroglu, J. Laaksonen, J. A. Jørgensen, V. Kyrki, D. Kragic, Assessing grasp stability based on learning and haptic data. *IEEE Trans. Robot.* **27**, 616–629 (2011). doi: 10.1109/TRO.2011.2132870
77. R. Detry, C. H. Ek, M. Madry, J. Piater, D. Kragic, "Learning a dictionary of prototypical grasp-predicting parts from grasping experience" in *Proceedings of the 2013 IEEE International Conference on Robotics and Automation* (IEEE, 2013), pp. 601–608.

78. A. Boularias, O. Kroemer, J. Peters, "Learning robot grasping from 3-D images with markov random fields" in *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2011), pp. 1548–1553.
79. S. Kim, A. Shukla, A. Billard, Catching Objects in Flight. *IEEE Trans. Robot.* **30**, 1049–1065 (2014). doi: 10.1109/TRO.2014.2316022
80. D. Nguyen-Tuong, J. Peters, Model learning for robot control: A survey. *Cogn. Process.* **12**, 319–340 (2011). doi: 10.1007/s10339-011-0404-1; pmid: 21487784
81. M. Andrychowicz, B. Baker, M. Chociej, R. Jozefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, J. Schneider, S. Sidor, J. Tobin, P. Welinder, L. Weng, W. Zaremba, Learning dexterous in-hand manipulation. arxiv:1808.00177 [cs.LG] (1 August 2018).
82. H. van Hoof, T. Hermans, G. Neumann, J. Peters, "Learning robot in-hand manipulation with tactile features" in *Proceedings of the 2015 IEEE/RAS International Conference on Humanoid Robots* (IEEE, 2015), pp 121–127.
83. V. Kumar, E. Todorov, S. Levine, "Optimal control with learned local models: Application to dexterous manipulation" in *Proceedings of the 2016 IEEE International Conference on Robotics and Automation* (IEEE, 2016), pp. 378–383.
84. Y. Yang, Y. Li, C. Fermüller, Y. Aloimonos, "Robot learning manipulation action plans by "watching" unconstrained videos from the World Wide Web" in *Proceedings of the 29th Conference on Artificial Intelligence* (AAAI Press, 2015), pp. 3686–3693.
85. J. Krüger, T. K. Lien, A. Verl, Cooperation of human and machines in assembly lines. *CIRP Ann. Manuf. Technol.* **58**, 628–646 (2009). doi: 10.1016/j.cirp.2009.09.009
86. A. Cherubini, R. Passama, A. Crosnier, A. Lasnier, P. Fraisse, Collaborative manufacturing with physical human-robot interaction. *Robot. Comput.-Integr. Manuf.* **40**, 1–13 (2016). doi: 10.1016/j.rcim.2015.12.007
87. A. Ajoudani *et al*., Progress and prospects of the human-robot collaboration. *Auton. Robots* **42**, 957–975 (2018). doi: 10.1007/s10514-017-9677-2
88. M. Huber, A. Knoll, T. Brandt, S. Glasauer, Handing over a cube: Spatial features of physical joint-action. *Ann. N. Y. Acad. Sci.* **1164**, 380–382 (2009). doi: 10.1111/j.1749-6632.2008.03743.x; pmid: 19645931
89. J. Mainprice, M. Gharbi, T. Simeon, R. Alami, "Sharing effort in planning human-robot handover tasks" in *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication* (IEEE, 2012), pp. 764–770.
90. C.-M. Huang, M. Cakmak, B. Mutlu, "Adaptive coordination strategies for human-robot handovers" in *Proceedings of the 2015 Conference on Robotics: Science and Systems* (2015); www.roboticsproceedings.org/rss11/p31.html.
91. M. Cakmak, S. S. Srinivasa, M. K. Lee, S. Kiesler, J. Forlizzi, "Using spatial and temporal contrast for fluent robot-human handovers" in *Proceedings of the International Conference on Human-Robot Interaction* (IEEE, 2011), p. 489.
92. P. Basili, M. Huber, T. Brandt, S. Hirche, S. Glasauer, "Investigating human-human approach and hand-over" in *Human Centered Robot Systems* (Cognitive Systems Monographs, Springer, 2009), vol. 6, pp. 151–160.
93. A. Edsinger, C. C. Kemp, "Human-robot interaction for cooperative manipulation: Handing objects to one another" in *Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication* (IEEE, 2007), pp. 1167–1172.
94. J. Aleotti, V. Micelli, S. Caselli, "Comfortable robot to human object hand-over" in *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication* (IEEE, 2012), pp. 771–776.
95. M. Pan, E. Croft, G. Niemeyer, "Exploration of geometry and forces occurring within human-to-robot handovers" in *2018 IEEE Haptics Symposium* (IEEE, 2018), pp. 327–333.
96. K. Yokoyama, H. Handa, T. Isozumi, Y. Fukase, K. Kaneko, F. Kanehiro, Y. Kawai, F. Tomita, H. Hirukawa, "Cooperative works by a human and a humanoid robot" in *Proceedings of the 2003 IEEE International Conference on Robotics and Automation* (IEEE, 2003), vol. 3, pp. 2985–2991.
97. E. Berger, D. Vogt, N. Haji-Ghassemi, B. Jung, H. B. Amor, "Inferring guidance information in cooperative human-robot tasks" in *Proceedings of the 2013 IEEE/RAS International Conference on Humanoid Robots* (IEEE, 2013), pp. 124–129.
98. M. Bellaccini, L. Lanari, A. Paolillo, M. Vendittelli, "Manual guidance of humanoid robots without force sensors: Preliminary experiments with NAO" in *Proceedings of the 2014 IEEE International Conference on Robotics and Automation* (IEEE, 2014), pp. 1184–1189.
99. A. Bussy, P. Gergondet, A. Kheddar, F. Keith, A. Crosnier, "Proactive behavior of a humanoid robot in a haptic transportation task with a human partner" in *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication* (IEEE, 2012), pp. 962–967.
100. K. Kosuge, M. Sato, N. Kazamura, "Mobile robot helper" in *Proceedings of the 2000 IEEE International Conference on Robotics and Automation* (IEEE, 2000), vol. 1, pp. 583–588.
101. J. Stuckler, S. Behnke, "Following human guidance to cooperatively carry a large object" in *Proceedings of the 2011 IEEE/RAS International Conference on Humanoid Robots* (IEEE, 2011), pp. 218–223.
102. M. Lawitzky, A. Mortl, S. Hirche, "Load sharing in human-robot cooperative manipulation" in *Proceedings of the 19th IEEE International Symposium on Robot and Human Interactive Communication* (IEEE, 2010), pp. 185–191.
103. Y. Karayiannidis, C. Smith, D. Kragic, "Mapping human intentions to robot motions via physical interaction through a jointly-held object" in *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication* (IEEE, 2014), pp. 391–397.
104. D. Kruse, R. J. Radke, J. T. Wen, "Collaborative human-robot manipulation of highly deformable materials" in *Proceedings of the 2015 IEEE International Conference on Robotics and Automation* (IEEE, 2015), pp. 3782–3787.
105. Y. Gao, H. J. Chang, Y. Demiris, "Iterative path optimization for personalized dressing assistance using vision and force information" in *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2016), pp. 4398–4403.
106. L. Lu, J. T. Wen, Human-directed coordinated control of an assistive mobile manipulator. *Int. J. Intell. Robot. Appl.* **1**, 104–120 (2017). doi: 10.1007/s41315-016-0005-3

ACKNOWLEDGMENTS

We thank the reviewers for many helpful comments to improve the article, A. Kheddar and J. Paik for providing images of their research, and L. Cohen for hand-drawn illustrations. **Funding:** We acknowledge funding from the European Research Council, the Knut and Alice Wallenberg Foundation, and the Swedish Foundation for Strategic Research. **Competing interests:** The authors declare no competing interests.

10.1126/science.aat8414

RUMINANT GENOMES

RESEARCH ARTICLES

RELATED ITEMS

► PERSPECTIVE P. 1130
► VIDEO

RESOLVING RUMINANTS

By **Laura M. Zahn**

Ruminants are one of the most diverse groups of mammals. All members have a specialized digestive organ known as the rumen, as well as other organs that permit them to subsist on plants. These adaptations allow ruminants to digest foliage more efficiently than other mammalian species and may underlie the success of this clade. Additionally, most ruminant species have impressive bony outgrowths, collectively referred to as headgear, that may aid in mate selection. These varied ornaments, some of which are exclusive to males, include the horns carried throughout the lives of sheep, antelope, and cattle, as well as the annually regenerating deer antlers, one of the fastest-growing tissues of any mammal. Here, *Science* presents three studies that use genomic analyses to resolve the evolutionary relationships among the major ruminant lineages, identify the genes involved in the evolution of headgear, and examine how reindeer adapted to the long days and nights of the Arctic region. These studies allow us to better appreciate the familiar domestic cows, sheep, and deer and their wild relatives and to peer into the adaptations underlying the antelopes, giraffes, bison, and pronghorn that occupy the world's grasslands.

An adult female reindeer (*Rangifer tarandus*) in Cairngorms, Scotland, UK. Reindeer are ruminant mammals that exhibit species-specific adaptations to their harsh Arctic distribution.

PHOTO: MARK HAMBLIN/NATUREPL.COM/GETTY IMAGES

Published by AAAS

RESEARCH ARTICLE SUMMARY

RUMINANT GENOMICS

Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits

Lei Chen*, Qiang Qiu*, Yu Jiang*, Kun Wang*, Zeshan Lin*, Zhipeng Li*, Faysal Bibi, Yongzhi Yang, Jinhuan Wang, Wenhui Nie, Weiting Su, Guichun Liu, Qiye Li, Weiwei Fu, Xiangyu Pan, Chang Liu, Jie Yang, Chenzhou Zhang, Yuan Yin, Yu Wang, Yue Zhao, Chen Zhang, Zhongkai Wang, Yanli Qin, Wei Liu, Bao Wang, Yandong Ren, Ru Zhang, Yan Zeng, Rute R. da Fonseca, Bin Wei, Ran Li, Wenting Wan, Ruoping Zhao, Wenbo Zhu, Yutao Wang, Shengchang Duan, Yun Gao, Yong E. Zhang, Chunyan Chen, Christina Hvilsom, Clinton W. Epps, Leona G. Chemnick, Yang Dong, Siavash Mirarab, Hans Redlef Siegismund, Oliver A. Ryder, M. Thomas P. Gilbert, Harris A. Lewin, Guojie Zhang†, Rasmus Heller†, Wen Wang†

INTRODUCTION: The ruminants are one of the most successful mammalian lineages, exhibiting extensive morphological and ecological diversity and containing several key livestock species, such as cattle, buffalo, yak, sheep, and goat. Ruminants have evolved several distinct characteristics such as a multichambered stomach, cranial appendages (headgear), specialized dentition, a highly cursorial locomotion, and a wide range of body size variations. Despite their biological prominence and value to human societies, the evolutionary history of ruminants has not been fully resolved, and the molecular mechanisms underlying their particular characteristics remains largely unknown.

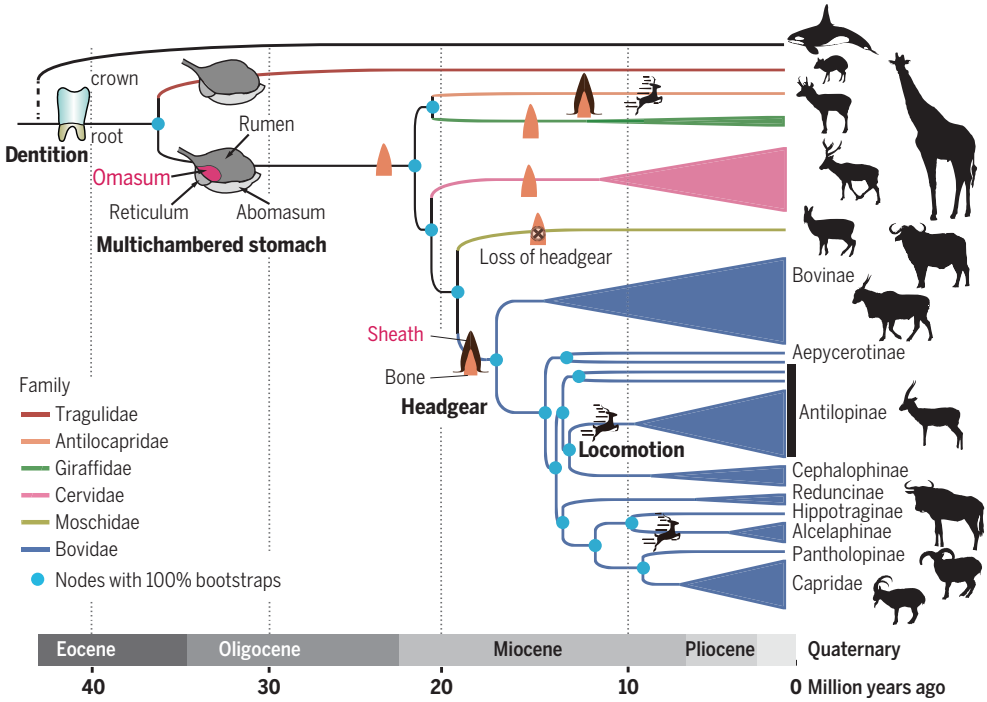
RATIONALE: We seek to resolve the controversies in the ruminant phylogeny and reveal the genetic basis underpinning the evolutionary innovations in ruminants. Here, we report the newly sequenced genomes of 44 ruminant species, covering about half the genera and all six extant Ruminantia families. We included seven published ruminant genomes (five bovids and two cervids) to reconstruct the phylogenetic tree by using improved time calibrations. We also reconstructed the Pleistocene demographic histories of these ruminant species using whole-genome heterozygosity information. Together with transcriptomic data of 516 samples from 68 tissues of four species, we conducted comparative genomic analyses to reveal candidate genes and regulatory elements that might have contributed to the evolution of the distinct ruminant characteristics.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aav6202

RESULTS: Using whole-genome orthologous sequences obtained from 51 ruminants, we have produced a new well-supported ruminant phylogenetic tree. The new tree resolves previous controversies over the deep branches of ruminant families, as well as the highly radiated Bovidae family. We estimated the emergence of crown Ruminantia to the late Oligocene (39.1 million to 32.3 million years ago) and that of Pecora to the Neocene (23.3 million to 20.8 million years ago). Investigations of demographic history revealed massive population decline events that occurred in most ruminant species, starting from ~100,000 to 50,000 years ago, which was temporally and spatially concurrent with the increased human activities on different continents during this period. We further identified many genomic changes that associate with important evolutionary innovations, such as the multichambered stomach, headgear, body size variation, cursorial locomotion, and dentition.

CONCLUSION: Our results demonstrate the power of using comparative phylogenomic approaches in resolving the deep branches of phylogeny that result from rapid radiations. The data and results presented in this study provide valuable resources and insights into the evolution of ruminant and mammalian biology. ■



Phylogeny and trait evolution of ruminants. The phylogenic tree of ruminants is presented with the species within same families and subfamilies collapsed. The ruminants have many textbook examples of distinct traits. The four-chambered stomach with omasum chamber is a key innovation evolved in pecoran ruminants. Headgear keratinous sheath only appear in Bovidae and Antilocapridae lineages. Many ruminants have evolved high-crowned or hypsodont teethes. The Antilocapridae and two bovid lineages are among the mammals with highest cursorial locomotion ability.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. E-mail: wwang@mail.kiz.ac.cn (W.W.); rheller@bio.ku.dk (R.H.); guojie.zhang@bio.ku.dk (G.Z.)
Cite this article as L. Chen *et al.*, *Science* **364**, eaav6202 (2019). DOI: 10.1126/science.aav6202

RESEARCH ARTICLE

RUMINANT GENOMICS

Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits

Lei Chen[1]*, Qiang Qiu[1]*, Yu Jiang[2]*, Kun Wang[1]*, Zeshan Lin[1]*, Zhipeng Li[3]*, Faysal Bibi[4], Yongzhi Yang[5], Jinhuan Wang[6], Wenhui Nie[6], Weiting Su[6], Guichun Liu[1,7], Qiye Li[8], Weiwei Fu[2], Xiangyu Pan[2], Chang Liu[1], Jie Yang[1], Chenzhou Zhang[1], Yuan Yin[1], Yu Wang[2], Yue Zhao[2], Chen Zhang[1], Zhongkai Wang[1], Yanli Qin[1], Wei Liu[7], Bao Wang[7], Yandong Ren[7], Ru Zhang[1], Yan Zeng[7], Rute R. da Fonseca[9,10], Bin Wei[2], Ran Li[2], Wenting Wan[1,7], Ruoping Zhao[7], Wenbo Zhu[1], Yutao Wang[11], Shengchang Duan[12], Yun Gao[12], Yong E. Zhang[13,14,15], Chunyan Chen[13,14], Christina Hvilsom[16], Clinton W. Epps[17], Leona G. Chemnick[18], Yang Dong[19,20], Siavash Mirarab[21], Hans Redlef Siegismund[9], Oliver A. Ryder[18,22], M. Thomas P. Gilbert[23,24], Harris A. Lewin[25], Guojie Zhang[7,8,15,26]†, Rasmus Heller[9]†, Wen Wang[1,7,15]†

The ruminants are one of the most successful mammalian lineages, exhibiting morphological and habitat diversity and containing several key livestock species. To better understand their evolution, we generated and analyzed de novo assembled genomes of 44 ruminant species, representing all six Ruminantia families. We used these genomes to create a time-calibrated phylogeny to resolve topological controversies, overcoming the challenges of incomplete lineage sorting. Population dynamic analyses show that population declines commenced between 100,000 and 50,000 years ago, which is concomitant with expansion in human populations. We also reveal genes and regulatory elements that possibly contribute to the evolution of the digestive system, cranial appendages, immune system, metabolism, body size, cursorial locomotion, and dentition of the ruminants.

Ruminantia is an important group of terrestrial herbivores, including at least 200 extant species (*1*) spanning six families: Tragulidae, Antilocapridae, Giraffidae, Moschidae, Cervidae, and Bovidae. The most species-rich of these families is Bovidae, which encompasses at least 143 species (*2*, *3*), including important livestock animals (such as cattle, buffalo, yak, sheep, and goat) (*4*, *5*). Ruminants possess several distinct and characteristic anatomical hallmarks, such as a multichambered stomach and cranial appendages (headgear). The acquisition of the rumen in the ruminants and of the omasum in pecorans (all ruminants except the Tragulidae) allow these animals to use plant material with a higher efficiency than other herbivorous mammals, such as equids (*6*–*8*). This effect is believed to be associated with the evolutionary success of these animals in terms of diversity, abundance, and geographic range (*9*). Ruminants have also evolved extreme morphological diversity, ranging in body weight from <2 kg to >1200 kg (*10*, *11*), and a wide variety of distinct behavioral and physiological traits. Ultimately, these adaptations in ruminants likely explain the remarkable abundance of these animals among domestic animals.

Despite their biological prominence and value to human civilization, much remains to be learned about the ruminants. For example, the phylogeny of ruminants is far from resolved, and inconsistencies remain even at the family level (*12*–*15*). Moreover, the genetic basis underlying many of their characteristic traits remains unknown. To fill in our gaps in knowledge, we performed de novo assembly of the genomes of 44 ruminant species, representing 36 genera that span all six families. In combination with five previously published bovid genomes, two published cervid genomes (*16*–*22*), and recently updated fossil information (*15*), we constructed a time-calibrated phylogenetic tree of the group, analyzed species population histories, and investigated the genomic evolution of these species. Our results not only provide data for understanding the origin and evolution of this important mammalian group and their particular traits but also have implications for placing ruminant livestock genomic resources into an evolutionary context and for conserving ruminant biodiversity.

Results

Genome sequencing, assembly, and annotation

We used Illumina sequencing technology (*23*) to generate more than 40 trillion base pairs (Tbp) of raw data and then de novo assembled genomes for 44 ruminant species (Table 1 and tables S1 and S2). Eleven of the species are listed as "vulnerable" or worse on the International Union for Conservation of Nature (IUCN) Red List (table S1). Forty of the genomes were assembled into large scaffolds (Table 1 and tables S3 and S4) with a series of mate-paired large-insert libraries, whereas four genomes—common eland (*Taurotragus oryx*), mountain nyala (*Tragelaphus buxtoni*), bongo (*Tragelaphus eurycerus*), and oribi (*Ourebia ourebi*)—were only assembled to the contig level because of degenerated samples but still qualified for most comparative genomic analyses (Table 1 and table S4). In addition, we improved the quality of the genome assembly of the black muntjac deer (*Muntiacus crinifrons*; contig N50 = 2.4 Mbp)

[1]Center for Ecological and Environmental Sciences, Northwestern Polytechnical University, Xi'an 710072, China. [2]Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling 712100, China. [3]Department of Special Animal Nutrition and Feed Science, Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, Changchun 130112, China. [4]Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Invalidenstrasse 43, 10115 Berlin, Germany. [5]State Key Laboratory of Grassland Agro-Ecosystem, School of Life Sciences, Lanzhou University, Lanzhou, 730000, China. [6]Kunming Cell Bank, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. [7]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. [8]China National GeneBank, BGI-Shenzhen, Shenzhen 518120, China. [9]Section for Computational and RNA Biology, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark. [10]Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark. [11]College of Life and Geographic Sciences, Kashgar University, Kashgar 844000, China. [12]Nowbio Biotechnology Company, Kunming 650201, China. [13]Key Laboratory of Zoological Systematics and Evolution and State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China. [14]University of Chinese Academy of Sciences, Beijing 100049, China. [15]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China. [16]Copenhagen Zoo, Frederiksberg, Denmark. [17]Department of Fisheries and Wildlife, Oregon State University, Corvallis, OR 97331, USA. [18]San Diego Zoo Institute for Conservation Research, Escondido, CA 92027, USA. [19]Yunnan Research Institute for Local Plateau Agriculture and Industry, Kunming 650201, China. [20]State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University, Kunming 650201, China. [21]Department of Electrical and Computer Engineering, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA. [22]Evolution, Behavior, and Ecology, Division of Biology, University of California, San Diego, La Jolla, CA 92093, USA. [23]EvoGenomics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark. [24]Norwegian University of Science and Technology, University Museum, 7491 Trondheim, Norway. [25]Department of Evolution and Ecology and the UC Davis Genome Center, University of California, Davis, CA, 95616, USA. [26]Section for Ecology and Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark.
*These authors contributed equally to this work.
†Corresponding author. E-mail: wwang@mail.kiz.ac.cn (W.W.); rheller@bio.ku.dk (R.H.); guojie.zhang@bio.ku.dk (G.Z.)

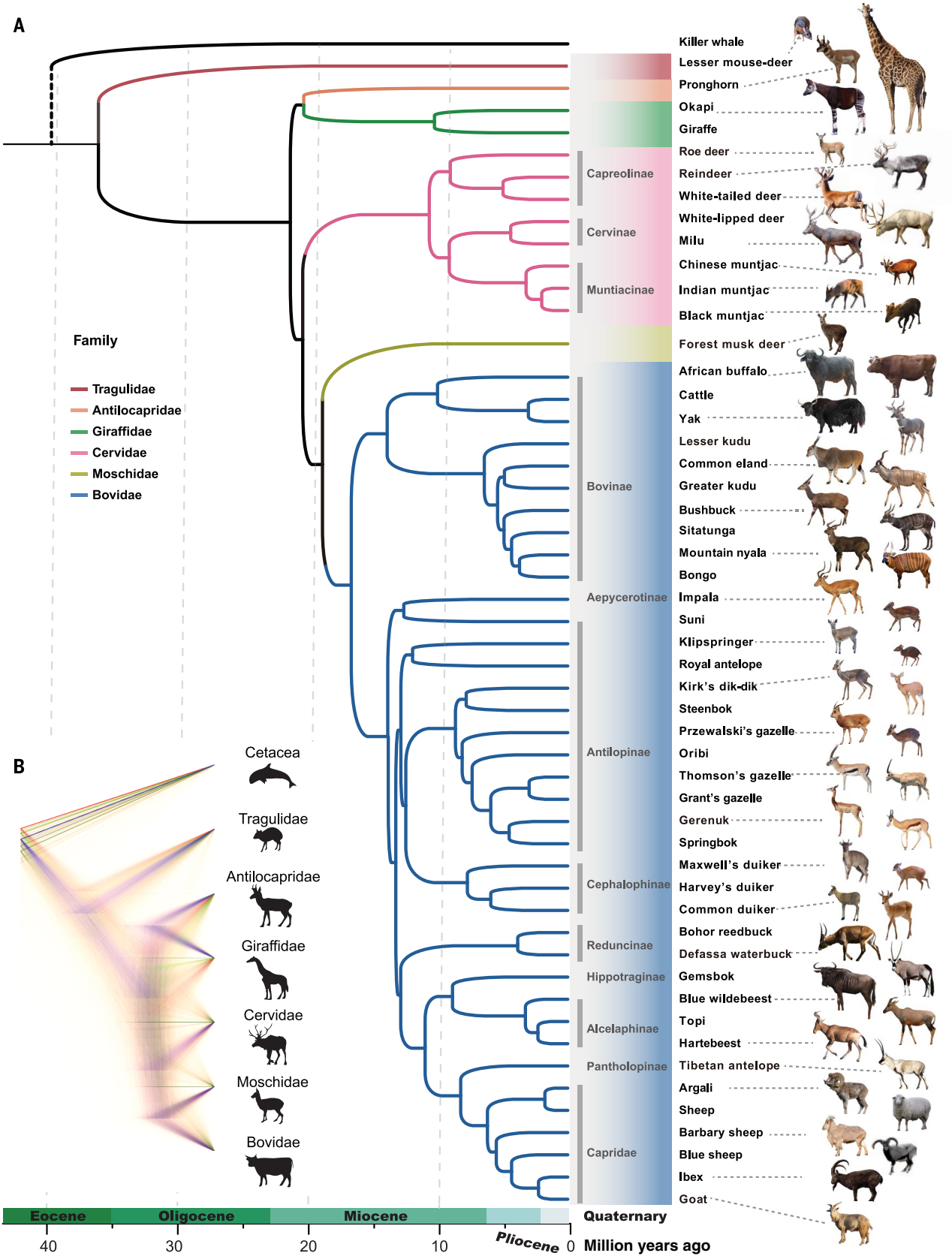


Fig. 1. Phylogeny of ruminants. (**A**) The maximum likelihood phylogenetic tree from whole-genome sequences of 51 ruminant species and 13 fossil calibrations. To compute the node supports, 200 bootstraps were used, and all nodes have 100% support. The origin and credit of the portraits of different species are listed in table S54. (**B**) Prevalent discordance among 10,000 random WGTs was observed across different families of ruminant.

using long reads generated from PacBio Single Molecule, Real-Time (SMRT) sequencing. This improved assembly of a Cervidae species facilitated synteny and other comparative analyses within the ruminants. Three genomes—reindeer (*Rangifer tarandus*), milu deer (*Elaphurus davidianus*), and Marco Polo sheep (*Ovis ammon*)—were released in data notes before this study (*24*–*26*).

To evaluate the quality of these genome assemblies (fig. S1 and tables S5 to S7), we performed BUSCO and synteny analyses and observed high BUSCO (*27*) scores (average 87%, from 75 to 93%) (table S6) and synteny continuity (table S7), showing that the majority of the assemblies were of high quality for downstream comparative analyses. The assembled genomes ranged in size from 2.52 to 3.25 Gbp, which was mostly consistent with the cytological *C* values (*28*) and *k*-mer–based estimations (table S5), indicating relatively complete genome coverages for all species.

After masking repeats (table S8), we used both de novo and homology-based gene prediction to annotate the genomes. The protein sequences of human (Ensembl 87 release), cattle (Ensembl 87 release), and sheep (Ensembl 87 release) were used as templates for homology-based gene prediction. The final annotated gene numbers range from 19,304 to 25,753 (table S9) among different species, with variations driven primarily by the quality of the assembled genomes. The gene length and exon length distributions were similar to those of other mammals (fig. S2 and table S9). We also identified 221,166 ruminant-specific conserved nonexonic elements (RSCNEs) by comparing all the ruminant genomes to 12 mammalian outgroup species (fig. S3 and table S10). These RSCNEs occupy ~0.61% of the genomes (~16.5 Mbp) (table S11) (*23*). A user-friendly, publicly available genome browser database was established (http://animal.nwsuaf.edu.cn/code/index.php/Ruminantia) for visualization of the genomic and transcriptomic data presented in this study.

Resolving the ruminant phylogeny

The lack of a fully resolved phylogeny for ruminants (*12*–*15*) still hinders an understanding of the evolution of ruminant diverse phenotypes. In particular, the phylogenetic positions of the Antilocapridae and Moschidae families have been strongly debated, and so have the relationships among subfamilies within the diverse Bovidae family [(*12*), review]. Furthermore, although "dwarf antelopes" have previously been grouped in the tribe Neotragini, recent molecular studies suggest that these animals are polyphyletic (*14*, *15*). These controversies are probably attributable to convergent evolution (challenging morphology-based approaches) and incomplete lineage sorting (ILS) in conjunction with the short internal branches in the ruminant radiation (challenging genetic inference) (*12*).

To resolve these phylogenetic challenges, we first estimated a whole-genome phylogenetic tree with ExaML under the GTR+GAMMA model (*23*) using the killer whale genome (*29*) as an outgroup. In total, 373 Mbp of orthologous syntenic sequences were obtained from whole-genome alignment by using the goat as a reference genome (*19*), yielding a whole-genome tree with 100% bootstrap support for all nodes (Fig. 1A). We also performed phylogenetic analyses with other genome partitions, including 6406 orthologous protein-coding genes identified in the 51 ruminant species and the killer whale, the fourfold degenerate sites in these genes, conserved nonexonic elements (CNEs), and complete mitochondrial genomes (mtDNA). With the exception of the mtDNA tree, the topologies of all other trees were identical with that of whole-genome tree (Fig. 1A and figs. S4 to S10).

Although the phylogenetic tree constructed with concatenated nuclear genome sequences (nDNA tree) was highly supported, phylogenetic discordance was pervasive across genomic regions (Fig. 1B). To further assess genome-wide tree discordance, 10,000 random genomic 1-Kbp windows at least 50 Kbp distant from each other were extracted. These 1-Kbp windows had enough segregating sites to generate window-based gene trees (WGTs) of high resolution (fig. S11). Although all individual WGTs differed from the species-level nDNA tree topology, 21.3% of WGTs exhibited the same family-level topology as that of the concatenated nDNA tree (table S12). This type of tree incongruence is usually caused by ILS, which has been widely observed in phylogenies of many animal groups—such as African cichlids (*30*), birds (*31*), and great apes (*32*, *33*)—and is known to be exacerbated by the effects of gene tree estimation error (*34*, *35*).

To ensure that the reconstructed phylogeny is robust in the presence of ILS, we performed several analyses with the coalescent-based phylogenetic method ASTRAL (*36*) using the 10,000 WGTs. We did not use entire genes as the unit of gene

Table 1. Assembly statistics of 44 ruminant species.

Species	Common name	Scaffold N50 (bp)	Contig N50 (bp)	Species	Common name	Scaffold N50 (bp)	Contig N50 (bp)
Tragulus javanicus	Lesser mouse-deer	243,250	6286	*Redunca redunca*	Bohor reedbuck	438,845	17,874
Antilocapra americana	Pronghorn	1,463,792	61,696	*Syncerus caffer*	African buffalo	2,316,376	11,115
Okapia johnstoni	Okapi	3,620,116	58,892	*Gazella thomsoni*	Thomson gazelle	1,581,717	36,935
Giraffa camelopardalis	Giraffe	3,197,404	22,538	*Tragelaphus strepsiceros*	Greater kudu	520,720	16,623
Muntiacus muntjak	Indian muntjac	1,398,591	10,925	*Nanger granti*	Grant's gazelle	520,131	6041
Muntiacus reevesi	Chinese muntjac	1,253,719	68,151	*Sylvicapra grimmia*	Common duiker	541,191	5209
Elaphurus davidianus	Milu	3,040,530	32,708	*Aepyceros melampus*	Impala	343,699	51,379
Rangifer tarandus	Reindeer	1,059,113	91,805	*Madoqua kirkii*	Kirk's dik-dik	489,835	26,372
Muntiacus crinifrons	Black muntjac	NA	1,458,913	*Oreotragus oreotragus*	Klipspringer	340,335	13,019
Gervus albirostris	White-lipped deer	3,567,448	22,599	*Antidorcas marsupialis*	Springbok	698,575	10,778
Moschus chrysogaster	Forest musk deer	2,509,225	57,721	*Tragelaphus imberbis*	Lesser kudu	1,774,691	6545
Oryx gazella	Gemsbok	1,583,972	18,132	*Tragelaphus spekii*	Sitatunga	78,973	5410
Litocranius walleri	Gerenuk	3,128,641	47,546	*Philantomba maxwellii*	Maxwell's duiker	390,552	4423
Damaliscus lunatus	Topi	1,172,125	25,829	*Raphicerus campestris*	Steenbok	474,033	5764
Ammotragus lervia	Barbary sheep	1,263,981	18,541	*Neotragus pygmaeus*	Royal antelope	365,736	5931
Pseudois nayaur	Blue sheep	2,076,308	23,854	*Alcelaphus buselaphus*	Hartebeest	11,258	4274
Ovis ammon	Argali	5,734,776	45,638	*Capra ibex*	Ibex	15,190,720	24,835
Kobus ellipsiprymnus	Defassa waterbuck	782,102	20,722	*Neotragus moschatus*	Suni	957,022	8233
Procapra przewalskii	Przewalski's gazelle	5,152,914	20,018	*Taurotragus oryx*	Common eland	no scaffold	1262
Connochaetes taurinus	Blue wildebeest	3,511,341	46,638	*Tragelaphus buxtoni*	Mountain nyala	no scaffold	1309
Cephalophus harveyi	Harvey's duiker	365,462	39,715	*Tragelaphus euryceros*	Bongo	no scaffold	1980
Tragelaphus scriptus	Bushbuck	890,554	9965	*Ourebia ourebi*	Oribi	no scaffold	1259

tree construction because genes can span over several recombination blocks in the genome (*37*). The topology of the obtained ASTRAL coalescent tree (fig. S12) is identical to the nDNA tree (Fig. 1A). We also used another coalescent-based phylogenetic method, MP-EST (*38*), to carry out additional phylogenetic analyses using both the 10,000 windows as well as the 6406 orthologous genes and obtained an identical topology as those of the ASTRAL and nDNA trees (fig. S13), further supporting the robustness of the inferred ruminant phylogeny. To minimize negative impacts of low tree resolution in the WGTs, we contracted low-supported branches [below 3, 10, and 20%, bootstrap support, as suggested in (*36*)] and still observed the same topology as that of the nDNA tree in all cases (fig. S14).

We further tested whether gene flow could have contributed to the observed topological discordances. The results from the ABBA-BABA test (*39*), the D_{FOIL} software (extend from 4 taxa to 5 taxa) (*40*), and admixturegraph (*41*) consistently suggested some possible ancient gene flows among Giraffidae, Cervidae, Bovidae, and Moschidae (fig. S15 and table S13). The strongest gene flow signal between Giraffidae and Cervidae-Bovidae-Moschidae coincides with an overrepresentation of this topology in the WGTs relative to the expectation of ILS, which requires that for a set of four species trees, alternative two phylogenetic topologies other than the species tree have equal proportion (fig. S16). By contrast, PhyloNet (*42*) and PhyloNetworks (*43*) were sensitive to model and parameter choices and did not generate consistent results (figs. S17 and S18). It is plausible that the high parameter space of the phylogenetic model precluded these methods from performing complete explorations of the parameter space in our large-time-scale phylogenomic data.

Our new ruminant tree supports a sister-group relationship for Antilocapridae and Giraffidae, representing the oldest branch among the extant pecoran families. This is different from the mtDNA-based phylogenies, which placed Antilocapridae as an outgroup to all other pecorans (figs. S9 and S10) (*12*, *14*, *15*). The sister relationship of these two groups has been proposed before (*44*, *45*) but now received 100% bootstrap support by the nDNA tree and a local posterior probability (*46*) of 1 in the ASTRAL tree and had 23% higher frequency of WGTs than those of the alternative topologies (fig. S16). Another contentious issue is on the placement of Moschidae, which was previously placed at the base of Pecora (*47*–*49*) or as a sister group to Bovidae (*50*) from morphological data. In some mtDNA studies, Moschidae was proposed as a sister group of Cervidae (*51*) or a sister taxon to Bovidae (*14*, *52*, *53*). Our nDNA tree confirmed that Moschidae is a sister group of Bovidae, and the frequency of WGTs further support this conclusion (fig. S16).

Our results also resolved some controversial subfamily relations in Bovidae, such as the

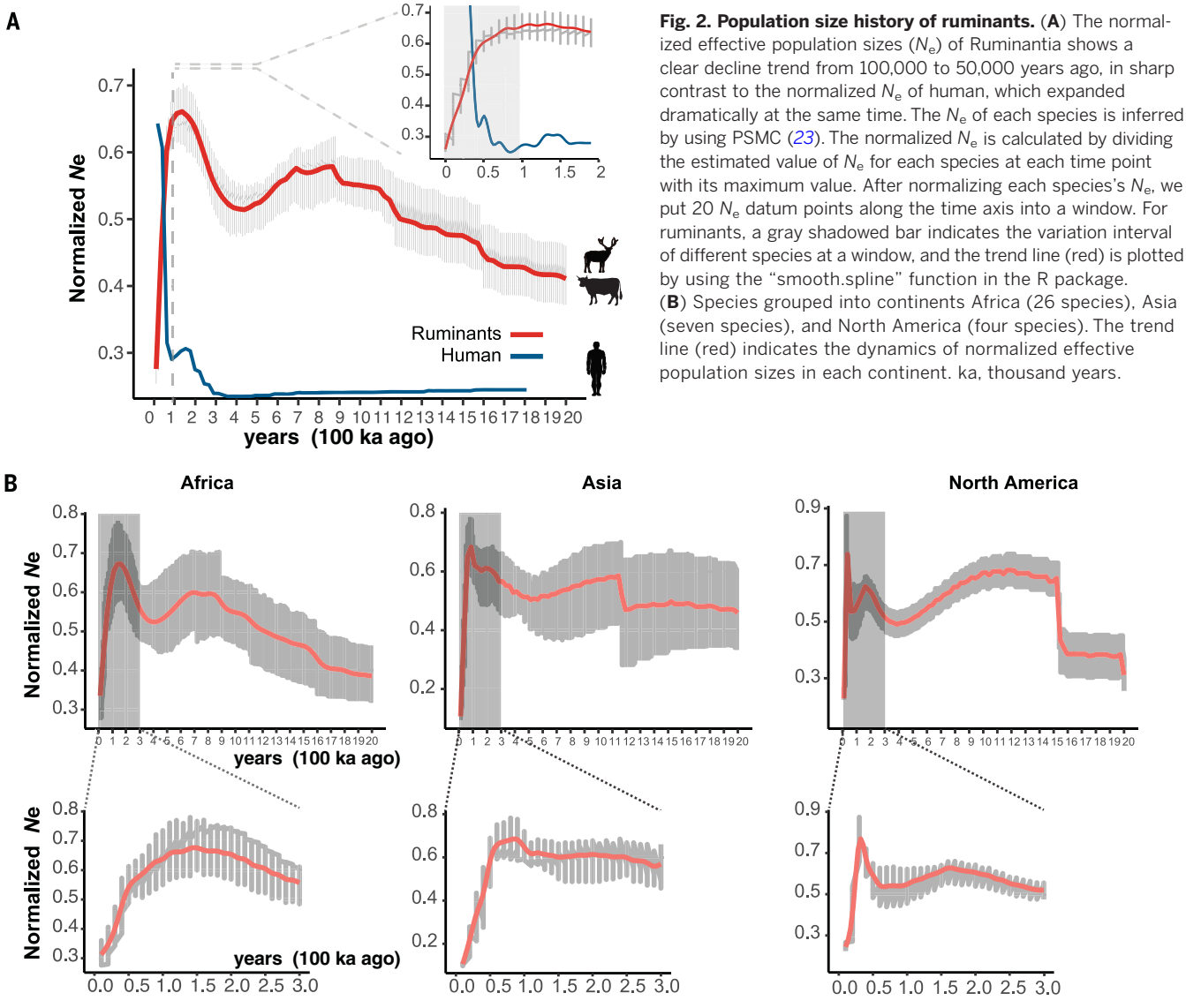


Fig. 2. Population size history of ruminants. (**A**) The normalized effective population sizes (N_e) of Ruminantia shows a clear decline trend from 100,000 to 50,000 years ago, in sharp contrast to the normalized N_e of human, which expanded dramatically at the same time. The N_e of each species is inferred by using PSMC (*23*). The normalized N_e is calculated by dividing the estimated value of N_e for each species at each time point with its maximum value. After normalizing each species's N_e, we put 20 N_e datum points along the time axis into a window. For ruminants, a gray shadowed bar indicates the variation interval of different species at a window, and the trend line (red) is plotted by using the "smooth.spline" function in the R package. (**B**) Species grouped into continents Africa (26 species), Asia (seven species), and North America (four species). The trend line (red) indicates the dynamics of normalized effective population sizes in each continent. ka, thousand years.

placement of the Reduncinae (12). With the whole-genome tree, Reduncinae were confirmed to be a sister group to the ancestral lineage of Caprinae, Alcelaphinae, and Hippotraginae. This topology is also supported by the ASTRAL tree and has a slightly higher WGT frequency than those of alternatives (fig. S19). We further confirmed the previously ambiguous sister-group relationship between impala (*Aepyceros melampus*) and suni (*Neotragus moschatus*) (54, 55) and found no phylogenetic support for the "waste-bucket" Neotragini tribe (56). These findings mostly agree with the topology from Decker *et al.* (57), which used single-nucleotide polymorphism (SNP) chip data but lacked samples from Moschidae and Tragulidae. Furthermore, our results provided higher bootstrap supports for the deep nodes and also refined some species positions within Bovidae—the position of bushbuck (*Tragelaphus scriptus*) and mountain nyala (*Tragelaphus buxtoni*).

Using fossil calibrations (tables S14 and S15) (15), we estimated the emergence of crown Ruminantia at 39.1 million to 32.3 million years ago (late Oligocene), and the emergence of Pecora at 23.3 million to 20.8 million years ago (Neocene) (Fig. 1A and figs. S20 to S27). The evolutionary rate in the ancestral ruminant lineage was ~1.5×10^{-9}, which was significantly higher than that in other mammals (Student's t test, $P < 0.01$) (fig. S28). Among the ruminant families, Tragulidae had the highest evolutionary rate, and Giraffidae had the lowest evolutionary rate (fig. S28). We found a significant negative correlation between evolutionary rate and log body size (fig. S29).

Pleistocene population dynamics in ruminants

We used our whole-genome dataset to investigate the demographic histories of different ruminant species using the pairwise sequentially Markovian coalescent (PSMC) method (58), which can infer changes in the effective population size (N_e) over the Pleistocene (figs. S30 and S31 and table S16). The analyses produced a species-specific demographic pattern with no clear grouping of patterns according to their habitat types or feeding types (fig. S32). This might imply that the ruminant species responded differentially to biotic and abiotic pressures associated with their different ecological niches (59). However, we found population declines for many species (25 out of the 40 species with scaffolded genome assemblies) starting from ~100,000 to 50,000 years ago (Fig. 2A), which suggests that late Pleistocene large mammal declines were much more severe than previously suspected, involving major declines in the populations of most species along with the mass extinction of large mammals at this time (60). We speculate that this community-wide decline might be at least partially attributable to human activities. This is supported by ruminant population declines coinciding with increasing human effective population size after the dispersal out of Africa during this period (Fig. 2A) (61). Intriguingly, the onset of ruminant decline coincided with the sequential population expansion of humans on different continents in this period (Fig. 2B), which would not be expected if the decline signal was a structural artefact. By contrast, ruminant population size dynamics showed no apparent correlation with the climatic changes dominated by serial glaciations during the Pleistocene (figs. S30 and S31). Taken together, these results highlight a possible role for humans in the massive decline of mammalian species in the late Pleistocene (60).

Structure and evolution of ruminant genomes

Synteny

We explored the general architecture of ruminant genomes through comparison with other high-quality mammalian genomes (23). Specifically, we compared the syntenic relationship between the goat (the most well-assembled ruminant genome) (19) and human (Primates) (62), dog (Carnivora) (63), horse (Perissodactyla) (64), pig (Suina) (65), camel (Tylopoda) (66), killer whale (Cetacea) (29), and black muntjac (Cervidae of Ruminantia). The high-quality assembly of the black muntjac (Table 1) by use of PacBio reads allowed us to assess the within-ruminant synteny by including both a Cervidae representative as well as the high-quality Bovidae representative (goat). We found that Primates and Perissodactyla have more prevalent genome rearrangements (>200 rearrangements per million base pairs) compared with that of Ruminantia (Fig. 3A and table S17). Within the Cetartiodactyla, pig (2n = 18, Suina) and camel (2n = 74, Tylopoda) experienced more genome rearrangements (~120 rearrangements per million base pairs) than the killer whale (2n = 44, Cetacea). Thus, the greater level of conserved synteny between Cetacea and Ruminantia (68 rearrangements per million base pairs) is consistent with their sister relationship within the Cetartiodactyla. Few rearrangements (19 rearrangements per million base pairs) were observed between the black muntjac (2n = 8/9) and the goat (2n = 58), suggesting that synteny has been conserved among different lineages of Ruminantia, despite large variation in chromosome numbers (table S18).

Genome size

Our ruminant genome assembly sizes ranged from 2.52 Gbp [oribi (*Ourebia ourebi*)] to 3.25 Gbp [klipspringer (*Oreotragus oreotragus*)] (table S5). The average assembled genome size of ruminants was (2.7 Gbp), which is larger than Carnivora (~2.3 Gbp), Perissodactyla (~2.4 Gbp), Suina (~2.5 Gbp), Tylopoda (~2.0 Gbp), and Cetacea (~2.4 Gbp) and smaller than Primates (~3.0 Gbp) (Fig. 3B). Analyses (23) confirm that transposable element (TE) content is the major cause of genome size variation (Fig. 3, B and C, and table S19).

When comparing the goat genome with the human, horse, pig, and killer whale genomes, we also observed and validated (23) large insertions and deletions (over 50 kbp in length) in ruminants (table S20). The largest insertion from segmental duplication in the goat contains a cluster of PAG (pregnancy-associated glycoprotein) genes, with 36 coding sequences and 32 pseudogenes (Fig. 3D). The main functions of PAGs are immune regulation and maintenance of pregnancy (67). We also found other insertions with important gene annotations in ruminants, including interferon and olfactory receptors (fig. S33 and table S21).

Evolution of genes and gene families in Ruminants

We obtained a high-confidence orthologous gene set for the full set of 51 ruminant species using camel, cat, dog, horse, human, minke whale, killer whale, and pig as outgroups (23). Using the resolved phylogeny of Ruminantia, we identified rapidly evolving genes (REGs), positively selected genes (PSGs), and newly evolved genes (Fig. 4A and tables S22 to S25) (23).

Functional enrichment analyses of the PSGs (tables S26 and S27), the REGs (table S28), and expanded gene families (table S29) in the ancestral ruminant branch all exhibit enrichment in immune functions. As many as 20 PSGs and 12 REGs are involved in the crossing of blood vessels by leukocytes, which respectively constitutes 17.9 and 10.7% of this key pathway in active immunization (Fig. 4B and tables S30 and S31) (68, 69). We also observed gene family expansions in the ruminant ancestor (table S32), including the interferon family (IFNs) (figs. S34 to S36), PAG family (fig. S34), and cathelicidin and serpin peptidase inhibitor families (figs. S34, S37, and S38), which are all involved in immune system pathways. In addition, we identified a rumen-specifically expressed newly evolved gene, *ENSBTAG00000038127* (23), which contains an immunoglobulin V-set domain, usually involved in mimicking the antibody variable domain of several diverse protein families (70). Although rapid evolution of immune genes was found in a wide range of animals, the number of PSGs from the leukocyte transendothelial pathway is highest in Ruminantia and could suggest a key role of this pathway in the evolution of ruminant immune system (table S33).

In addition to immune system–related genes, we also observed a series of PSGs, REGs, and expanded gene families in ruminants involved in lipid metabolism, glycolysis, oxidative phosphorylation, and amino acid metabolism (Fig. 4, C and D, and table S34) (23). Ruminants have a distinct digestive system, in which the main source of energy comes from volatile fatty acids (VFAs) and the rate of glycolysis is low (71). The genomic changes associated with metabolism may reflect the adaptation of the digestive system in ruminants and may therefore have played important parts in the success of the ruminants.

Genomic variations related to ruminant morphological characteristics

We were able to leverage our phylogenetic tree and genomic data to conduct evolutionary genomic analyses aimed at identifying genomic variations correlated with particular ruminant characteristics. Specifically, we investigated the evolution of the multichambered stomach, headgear, body size, cursorial locomotion, and dentition.

Evolution of the multichambered stomach

Whereas pecoran ruminants have a four-chambered stomach composed of the rumen, reticulum, omasum, and abomasum, the basal group of Ruminantia, the Tragulidae, lacks the omasum (*72*).

Using a transcriptomic dataset from 516 samples covering 50 tissues of sheep (table S35) (*23*), we found that the gene expression profiles of rumen, reticulum, and omasum are closest to that of esophagus (Fig. 5A), implying that the three stomach chambers might have originated from the esophagus, as suggested by Warner (*73*) and Xiang *et al.* (*74*). Regarding the distinct rumen organ, we found that several newly evolved genes played a role in its function. Among the 295 newly evolved genes identified in the ancestor of ruminants (Fig. 4A and fig. S39), seven were highly or specifically expressed in the rumen (fig. S40). Two genes, *PRD-SPRRII* and *TCHHL2*, are important structural genes in the rumen (*16*). Three newly evolved *serpin* genes may have inhibitory functions of different proteases or through anti-inflammatory functions (*75*, *76*), and two *KRT6A* genes are likely involved in the activation of follicular keratinocytes (*77*, *78*).

Although the abomasum is analogous to the true stomach in other mammals, it is hypothesized that the ruminant abomasum has evolved particular adaptations to digest the microbe-rich content from upstream chambers (*72*). The lysozyme *c* family, which degrades bacterial and microbial cells as members enter the abomasum to extract nutrients (*79*), has expanded to 10 or more copies in ruminants, whereas other mammals have only one or a few copies of this gene family (*79*). Our comparative genomics analyses revealed that the duplication of lysozyme *c* genes began in the ancestors of Ruminantia, continued to

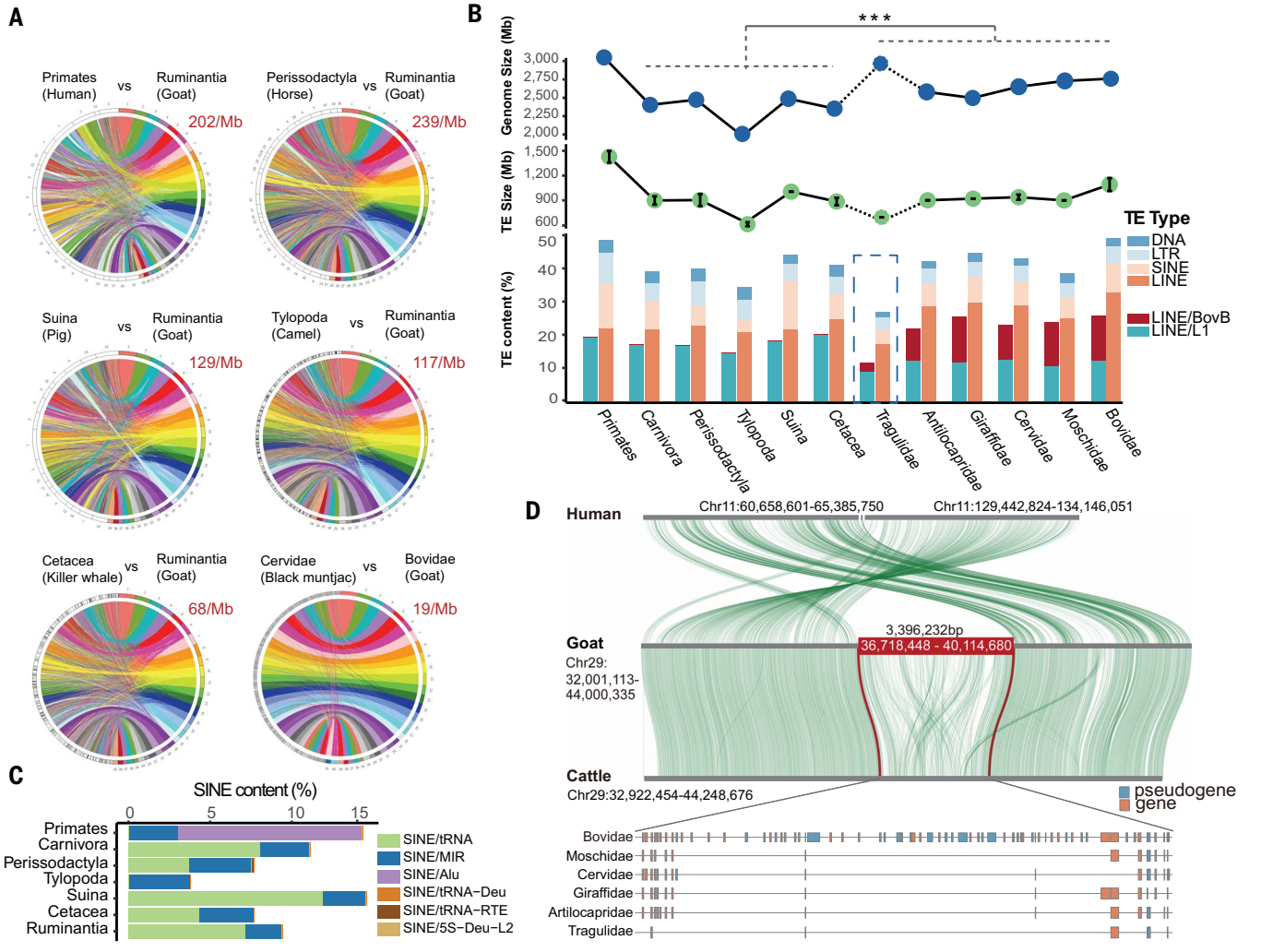


Fig. 3. Structural characteristics and evolution of ruminant genomes. (**A**) Goat (*19*) was used to represent Ruminantia-specific genomic rearrangements in comparisons with Primates (human), Perissodactyla (horse), Suina (pig), Tylopoda (camel), and Cetacea (Killer whale). Syntenic blocks are linked between genomes in a circos plot. The red number beside each circos quantifies the occurrences of rearrangement events per aligned megabase (Mb) sequence. The high-quality de novo assembly of the black muntjac was included for within-ruminant synteny inference. (**B**) The average genome sizes, TE sizes, and contents of different TE types of Primates (human, chimpanzee, gorilla, and orangutan), Carnivora (dog, cheetah, and polar bear), Perissodactyla (horse, przewalski's horse, and rhinoceros), Tylopoda (dromedary camel, bactrian camel, and alpaca), Suina (pig), and Cetacea (minke whale, killer whale, beluga whale, sperm whale, and yangtze finless porpoise). Overall, the average genome size of Ruminantia is significantly (Student's *t* test, $P < 0.01$) larger than those of Canivora, Perissodactyla, Tylopoda, Suina, and Cetacea. Tragulidae is marked with dashed lines because the genome assembly contains more gaps, which hindered the annotation of TEs. The proportions of LINE, SINE, LTR, and DNA transposons are presented in the stacked bar plot. LINE/BovB and LINE/L1 are highlighted here to present their dynamic changes among ruminants. ***$P < 0.01$. (**C**) The average contents of different SINE types are plotted in the stacked bar plot across mammalian orders and suborders of Cetartiodactyla, with different colors. (**D**) A large fragment insertion of 3,396,232 bp is observed in the goat genome, which is also validated in other ruminant families, containing a cluster of PAG genes, specifically containing 36 coding genes and 32 pseudogenes in goat.

expand along the diversification of pecoran families, and eventually culminated with the highest copy numbers in Bovidae (Fig. 4D). This accumulated evolution of lysozyme *c* gene copy number in ruminants may be associated with an improved harvesting of nutrients from the rumen microbiome.

As a newly evolved organ in the pecorans, the omasum is adjacent to the reticulum, whereas its expression profile is closest to that of the rumen. This expression overlap may be linked to the resemblance in structure and function of the rumen and omasum, which are both involved in the absorption of water, minerals, electrolytes, and VFAs, whereas the reticulum mainly has a different function as a filter for the fermentation products from the rumen (*6*). Anatomically, the omasum is composed of the same stratified squamous epithelium of mucosal-layered tissue as the rumen and reticulum, which is different from the abomasum and intestines (*6*). To further reveal the genetic basis underlying the evolution of the omasum in pecorans, we identified 75 genes that were specifically highly expressed in the omasum compared with other organs (fig. S41). Among these, one gene was newly evolved (*LOC101107119*), and another (*SCNN1D*) exhibited pecoran-specific amino acid changes relative to Tragulidae and killer whale genomes (fig. S42). *LOC101107119* is annotated as a prostaglandin F synthase 1–like gene and might have a similar function to that of the prostaglandin F synthase 1 (PGDF1), which is involved in ketone metabolism (*80*). *SCNN1D* encodes the δ subunit of the epithelial sodium channel, which mediates Na^+ reabsorption and water absorption in the digestive tract (*81*).

Transcriptional factors and regulatory elements may have played important roles in the evolution of the omasum by changing the expression patterns of their host genes to be recruited in omasum functions. We found four genes with pecoran-specific CNEs within their immediate upstream/downstream 10-kbp regions (*SIM2*, *PAX9*, *KCNK5*, and *DENND2C*) (table S36), of which PAX9 and SIM2 are important transcriptional factors. PAX9 regulates squamous cell differentiation in the esophageal epithelium (*82*). A CNE with two pecoran-specific mutations was found at the 5′ upstream of *PAX9* in pecorans, which might play a role in regulating *PAX9* expression in the omasum. *SIM2* is highly expressed in the omasum but inactive in the rumen [fragments per kilobase of exon per million fragments mapped (FPKM) < 1], and it works as a suppressor of cell proliferation in the epithelium (*83*, *84*). The differential expression patterns of *SIM2* in the omasum and rumen is consistent with the epithelial cells in the omasum not being completely renewed, as is the case in the rumen (*85*).

Evolution of headgear

The ruminant families exhibit spectacular variation in headgear morphology (*86*). To reveal the genetic basis of headgear origin in ruminants, its secondary loss in two independent lineages and the biologically exceptional, rapid regeneration ability of cervid antlers, we performed large-scale comparative transcriptomics and functional experiments in an accompanying paper (*87*). Substantial similarities in the transcriptome profiling of different headgear types (*87*) are consistent with a single origin of headgear in the pecoran ancestor. This ancestral headgear subsequently diversified into horns, antlers, ossicones, and pronghorns in the different pecoran families and was lost independently in the lineages of Moschidae and Hydropotinae perhaps because of the convergent pseudogenization of the horn-development *RXFP2* gene (*87*). These complex patterns, along with the lack of a well-resolved phylogeny, have confounded a synthesis of headgear evolution.

We further examined the 201 highly expressed genes in the headgear of both bovids and cervids (*87*) and identified 36 of these genes harboring pecoran-specific CNEs (table S36). Nine genes (*ALX1*, *VCAN*, *COL1A1*, *SATB2*, *RUNX2*, *POSTN*, *SP7*, *TNC*, and *COL4A2*) were classified in Gene Ontology (GO) categories related to bone development. *RUNX2* is a key transcriptional factor in the regulation of bone development (*88*) and has four pecoran-specific CNEs in its intronic regions (fig. S43A and table S36), potentially causing high *RUNX2* expression in the headgear sprouts. *SP7* encodes an important regulatory factor of the biomineralization and formation of bones (*89*) and has specific mutations in the pecoran 3′ untranslated region, one of which was located in the binding region of the microRNA bta-mir-145 (fig. S43B). These genes were possibly rewired for the formation of bony headgears of ruminants, laying out a hypothesis that can be tested experimentally in the future.

In addition, we explored the genetic basis of the keratinous sheath found convergently in Bovidae and Antilocapridae headgear (Fig. 5B). Transcriptomic data from horn sprouts of sheep and goat (*87*), both of which belong to Bovidae, identified seven highly expressed keratin genes: *KRT1*, *KRT2*, *KRT3*, *KRT5*, *KRT10*, *KRT14*, and *KRT84* (fig. S44A). Except for *KRT10* and *KRT14*, the above keratin genes encode Type II α-keratin proteins, suggesting an essential role for Type II α-keratin genes in the formation of the keratinous sheath of bovids. We further examined convergent amino acid substitutions between the pronghorn (*Antilocapra americana*) and bovids. Using the 12 mammalian species and other ruminant species as outgroups, we identified 106 proteins (table S37) that contain at least two convergent amino acid changes in these two ruminant families (*23*). Among these proteins, a horn-specific Type II α-keratin protein, KRT82, contained two convergently changed amino acid sites, $Ala^{82}Thr$ and $Ser^{509}Ala$, of which the $Ala^{82}Thr$ is located in the keratin head domain (Fig. 5B and fig. S44B). This suggests that Type II α-keratin may be important in forming the keratinous sheath that evolved convergently in Bovidae and Antilocapridae.

Evolution of body size in ruminants

Ruminants, especially bovids, encompass a wide body size range that spans four orders of magnitude, from as little as 2 kg to as high as 1200 kg (Fig. 5C) (*10*, *11*). We retrieved 642 genes related to the development of body size and estimated the ratios between the nonsynonymous substitution rate (d*N*) and synonymous substitution rate (d*S*) of these genes on branches that exhibit substantially increased body sizes and on branches that exhibit decreased body sizes (fig. S45). Compared with the background, these genes showed significantly higher d*N*/d*S* ratios in branches with large-sized and small-sized species (Student's *t* test, $P < 0.01$) but similar ratios in branches with medium-sized species (Fig. 5D). We observed six genes (*CXCL13*, *RNF115*, *NPNT*, *KL*, *SLC9A3R1*, and *MSTN*) that had significantly elevated d*N*/d*S* ratios along with the increasing of body size (Student's *t* test, $P < 0.01$) (table S38), and *SLC9A3R1* and *MSTN* have d*N*/d*S* values > 1, an indication of positive selection. *SLC9A3R1* affects osteogenesis by mineralizing osteoblasts, and disrupting this gene resulted in reduced body weights in mice (*90*). *MSTN* is an important gene in the regulation of muscle cell growth and differentiation (*91*), and mutations in *MSTN* affect the muscle mass of goat (*92*), sheep (*93*), and cattle (*94*, *95*) as well as other mammals (*96*, *97*). These results indicate that *SLC9A3R1* and *MSTN* might be targets of natural selection favoring increased body sizes (fig. S45) by regulating the development of bone and muscle. In reduced body size branches, five genes (*SBDS*, *BMP3*, *LRRN3*, *NFATC3*, and *SMARCAL1*) had significantly elevated d*N*/d*S* ratios (table S38), and the d*N*/d*S* ratio of *SBDS* was larger than 1. *SBDS* is an important gene in cell proliferation and is the causal gene of Shwachman-Diamond syndrome in humans, which is characterized by skeletal abnormalities and short stature (*98*). *BMP3* is a well-known gene involved in the regulation of osteogenesis in mammals (*99*). The genes mentioned above may therefore explain the body size variation in ruminants and may be relevant to livestock breeding application.

As the tallest terrestrial animal, giraffes have a distinct stature and body morphology, which likely are adaptions to their savanna habitat (*100*). Among the 366 genes related to bone development in the KEGG annotation, 115 genes had giraffe-specific mutations (table S39), including genes in the transforming growth factor–β (TGF- β), Hedgehog, Notch, Wnt, and FGF signaling pathways. Eleven genes with more than four nonsynonymous mutations may be related to the extreme elongation of body structures in giraffe because they are part of bone development pathways: TGF-β (*CHRD* and *LTBP1*), Wnt (*APC*, *APC2*, and *CREBBP*), Notch (*NOTCH1*, *NOTCH3*, and *NOTCH4*), Hedgehog (*GLI2* and *GLI3*), and FGF (*FGFRL1*) (Fig. 5C). Three such genes (*FGFRL1*, *NOTCH4*, and *CREBBP*) had been identified in a previous comparative study by using a low-coverage genome assembly of giraffe (*Giraffa camelopardalis*) (table S39) (*100*).

Adaptations to cursorial locomotion

Although early ruminants were probably forest-dwelling (*101*), many lineages have adapted to open habitats by adopting more cursorial body

plans designed for rapid and/or sustained movement in an open habitat (*102*). These include adaptations in limb morphology (*103*) and possibly also physiology. The pronghorn (*Antilocapra americana*) and members of Antilopinae such as the springbok (*Antidorcas marsupialis*) are among the fastest-running animals on Earth (*104*). Both lineages exhibit strong selection in their mitochondrial electron transport chains, although the targeted genes of selection differ between these species. In Antilopinae, two genes, *SUOX* (sulfite oxidase) and *NLN* (neurolysin), were under positive selection (χ^2 test, $P < 0.05$) (table S24). In the pronghorn, two mitochondrial electron transport chain enzyme encoding genes [*COX5A* (cytochrome c oxidase subunit 5A) and *PPOX* (protoporphyrinogen oxidase)] were subjected to positive selection (χ^2 test, P value < 0.05) (table S24). In addition, four proteins in the mitochondrial electron transport chain—NDUFA10 (reduced nicotinamide adenine dinucleotide dehydrogenase 1 α subcomplex subunit 10), SDHB (succinate dehydrogenase iron-sulfur subunit), UQCRC2 (cytochrome b-c1 complex subunit 2), and ATP5B (adenosine 5′-triphosphate synthase subunit β), functioning in oxidative phosphorylation—exhibited convergent amino acid changes (Fig. 5E and tables S40 and S41). Furthermore, all species of the bovid tribe Alcelaphini, which are adapted to open and seasonal

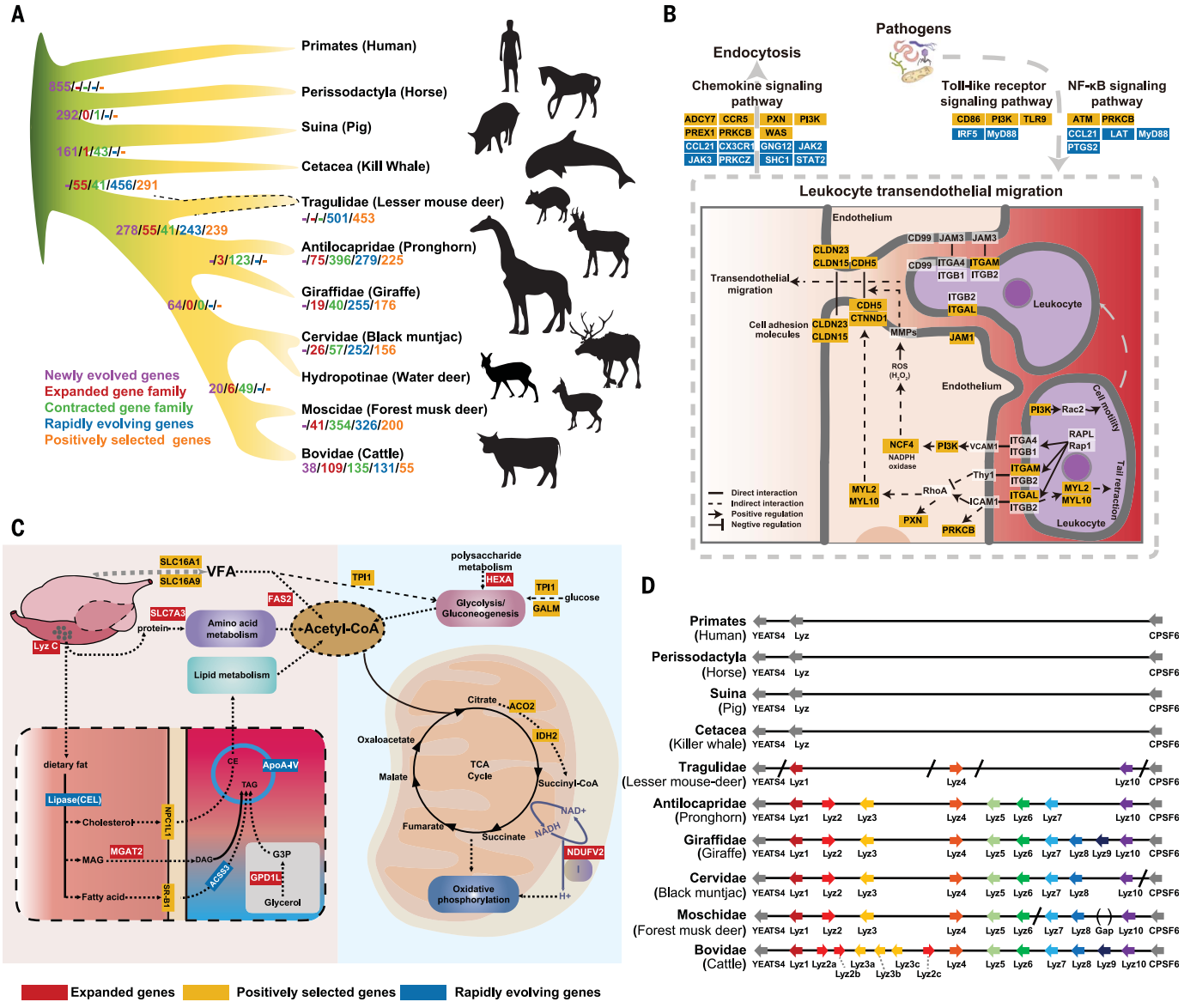


Fig. 4. Newly evolved genes, expanded and contracted gene families, REGs, and PSGs in ruminants. (**A**) Positively selected genes (PSGs), rapidly evolving genes (REGs), expanded genes, and newly evolved genes are shown along the phylogenetic tree. PSGs and REGs are identified with PAML. Newly evolved genes are identified based on the synteny relationships, with goat used as a reference. Expanded gene families are identified for the pecoran families. Tragulidae are excluded because of the relatively low quality of the assembled tragulid genome. (**B**) Diagram of the process of leukocytes crossing blood vessels, highlighting PSGs (orange) and REGs (blue). The orange rectangles indicate PSGs, and the blue rectangles represent the REGs. The solid lines represent direct interaction, and the dotted lines represent indirect interactions. (**C**) Diagram of nutrient metabolism evolution in Ruminantia displaying genes involved in nutrition metabolism. Expanded gene families are marked red, PSGs are marked yellow, and REGs are marked blue in (C) and (D). (**D**) Expansion of the lysozyme *c* gene family throughout the Ruminantia. The different expanded copies are presented with colors. Each line corresponds to the order or family in (A), and we chose the best assembled species genome sequences to draw the region. The slashes on each line indicate the ends of scaffolds in the assembled genomes.

grasslands and have cursorial locomotion, shared three specific mutations (Tyr[211]Phe, Gln[481]Arg, and Glu[622]Ala) in the *ACE* (angiotensin I converting enzyme) gene (fig. S46) and one specific mutation (Glu[97]Asp) in the *EPO* (erythropoietin) gene (fig. S47), both of which are important for endurance (*105*, *106*). These observations imply that shared or distinct molecular pathways might have played roles in convergent phenotypic evolution during the radiation of ruminants.

Evolution of dentition

Ruminants have specialized dentition patterns, lacking upper incisors and having a high prevalence of high-crowned or hypsodont teeth—a likely adaptation to abrasive diets such as grass or

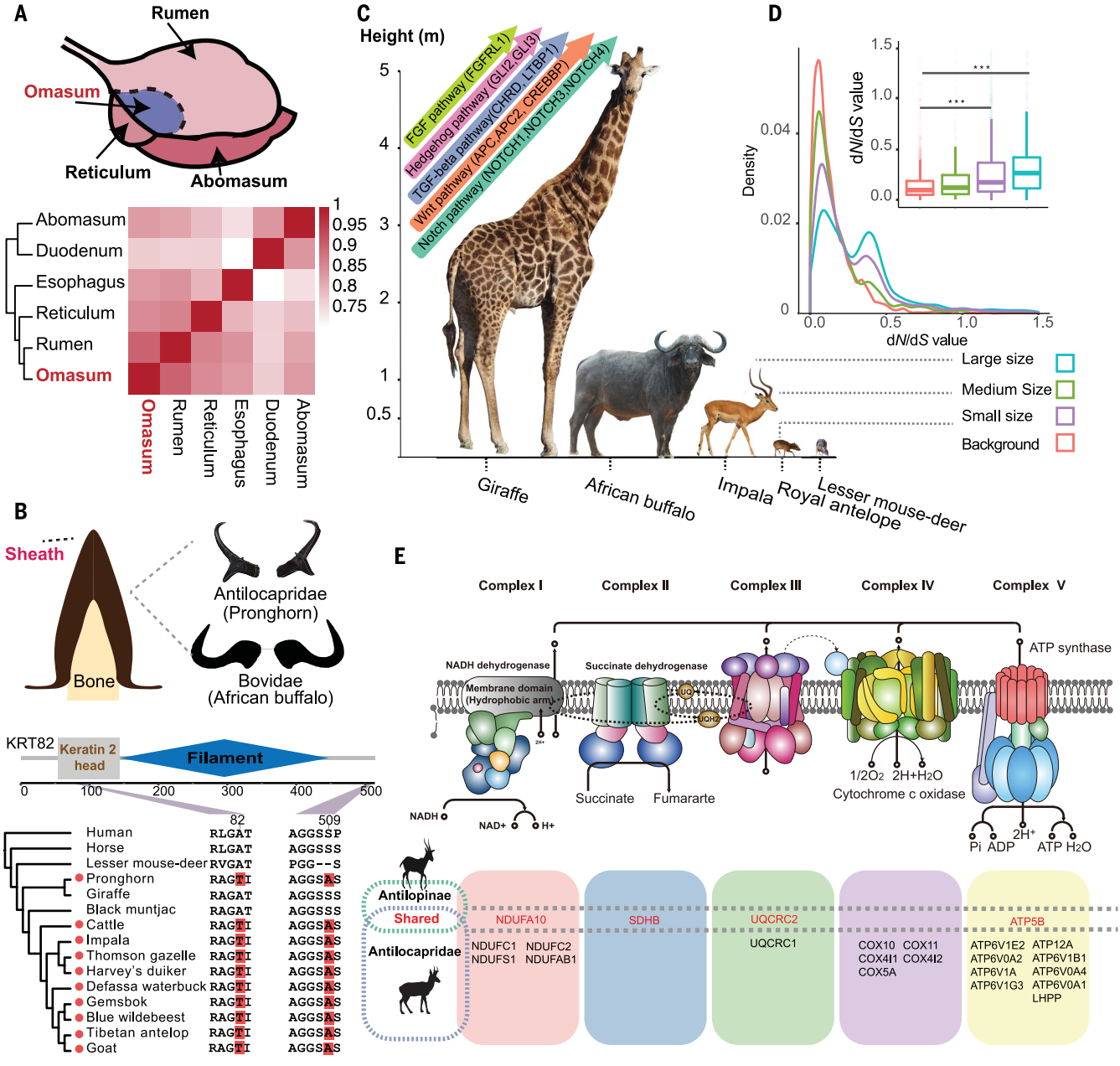


Fig. 5. Genomic features related to ruminant characteristics. (**A**) Pairwise Spearman correlations of gene expressions indicated that the omasum, rumen, and reticulum evolved as an extension of the esophagus, whereas the abomasum evolved as an extension of the duodenum. Although it is anatomically closer to the reticulum, the omasum has a more similar gene expression atlas with that of the rumen than the reticulum, which mirrors the similar functions between them. (**B**) Diagram of the convergent feature of keratinous sheath in the Bovidae and Antilocapridae. The two red amino acids indicate convergent mutations of the KRT82 between Antilocapridae and Bovidae, and the red dots indicate the included species of Antilocapridae and Bovidae. (**C**) Body size contrast among ruminants and 11 genes related with bone development that have at least four specific mutations in giraffe and are involved in TGF-β, Hedgehog, Notch, Wnt, and FGF pathways. (**D**) A total of 642 genes in GO related to the development of body size are retrieved, and we calculated the d*N*/d*S* ratios on the branches, leading to large, medium, and small body sizes, under the two-ratio branch model (model 2) of PAML. Background d*N*/d*S* ratios are calculated under one-ratio branch model (model 1) of PAML. The distribution densities of the d*N*/d*S* values are shown. A box plot of d*N*/d*S* values in different body size categories is shown. The mean d*N*/d*S* values at the branches of large body size and small body size are significantly larger than background. ***$P < 0.01$. (**E**) Antilocapridae and Antilopinae species are among the most mobile land mammals, and both of them have specific mutations in several genes of the mitochondrial electron transport chain.

feeding under abrasive conditions, as encountered in dry habitats (*107*). The distinct dentition is believed to have contributed to the evolutionary success of the ruminants by expanding their food sources (*108*). We observed that 11 genes related to dentition had RSCNEs in the 10-kbp downstream or upstream or in introns (table S42). Among these genes, *ENAM* functions in the mineralization and structural organization of enamel (*109*) and is possibly associated with enamel thickness in humans (*110*). Furthermore, *ENAM* is an important factor in the adaptation of dentition to specific diets (*111*). We observed seven RSCNEs in this gene region (fig. S48A). In addition to these RSCNEs, *ENAM* had a two–amino acid insertion and several mutations that were specific to ruminants (fig. S48, B and C). A genome-wide association analysis using the hypsodonty index of ruminants [data from (*112*)] identified one SNP significantly associated (Fisher's exact test, $P < 0.01$) with the hypsodonty index located in the binding site of the *EBF1* transcription factor in the intron of *FGF14* (fig. S49), an important factor controlling tooth development (*113*). These genes provide candidates for future experimental studies on mammal dentition.

Conclusion

A well-resolved phylogeny of Ruminantia with full genome data provides an opportunity to understand the evolutionary processes and genetic basis of the distinct structure and evolutionary patterns of ruminants. Our comprehensive evolutionary and comparative analyses have revealed numerous genetic variations correlated with specific traits in ruminants. This study provides valuable genomic resources as well as insights into not only the evolution and diversification of ruminants but also our understanding of mammalian biology.

Materials and methods

We sequenced and assembled 44 ruminant genomes using Illumina reads, with the black muntjac (*Muntiacus crinifrons*) further sequenced with PacBio SMART long reads. The assemblies were performed with SOAPdenovo v1.05 (*114*), Platanus v1.2.4 (*115*), and Supernova assembler (*116*), and some contigs were further scaffolded with SSPACE v3.0 (*117*) and "cross_genome" commands in the Phusion2 package (*118*). Repeat elements were annotated by combining results of Tandem Repeat Finder v4.07b (*119*), RepeatMasker v4.0.5 (*120*), RepeatModeller v1.0.4 (*120*), and LTR_FINDER v1.0.6 (*121*). Genes were annotated with homologous TBLASTN (*122*) protein searches of human (Ensembl 87 release), cattle (Ensembl 87 release), and sheep (Ensembl 87 release), combining de novo prediction of SNAP (*123*), GENSCAN v1.0 (*124*), glimmerHMM v3.0.3 (*125*), and AUGUSTUS v2.5.5 (*126*). Whole-genome alignments were constructed with LAST v885 (*127*) and MULTIZ v11.2 (*128*), using goat as the reference. The maximum likelihood phylogenic trees were constructed with RAxML v8.2.9 (*129*) and ExaML v3.0.17 (*130*) by using a general time-reversible model ("GTR+GAMMA"). ASTRAL III v5.1.1 (*36*) and MP-EST v2.0 (*38*) were used to perform coalescent species tree estimations. DiscoVista v1.0 (*131*) was used to quantify and visualize gene tree discordance for alternative topologies. Gene flows among ruminant families were performed with ABBA-BABA test (*39*), D_{FOIL} (*40*), admixturegraph v1.0.2 (*41*), PhyloNet v3.6.9 (*42*), and PhyloNetworks v0.9.0 (*43*). Time calibration was conducted with r8s v1.70 (*132*), BEAST v1.8.4 (*133*), MCMCTREE in PAML v4.8 (*134*), and Multidivtime (*135*). Demographic history reconstruction was carried out by means of PSMC analysis (*56*). PhyloFit v1.4 (*136*) and phastCons v1.4 (*137*) were used to infer conserved nonexonic elements in ruminants. PSGs and REGs were identified by use of branch and branch-site models in PAML v4.8 (*134*). The gene family expansion or contraction analysis was performed by using CAFÉ v4.7 (*138*). In-house scripts and pipelines are deposited in Zenodo (https://zenodo.org/record/2549147). Detailed methods and materials are described in the supplementary materials (*23*).

REFERENCES AND NOTES

1. IUCN, *The IUCN Red List of Threatened Species*, Version 2017-3. (2017); www.iucn.org.
2. D. E. Wilson, D. M. Reeder, *Mammal Species of the World: A Taxonomic and Geographic Reference* (JHU Press, 2005).
3. R. Heller, P. Frandsen, E. D. Lorenzen, H. R. Siegismund, Are there really twice as many bovid species as we thought? *Syst. Biol.* **62**, 490–493 (2013). doi: 10.1093/sysbio/syt004; pmid: 23362112
4. T. F. Randolph *et al.*, Invited review: Role of livestock in human nutrition and health for poverty reduction in developing countries. *J. Anim. Sci.* **85**, 2788–2800 (2007). doi: 10.2527/jas.2007-0467; pmid: 17911229
5. P. H. Hemsworth, G. J. Coleman, *Human-Livestock Interactions: The Stockperson and the Productivity of Intensively Farmed Animals*. (CABI, 2010).
6. C. Janis, The evolutionary strategy of the Equidae and the origins of Rumen and Cecal digestion. *Evolution* **30**, 757–774 (1976). doi: 10.1111/j.1558-5646.1976.tb00957.x; pmid: 28563331
7. J. B. Russell, J. L. Rychlik, Factors that alter rumen microbial ecology. *Science* **292**, 1119–1122 (2001). doi: 10.1126/science.1058830; pmid: 11352069
8. M. Clauss, G. E. Rössner, Old world ruminant morphophysiology, life history, and fossil record: Exploring key innovations of a diversification sequence. *Ann. Zool. Fenn.* **51**, 80–94 (2014). doi: 10.5735/086.051.0210
9. T. J. Hackmann, J. N. Spain, Invited review: ruminant ecology and evolution: perspectives useful to ruminant livestock research and production. *J. Dairy Sci.* **93**, 1320–1334 (2010). doi: 10.3168/jds.2009-2071; pmid: 20338409
10. J. R. Castelló, *Bovids of the World: Antelopes, Gazelles, Cattle, Goats, Sheep, and Relatives*. (Princeton Univ. Press, 2016).
11. J. T. du Toit, N. Owen-Smith, Body size, population metabolism, and habitat specialization among large African herbivores. *Am. Nat.* **133**, 736–740 (1989). doi: 10.1086/284949
12. M. Hernández Fernández, E. S. Vrba, A complete estimate of the phylogenetic relationships in Ruminantia: A dated species-level supertree of the extant ruminants. *Biol. Rev. Camb. Philos. Soc.* **80**, 269–302 (2005). doi: 10.1017/S1464793104006670; pmid: 15921052
13. J. E. Decker *et al.*, Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18644–18649 (2009). doi: 10.1073/pnas.0904691106; pmid: 19846765
14. A. Hassanin *et al.*, Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. *C. R. Biol.* **335**, 32–50 (2012). doi: 10.1016/j.crvi.2011.11.002; pmid: 22226162
15. F. Bibi, A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla, Ruminantia) and the importance of the fossil record to systematics. *BMC Evol. Biol.* **13**, 166 (2013). doi: 10.1186/1471-2148-13-166; pmid: 23927069
16. Y. Jiang *et al.*, The sheep genome illuminates biology of the rumen and lipid metabolism. *Science* **344**, 1168–1173 (2014). doi: 10.1126/science.1252806; pmid: 24904168
17. A. V. Zimin *et al.*, A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biol.* **10**, R42 (2009). doi: 10.1186/gb-2009-10-4-r42; pmid: 19393038
18. Q. Qiu *et al.*, The yak genome and adaptation to life at high altitude. *Nat. Genet.* **44**, 946–949 (2012). doi: 10.1038/ng.2343; pmid: 22751099
19. D. M. Bickhart *et al.*, Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. *Nat. Genet.* **49**, 643–650 (2017). doi: 10.1038/ng.3802; pmid: 28263316
20. R. L. Ge *et al.*, Draft genome sequence of the Tibetan antelope. *Nat. Commun.* **4**, 1858 (2013). doi: 10.1038/ncomms2860; pmid: 23673643
21. C. M. Seabury *et al.*, Genome-wide polymorphism and comparative analyses in the white-tailed deer (Odocoileus virginianus): A model for conservation genomics. *PLOS ONE* **6**, e15811 (2011). doi: 10.1371/journal.pone.0015811; pmid: 21283515
22. R. Kropatsch *et al.*, SOX9 duplication linked to intersex in deer. *PLOS ONE* **8**, e73734 (2013). doi: 10.1371/journal.pone.0073734; pmid: 24040047
23. Materials and methods are available as supplementary materials.
24. Z. Li *et al.*, Draft genome of the reindeer (*Rangifer tarandus*). *Gigascience* **6**, 1–5 (2017). doi: 10.1093/gigascience/gix102; pmid: 29099922
25. C. Zhang *et al.*, Draft genome of the milu (*Elaphurus davidianus*). *Gigascience* **7**, gix130 (2018). doi: 10.1093/gigascience/gix130; pmid: 29267854
26. Y. Yang *et al.*, Draft genome of the Marco Polo Sheep (*Ovis ammon polii*). *Gigascience* **6**, 1–7 (2017). doi: 10.1093/gigascience/gix106; pmid: 29112761
27. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015). doi: 10.1093/bioinformatics/btv351; pmid: 26059717
28. T. R. Gregory, *Animal Genome Size Database* (2017); www.genomesize.com.
29. A. D. Foote *et al.*, Convergent evolution of the genomes of marine mammals. *Nat. Genet.* **47**, 272–275 (2015). doi: 10.1038/ng.3198; pmid: 25621460
30. D. Brawand *et al.*, The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**, 375–381 (2014). doi: 10.1038/nature13726; pmid: 25186727
31. E. D. Jarvis *et al.*, Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**, 1320–1331 (2014). doi: 10.1126/science.1253451; pmid: 25504713
32. A. Hobolth, J. Y. Dutheil, J. Hawks, M. H. Schierup, T. Mailund, Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. *Genome Res.* **21**, 349–356 (2011). doi: 10.1101/gr.114751.110; pmid: 21270173
33. K. Prüfer *et al.*, The bonobo genome compared with the chimpanzee and human genomes. *Nature* **486**, 527–531 (2012). doi: 10.1038/nature11128; pmid: 22722832
34. S. Mirarab, M. S. Bayzid, B. Boussau, T. Warnow, Statistical binning enables an accurate coalescent-based estimation of the avian tree. *Science* **346**, 1250463 (2014). doi: 10.1126/science.1250463; pmid: 25504728
35. S. Patel, R. Kimball, E. Braun, Error in phylogenetic estimation for bushes in the tree of life. *J. Phylogenetics Evol. Biol.* **1**, 110 (2013). doi: 10.4172/2329-9002.1000110
36. C. Zhang, M. Rabiee, E. Sayyari, S. Mirarab, ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153 (2018). doi: 10.1186/s12859-018-2129-y; pmid: 29745866
37. M. S. Springer, J. Gatesy, The gene tree delusion. *Mol. Phylogenet. Evol.* **94** (Pt A), 1–33 (2016). doi: 10.1016/j.ympev.2015.07.018; pmid: 26238460
38. L. Liu, L. Yu, S. V. Edwards, A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. *BMC Evol. Biol.* **10**, 302 (2010). doi: 10.1186/1471-2148-10-302; pmid: 20937096
39. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011). doi: 10.1093/molbev/msr048; pmid: 21325092

40. J. B. Pease, M. W. Hahn, Detection and polarization of introgression in a five-taxon phylogeny. *Syst. Biol.* **64**, 651–662 (2015). doi: 10.1093/sysbio/syv023; pmid: 25888025
41. K. Leppälä, S. V. Nielsen, T. Mailund, admixturegraph: An R package for admixture graph manipulation and fitting. *Bioinformatics* **33**, 1738–1740 (2017). doi: 10.1093/bioinformatics/btx048; pmid: 28158333
42. D. Wen, Y. Yu, J. Zhu, L. Nakhleh, Inferring phylogenetic networks using PhyloNet. *Syst. Biol.* **67**, 735–740 (2018). doi: 10.1093/sysbio/syy015; pmid: 29514307
43. C. Solís-Lemus, P. Bastide, C. Ané, PhyloNetworks: A package for phylogenetic networks. *Mol. Biol. Evol.* **34**, 3292–3298 (2017). doi: 10.1093/molbev/msx235; pmid: 28961984
44. O. R. Bininda-Emonds *et al*., The delayed rise of present-day mammals. *Nature* **446**, 507–512 (2007). doi: 10.1038/nature05634; pmid: 17392779
45. R. W. Meredith *et al*., Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* **334**, 521–524 (2011). doi: 10.1126/science.1211028; pmid: 21940861
46. E. Sayyari, S. Mirarab, Fast coalescent-based computation of local branch support from quartet frequencies. *Mol. Biol. Evol.* **33**, 1654–1668 (2016). doi: 10.1093/molbev/msw079; pmid: 27189547
47. S. D. Webb, B. E. Taylor, *The Phylogeny of Hornless Ruminants and a Description of the Cranium of Archaeomeryx* (AMNH, 1980).
48. C. M. Janis, J. M. Theodor, Cranial and postcranial morphological data in ruminant phylogenetics. *Zitteliana* **32**, 15–31 (2014).
49. C. M. Janis, Grades and clades in hornless ruminant evolution: The reality of the Gelocidae and the systematic position of Lophiomeryx and Bachitherium. *J. Vertebr. Paleontol.* **7**, 200–216 (1987). doi: 10.1080/02724634.1987.10011653
50. I. M. Sánchez, M. S. Domingo, J. Morales, The genus Hispanomeryx (Mammalia, Ruminantia, Moschidae) and its bearing on musk deer phylogeny and systematics. *Palaeontology* **53**, 1023–1047 (2010). doi: 10.1111/j.1475-4983.2010.00992.x
51. A. Hassanin, E. J. Douzery, Molecular and morphological phylogenies of ruminantia and the alternative position of the moschidae. *Syst. Biol.* **52**, 206–228 (2003). doi: 10.1080/10635150390192726; pmid: 12746147
52. M. dos Reis *et al*., Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. Biol. Sci.* **279**, 3491–3500 (2012). doi: 10.1098/rspb.2012.0683; pmid: 22628470
53. F. Bibi, Assembling the ruminant tree: Combining morphology, molecules, extant taxa, and fossils. *Zitteliana* **32**, 197–211 (2014).
54. C. A. Matthee, S. K. Davis, Molecular insights into the evolution of the family Bovidae: A nuclear DNA perspective. *Mol. Biol. Evol.* **18**, 1220–1230 (2001). doi: 10.1093/oxfordjournals.molbev.a003908; pmid: 11420362
55. C. Z. Yang *et al*., Phylogenetic analyses and improved resolution of the family Bovidae based on complete mitochondrial genomes. *Biochem. Syst. Ecol.* **48**, 136–143 (2013). doi: 10.1016/j.bse.2012.12.005
56. C. A. Matthee, T. J. Robinson, Cytochrome b phylogeny of the family bovidae: Resolution within the alcelaphini, antilopini, neotragini, and tragelaphini. *Mol. Phylogenet. Evol.* **12**, 31–46 (1999). doi: 10.1006/mpev.1998.0573; pmid: 10222159
57. A. García-Ruiz *et al*., Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3995–E4004 (2016). doi: 10.1073/pnas.1519061113; pmid: 27354521
58. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011). doi: 10.1038/nature10231; pmid: 21753753
59. E. D. Lorenzen *et al*., Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* **479**, 359–364 (2011). doi: 10.1038/nature10574; pmid: 22048313
60. A. D. Barnosky, P. L. Koch, R. S. Feranec, S. L. Wing, A. B. Shabel, Assessing the causes of late Pleistocene extinctions on the continents. *Science* **306**, 70–75 (2004). doi: 10.1126/science.1101476; pmid: 15459379
61. C. J. Bae, K. Douka, M. D. Petraglia, On the origin of modern humans: Asian perspectives. *Science* **358**, eaai9067 (2017). doi: 10.1126/science.aai9067; pmid: 29217544
62. J. C. Venter *et al*., The sequence of the human genome. *Science* **291**, 1304–1351 (2001). doi: 10.1126/science.1058040; pmid: 11181995
63. E. F. Kirkness *et al*., The dog genome: Survey sequencing and comparative analysis. *Science* **301**, 1898–1903 (2003). doi: 10.1126/science.1086432; pmid: 14512627
64. C. M. Wade *et al*., Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* **326**, 865–867 (2009). doi: 10.1126/science.1178158; pmid: 19892987
65. M. A. M. Groenen *et al*., Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491**, 393–398 (2012). doi: 10.1038/nature11622; pmid: 23151582
66. H. Wu *et al*., Camelid genomes reveal evolution and adaptation to desert environments. *Nat. Commun.* **5**, 5188 (2014). doi: 10.1038/ncomms6188; pmid: 25333821
67. B. P. Telugu, A. M. Walker, J. A. Green, Characterization of the bovine pregnancy-associated glycoprotein gene family—Analysis of gene sequences, regulatory regions within the promoter and expression of selected genes. *BMC Genomics* **10**, 185 (2009). doi: 10.1186/1471-2164-10-185; pmid: 19393060
68. W. A. Muller, Mechanisms of leukocyte transendothelial migration. *Annu. Rev. Pathol.* **6**, 323–344 (2011). doi: 10.1146/annurev-pathol-011110-130224; pmid: 21073340
69. T. H. Mogensen, Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin. Microbiol. Rev.* **22**, 240–273 (2009). doi: 10.1128/CMR.00046-08; pmid: 19366914
70. R. D. Finn *et al*., Pfam: The protein families database. *Nucleic Acids Res.* **42** (D1), D222–D230 (2014). doi: 10.1093/nar/gkt1223; pmid: 24288371
71. R. A. Nafikov, D. C. Beitz, Carbohydrate and lipid metabolism in farm animals. *J. Nutr.* **137**, 702–705 (2007). doi: 10.1093/jn/137.3.702; pmid: 17311965
72. C. M. B. Membrive, *Rumenology* (Springer, 2016).
73. E. D. Warner, The organogenesis and early histogenesis of the bovine stomach. *Am. J. Anat.* **102**, 33–63 (1958). doi: 10.1002/aja.1001020103; pmid: 13545182
74. R. Xiang, V. H. Oddy, A. L. Archibald, P. E. Vercoe, B. P. Dalrymple, Epithelial, metabolic and innate immunity transcriptomic signatures differentiating the rumen from other sheep and mammalian gastrointestinal tract tissues. *PeerJ* **4**, e1762 (2016). doi: 10.7717/peerj.1762; pmid: 26989612
75. F. Lunardi *et al*., Overexpression of SERPIN B3 promotes epithelial proliferation and lung fibrosis in mice. *Lab. Invest.* **91**, 945–954 (2011). doi: 10.1038/labinvest.2011.1; pmid: 21403642
76. M. Gatto *et al*., Serpins, immunity and autoimmunity: Old molecules, new functions. *Clin. Rev. Allergy Immunol.* **45**, 267–280 (2013). doi: 10.1007/s12016-013-8353-3; pmid: 23325331
77. P. Martin, Wound healing—Aiming for perfect skin regeneration. *Science* **276**, 75–81 (1997). doi: 10.1126/science.276.5309.75; pmid: 9082989
78. P. Wong, P. A. Coulombe, Loss of keratin 6 (K6) proteins reveals a function for intermediate filaments during wound repair. *J. Cell Biol.* **163**, 327–337 (2003). doi: 10.1083/jcb.200305032; pmid: 14568992
79. D. M. Irwin, Genomic organization and evolution of ruminant lysozyme c genes. *Dongwuxue Yanjiu* **36**, 1–17 (2015). pmid: 25730456
80. N. Hevir, K. Vouk, J. Šinkovec, M. Ribič-Pucelj, T. L. Rižner, Aldo-keto reductases AKR1C1, AKR1C2 and AKR1C3 may enhance progesterone metabolism in ovarian endometriosis. *Chem. Biol. Interact.* **191**, 217–226 (2011). doi: 10.1016/j.cbi.2011.01.003; pmid: 21232532
81. I. Hanukoglu, A. Hanukoglu, Epithelial sodium channel (ENaC) family: Phylogeny, structure-function, tissue distribution, and associated inherited diseases. *Gene* **579**, 95–132 (2016). doi: 10.1016/j.gene.2015.12.061; pmid: 26772908
82. Z. Xiong *et al*., PAX9 regulates squamous cell differentiation and carcinogenesis in the oro-oesophageal epithelium. *J. Pathol.* **244**, 164–175 (2018). doi: 10.1002/path.4998; pmid: 29055049
83. X. Meng, J. Shi, B. Peng, X. Zou, C. Zhang, Effect of mouse Sim2 gene on the cell cycle of PC12 cells. *Cell Biol. Int.* **30**, 349–353 (2006). doi: 10.1016/j.cellbi.2005.11.012; pmid: 16530433
84. B. Laffin *et al*., Loss of singleminded-2s in the mouse mammary gland induces an epithelial-mesenchymal transition associated with up-regulation of slug and matrix metalloprotease 2. *Mol. Cell. Biol.* **28**, 1936–1946 (2008). doi: 10.1128/MCB.01701-07; pmid: 18160708
85. G. B. Penner, M. A. Steele, J. R. Aschenbach, B. W. McBride, Ruminant Nutrition Symposium: Molecular adaptation of ruminal epithelia to highly fermentable diets. *J. Anim. Sci.* **89**, 1108–1119 (2011). doi: 10.2527/jas.2010-3378; pmid: 20971890
86. E. B. Davis, K. A. Brakora, A. H. Lee, Evolution of ruminant headgear: A review. *Proc. Biol. Sci.* **278**, 2857–2865 (2011). doi: 10.1098/rspb.2011.0938; pmid: 21733893
87. Y. Wang *et al*., Genetic basis of ruminant headgear and rapid antler regeneration. *Science* **364**, eaav6335 (2019).
88. T. Komori, Regulation of bone development and maintenance by Runx2. *Front. Biosci.* **13**, 898–903 (2008). doi: 10.2741/2730; pmid: 17981598
89. C. Zhang, Transcriptional regulation of bone formation by the osteoblast-specific transcription factor Osx. *J. Orthop. Surg. Res.* **5**, 37 (2010). doi: 10.1186/1749-799X-5-37; pmid: 20550694
90. S. Shenolikar, J. W. Voltz, C. M. Minkoff, J. B. Wade, E. J. Weinman, Targeted disruption of the mouse NHERF-1 gene promotes internalization of proximal tubule sodium-phosphate cotransporter type IIa and renal phosphate wasting. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11470–11475 (2002). doi: 10.1073/pnas.162232699; pmid: 12169661
91. R. Sartori *et al*., BMP signaling controls muscle mass. *Nat. Genet.* **45**, 1309–1318 (2013). doi: 10.1038/ng.2772; pmid: 24076600
92. X. Wang *et al*., CRISPR/Cas9-mediated MSTN disruption and heritable mutagenesis in goats causes increased body mass. *Anim. Genet.* **49**, 43–51 (2018). doi: 10.1111/age.12626; pmid: 29446146
93. H. Li *et al*., Generation of biallelic knock-out sheep via gene-editing and somatic cell nuclear transfer. *Sci. Rep.* **6**, 33675 (2016). doi: 10.1038/srep33675; pmid: 27654750
94. L. Grobet *et al*., A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. *Nat. Genet.* **17**, 71–74 (1997). doi: 10.1038/ng0997-71; pmid: 9288100
95. J. Luo *et al*., Efficient generation of myostatin (MSTN) biallelic mutations in cattle using zinc finger nucleases. *PLOS ONE* **9**, e95225 (2014). doi: 10.1371/journal.pone.0095225; pmid: 24743319
96. H. Gu *et al*., Establishment and phenotypic analysis of an Mstn knockout rat. *Biochem. Biophys. Res. Commun.* **477**, 115–122 (2016). doi: 10.1016/j.bbrc.2016.06.030; pmid: 27289021
97. Y. Bi *et al*., Isozygous and selectable marker-free MSTN knockout cloned pigs generated by the combined use of CRISPR/Cas9 and Cre/LoxP. *Sci. Rep.* **6**, 31729 (2016). doi: 10.1038/srep31729; pmid: 27530319
98. G. R. Boocock *et al*., Mutations in SBDS are associated with Shwachman-Diamond syndrome. *Nat. Genet.* **33**, 97–101 (2003). doi: 10.1038/ng1062; pmid: 12496757
99. A. Daluiski *et al*., Bone morphogenetic protein-3 is a negative regulator of bone density. *Nat. Genet.* **27**, 84–88 (2001). doi: 10.1038/83810; pmid: 11138004
100. M. Agaba *et al*., Giraffe genome sequence reveals clues to its unique morphology and physiology. *Nat. Commun.* **7**, 11519 (2016). doi: 10.1038/ncomms11519; pmid: 27187213
101. D. R. Prothero, S. E. Foss, *The Evolution of Artiodactyls* (JHU Press, 2007).
102. M. Köhler, *Skeleton and Habitat of Recent and Fossil Ruminants* (F. Pfeil, 1993).
103. J. Kappelman, Morphology and locomotor adaptations of the bovid femur in relation to habitat. *J. Morphol.* **198**, 119–130 (1988). doi: 10.1002/jmor.1051980111; pmid: 3199446
104. L. Baskin, K. Danell, *Ecology of ungulates: a handbook of species in Eastern Europe and Northern and Central Asia* (Springer Science & Business Media, 2003).
105. F. Ma *et al*., The association of sport performance with ACE and ACTN3 genetic polymorphisms: A systematic review and meta-analysis. *PLOS ONE* **8**, e54685 (2013). doi: 10.1371/journal.pone.0054685; pmid: 23358679
106. E. A. Ostrander, H. J. Huson, G. K. Ostrander, Genetics of athletic performance. *Annu. Rev. Genomics Hum. Genet.* **10**, 407–429 (2009). doi: 10.1146/annurev-genom-082908-150058; pmid: 19630564
107. O. Toljagic, *Macroevolution With a Bite: Teeth Evolution and Diversification in Ruminants* (Univ. Oslo Press, 2018).
108. J. L. Cantalapiedra *et al*., Dietary innovations spurred the diversification of ruminants during the Caenozoic. *Proc. Biol. Sci.* **281**, 20132746 (2013). doi: 10.1098/rspb.2013.2746; pmid: 24352949

109. J. C. Hu *et al*., Enamel defects and ameloblast-specific expression in Enam knock-out/lacz knock-in mice. *J. Biol. Chem.* **283**, 10858–10871 (2008). doi: 10.1074/jbc.M710565200; pmid: 18252720
110. D. M. Daubert *et al*., Human enamel thickness and ENAM polymorphism. *Int. J. Oral Sci.* **8**, 93–97 (2016). doi: 10.1038/ijos.2016.1; pmid: 27357321
111. R. P. Shellis, A. D. Beynon, D. J. Reid, K. M. Hiiemae, Variations in molar enamel thickness among primates. *J. Hum. Evol.* **35**, 507–522 (1998). doi: 10.1006/jhev.1998.0238; pmid: 9774508
112. O. Toljagic, K. L. Voje, M. Matschiner, L. H. Liow, T. F. Hansen, Millions of Years Behind: Slow adaptation of ruminants to grasslands. *Syst. Biol.* **67**, 145–157 (2018). doi: 10.1093/sysbio/syx059; pmid: 28637223
113. C. Y. Li, J. Prochazka, A. F. Goodwin, O. D. Klein, Fibroblast growth factor signaling in mammalian tooth development. *Odontology* **102**, 1–13 (2014). doi: 10.1007/s10266-013-0142-1; pmid: 24343791
114. R. Li *et al*., De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20**, 265–272 (2010). doi: 10.1101/gr.097261.109; pmid: 20019144
115. R. Kajitani *et al*., Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res.* **24**, 1384–1395 (2014). doi: 10.1101/gr.170720.113; pmid: 24755901
116. N. I. Weisenfeld, V. Kumar, P. Shah, D. M. Church, D. B. Jaffe, Direct determination of diploid genome sequences. *Genome Res.* **27**, 757–767 (2017). doi: 10.1101/gr.214874.116; pmid: 28381613
117. M. Boetzer, C. V. Henkel, H. J. Jansen, D. Butler, W. Pirovano, Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578–579 (2011). doi: 10.1093/bioinformatics/btq683; pmid: 21149342
118. J. C. Mullikin, Z. Ning, The phusion assembler. *Genome Res.* **13**, 81–90 (2003). doi: 10.1101/gr.731003; pmid: 12529309
119. G. Benson, Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999). doi: 10.1093/nar/27.2.573; pmid: 9862982
120. A. Smit, R. Hubley, P. Green, RepeatMasker Open-4.0 (2015); www.repeatmasker.org
121. Z. Xu, H. Wang, LTR_FINDER: An efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265-8 (2007). doi: 10.1093/nar/gkm286; pmid: 17485477
122. S. F. Altschul *et al*., Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997). doi: 10.1093/nar/25.17.3389; pmid: 9254694
123. I. Korf, Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004). doi: 10.1186/1471-2105-5-59; pmid: 15144565
124. C. Burge, S. Karlin, Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94 (1997). doi: 10.1006/jmbi.1997.0951; pmid: 9149143
125. W. H. Majoros, M. Pertea, S. L. Salzberg, TigrScan and GlimmerHMM: Two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20**, 2878–2879 (2004). doi: 10.1093/bioinformatics/bth315; pmid: 15145805
126. M. Stanke *et al*., AUGUSTUS: Ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435-9 (2006). doi: 10.1093/nar/gkl200; pmid: 16845043
127. S. M. Kietbasa, R. Wan, K. Sato, P. Horton, M. C. Frith, Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011). doi: 10.1101/gr.113985.110; pmid: 21209072
128. M. Blanchette *et al*., Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14**, 708–715 (2004). doi: 10.1101/gr.1933104; pmid: 15060014
129. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014). doi: 10.1093/bioinformatics/btu033; pmid: 24451623
130. A. M. Kozlov, A. J. Aberer, A. Stamatakis, ExaML version 3: A tool for phylogenomic analyses on supercomputers. *Bioinformatics* **31**, 2577–2579 (2015). doi: 10.1093/bioinformatics/btv184; pmid: 25819675
131. E. Sayyari, J. B. Whitfield, S. Mirarab, DiscoVista: Interpretable visualizations of gene tree discordance. *Mol. Phylogenet. Evol.* **122**, 110–115 (2018). doi: 10.1016/j.ympev.2018.01.019; pmid: 29421312
132. M. J. Sanderson, r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**, 301–302 (2003). doi: 10.1093/bioinformatics/19.2.301; pmid: 12538260
133. A. J. Drummond, A. Rambaut, BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007). doi: 10.1186/1471-2148-7-214; pmid: 17996036
134. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007). doi: 10.1093/molbev/msm088; pmid: 17483113
135. J. L. Thorne, H. Kishino, Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* **51**, 689–702 (2002). doi: 10.1080/10635150290102456; pmid: 12396584
136. M. J. Hubisz, K. S. Pollard, A. Siepel, PHAST and RPHAST: Phylogenetic analysis with space/time models. *Brief. Bioinform.* **12**, 41–51 (2011). doi: 10.1093/bib/bbq072; pmid: 21278375
137. A. Siepel *et al*., Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005). doi: 10.1101/gr.3715005; pmid: 16024819
138. T. De Bie, N. Cristianini, J. P. Demuth, M. W. Hahn, CAFE: A computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269–1271 (2006). doi: 10.1093/bioinformatics/btl097; pmid: 16543274
139. L. Chen *et al*., Data for "Large-scale ruminant genome sequencing provides insights into their evolution and unique traits". Dryad (2019); doi:10.5061/dryad.52213gc

ACKNOWLEDGMENTS

We thank J. Zhou, L. Wong, and X. Ren from the China Biodiversity Conservation and Green Development Foundation (CBCGDF) for providing the sample of Przewalski's gazelle, A. al-Cher for help with extracting DNA from tissue, the Givskud Zoo for providing a skin sample for a bongo, R. Crowhurst for assistance with DNA extraction for the pronghorn, the members of the FANNG project for sharing their transcriptome data, and L. Nakeleh and J. Zhu for their patient instruction when doing PhyloNet analysis. **Funding:** This work was supported by the Talents Team Construction Fund of Northwestern Polytechnical University (NWPU); the Strategic Priority Research Program of CAS (XDB13000000) to W.W.; the Talents Team Construction Fund of Northwestern Polytechnical University (NWPU) and the National Program for Support of Top-notch Young Professionals to Q.Q.; the Villum Foundation Young Investigator grant (VKR023447) to R.H.; the Lundbeck Fellowship (R190-2014-2827), the Carlsberg Foundation (CF16-0663), and the Villum Investigator Grant to G.Z.; the National Natural Science Foundation of China (31822052) to Y.J.; the 1000 Talent Project of Shaanxi Province to W.W., Q.Q., and W.K.; the Fundamental Research Funds for the Central Universities (3102019JC007) to W.W. and Q.Q.; the Natural Science Foundation of China (NSFC) (31501984); the Natural science foundation from Jilin province (20170101158JC) and the Central Public-interest Scientific Institution Basal Research Fund (Y2019GH13) to Z.P.L.; the Danish National Research Foundation (grant DNRF96) to R.R.F.; the Robert and Rosabel Osborne Endowment and the U.S. Department of Agriculture Grants (538 AG2009-34480-19875 and 538 AG 58-1265-0-031) to H.A.L.; the National Science Foundation (NSF III-1845967) to S.M.; and the San Diego Zoo Global John and Beverly Stauffer Foundation to O.A.R. **Author contributions:** W.W., R.H., and G.Z. conceived the project and designed research. L.C., W.W., and K.W. drafted the manuscripts. W.W., R.H., G.Z., F.B., Y.J., H.R.S., M.T.P.G., Q.Q., H.A.L., and S.M. revised the manuscript. L.C, Z.S.L., K.W., Q.Q., C.Z.Z., Y.Y., C.L., Y.Zh., Y.G., R.Z., Y.R., R.L., and Y.Ze. performed genome assembly and genome annotation analysis. K.W., L.C., Z.S.L, S.M., and W.L. performed phylogeny analyses. J.Y., L.C., and R.L. assembled mitochondrial genomes. L.C., Z.S.L., Y.J., F.B., and B.We. performed time calibration, PSMC analyses, and interpretation of results. L.C., Z.S.L., X.P., W.F., Y.W., Y.E.Z., C.C., W.L., and B.Wa. conducted the analysis in identifying conserved elements and new genes. L.C., C.L., Z.S.L., Y.Z.Y., and Q.Q. performed gene family analysis. Q.Q., Y.Z.Y., and C.L. conducted positive selection and rapidly evolving gene analysis. L.C., Z.S.L., X.P., Y.W., Y.Zh.., B.We., C.Z., Z.W., and Y.Q. performed analysis relating to genomic features of ruminant characters. L.C., Z.S.L., C.L., and Z.P.L. performed metabolism pathway analysis. H.A.L. contributed the sequence data of gemsbok and indian muntjac. G.L., W.T.W., W.Z., and R.P.Z. prepared and extracted DNA samples. J.W., W.S., W.N., Y.T.W., C.H., C.W.E., L.G.C., O.R., R.H., R.R.F., and H.R.S. provided tissue samples. Y.D., W.T.W., S.D., and Y.G. prepared the library and sequencing. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** Genome assembles and DNA sequencing data have been deposited into NCBI with project no. PRJNA438286 and can also be accessed through the browser of database we constructed: http://animal.nwsuaf.edu.cn/code/index.php/Ruminantia. Transcriptome data of newly sequenced seven sheep samples are deposited into NCBI with accession IDs SRR7755234, SRR7755240, SRR7755450, SRR7755493, SRR7755310, SRR7755366, and SRR7755370. The alignment datasets of whole genomes, mitogenomes, orthologous genes, CNEs, and 1-kbp windows and the alternative trees based on these data are deposited in the Dryad Digital Repository (*139*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/eaav6202/suppl/DC1
Materials and Methods
Figs. S1 to S54
Tables S1 to S54
References (*140*–*199*)

8 October 2018; accepted 16 May 2019
10.1126/science.aav6202

RESEARCH ARTICLE SUMMARY

RUMINANT GENOMICS

Genetic basis of ruminant headgear and rapid antler regeneration

Yu Wang*, Chenzhou Zhang*, Nini Wang*, Zhipeng Li*, Rasmus Heller*, Rong Liu*, Yue Zhao*, Jiangang Han*, Xiangyu Pan, Zhuqing Zheng, Xueqin Dai, Ceshi Chen, Mingle Dou, Shujun Peng, Xianqing Chen, Jing Liu, Ming Li, Kun Wang, Chang Liu, Zeshan Lin, Lei Chen, Fei Hao, Wenbo Zhu, Chengchuang Song, Chen Zhao, Chengli Zheng, Jianming Wang, Shengwei Hu, Cunyuan Li, Hui Yang, Lin Jiang, Guangyu Li, Mingjun Liu, Tad S. Sonstegard, Guojie Zhang, Yu Jiang†, Wen Wang†, Qiang Qiu†

INTRODUCTION: All pecoran families, except the Moschidae, have cranial appendages or headgear, a unique structure among mammals that has a different morphology in each family (ossicones in giraffids, pronghorns in pronghorn, antlers in cervids, and horns in bovids). Moreover, the deer antler is the only completely regenerable organ found in mammals, thus providing a unique model for regenerative biology. Antlers also have extremely rapid growth rates (~1.7 cm/day in red deer), with rates of cell proliferation that surpass even cancerous tissue growth. Cervids also have low cancer rates. The relation between a tight regulation of antler growth and inhibition of oncogenesis in deer may provide insights for cancer prevention and therapy in humans and other organisms.

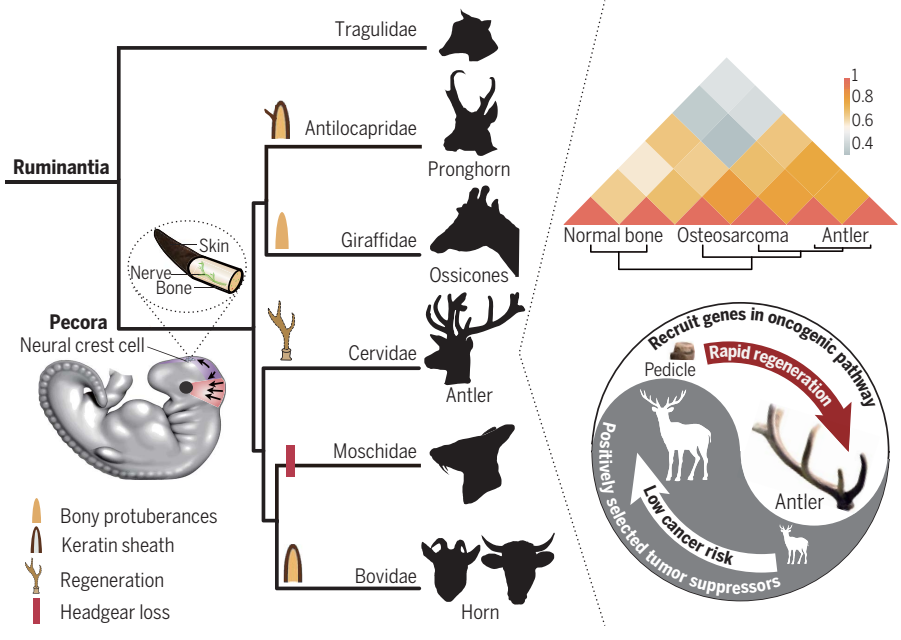
RATIONALE: We obtained 221 transcriptomes from bovids and cervids and sequenced three genomes representing the two pecoran lineages that convergently lack headgear. Comparing the data with a large set of ruminant genomes, including nine cervids, we detect genetic changes (lineage-specific positively selected genes and conserved elements) in pecorans with headgear (PWH), particularly cervids. Using the observed genetic changes and gene expression in headgear, we explore the genomic basis of ruminant headgear origin and antler regeneration.

RESULTS: We find that highly or specifically expressed genes in horns and antlers are most frequently coexpressed in bone, skin, nerve tissues, and testis. Many genes under positive selection in PWH (e.g., *OLIG1*, *OTOP3*), PWH-specific genes associated with highly conserved elements (e.g., *HOXD* gene cluster, *SNAI2*, *TWIST1*, *SOX9)*, and genes highly or specifically expressed in headgear (*RXFP2*, *SOX10*, *NGFR*) are involved in neural functions. In addition, *RXFP2*, which is specifically expressed in headgear and testis, was convergently pseudogenized in the headgearless lineages of Moschidae and Hydropotinae. The expression profile of antler is more correlated with osteocarcinoma than with normal bone tissue expression profiles. A number of proto-oncogenes (*FOS*, *REL*, *FAM83A*) and tumor suppression genes have been positively selected in cervids, especially several cofactor genes (*PML*, *NMT2*, and *CD2AP)* and regulator genes (*ELOVL6*, *S100A8*, *ISG15*, *CNOT3*, and *CCDC69*) of the p53 tumor suppressor, suggesting that these adaptive changes may enhance cancer resistance in deer.

CONCLUSION: Together, the phylogeny, gene expression profiles, and convergent headgear losses support a single evolutionary origin of the ruminant headgear. Pecoran headgear likely share a common cellular origin from neural crest stem cells, and the determination of the chondrogenic and neural lineages is important for headgear development. In addition, cervid-specific genetic changes in tumor suppressor and proto-oncogenes imply that the regenerative properties of antler tissue exploit oncogenesis pathways. Our study reveals genetic mechanisms underlying the evolutionary, developmental, and histological origin of ruminant headgear, as well as antler regeneration. The identified genes and their unique mutations provide guidelines for future functional studies of headgear development, regeneration of mammalian organs, and oncogenesis. ■

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aav6335



Neural crest cellular origin of ruminant headgear and the tight control of rapid antler regeneration and low cancer risk in cervids. (Left) Phylogenomic relationships of the six ruminant families. Anatomic features of family-specific headgear are depicted, showing that headgear of ruminants share tissue and cellular origins. (Upper right) The gene expression profile of antler correlates more strongly with osteocarcinoma than with normal bone tissue. (Lower right) The balance between rapid antler regeneration, which depends on genes in the oncogenic pathway, and reduced cancer risk, which may involve adaptive evolution of tumor suppressor genes.

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: qiuqiang@lzu.edu.cn (Q.Q.); wwang@mail.kiz.ac.cn (W.W.); yu.jiang@nwafu.edu.cn (Y.J.)
Cite this article as Y. Wang *et al.*, *Science* **364**, eaav6335 (2019). DOI: 10.1126/science.aav6335

RESEARCH ARTICLE

RUMINANT GENOMICS

Genetic basis of ruminant headgear and rapid antler regeneration

Yu Wang[1]*, Chenzhou Zhang[2]*, Nini Wang[1]*, Zhipeng Li[3]*, Rasmus Heller[4]*, Rong Liu[5,6]*, Yue Zhao[1]*, Jiangang Han[7]*, Xiangyu Pan[1], Zhuqing Zheng[1], Xueqin Dai[5,6], Ceshi Chen[5,6], Mingle Dou[1], Shujun Peng[1], Xianqing Chen[2], Jing Liu[1], Ming Li[1], Kun Wang[2], Chang Liu[2], Zeshan Lin[2], Lei Chen[2], Fei Hao[8], Wenbo Zhu[2], Chengchuang Song[1], Chen Zhao[1], Chengli Zheng[9], Jianming Wang[9], Shengwei Hu[10], Cunyuan Li[10], Hui Yang[8], Lin Jiang[7], Guangyu Li[3], Mingjun Liu[11], Tad S. Sonstegard[12], Guojie Zhang[6,13,14,15], Yu Jiang[1]†, Wen Wang[2,6,14]†, Qiang Qiu[2]†

Ruminants are the only extant mammalian group possessing bony (osseous) headgear. We obtained 221 transcriptomes from bovids and cervids and sequenced three genomes representing the only two pecoran lineages that convergently lack headgear. Comparative analyses reveal that bovid horns and cervid antlers share similar gene expression profiles and a common cellular basis developed from neural crest stem cells. The rapid regenerative properties of antler tissue involve exploitation of oncogenetic pathways, and at the same time some tumor suppressor genes are under strong selection in deer. These results provide insights into the evolutionary origin of ruminant headgear as well as mammalian organ regeneration and oncogenesis.

Ruminants are the only group of extant mammals with osseous cranial appendages, which are collectively termed headgear (*1*). Osseous headgear are exclusively found in the pecorans (all ruminants, excluding Tragulidae), a group that radiated ~23.3 million to 20.8 million years (Ma) ago (*2*) (Fig. 1). Each family in the pecoran group exhibits a distinct headgear morphology (*3*). The ossicones of Giraffidae consist of bony protuberances covered only by skin and hair. The pronghorns of Antilocapridae are composed of bone covered by skin, hair, and an annually deciduous forked keratinous sheath. The horns of Bovidae also have a bony core but are covered by a nondeciduous, nonforked keratinous sheath. The antlers of Cervidae are wholly deciduous, regenerating annually as an outgrowth of bone from the frontal skull. Despite this variation, all these types of pecoran headgear (including those of extinct species) share characteristic features, such as their frontal cranial position and a bony core covered by integument (fig. S1). The evolutionary origin of headgear, specifically whether headgear evolved only once or multiple times, has been a matter of considerable scientific discussion (*1*). Resolving this has proved challenging because of the lack of consensus regarding the family-level phylogeny of the ruminants.

[1]Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling 712100, China. [2]Center for Ecological and Environmental Sciences, Northwestern Polytechnical University, Xi'an 710072, China. [3]Department of Special Animal Nutrition and Feed Science, Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, Changchun 130112, China. [4]Section for Computational and RNA Biology, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark. [5]Key Laboratory of Animal Models and Human Disease Mechanisms of Chinese Academy of Sciences and Yunnan Province, Chinese Academy of Sciences, Kunming Institute of Zoology, Kunming, Yunnan 650223, China. [6]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China. [7]Institute of Animal Science (IAS), Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China. [8]Center of Special Environmental Biomechanics and Biomedical Engineering, School of Life Sciences, Northwestern Polytechnical University, Xi'an 710072, China. [9]Sichuan Institute of Musk Deer Breeding, Sichuan 610000, China. [10]College of Life Sciences, Shihezi University, Shihezi, Xinjiang 832003, China. [11]The Key Laboratory of Animal Biotechnology of Xinjiang, Xinjiang Academy of Animal Science, Xinjiang, Urumqi 830026, China. [12]Recombinetics, Inc., St. Paul, MN 55104, USA. [13]China National GeneBank, BGI-Shenzhen, Shenzhen 518083, China. [14]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. [15]Section for Ecology and Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark.
*These authors contributed equally to this work.
†Corresponding author. Email: qiuqiang@lzu.edu.cn (Q.Q.); wwang@mail.kiz.ac.cn (W.W.); yu.jiang@nwafu.edu.cn (Y.J.)

Results

A comprehensive phylogenetic analysis of the ruminants (*2*) resolved the family-level topology and shows that the most phylogenetically parsimonious hypothesis is a single origin of headgear in pecorans followed by two independent losses (Fig. 1). Multiple independent origins would be at odds with the rapid radiation of the five Pecora families, which took place during an ~2.5-million-year interval (23.3 to 20.8 Ma ago), and furthermore it is difficult to explain why headgear would have evolved multiple times in pecorans yet be absent in all other mammalian taxa. The Ruminant Genome Project (*2*) provides an opportunity to investigate the genomic background of ruminant headgear evolution and address its implications in organ regeneration.

Shared gene recruitment in horns and antlers

The evolutionary origin of any new organ typically depends on the recruitment of genes that were originally expressed in other tissues (*4*). To identify genes recruited in bovid horns, we compared 181 transcriptomes obtained in this study—representing 16 tissues and including 7 transcriptomes of goat horn sprouts, 61 of other goat tissues, 3 of sheep horn sprouts, and 110 of other sheep tissues—and added 49 published sheep transcriptomes (*5*) (table S1). In addition, transcriptomes from two fetal sheep horn buds and adjacent frontal skin tissues were sequenced to identify differentially expressed genes (DEGs) between these two tissue types at this important developmental stage. For cervid antlers, we sequenced 20 roe deer (*Capreolus capreolus*) and 20 sika deer (*Cervus nippon*) samples representing 16 tissues, including neonatal antlers (table S1). Genes specifically expressed in headgear tissues (hereafter headgear-specific genes) were defined as those that have a τ index exceeding 0.8 and are expressed most strongly or second most strongly in headgear tissues (*5*). We identified 624 horn-specific genes (table S2), and these were most highly coexpressed in bone, skin, testis, and brain tissues (Fig. 2A and fig. S2A). We also identified 761 antler-specific genes that were most highly coexpressed in the same four tissues (Fig. 2A, fig. S2B, and table S3). In addition, 201 headgear-specific expression genes were shared by both antler and horn tissues (fig. S3A), and these genes were enriched in bone development, skin development, and neurogenesis pathways (Fisher's exact test, adjusted P value < 1×10^{-5}) (fig. S3B and table S4). The DEGs (table S5) between fetal sheep horn bud and adjacent frontal skin tissues were enriched in nerve development pathways (fig. S4 and table S6). Histological analysis of cattle fetal horn buds suggested that, in contrast to the frontal skin of polled fetuses, neural tissue exists only in the horn buds (*6*). Overall, the gene expression results suggest that the development of both kinds of headgear considered here (horns and antlers) depends on similar gene expression profiles, largely recruited from nerve, bone, and skin tissues.

A common genetic and cellular basis of headgear

To identify the genetic basis of headgear evolution, we used the branch-site likelihood ratio test (*5*) to detect positively selected genes shared only by pecorans with headgear (PWH). A total of 240 genes were identified as positively selected in PWH (table S7). Enriched functional Gene Ontology (GO) categories of these genes included biomineral tissue development (Fisher's exact test, adjusted P value = 4.9×10^{-2}) (table S8), providing further evidence that the evolution of osseous headgear depends on bone development.

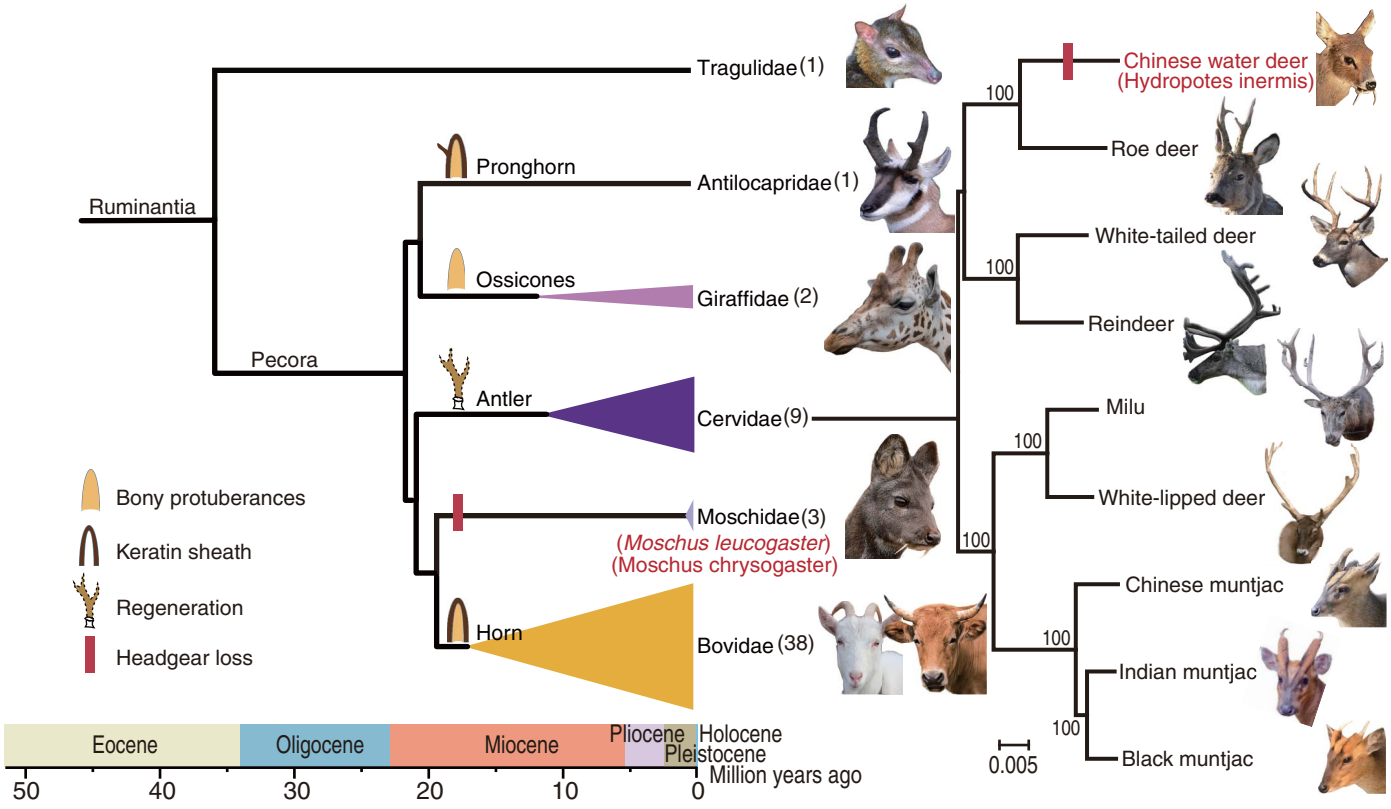


Fig. 1. Phylogenomic placement of species without headgear in the Ruminantia and Cervidae. (**Left**) The phylogenomic relationships of the six ruminant families from (*2*). (**Right**) Maximum-likelihood tree for the nine studied cervid species obtained using 3,316,385 four-fold degenerate sites. The anatomic structure of headgear in each family is depicted, including the keratin sheath of Antilocapridae and Bovidae and the regenerable antler of Cervidae. The red bar indicates secondary headgear loss and the red text highlights the de novo assembled species in this study. The number of species in each family used in this study is indicated in parentheses. Sources and credits for species photos are listed in table S26.

Among these genes, we found that *OTOP3* was under positive selection in PWH along with a headgear-specific gene (fig. S5 and tables S2, S3, and S7). *OTOP3* is a member of the otopetrin gene family, which regulates biomineralization processes (*7*). Of the eight Pecora-specific amino acid mutations in the OTOP3 protein, four are located in the otopetrin functional domain (PF03189) (fig. S5). *OTOP3* is expressed in the neural crest of *Xenopus* embryos, suggesting a role in neural crest function (*8*). *OLIG1* (fig. S6), another positively selected gene in PWH, is also related to neural crest differentiation pathways (*9*). A previous study showed that a 212–base pair (bp) duplication (~65 kbp) flanking the *OLIG1* gene region is the causal mutation of polled cattle, i.e., cattle that completely lack horns (*10*). We detected nine distinct amino acid changes in *OLIG1* in the inferred common ancestor of Pecora, resulting in domain structure changes as revealed by protein structure homology modeling (fig. S6). Given that the frontal cranial bones are derived from cells of the cranial neural crest (*11*), changes in the *OTOP3* and *OLIG1* genes likely played crucial roles in the evolution and development of pecoran headgear (Fig. 2B).

We also identified 8732 lineage-specific highly conserved elements (HCEs) (≥20 bp) (table S9) in PWH using the phylogenetic hidden Markov model approach (*5*). The PWH-specific HCE-associated genes are enriched in the signaling pathway regulating the pluripotency of stem cells (Fisher's exact test, adjusted P value = 2.5×10^{-9}) and transforming growth factor–β (TGF-β) (Fisher's exact test, adjusted P value = 3.5×10^{-7}) (table S10). The TGF-β signaling pathway plays an important role in bone formation and the regulation of cranial neural crest cell proliferation during frontal bone development (*12*). The PWH-specific HCE-associated genes *SNAI2*, *TWIST1*, *SOX9*, and the *HOXD* gene cluster are involved in neural crest cell migration (*9*) (Fig. 2B and table S9). Specifically, we identified a PWH-specific 25-bp HCE 15 kbp downstream from the *HOXD* gene cluster (fig. S7A) that serves as a master regulator in neural crest patterning, particularly in the cranial region (*13*). Further analysis indicates that this element resulted from a 3.6-kbp Pecora-specific transposable element insertion (fig. S7, B and C) located in the candidate region causing the four-horned phenotype in sheep (*14*). These results suggest that the evolution of PWH-specific regulatory elements may also play a role in reprogramming neural crest cells to develop into headgear.

In addition, we found that six neural crest cell migration-related genes (*SOX10*, *SNAI1*, *SNAI2*, *TFAP2A*, *NGFR*, and *COL11A2*) are specifically expressed in headgear (tables S2 and S3). *SOX10*, *SNAI1*, and *TFAP2A* are highly expressed in fetal sheep horn buds but not in adjacent skin tissue (fig. S4 and table S5). Notably, *SOX10* and *NGFR* are used as marker genes of neural crest cells (*15*), and our immunohistochemical analysis confirmed that *SOX10*- and *NGFR*-positive cells are present in the embryonic horn bud of sheep (fig. S8). In addition, two headgear-specific genes, *FOXL2* and *TWIST1*, are related to horn abnormalities (*16*, *17*) (tables S2 and S3) and both interact with *SOX9*, a marker gene for the determination of the chondrogenic lineage in the cranial neural crest (*18*). From the results of our comparative genomic analysis and previous findings regarding horn-related genes (*10*, *14*, *16*, *17*), we conclude that pecoran headgear likely share a common cellular origin in neural crest stem cells (Fig. 2B). This supports a single evolutionary origin for pecoran headgear despite their morphological diversity.

Convergent pseudogenization led to secondary loss of headgear

We supplemented the 51 ruminant genomes in the Ruminant Genome Project (*2*) with a high-quality reference genome from a secondarily antlerless species, the Chinese water deer (*Hydropotes inermis*) of the cervid subfamily Hydropotinae

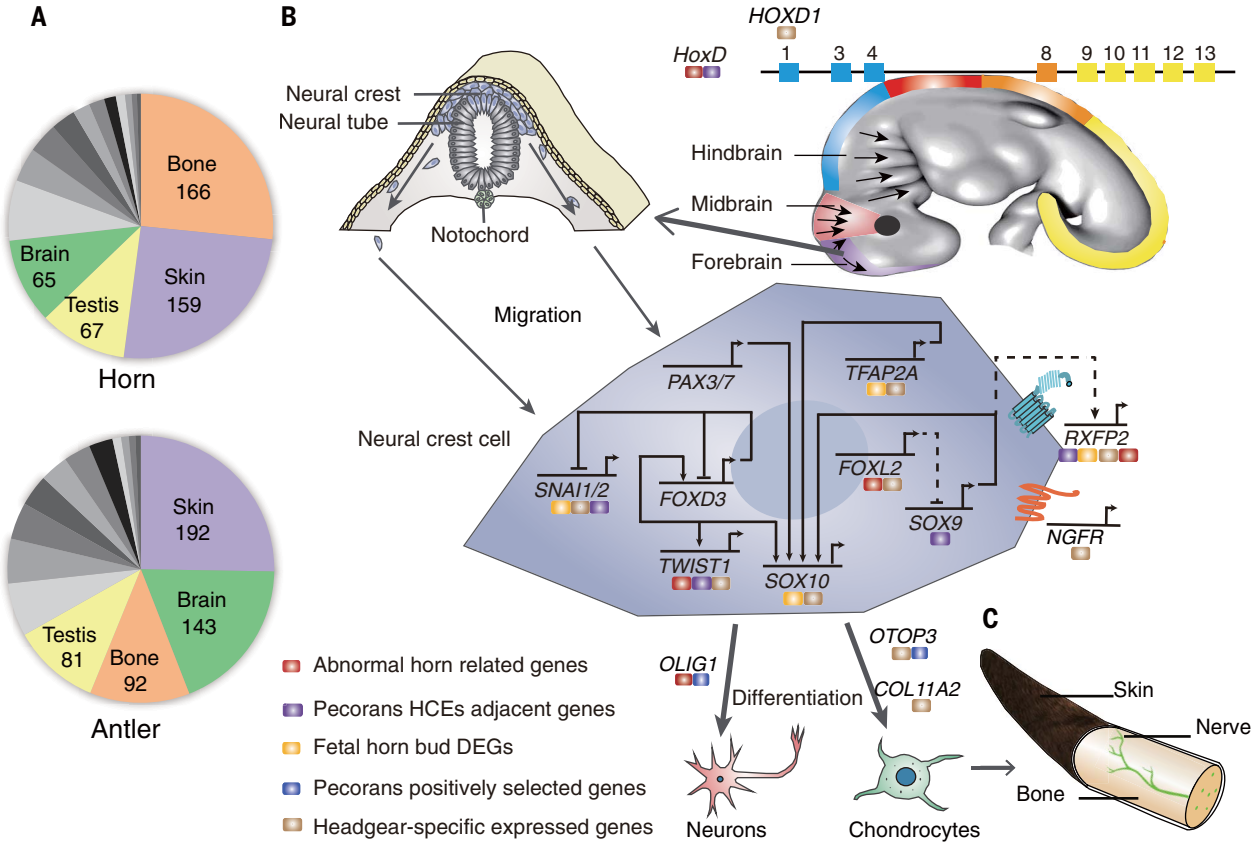


Fig. 2. Gene recruitment and cellular origin of ruminant headgear. (**A**) Genes recruited to headgear from different organs. Bone, skin, testis, brain, and others are marked as orange, purple, yellow, green, and various shades of gray, respectively. (**B**) Diagram of neural crest cell genes involved in headgear development. The *HOXD* gene cluster is depicted as the hypothesized master regulator of headgear. Genes annotated in the neural crest cell migration and differentiation pathway are labeled with different colors to indicate positively selected genes (PSGs), HCEs, fetal horn bud DEGs, and genes related to abnormal horn development. The solid arrows represent known neural crest cell pathways. The dashed arrows indicate the pathway known to be related to other cell types that has not been recorded in neural crest cells. (**C**) Diagram of the headgear of the ruminant ancestor, mainly containing bone, skin, and nerve tissues.

(fig. S9 and tables S11 to S13), and two contig-level genomes of Moschidae species (*Moschus chrysogaster* and *M. leucogaster)* (table S14). From the high-confidence phylogenetic tree obtained with whole-genome data (Fig. 1), it is clear that Moschidae and Hydropotinae have independently lost their headgear. To explain this convergent feature and learn more about genes controlling headgear, we investigated pseudogenization in the shape of premature stop codons and frameshifts in both of these lineages (Fig. 1 and tables S15 and S16). Of 289 pseudogenes identified in these two distant lineages, *RXFP2* was the only headgear-specific gene that was convergently pseudogenized, although the mutations occur at different sites of the *RXFP2* gene in these lineages (fig. S10). *RXFP2*, which our transcriptomic data indicate was recruited from testis tissue into horn and antler, is highly expressed in the fetal sheep horn bud (fig. S10) and has two PWH-specific HCEs in its intron region, one of them with a binding motif for the horn-specific gene *ARNT* (fig. S11). Previous studies identified *RXFP2* as a sexually selected gene associated with horn morphology in bighorn sheep (*Ovis canadensis*) (*19*, *20*). Additionally, a 1.8-kbp insertion in the 3′ untranslated region (3′UTR) of *RXFP2* is associated with lack of horns in sheep (*21*). Collectively, these results indicate that pseudogenization of *RXFP2* is the most likely functional mechanism behind the convergent secondary loss of headgear in the Moschidae and Hydropotinae lineages.

Neural processes involved in antler regeneration

The deer antler is the only completely regenerable organ found in mammals and thus provides a unique model for regenerative biology. Antler regeneration is a stem cell–based process (*22*), and we demonstrated that antler stem cells may originate from cranial neural crest cells, which have the potential to rapidly proliferate and differentiate into cartilage and neural cells (Fig. 2B). Notably, growing antlers are richly innervated by sensory fibers, and resection of the antler pedicle sensory nerve markedly reduces antler regeneration and stunts antler size (*23*). Both antler-specific expression genes (table S3) and cervid-specific HCE-associated genes (table S17) are enriched in annotation terms associated with the axon guidance pathway, particularly the genes coding for key guidance molecules for axon growth: slits, ephrins, and semaphorins (Fig. 3A and tables S18 and S19). In addition, the top eight rapidly evolving genes in cervids with adjacent HCEs are all related to neural functions (fig. S12 and table S20). We also found that the nerve growth factor receptor (*NGFR*) gene is strongly and specifically expressed in antlers (table S3). This is consistent with findings that neural growth factors promote the growth of antler nerves (*24*), and these lines of evidence corroborate the involvement of neural processes in annual antler regeneration.

Similarities between antler growth and cancer cell growth programs

Antlers grow extremely quickly, as exemplified by red deer antlers, which have average growth rates of 1.7 cm/day and can reach a weight of up to 30 kg (*25*). These antlers regrow annually (from spring to summer) owing to very fast cell proliferation (Fig. 3B) that surpasses even cancerous tissue growth (*25*). Antler growth mainly

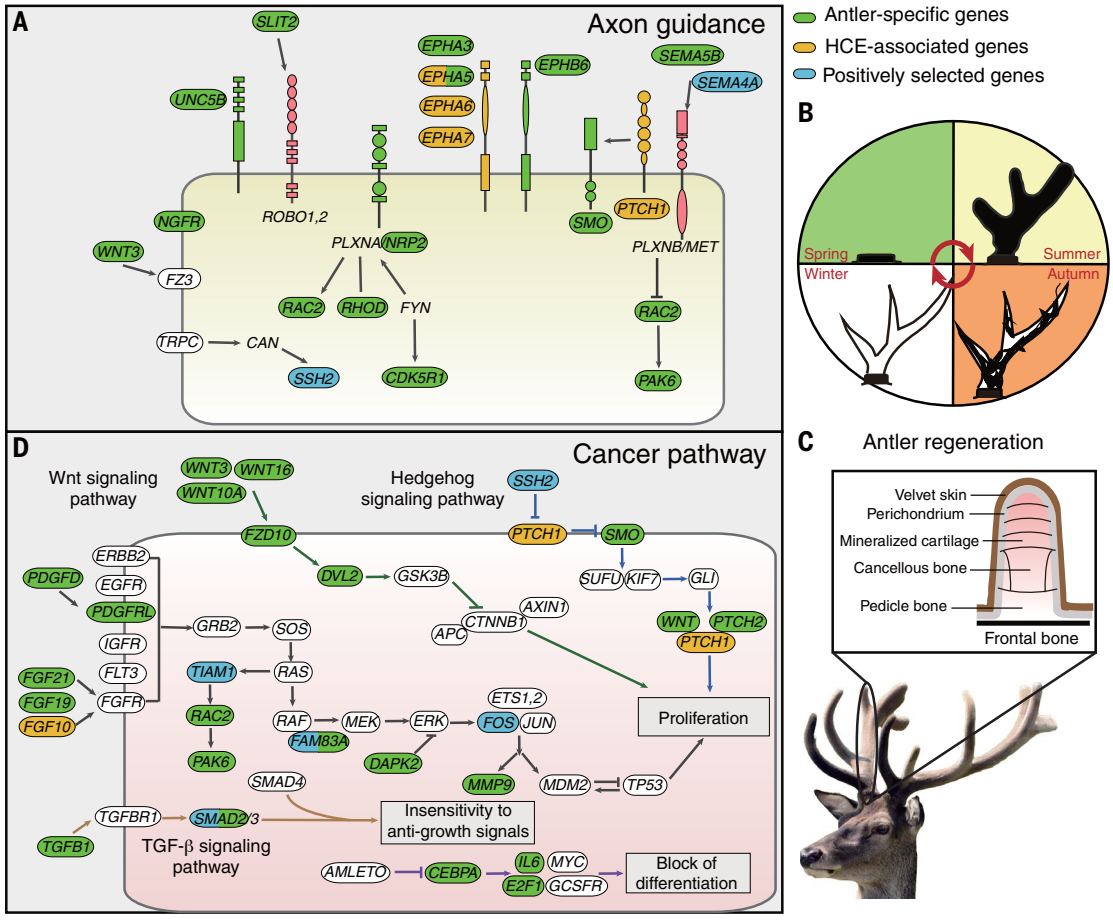


Fig. 3. Oncogenesis and neurogenesis pathways involved in rapid antler regeneration. (**A** and **D**) Positively selected genes in cervids (blue), cervid-specific HCE-associated genes (yellow), and antler-specific expression genes (green) in axon guidance (A) and oncogenesis (D) pathways. (**B**) Annual antler regeneration cycle. Antlers are shed in winter and regenerate in spring, growing most rapidly in summer. Antlers then calcify and shed their velvet in autumn. (**C**) The anatomy of the antler. The source and credit for the red deer photo are listed in table S26.

proceeds by chondrocyte proliferation and ossification (Fig. 3C) and thus provides a research model for osteosarcoma. We found a higher correlation between the gene expression profiles of antler and osteosarcoma (r = 0.67 to 0.78) than between those of antler and normal bone tissues (r = 0.33 to 0.47) (fig. S13), showing similar patterns of developmental programs in antler growth and oncogenesis.

We found evidence that three proto-oncogenes (*FOS*, *FAM83A*, and *REL*) were under positive selection in the cervid ancestor (fig. S14 and table S21). Of these, *FOS* acts as a downstream growth factor signaling pathway that regulates cell proliferation and differentiation. Overexpression of *FOS* induces osteosarcoma formation in mice via the transformation of chondroblasts and osteoblasts (*26*). Additionally, *FAM83A* has been identified as an oncogene involved in the epidermal growth factor receptor (EGFR) signaling pathway (*27*). We also observed antler-specific expression of five growth factor and receptor genes (*FGF19*, *FGF21*, *FGFBP3*, *PDGFD*, and *PDGFRL*) that play important roles in driving cancer cell proliferation and survival (*28*) (table S3). In addition, a cervid-specific HCE is located in the 3′UTR of *NOVA1*, which is believed to activate telomerase and promote tumor growth in vivo (*29*) (figs. S15 and S16 and table S17). Taken together, these cell growth–associated cervid-specific changes and the expression of tumor promoters in antlers (Fig. 3D) indicate that the rapid proliferation of cells required for rapid antler growth has similarities with cancer cell growth programs.

Regulation of antler growth may confer cancer resistance

Cancer frequency records from both the Philadelphia and San Diego zoos indicate that cancer incidence rates are ~5 times lower in cervids than in other mammals (0.4 to 0.8% and 2.1 to 4.6%, respectively) (*30*, *31*). This tentatively suggests that the precisely regulated cell growth regulators required for controlled rapid antler regeneration may confer protection against the development of cancers in cervids because of specific genetic changes relevant to cancer avoidance. Accordingly, Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of DEGs between antler and osteosarcoma shows enrichment for cancer- and metabolism-related pathways (fig. S13B and table S22).

We also observed that many tumor suppressor genes are under positive selection (table S21) and strongly expressed in antlers. Among these, the tumor suppressor gene *PML* has 11 cervid-specific nonsynonymous changes and carries the strongest signal of positive selection detected in cervids (Fig. 4A and table S21). PML is a transcriptional coactivator of p53, and its overexpression enhances p53 transcriptional activation and leads to cell growth arrest (*32*). The *TP53* (encoding the protein p53) signaling pathway plays a central role in regulating cell division and preventing tumor formation (*33*). We observed that three p53 cofactor genes (*PML*, *NMT2*, and *CD2AP)* and five p53 regulator genes (*ELOVL6*, *S100A8*, *ISG15*, *CNOT3*, and *CCDC69*) were under positive selection in the cervid lineage and expressed in antlers (Fig. 4B and table S21). We also noticed that *TP53* itself was identified as a rapidly evolving gene in the cervid lineage from the evolutionary analysis of 51 ruminant genomes (*2*).

In addition to the *TP53* pathway–related genes, we also observed several other tumor suppressor genes that were under selection in the cervid lineage and expressed in antlers. One such gene, *ADAMTS18* (Fig. 4A), belongs to the *ADAMTS* family, which encompasses disintegrin and metalloproteinase-like proteinases that inhibit growth of carcinoma cells by controlling the structure and function of the extracellular matrix (ECM) and regulating the tumor microenvironment (*34*). The components and function of the ECM are known to play an important role in cancer resistance in the naked mole-rat (*35*). In addition, the *ADAMTS* family members *ADAMTS2*, *ADAMTS4*, *ADAMTS12*, *ADAMTS14*, *ADAMTS17*, and *ADAMTS18* are not only all antler-specific genes (fig. S17) but are also more highly expressed in antler than in osteosarcoma (fig. S13C).

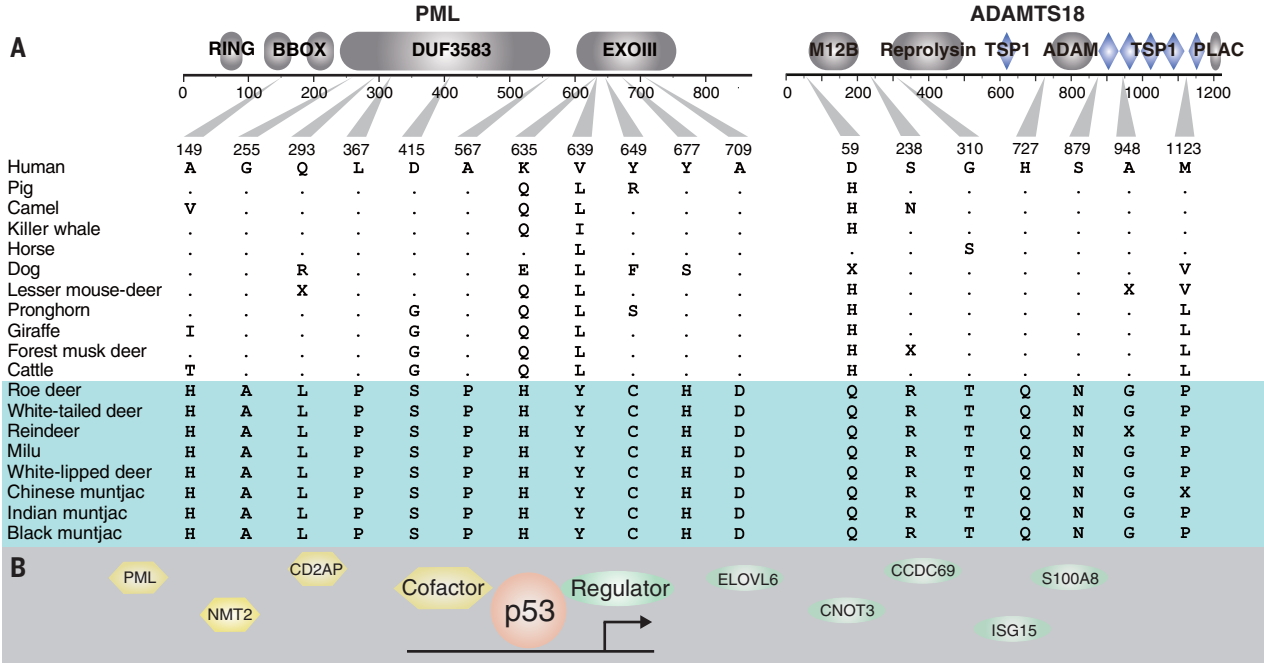


Fig. 4. Examples of positively selected tumor suppressor genes in cervids. (**A**) Gene models showing cervid-specific mutations of two positively selected tumor suppressor genes, *PML* and *ADAMTS18*. *PML* has the strongest selection signals detected in cervids (likelihood ratio test, P value = 1.63×10^{-6}) (table S21). (**B**) Genes positively selected in the p53 signaling pathway of cervids. Selected p53 cofactors are highlighted with yellow hexagons, and selected regulators are marked with green ovals.

Finally, several genes involved in DNA damage response pathways showed signatures of cervid-specific evolution in our transcriptomic and comparative genomic analyses. *TP73* and *TP53I13* suppress tumors through their roles in the p53-mediated DNA damage response pathway (*36*) and are specifically expressed in antlers (table S3). Moreover, we found that three more DNA damage response genes (*SLF1*, *RHNO1*, and *DDB2*) were under positive selection in the cervid lineage (table S21).

The cervid-specific expression and genetic changes in these tumor suppressor and DNA repair genes may play important roles in the fine-tuned regulation of rapid antler regeneration, while at the same time preventing the onset of cancers. Further detailed functional studies may therefore be of great scientific significance in demonstrating the mechanisms underlying rapid but controlled cell growth and exploring the potential of cervids as a cancer model.

Discussion

Headgear, or cranial appendages (antlers in cervids, horns in bovids, pronghorns in pronghorn, and ossicones in giraffids), are conspicuous and diverse features of the Pecora lineage within Ruminantia and are unique among all mammals. Headgear evolution is still debated, partly because of differences in the evolutionary scenarios suggested by different phylogenic trees, coupled with the notable differences in headgear anatomy and development (*1*). Therefore, it has been proposed that pecoran headgear could have multiple independent origins. We show that the most parsimonious explanation is a single headgear origin (*2*). Our comparative genomic and transcriptomic analyses indicate that horns and antlers are very similar in their gene expression profiles and that pecoran headgear share cellular origins from neural crest stem cells. Moreover, the headgear-specific gene *RXFP2* was convergently pseudogenized in secondarily headgearless lineages, corroborating the single evolutionary origin hypothesis by providing a simple genetic mechanism for the otherwise puzzling convergent loss of headgear. Notably, the extremely rapid growth and peculiar regeneration of cervid antlers are biological features of interest. For instance, an incision in an antler tine results in a slight scar that in the following year leads to a small tine, whereas injury to the pedicle leads to permanent inhibition of either pedicle or antler growth (*37*, *38*). Furthermore, it has been shown that electrical stimulation of antler nerves increases antler length and weight (*39*), suggesting that neural tissue plays an important role in antler regeneration. This link was corroborated by the neural associations of several antler-specific genes and cervid-specific HCE-associated genes identified in this study. Further integrated functional, physiological, and transcriptomic analyses of a wider range of samples during various antler growth stages are warranted to clarify the roles of these antler-specific neural genes in development and their potential utility in regenerative medicine and in vitro organ regeneration.

The antler is an exclusively male trait in cervids (except reindeer), which has been strongly selected by sexual selection, and some species most likely became extinct as a result of exaggerated antler size (*40*). Our data suggest that fast-growing cervid antlers have expression profiles that are more similar to those of osteosarcoma than to those of normal bone tissues (fig. S13). On the other hand, cervids have much lower cancer incidence than other mammals (*30*, *31*). It is conceivable that natural selection might have selected for efficient cancer-defense mechanisms in deer. It has previously been shown that elephants have a reduced cancer risk because of functional duplicates of the master tumor suppressor *TP53* (*41*). In contrast, cervids have a single copy of *TP53*, but other genes (*PML*, *NMT2*, *CD2AP*, *ELOVL6*, *S100A8*, *ISG15*, *CNOT3*, and *CCDC69*) functioning in the p53 pathway are under positive selection, suggesting that cervids may have evolved an enhanced *TP53* signaling pathway to constrain tumor growth. Elephants and deer may therefore have independently evolved different strategies to avoid cancer by targeting the same central tumor-controlling p53 regulatory pathway. We also found evidence that other tumor suppressor genes and proto-oncogenes have been under strong positive selection in cervids and/or are strongly and specifically expressed in the antler (e.g., *ADAMTS18*, *FOS*, *REL*, and *FAM83A*). Our study reveals the genetic mechanisms underlying the evolutionary, developmental, and histological origins of pecoran headgear and provides insights into the molecular mechanisms of regeneration of deer antler and its relevance to cancer resistance. The identified genes and their specific mutations provide a starting point for future functional studies of headgear

development, regeneration of mammalian organs, and oncogenesis.

Materials and methods

We sequenced and collected 221 and 49 transcriptomes, respectively, including 20 of roe deer, 20 of sika deer, 68 of goat, and 162 of sheep. RNA-sequencing reads were aligned with HISAT2 v2.0.3 (*42*), and the gene expression levels were quantified with StingTie v1.2.2 (*43*). Tissue-specific expressed genes were calculated with the τ index method (*44*). The DEGs in the embryonic horn bud were identified with DESeq2 v1.20.0 (*45*). Immunohistochemical analysis was performed to detect the expression of neural crest marker genes *NGFR* and *SOX10* in the embryonic horn bud. The genome of the Chinese water deer (*H. inermis*) was sequenced using 10X genomic platform and assembled with Supernova software v1.2.1 (*46*). Two contig-level assemblies of Moschidae species were obtained by Illumina HiSeq 2000 sequencing and assembled using the SOAPdenovo software v2.04 (*47*). Pseudogenes of the three headgearless genomes were identified with GeneWise v4.0 (*48*). Whole-genome alignments of 54 ruminant species were carried out with LAST vlast867 (*49*) and multiz v11.2 (*50*) using goat as the reference genome. Positively selected genes were identified by using the branch-site model in PAML v4.9e (*51*). PWH-specific HCEs were identified with phastCons v1.4 (*52*). Elements showing accelerated evolution in Cervidae were detected using phyloP v1.4 with the CONACC mode and LRT test method (*52*).

REFERENCES AND NOTES

1. E. B. Davis, K. A. Brakora, A. H. Lee, Evolution of ruminant headgear: A review. *Proc. Biol. Sci.* **278**, 2857–2865 (2011). doi: 10.1098/rspb.2011.0938; pmid: 21733893
2. L. Chen *et al*., Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science* **364**, eaav6202 (2019).
3. G. A. Bubenik, A. B. Bubenik, Eds., *Horns, Pronghorns, and Antlers: Evolution, Morphology, Physiology, and Social Significance* (Springer, 1990).
4. O. W. Griffith, G. P. Wagner, The placenta as a model for understanding the origin and evolution of vertebrate organs. *Nat. Ecol. Evol.* **1**, 0072 (2017). doi: 10.1038/s41559-017-0072; pmid: 28812655
5. Materials and methods are available as supplementary materials.
6. D. J. Wiener, N. Wiedemar, M. M. Welle, C. Drögemüller, Novel features of the prenatal horn bud development in cattle (*Bos taurus*). *PLOS ONE* **10**, e0127691 (2015). doi: 10.1371/journal.pone.0127691; pmid: 25993643
7. I. Hughes, M. Saito, P. H. Schlesinger, D. M. Ornitz, Otopetrin 1 activation by purinergic nucleotides regulates intracellular calcium. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12023–12028 (2007). doi: 10.1073/pnas.0705182104; pmid: 17606897
8. V. V. Novoselov, E. M. Alexandrova, G. V. Ermakova, A. G. Zaraisky, Expression zones of three novel genes abut the developing anterior neural plate of *Xenopus* embryo. *Gene Expr. Patterns* **3**, 225–230 (2003). doi: 10.1016/S1567-133X(02)00077-7; pmid: 12711553
9. P. Betancur, M. Bronner-Fraser, T. Sauka-Spengler, Assembling neural crest regulatory circuits into a gene regulatory network. *Annu. Rev. Cell Dev. Biol.* **26**, 581–603 (2010). doi: 10.1146/annurev.cellbio.042308.113245; pmid: 19575671
10. D. F. Carlson *et al*., Production of hornless dairy cattle from genome-edited cell lines. *Nat. Biotechnol.* **34**, 479–481 (2016). doi: 10.1038/nbt.3560; pmid: 27153274
11. G. Couly, S. Creuzet, S. Bennaceur, C. Vincent, N. M. Le Douarin, Interactions between Hox-negative cephalic neural crest cells and the foregut endoderm in patterning the facial skeleton in the vertebrate head. *Development* **129**, 1061–1073 (2002). pmid: 11861488
12. T. Sasaki *et al*., TGFβ-mediated FGF signaling is crucial for regulating cranial neural crest cell proliferation during frontal bone development. *Development* **133**, 371–381 (2006). doi: 10.1242/dev.02200; pmid: 16368934
13. M. Gouti, J. Briscoe, A. Gavalas, Anterior *Hox* genes interact with components of the neural crest specification network to induce neural crest fates. *Stem Cells* **29**, 858–870 (2011). doi: 10.1002/stem.630; pmid: 21433221
14. J. W. Kijas, T. Hadfield, M. Naval Sanchez, N. Cockett, Genome-wide association reveals the locus responsible for four-horned ruminant. *Anim. Genet.* **47**, 258–262 (2016). doi: 10.1111/age.12409; pmid: 26767438
15. D. E. Wagner *et al*., Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018). doi: 10.1126/science.aar4362; pmid: 29700229
16. E. Pailhoux *et al*., A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nat. Genet.* **29**, 453–458 (2001). doi: 10.1038/ng769; pmid: 11726932
17. A. Capitan *et al*., A newly described bovine type 2 scurs syndrome segregates with a frame-shift mutation in *TWIST1*. *PLOS ONE* **6**, e22242 (2011). doi: 10.1371/journal.pone.0022242; pmid: 21814570
18. Y. Mori-Akiyama, H. Akiyama, D. H. Rowitch, B. de Crombrugghe, Sox9 is required for determination of the chondrogenic cell lineage in the cranial neural crest. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9360–9365 (2003). doi: 10.1073/pnas.1631288100; pmid: 12878728
19. M. Kardos *et al*., Whole-genome resequencing uncovers molecular signatures of natural and sexual selection in wild bighorn sheep. *Mol. Ecol.* **24**, 5616–5632 (2015). doi: 10.1111/mec.13415; pmid: 26454263
20. S. E. Johnston *et al*., Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**, 93–95 (2013). doi: 10.1038/nature12489; pmid: 23965625
21. N. Wiedemar, C. Drögemüller, A 1.8-kb insertion in the 3′-UTR of *RXFP2* is associated with polledness in sheep. *Anim. Genet.* **46**, 457–461 (2015). doi: 10.1111/age.12309; pmid: 26103004
22. C. Li, F. Yang, A. Sheppard, Adult stem cells and mammalian epimorphic regeneration-insights from studying annual renewal of deer antlers. *Curr. Stem Cell Res. Ther.* **4**, 237–251 (2009). doi: 10.2174/157488809789057446; pmid: 19492976
23. C. Li, P. W. Sheard, I. D. Corson, J. M. Suttie, Pedicle and antler development following sectioning of the sensory nerves to the antlerogenic region of red deer (*Cervus elaphus*). *J. Exp. Zool.* **267**, 188–197 (1993). doi: 10.1002/jez.1402670212; pmid: 8409900
24. C. Li *et al*., Nerve growth factor mRNA expression in the regenerating antler tip of red deer (*Cervus elaphus*). *PLOS ONE* **2**, e148 (2007). doi: 10.1371/journal.pone.0000148; pmid: 17215957
25. R. J. Goss, *Deer Antlers: Regeneration, Function, and Evolution* (Academic Press, 1983).
26. Z. Q. Wang, J. Liang, K. Schellander, E. F. Wagner, A. E. Grigoriadis, c-fos-induced osteosarcoma formation in transgenic mice: Cooperativity with c-jun and the role of endogenous c-fos. *Cancer Res.* **55**, 6244–6251 (1995). pmid: 8521421
27. S. Chen *et al*., *FAM83A* is amplified and promotes cancer stem cell-like traits and chemoresistance in pancreatic cancer. *Oncogenesis* **6**, e300 (2017). doi: 10.1038/oncsis.2017.3; pmid: 28287611
28. N. Turner, R. Grose, Fibroblast growth factor signalling: From development to cancer. *Nat. Rev. Cancer* **10**, 116–129 (2010). doi: 10.1038/nrc2780; pmid: 20094046
29. A. T. Ludlow *et al*., NOVA1 regulates *hTERT* splicing and cell growth in non-small cell lung cancer. *Nat. Commun.* **9**, 3112 (2018). doi: 10.1038/s41467-018-05582-x; pmid: 30082712
30. L. A. Griner, *Pathology of Zoo Animals: A Review of Necropsies Conducted Over a Fourteen-Year Period at the San Diego Zoo and San Diego Wild Animal Park* (Zoological Society of San Diego, 1983).
31. L. S. Lombard, E. J. Witte, Frequency and types of tumors in mammals and birds of the Philadelphia Zoological Garden. *Cancer Res.* **19**, 127–141 (1959). pmid: 13629476
32. M. Pearson, P. G. Pelicci, PML interaction with p53 and its role in apoptosis and replicative senescence. *Oncogene* **20**, 7250–7256 (2001). doi: 10.1038/sj.onc.1204856; pmid: 11704853
33. E. R. Kastenhuber, S. W. Lowe, Putting p53 in context. *Cell* **170**, 1062–1078 (2017). doi: 10.1016/j.cell.2017.08.028; pmid: 28886379
34. H. Jin *et al*., Epigenetic identification of ADAMTS18 as a novel 16q23.1 tumor suppressor frequently silenced in esophageal, nasopharyngeal and multiple other carcinomas. *Oncogene* **26**, 7490–7498 (2007). doi: 10.1038/sj.onc.1210559; pmid: 17546048
35. X. Tian *et al*., High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. *Nature* **499**, 346–349 (2013). doi: 10.1038/nature12234; pmid: 23783513
36. D. W. Meek, Tumour suppression by p53: A role for the DNA damage response? *Nat. Rev. Cancer* **9**, 714–723 (2009). doi: 10.1038/nrc2716; pmid: 19730431
37. A. B. Bubenik, R. Pavlansky, Trophic responses to trauma in growing antlers. *J. Exp. Zool.* **159**, 289–302 (1965). doi: 10.1002/jez.1401590302; pmid: 5883952
38. R. J. Goss, C. W. Severinghaus, S. Free, Tissue relationships in the development of pedicles and antlers in the Virginia Deer. *J. Mammal.* **45**, 61–68 (1964). doi: 10.2307/1377295
39. G. A. Bubenik, A. B. Bubenik, E. D. Stevens, A. G. Binnington, The effect of neurogenic stimulation on the development and growth of bony tissues. *J. Exp. Zool.* **219**, 205–216 (1982). doi: 10.1002/jez.1402190210; pmid: 7061972
40. R. A. Moen, J. Pastor, Y. Cohen, Antler growth and extinction of Irish elk. *Evol. Ecol. Res.* **1**, 235–249 (1999).
41. L. M. Abegglen *et al*., Potential mechanisms for cancer resistance in elephants and comparative cellular response to DNA damage in humans. *JAMA* **314**, 1850–1860 (2015). doi: 10.1001/jama.2015.13134; pmid: 26447779
42. D. Kim, B. Langmead, S. L. Salzberg, HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015). doi: 10.1038/nmeth.3317; pmid: 25751142
43. M. Pertea *et al*., StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015). doi: 10.1038/nbt.3122; pmid: 25690850
44. N. Kryuchkova-Mostacci, M. Robinson-Rechavi, A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.* **18**, 205–214 (2017). doi: 10.1093/bib/bbw008; pmid: 26891983
45. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). doi: 10.1186/s13059-014-0550-8; pmid: 25516281
46. N. I. Weisenfeld, V. Kumar, P. Shah, D. M. Church, D. B. Jaffe, Direct determination of diploid genome sequences. *Genome Res.* **27**, 757–767 (2017). doi: 10.1101/gr.214874.116; pmid: 28381613
47. R. Luo *et al*., SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1**, 18 (2012). doi: 10.1186/2047-217X-1-18; pmid: 23587118
48. E. Birney, M. Clamp, R. Durbin, GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004). doi: 10.1101/gr.1865504; pmid: 15123596
49. S. M. Kiełbasa, R. Wan, K. Sato, P. Horton, M. C. Frith, Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011). doi: 10.1101/gr.113985.110; pmid: 21209072
50. M. Blanchette *et al*., Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14**, 708–715 (2004). doi: 10.1101/gr.1933104; pmid: 15060014
51. Z. Yang, PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997). doi: 10.1093/bioinformatics/13.5.555; pmid: 9367129
52. Y. Zhou, Y. Liang, K. H. Lynch, J. J. Dennis, D. S. Wishart, PHAST: A fast phage search tool. *Nucleic Acids Res.* **39** (suppl.), W347–W352 (2011). doi: 10.1093/nar/gkr485; pmid: 21672955

ACKNOWLEDGMENTS

We thank the High-Performance Computing Center (HPC) of Northwest A&F University (NWAFU) for providing computing resources. **Funding:** This project was supported by the National Natural Science Foundation of China (31822052, 31572381) and National Thousand Youth Talents Plan to Y.J.; Talents Team Construction Fund of Northwestern Polytechnical University (NWPU) to W.W. and Q.Q.; the National Program for Support of Top-notch Young Professionals to Q.Q.; Strategic Priority Research Program of CAS to W.W. (XDB13000000) and C.C. (XDA16010405); 1000 Talent Project of Shaanxi Province to W.W. and Q.Q.; Villum Foundation Young Investigator grant (VKR023447) to R.H.; and Natural Science Foundation of China (31501984), Natural Science Foundation of Jilin Province (20170101158JC), and Central Public-Interest Scientific Institution Basal Research Fund (Y2019GH13) to Z.L. **Author contributions:** Q.Q., W.W., and Y.J. conceived of the project and designed research; Y.W., C.Z., N.W., Y.Z., and J.H. performed the majority of analysis with contributions from Z.L., R.H., X.P.,

Z.Z., J.L., M.L., K.W., C.L., Z.Lin, L.C., and C.S.; Z.L., C.Zheng, J.W., S.H., C.Li, L.J., G.L., and M.Liu prepared the samples for transcriptome and genome sequencing; R.L., X.C., F.H., X.D., C.C., M.D., S.P., W.Z., C.Zhao, and H.Y. took part in the cancer gene analysis. Y.W. drafted the manuscript with input from all authors, and Q.Q., W.W., Y.J., R.H., Z.L., G.Z., and T.S.S. revised the manuscript. **Competing interests:** A provisional Chinese patent application on potential application in the treatment and prevention of cancer by way of the deer *PML* gene has been filed by Northwest A&F University (application number 201910266652.2), where Y.W., Q.Q., Y.J., W.W., Z.L., and R.L. are listed as inventors. All authors declare that they have no other competing interests. **Data and materials availability:** All the raw reads of transcriptomes have been deposited in the NCBI under project number PRJNA438286 (the detailed SRA numbers are provided in table S23). The assemblies for the Chinese water deer (*Hydropotes inermis*) and two Moschidae species have been deposited in the NCBI under project number PRJNA438286.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/eaav6335/suppl/DC1
Materials and Methods
Figs. S1 to S18
Tables S1 to S26
References (*53*–*95*)

8 October 2018; accepted 16 May 2019
10.1126/science.aav6335

RESEARCH ARTICLE SUMMARY

RUMINANT GENOMICS

Biological adaptations in the Arctic cervid, the reindeer (*Rangifer tarandus*)

Zeshan Lin*, Lei Chen*, Xianqing Chen*, Yingbin Zhong*, Yue Yang*, Wenhao Xia*, Chang Liu, Wenbo Zhu, Han Wang, Biyao Yan, Yifeng Yang, Xing Liu, Kjersti Sternang Kvie, Knut Håkon Røed, Kun Wang, Wuhan Xiao, Haijun Wei, Guangyu Li, Rasmus Heller, M. Thomas P. Gilbert, Qiang Qiu†, Wen Wang†, Zhipeng Li†

INTRODUCTION: Reindeer (*Rangifer tarandus*) are naturally distributed across the Arctic and subarctic regions. Consequently, these animals have evolved to face numerous challenges, including exposure to severe cold, limited food availability in winter, and extremely prolonged light or dark periods. Unlike all other cervid species, both male and female reindeer annually grow deciduous antlers. Furthermore, reindeer are the only fully domesticated species among the Cervidae. However, little is known about the underlying genetic causes of these traits.

RATIONALE: We performed comparative genomic analyses between reindeer, other ruminant species, and a number of mammalian outgroups to identify rapidly evolving genes, positively selected genes, and reindeer-specific mutants. We further resequenced the genomes of three domestic reindeer from northern China and three wild reindeer from Northern Europe to validate that the reindeer-specific mutations are fixed in the species rather than individual polymorphisms. To support our computationally derived insights, we subsequently conducted in vitro functional experiments to investigate possible functional consequences of some of the reindeer-specific mutated genes.

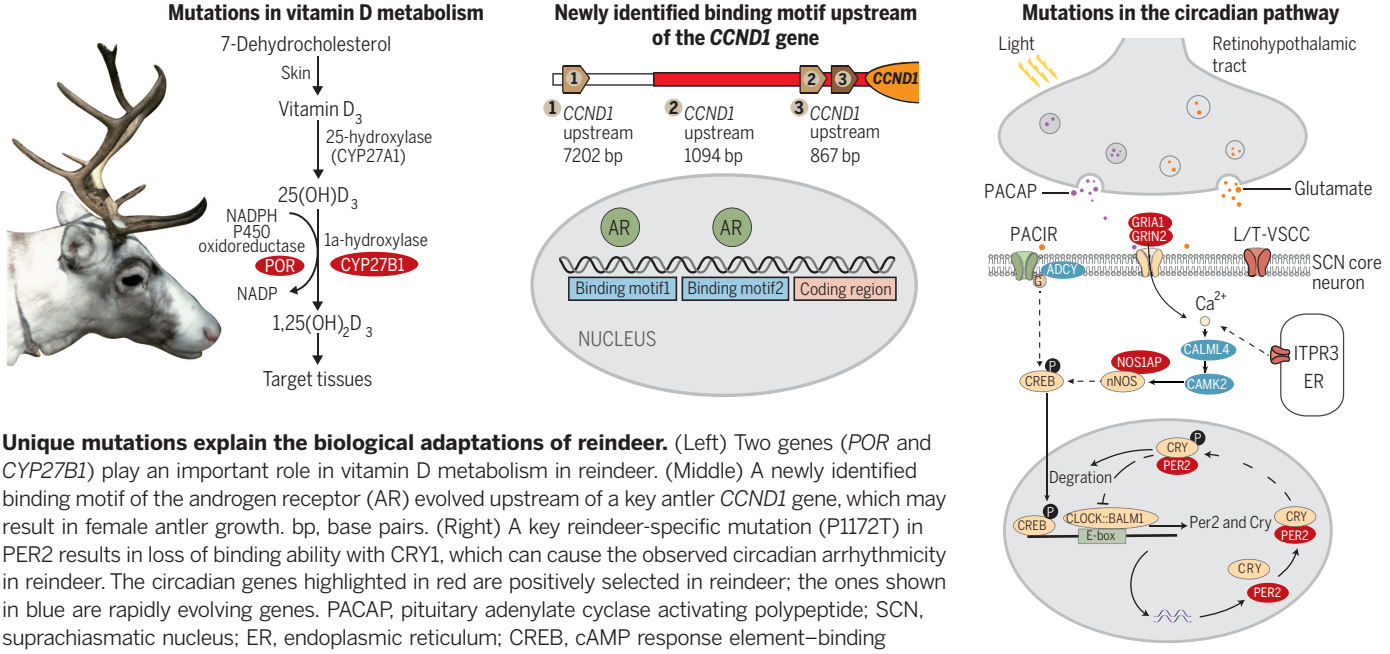
RESULTS: We found two genes (*CYP27B1* and *POR*) involved in the vitamin D metabolism pathway to be under positive selection in reindeer. Furthermore, our functional experiments validated that the two key enzymes (CYP27B1 and POR) exhibit much higher catalytic activity than that of the orthologs in goats and roe deer. We also identified fixed reindeer-specific mutations in genes that play a role in fat metabolism, including *APOB* and *FASN*. We showed that a mutation upstream of the reindeer *CCND1* gene endows an extra functional binding motif to the androgen receptor and thus may result in female antler growth. In the circadian rhythm pathway, we observed that eight genes have reindeer-specific mutations and that four genes have been rapidly evolving. Among them, the Pro^{1172}→Thr (P1172T) mutation in the reindeer PER2 causes loss of binding ability with CRY1, which can cause arrhythmicity. Finally, we found reindeer-specific mutations in 11 genes relating to development, migration, and differentiation of neural crest cells, probably accounting for the tameness of reindeer.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aav6312

CONCLUSION: Our results reveal the genetic basis of a broad spectrum of the Arctic deer's traits and provide a basis for understanding mammalian adaptive strategies to the Arctic. Our comparative genomic studies and functional assays identify a number of genes that exhibit functionality related to circadian arrhythmicity, vitamin D metabolism, docility, and antler growth, as well as genes that are uniquely mutated and/or are under positive selection. Our results may provide insights relevant to human health, including how the genetic response of vitamin D in reindeer affects bone and fat metabolism and how genes can affect circadian arrhythmicity. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author: lizhipeng01@caas.cn (Z.L.); wwang@mail.kiz.ac.cn (W.W.); qiuqiang@lzu.edu.cn (Q.Q.)
Cite this article as Z. Lin *et al.*, *Science* **364**, eaav6312 (2019). DOI: 10.1126/science.aav6312



Unique mutations explain the biological adaptations of reindeer. (Left) Two genes (*POR* and *CYP27B1*) play an important role in vitamin D metabolism in reindeer. (Middle) A newly identified binding motif of the androgen receptor (AR) evolved upstream of a key antler *CCND1* gene, which may result in female antler growth. bp, base pairs. (Right) A key reindeer-specific mutation (P1172T) in PER2 results in loss of binding ability with CRY1, which can cause the observed circadian arrhythmicity in reindeer. The circadian genes highlighted in red are positively selected in reindeer; the ones shown in blue are rapidly evolving genes. PACAP, pituitary adenylate cyclase activating polypeptide; SCN, suprachiasmatic nucleus; ER, endoplasmic reticulum; CREB, cAMP response element–binding protein; nNOS, neuronal nitric oxide synthase; P, phosphorus; G, G protein.

PHOTO: YIFENG YANG

RESEARCH ARTICLE

RUMINANT GENOMICS

Biological adaptations in the Arctic cervid, the reindeer (*Rangifer tarandus*)

Zeshan Lin[1]*, Lei Chen[1]*, Xianqing Chen[1]*, Yingbin Zhong[2,3]*, Yue Yang[1]*, Wenhao Xia[1]*, Chang Liu[1], Wenbo Zhu[1], Han Wang[2,3], Biyao Yan[1], Yifeng Yang[4], Xing Liu[5], Kjersti Sternang Kvie[6], Knut Håkon Røed[6], Kun Wang[1], Wuhan Xiao[5], Haijun Wei[4], Guangyu Li[4], Rasmus Heller[7], M. Thomas P. Gilbert[8,9], Qiang Qiu[1]†, Wen Wang[1,10,11]†, Zhipeng Li[4]†

The reindeer is an Arctic species that exhibits distinctive biological characteristics, for which the underlying genetic basis remains largely unknown. We compared the genomes of reindeer against those of other ruminants and nonruminant mammals to reveal the genetic basis of light arrhythmicity, high vitamin D metabolic efficiency, the antler growth trait of females, and docility. We validate that two reindeer vitamin D metabolic genes (*CYP27B1* and *POR*) show signs of positive selection and exhibit higher catalytic activity than those of other ruminants. A mutation upstream of the reindeer *CCND1* gene endows an extra functional binding motif of the androgen receptor and thereby may result in female antlers. Furthermore, a mutation (proline-1172→threonine) in reindeer PER2 results in loss of binding ability with CRY1, which may explain circadian arrhythmicity in reindeer.

Reindeer (*Rangifer tarandus*) are naturally distributed across the Arctic and subarctic regions. They face challenges such as severe cold and limited food availability during the winter and prolonged periods of both light and darkness during the year (*1*) and have thus had to evolve strategies and features to address these environmental obstacles. For example, reindeer do not exhibit 24-hour activity rhythms, and evidence suggests weak or even absent circadian organization (*2*–*4*), indicating that their internal biological clock may be adapted to the extreme light periodicity of the Arctic. These animals have also developed a particular fat metabolism process, limit heat loss by peripheral vasoconstriction, and maintain a low resting metabolic rate (*1*). A further challenge is that despite having to cope with low levels of solar energy, reindeer need to meet the calcium demands of rapid antler growth, therefore implying that they must have evolved a particularly high capacity for metabolizing vitamin D (*5*). Additionally, whereas archaeological evidence indicates that humans have exploited cervid species for millennia (*6*), *R. tarandus* is the only fully domesticated species within the Cervidae family (*7*). In this regard, the emergence of extensive reindeer husbandry during the last centuries was of considerable importance for the development of human civilization in the high Arctic of Asia and Europe. Reindeer are maintained in large domestic herds (*7*) that exhibit differences in activity patterns, tameness, and mitochondrial sequences from those of their wild relatives (*7*–*9*). However, despite the combination of these features, little is currently known about their underlying genetic causes, something we aimed to address in this study.

[1]Center for Ecological and Environmental Sciences, Northwestern Polytechnical University, Xi'an 710072, China. [2]Center for Circadian Clocks, Soochow University, Suzhou 215123, China. [3]School of Biology and Basic Medical Sciences, Medical College, Soochow University, Suzhou 215123, China. [4]Department of Special Animal Nutrition and Feed Science, Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, Changchun 130112, China. [5]State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China. [6]Department of Basic Sciences and Aquatic Medicine, Norwegian University of Life Sciences, Oslo 0102, Norway. [7]Section for Computational and RNA Biology, Department of Biology, University of Copenhagen, Copenhagen N 2200, Denmark. [8]Section for Evolutionary Genomics, Department of Biology, University of Copenhagen, Copenhagen N 2200, Denmark. [9]Norwegian University of Science and Technology, University Museum, Trondheim 7491, Norway. [10]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. [11]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China.
*These authors contributed equally to this work.
†Corresponding author: lizhipeng01@caas.cn (Z.L.); wwang@mail.kiz.ac.cn (W.W.); qiuqiang@lzu.edu.cn (Q.Q.)

Reindeer genomes exhibit vitamin D–specific mutations

Although the Ruminantia is an important group of terrestrial herbivores and its members have adapted to a wide range of terrestrial habitats, *R. tarandus* is the only species within the Cervidae that is widely distributed across the Arctic and subarctic regions. Because both reindeer sexes grow large antlers while inhabiting regions that undergo extended periods of low or no solar energy, it is conceivable that this species requires efficient calcium metabolism and reabsorption. Levels of active vitamin D [1α,25-$(OH)_2D_3$] in deer blood have been associated with antler growth rate (*10*, *11*). We therefore recovered all 28 genes from the reindeer reference genome (*12*) that are involved in the vitamin D metabolism pathway and compared them to their orthologs in other ruminant species from our ruminant genome project (*13*) as well as other mammalian outgroups (*14*), including the pig (*Sus scrofa domesticus*) (*15*), horse (*Equus caballus*) (*16*), and human (*Homo sapiens*) (*17*).

We found two genes (*CYP27B1* and *POR*) to be under positive selection in reindeer [branch-site model in phylogenetic analysis by maximum likelihood (PAML), chi-square test, $P < 0.05$] (Fig. 1A, fig. S1, and tables S1 and S2) (*14*). Furthermore, six genes (*ACAT2*, *CYP51A1*, *EBP*, *CYP2R1*, *BK2*, and *TRPV5*) exhibited reindeer-specific mutations (Fig. 1A, fig. S2, and table S3), and *APOB* was identified as a rapidly evolving gene [branch model in PAML, chi-square test, $P < 0.05$, a higher K_a/K_s ratio (i.e., the ratio of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site)] (Fig. 1A and tables S4 and S5) (*14*). To explore whether these alleles were fixed in reindeer, we resequenced three domesticated reindeer from the Greater Khingan Mountains of the Inner Mongolia Autonomous Region, China, and three wild reindeer from the Snøhetta mountain area of Norway (*14*) and found that the mutant alleles were identical in all seven individuals (fig. S2). Note that all mutations described hereafter in other genes are also fixed in all seven reindeer.

CYP27B1 and POR are enzymes that play key roles in triggering the active 1α,25-$(OH)_2D_3$ from vitamin D (*18*). Our genomic comparisons of reindeer against other ruminants [including the closest cervid genome available, which is from the Siberian roe deer (*Capreolus pygargus*), GenBank Assembly accession: GCA 000751575.1] and outgroups revealed reindeer-specific mutations at three positions in both the *CYP27B1* (Fig. 1B and table S3) and *POR* genes (fig. S2 and table S3). We also conducted three-dimensional (3D) structure simulations to examine the possible effects of these mutations on the enzyme structure using Phyre2 (*19*). We found that only the K282→N (K282N) mutation in CYP27B1 was close to the P450 functional domain (Fig. 1C), whereas the other mutated amino acids are not (Fig. 1C and fig. S3). Although these mutations are neither located in nor close to known functional domains, they may also affect the catalytic activities (*20*).

To explore the functional relevance of these mutations, we synthesized reindeer, roe deer, and goat (*Capra hircus*) orthologs in vitro and tested their enzyme catalytic activities by measuring the activities in a reconstituted system consisting of the enzyme, substrate, and NADPH (the reduced form of $NADP^+$) (*14*). For CYP27B1, the reindeer protein variant exhibited significantly higher metabolic efficiency than that of goat and roe deer proteins (~2- and ~1.5-fold increase,

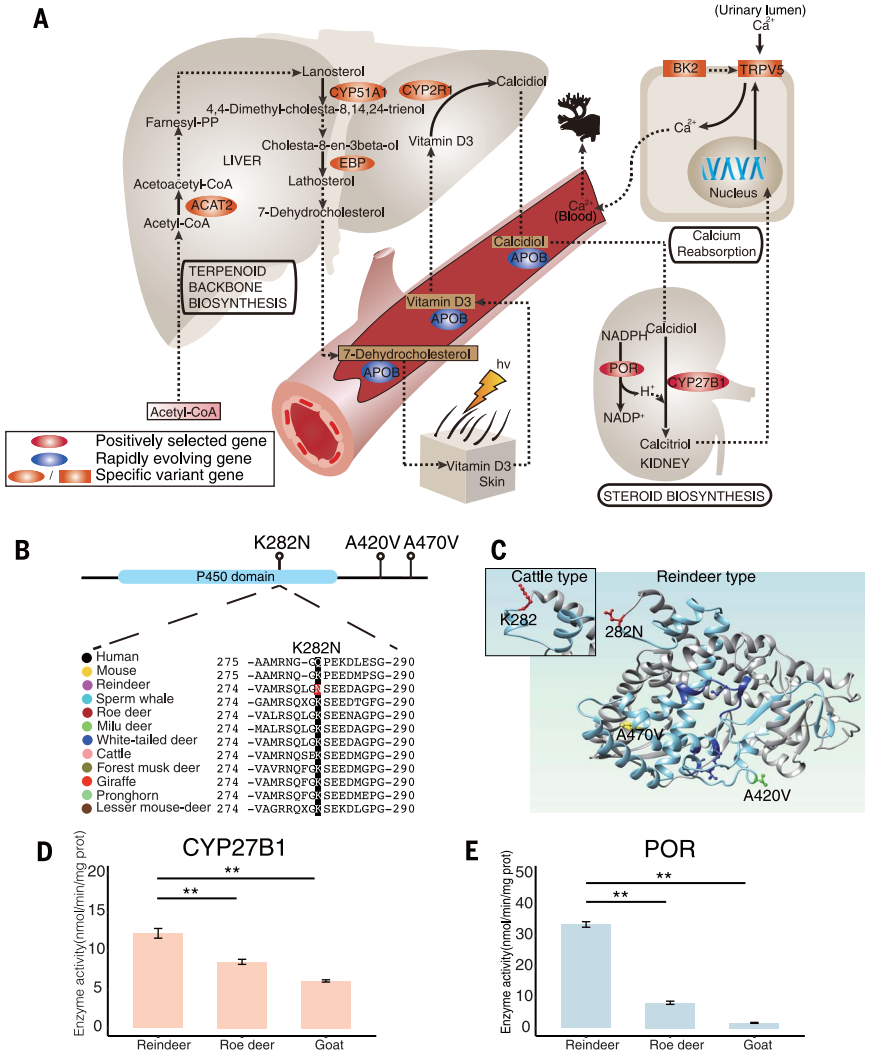


Fig. 1. Vitamin D metabolism in reindeer. (**A**) Alterations in the vitamin D metabolism pathways of reindeer. The genes involved in this pathway are labeled with different colors. Red genes are under positive selection, orange genes exhibit specific mutations in reindeer, and blue genes exhibit increased K_a/K_s values in reindeer. CoA, coenzyme A; h, Planck's constant; ν, frequency. (**B**) Specific mutations in the *CYP27B1* gene. There are three specific mutations (table S3), including one (K282N) in the P450 domain of CYP27B1. The CYP27B1 protein sequences of multiple species (indicated with different colors) were aligned, and the alignments of the K282N adjacent region are shown here. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (**C**) 3D structure simulation of reindeer CYP27B1 compared with that of cattle. The 3D structure of POR is provided in fig. S3. (**D** and **E**) Enzyme activities of reindeer CYP27B1 and POR compared with those of roe deer and goat in vitro. $^{**}P < 0.01$ calculated from the *t* test. Error bars indicate SD. prot, protein.

respectively; *t* test, $P < 0.01$) (Fig. 1D). Surprisingly, the efficiency of the reindeer POR was ~20 and ~6 times that of the goat and roe deer POR, respectively (Fig. 1E; *t* test, $P < 0.01$). These results indicate that reindeer have evolved a more efficient vitamin D metabolism pathway than that of other mammals. This change likely enables reindeer to procure the high levels of active vitamin D needed to sustain their metabolism and calcium absorption and to promote body fat oxidation required to survive in the Arctic and subarctic regions (*21*). We also noted that the Ankyrin repeat region of the calcium receptor coding gene *TRPV5*, which is activated by active vitamin D and thus affects the reabsorption of calcium, exhibited reindeer-specific mutations (fig. S2). This indicates that both the up- and downstream pathways associated with vitamin D metabolism evolved in reindeer to secure the supply of calcium and energy.

Reindeer-specific mutations related to fat metabolism

We further identified fixed reindeer-specific mutations in genes that play a role in fat metabolism. These genes include *APOB*, which participates in the transport of low-density lipoproteins, and *FASN*, which encodes fatty acid synthase, an essential enzyme for de novo lipogenesis (*22*) (fig. S2 and table S3). These genes have previously been reported as targets of positive selection in both polar bears (*Ursus maritimus*; *APOB*) and Adélie penguins (*Pygoscelis adeliae*; *FASN*) (*23*). Although the selected sites were different from that seen in polar bears and Adélie penguins, this observation suggests that fat metabolism pathways have undergone convergent evolution across homothermal polar animals via the independent modification of key fat metabolism genes in a species-specific manner.

Mutations related to female antler growth

It is believed that the growth of cervid antlers is either driven by or strongly regulated by androgens (*24*). Notably, the reindeer is the only cervid species in which females (Fig. 2A) as well as males grow antlers, and in contrast to other species of deer, the removal of the gonads of either sex soon after birth does not prevent the growth and subsequent seasonal replacement of the antlers (*5*). We therefore hypothesized that this phenomenon may be related to the increased sensitivity of some genes that regulate antler growth to low-level androgens. The characteristic 5′-TGTTCT-3′ motif has been identified as the androgen receptor binding site in mammals (*25*). We thus examined the promoter regions of 30 highly expressed genes associated with cervid antlers (fig. S4) (*26*) and identified three 5′-TGTTCT-3′ motifs upstream of the antler-specific *CCND1* gene (Fig. 2B), which regulates cell cycles (*27*) and is required for chondrocyte proliferation (*28*) and thus is closely associated with antler growth (*29*) (Fig. 2B). In particular, we noted that reindeer contain a third reindeer-specific motif. To validate its role in gene regulation, we used chromatin immunoprecipitation coupled with quantitative polymerase chain reaction (ChIP-qPCR) to determine the binding capacity of the androgen receptor to the motifs in the blood samples of five female reindeer. Samples from three male roe deer served as controls. The results show that the androgen receptor binds to the second and third motif upstream of the antler-specific *CCND1* gene in reindeer, but only to the second motif in roe deer (Fig. 2C). These findings suggest that this extra motif might enhance the expression of *CCND1* in the presence of low androgen levels, thus enabling female reindeer to grow antlers.

Circadian rhythm regulation in Arctic light conditions

Reindeer experience extended daylight fluctuations and a markedly different seasonality than do cervids from temperate, subtropical, or tropical habitats (*2*, *3*). The effect of the environment on reindeer circadian rhythmicity leads these animals to lose daily rhythmic activity in winter and summer (*2*) and exhibit an unusual internal rhythm of melatonin secretion (*2*–*4*). However,

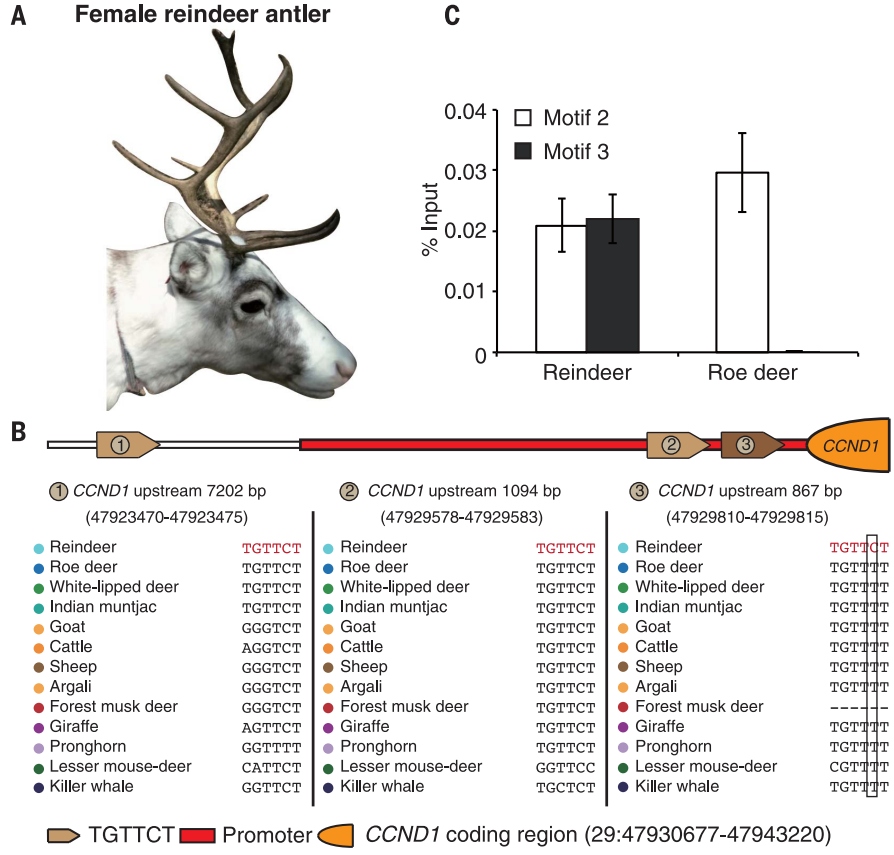


Fig. 2. Androgen receptor affinity sequence 5′-TGTTCT-3′ upstream of *CCND1*. (**A**) Characteristic antler of a female reindeer. (**B**) Three 5′-TGTTCT-3′ motifs were identified upstream of the reindeer *CCND1* gene. Motif 1 exists only in deer, whereas motif 3 exists only in reindeer. bp, base pairs. (**C**) ChIP-qPCR assays validated that the androgen receptor binds to both motif 2 and motif 3 of reindeer, but only to motif 2 of roe deer. Error bars indicate SD.

the genetic mechanism underlying this loss of a robust circadian clock remains unresolved. In other mammals, light regulates the biological clock in the suprachiasmatic nucleus (SCN) of the hypothalamus via a photic signal transduction pathway (*30*) by which light-induced neurotransmitters [PACAP (pituitary adenylate cyclase activating polypeptide) and glutamate] alter Ca^{2+} concentrations and trigger the phosphorylation of cAMP response element–binding proteins (CREBs) (Fig. 3A). We therefore retrieved 165 genes in reindeer that have been identified associated with the circadian rhythm pathway in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (*31*) and Reactome databases (*32*) and found eight genes that exhibit reindeer-specific mutations in functional domains (*PER2*, *NOCT*, *GRIA1*, *GRIN2B*, *GRIN2C*, *ITPR3*, *ADCY5*, and *NOS1AP*) (Fig. 3A and fig. S5). Moreover, among four genes (*ADCY2*, *ADCY8*, *CALML4*, and *CAMK2*) identified as rapidly evolving (table S4), reindeer-specific mutations were also observed in *ADCY8* and *CALML4* (fig. S5).

Phosphorylated CREB activates the transcription of period (*PER1*, *PER2*, and *PER3*) and cryptochrome (*CRY1* and *CRY2*) genes, which are the central elements for maintaining circadian rhythms (*33*). Three reindeer-specific mutations were found in the *Per2* gene of reindeer (fig. S5); of these, the P1172T mutation was predicted, through 3D structure simulations, to affect the binding domain between PER2 and CRY (Fig. 3B). To test the effect of this P1172T reindeer-specific mutation, T1172 was mutated to the corresponding wild-type (WT) amino acid (P), and then a coimmunoprecipitation (Co-IP) experiment was performed (*14*). Although WT PER2 (P1172, P-type) bound to CRY1, the reindeer-mutated PER2 (1172T, T-type) could not bind to CRY1 (Fig. 3C). Studies have documented that the binding of PER2 and CRY1 is required for sustained photic-induced circadian rhythms in peripheral tissues and cells, whereas CRY2 only modestly weakens rhythms (*34*). Together, these results show that the reindeer P1172T mutation caused the loss of the binding ability of PER2 with CRY1, which can lead to circadian arrhythmicity in reindeer.

Among the other seven circadian rhythm genes with reindeer-specific mutations, the *NOCT* gene serves as a key posttranscriptional regulator in the circadian control of many metabolic processes, including lipid metabolism (*35*). Notably, we observed several reindeer-specific mutations in this gene (Fig. 3D). Furthermore, among reindeer-specific mutated genes, three genes (*GRIA1*, *GRIN2*, and *ITPR3*) affect Ca^{2+} concentrations in neuron cells, with another four genes (*ADCY5*, *ADCY8*, *CAMK2*, and *NOS1AP*) affecting the phosphorylation of CREB. The *GRIA1* and *GRIN2* genes encode glutamate receptors, the activation of which is a critical step in the transmission of photic information to the SCN (*36*). The *ITPR3* gene encodes an intracellular calcium channel receptor (*37*). CREB is phosphorylated via ADCY-mediated pathways or CAMK2-, CALML4-, and NOS1AP-mediated pathways (*38*). Together, these results indicate not only that several core proteins and genes in the circadian clock are altered in reindeer but also that changes have occurred in their associated upstream and downstream pathways. We therefore hypothesize that these changes may be important components in facilitating the adaptation of reindeer to the Arctic's arrhythmic light conditions.

Mutations related to the docility of reindeer

Although most cervids have evolved sensitive and vigilant behavior to avoid predators, including humans (*39*), reindeer are quite docile. Since approximately the start of the first millennium CE, reindeer have been used by humans for transport, as decoy animals, and even for milking. This characteristic docility potentially explains why the reindeer is the most successfully domesticated cervid (*7*). It has been hypothesized that neural crest cells (NCCs) may play an important role in docility and domestication (*40*). Evidence in support of this hypothesis has been found in cats (*Felis catus*) (*41*) and dogs (*Canis lupus familiaris*) (*42*). We investigated reindeer-specific mutations in genes related to NCCs and found reindeer-specific mutations in 11 genes that relate to development of NCCs (*MSX2*, *ID3*, *BCAT1*, *CAD6*, and *CAD11*) (Fig. 4 and fig. S6), migration of NCCs (*TCOF1*, *BCAT1*, *NOTCH2*, and *NOTCH3*), and differentiation of NCCs (*COL2A1*, *KIT*, and *SI*) (Fig. 4 and figs. S6 and S7). In addition, *GEM* (*43*) and *ASCL1* (*44*) were identified as rapidly evolving genes that have been reported to be associated with the development and differentiation of NCCs, respectively. The *COL2A1* gene related to the differentiation of NCCs was also positively selected in reindeer (Fig. 4). Studies have reported mutations in *KIT* genes, which are associated with white spotting on the body (*40*, *45*), in domestic cats (*41*), horses (*46*), and pigs (*47*). Notably, white-spotted reindeer generally exhibit elevated levels of docility and can be easily approached by herders even when the rest of the herd is in a state of general excitement (*48*). Thus, we speculate that these specific NCC gene mutations may be related to the particular docility trait of reindeer.

Discussion

Our results reveal the genetic basis of a broad spectrum of Arctic deer biology and provide a foundation for understanding the adaptive strategies of mammals to the Arctic environment. Our comparative genomic studies and functional assays identify a number of genes that exhibit functionality related to circadian arrhythmicity,

PHOTO: YIFENG YANG

vitamin D metabolism, docility, and antler growth; are uniquely mutated; and/or are under positive selection in reindeer. Furthermore, these studies may provide insights relevant to human health, such as how the genetic response of vitamin D in reindeer affects bone and fat metabolism and how the genes related to circadian arrhythmicity could enhance our knowledge of seasonal affective disorders, insomnia, and depression.

Materials and methods summary

We aligned the genome sequences of 51 ruminant species (*13*) and 12 outgroup species to the goat (*C. hircus*) chromosomal genome (*49*) using lastal (version 867) (*50*). We then merged aligned sequences using Multiz (version 11.2) (*51*) and identified orthologous genes on the basis of synteny with the cattle genome (*52*), which is so far the best-annotated genome in ruminants. The whole-genome tree was obtained from our ruminant genome project (*13*), which was generated by ExaML (version 3.0.17) (*44*) and RaxML (version 8.2.10) (*53*). We used the orthologous gene sets of 11 species: white lipped deer (*Cervus albirostris*), milu deer (*Elaphurus davidianus*) (*54*), black muntjac (*Muntiacus crinifrons*), reindeer (*R. tarandus*) (*12*), white-tailed deer (*Odocoileus virginianus*), roe deer (*C. pygargus*), cattle (*Bos taurus*) (*52*), giraffe (*Giraffa camelopardalis tippelskirchi*) (*55*), pig (*S. scrofa domesticus*) (*15*), horse (*E. caballus*) (*16*), and human (*H. sapiens*) (*17*), in the following analyses. The Codeml module in the PAML software package (*56*) was used to identify the rapidly evolving genes and positively selected genes. KOBAS (version 2.0) was used to conduct the KEGG and Gene Ontology (GO) enrichment analysis (*57*). Protein sequences of selected genes were compared between reindeer and other animals to identify reindeer-specific mutations. Pfam (version 1.6) (*58*) was applied to determine whether the mutations were located in the domain region of the protein. To validate whether these identified mutations are specific for reindeer, three domestic reindeer samples from the Greater Khingan Mountains, Inner Mongolia Autonomous Region, China, and three wild reindeer samples from the Snøhetta mountain area in south-central Norway (*7*) were whole-genome resequenced using an Illumina HiSeq 2500 instrument. The high-quality reads were aligned to our previously released reindeer reference genome (*12*) using BWA-MEM (version 0.7.12) (*59*). SAMtools (version 0.1.19) (*60*) was used to sort and merge the alignment results. Reads around indels were realigned by the Genome Analysis Toolkit (GATK version 3.5) with the Realigner Target Creator and Indel Realigner programs (*61*). The protein 3D structure of interaction between PER2 and CRY1 was downloaded from the Protein Data Bank (*62*). The

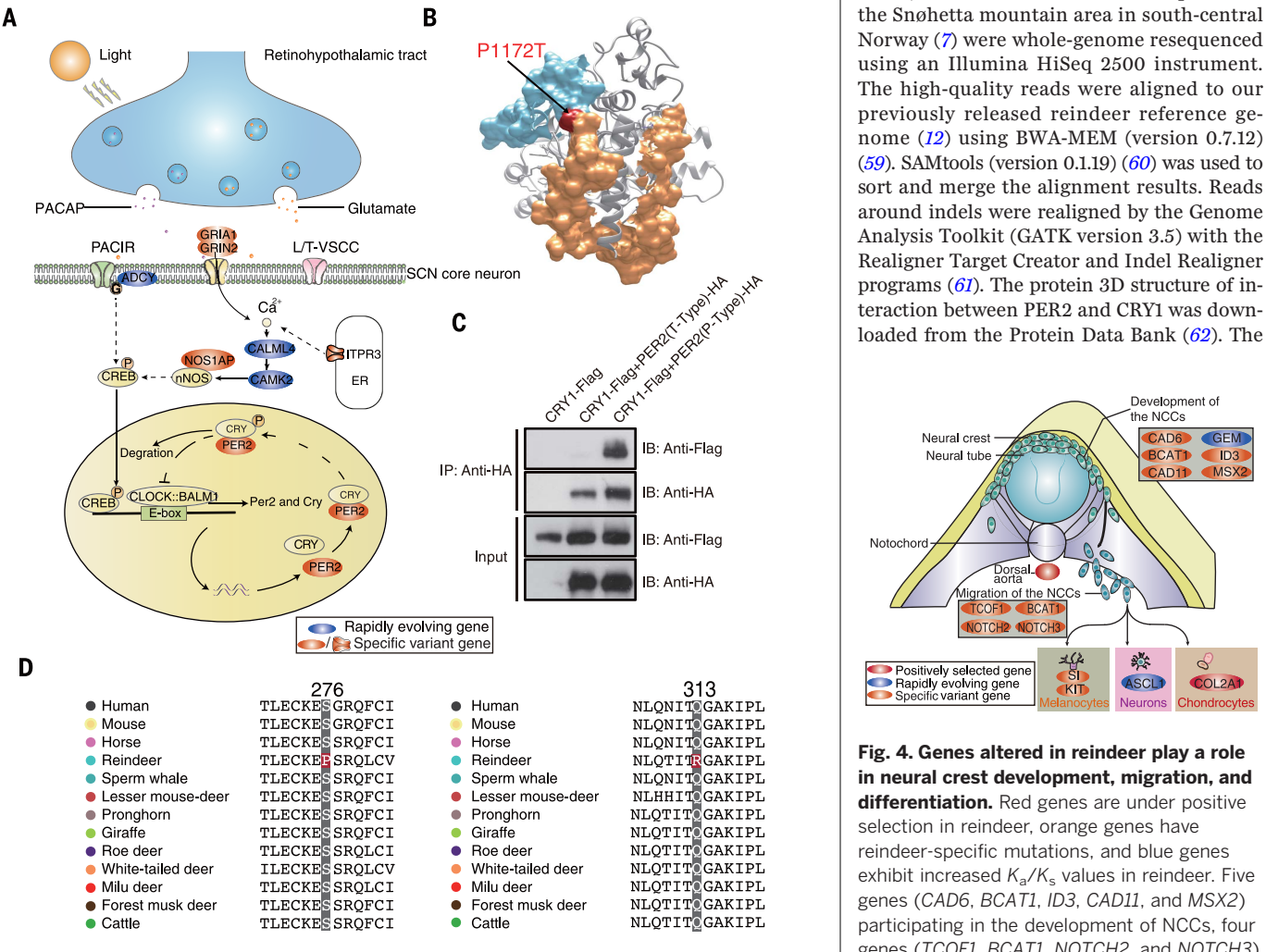


Fig. 3. Mutations in the reindeer circadian rhythm genes and pathways. Orange genes have reindeer-specific mutations; blue genes exhibit increased K_a/K_s values in reindeer. (**A**) Light regulates the molecular clockwork in reindeer SCN neurons. G, G protein; P, phosphorus; nNOS, neuronal nitric oxide synthase; ER, endoplasmic reticulum. (**B**) The reindeer-specific mutation (P1172T) of PER2 is located within the binding domain of the PER2-CRY complex, as determined by 3D modeling. The site of this mutation is marked in red. (**C**) Co-IP assays show that reverted PER2-T1172P (P-type) proteins can bind to CRY1, but reindeer PER2-P1172T (T-type) cannot bind. CRY1-Flag was transfected with or without PER2–hemagglutinin (HA) or PER2-T1172P, and Co-IP was conducted. IB, immunoblot. (**D**) The *NOCT* gene functional domain has specific mutations (fig. S5). The reindeer *NOCT* gene has two specific mutations (S276P and Q313R) in the C-terminal deadenylase domain. The amino acid substitution in reindeer is indicated in red.

Fig. 4. Genes altered in reindeer play a role in neural crest development, migration, and differentiation. Red genes are under positive selection in reindeer, orange genes have reindeer-specific mutations, and blue genes exhibit increased K_a/K_s values in reindeer. Five genes (*CAD6*, *BCAT1*, *ID3*, *CAD11*, and *MSX2*) participating in the development of NCCs, four genes (*TCOF1*, *BCAT1*, *NOTCH2*, and *NOTCH3*) involved in the migration of NCCs, and two genes (*SI* and *KIT*) related to the differentiation of NCCs into melanocytes were found to have specific mutations in reindeer. In addition, one gene (*ASCL1*) related to the differentiation of NCCs into neurons and one gene (*GEM*) involved in the development of NCCs showed increased K_a/K_s values in reindeer. One gene (*COL2A1*) related to the differentiation of NCCs into chondrocytes was under positive selection in reindeer.

3D structures of the CYP27B1 and POR proteins were predicated using Phyre2 (*19*) and then visualized using UCSF Chimera (*63*). The *CYP27B1* and *POR* genes were synthesized and then cloned into the vector pET-28a by Gibson assembly methods (*64*). The plasmid was transformed into *Escherichia coli* BL21 to express these two genes. The cultured cells were purified and then used to measure enzyme activity as described by Akiyoshi-Shibata *et al.* (*65*). The reindeer *PER2* (P1172T), reindeer *CRY1*, and reverted *PER2* (T1172P) genes were synthesized and then ligated into vector pET-28a by Gibson assembly methods (*64*). Co-IP and immunoblotting were performed as described previously (*66*) to test the interaction between PER2 and CRY1. ChIP-qPCR was applied to the blood samples of five female reindeer and three male roe deer to confirm that androgen receptor binds to the reindeer motif 3 with the reindeer-specific mutation (for details, see the full materials and methods in the supplementary materials) (*14*).

REFERENCES AND NOTES

1. A. S. Blix, Adaptations to polar life in mammals and birds. *J. Exp. Biol.* **219**, 1093–1105 (2016). doi: 10.1242/jeb.120477; pmid: 27103673
2. B. E. H. van Oort *et al.*, Circadian organization in reindeer. *Nature* **438**, 1095–1096 (2005). doi: 10.1038/4381095a; pmid: 16371996
3. W. Lu, Q. J. Meng, N. J. C. Tyler, K.-A. Stokkan, A. S. I. Loudon, A circadian clock is not required in an Arctic mammal. *Curr. Biol.* **20**, 533–537 (2010). doi: 10.1016/j.cub.2010.01.042; pmid: 20226667
4. K.-A. Stokkan, B. E. H. van Oort, N. J. C. Tyler, A. S. I. Loudon, Adaptations for life in the Arctic: Evidence that melatonin rhythms in reindeer are not driven by a circadian oscillator but remain acutely sensitive to environmental photoperiod. *J. Pineal Res.* **43**, 289–293 (2007). doi: 10.1111/j.1600-079X.2007.00476.x; pmid: 17803527
5. R. D. Brown, "Antler growth in male and female reindeer calves occurs in the absence of the gonads" in *The Biology of Deer* (Springer, 1992), pp. 493–498.
6. J. Clutton-Brock, *A Natural History of Domesticated Mammals* (Univ. of Texas Press, 1987).
7. K. H. Røed *et al.*, Genetic analyses reveal independent domestication origins of Eurasian reindeer. *Proc. R. Soc. B* **275**, 1849–1855 (2008). doi: 10.1098/rspb.2008.0332; pmid: 18460427
8. G. Bjørnstad, Ø. Flagstad, A. K. Hufthammer, K. H. Røed, Ancient DNA reveals a major genetic change during the transition from hunting economy to reindeer husbandry in northern Scandinavia. *J. Archaeol. Sci.* **39**, 102–108 (2012). doi: 10.1016/j.jas.2011.09.006
9. E. Reimers, K. H. Røed, J. E. Colman, Persistence of vigilance and flight response behaviour in wild reindeer with varying domestic ancestry. *J. Evol. Biol.* **25**, 1543–1554 (2012). doi: 10.1111/j.1420-9101.2012.02538.x; pmid: 22587024
10. K. L. Van der Eems, R. D. Brown, C. M. Gundberg, Circulating levels of 1,25 dihydroxyvitamin D, alkaline phosphatase, hydroxyproline, and osteocalcin associated with antler growth in white-tailed deer. *Acta Endocrinol.* **118**, 407–414 (1988). doi: 10.1530/acta.0.1180407; pmid: 3260707
11. A. J. Sempere, R. Grimberg, C. Silve, C. Tau, M. Garabedian, Evidence for extrarenal production of 1,25-dihydroxyvitamin during physiological bone growth: In vivo and in vitro production by deer antler cells. *Endocrinology* **125**, 2312–2319 (1989). doi: 10.1210/endo-125-5-2312; pmid: 2791992
12. Z. Li *et al.*, Draft genome of the reindeer (*Rangifer tarandus*). *Gigascience* **6**, 1–5 (2017). doi: 10.1093/gigascience/gix102; pmid: 29099922
13. L. Chen *et al.*, Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science* **364**, eaav6202 (2019). doi: 10.1126/science.aav6202
14. Materials and methods are available as supplementary materials.
15. M. A. M. Groenen *et al.*, Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491**, 393–398 (2012). doi: 10.1038/nature11622; pmid: 23151582
16. C. M. Wade *et al.*, Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* **326**, 865–867 (2009). doi: 10.1126/science.1178158; pmid: 19892987
17. J. C. Venter *et al.*, The sequence of the human genome. *Science* **291**, 1304–1351 (2001). doi: 10.1126/science.1058040; pmid: 11181995
18. D. D. Bikle, Vitamin D metabolism, mechanism of action, and clinical applications. *Chem. Biol.* **21**, 319–329 (2014). doi: 10.1016/j.chembiol.2013.12.016; pmid: 24529992
19. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015). doi: 10.1038/nprot.2015.053; pmid: 25950237
20. N. Huang, V. Agrawal, K. M. Giacomini, W. L. Miller, Genetics of P450 oxidoreductase: Sequence variation in 842 individuals of four ethnicities and activities of 15 missense mutations. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1733–1738 (2008). doi: 10.1073/pnas.0711621105; pmid: 18230729
21. M. J. Soares, L. L. Murhadi, A. V. Kurpad, W. L. Chan She Ping-Delfos, L. S. Piers, Mechanistic roles for calcium and vitamin D in the regulation of body weight. *Obes. Rev.* **13**, 592–605 (2012). doi: 10.1111/j.1467-789X.2012.00986.x; pmid: 22385576
22. T. Maier, S. Jenni, N. Ban, Architecture of mammalian fatty acid synthase at 4.5 Å resolution. *Science* **311**, 1258–1262 (2006). doi: 10.1126/science.1123248; pmid: 16513975
23. C. Li *et al.*, Two Antarctic penguin genomes reveal insights into their evolutionary history and molecular changes related to the Antarctic environment. *Gigascience* **3**, 27 (2014). doi: 10.1186/2047-217X-3-27; pmid: 25671092
24. J. Price, C. Faucheux, S. Allen, in *Curent. Topics in Developmental Biology*, G. P. Schatten, Ed. (Academic Press, 2005), vol. 67, pp. 1–48.
25. F. Claessens *et al.*, Selective DNA binding by the androgen receptor as a mechanism for hormone-specific gene regulation. *J. Steroid Biochem. Mol. Biol.* **76**, 23–30 (2001). doi: 10.1016/S0960-0760(00)00154-0; pmid: 11384860
26. Y. Wang *et al.*, Genetic basis of ruminant headgear and rapid antler regeneration. *Science* **364**, eaav6335 (2019). doi: 10.1126/science.aav6335
27. J. LaBaer *et al.*, New functional activities for the p21 family of CDK inhibitors. *Genes Dev.* **11**, 847–862 (1997). doi: 10.1101/gad.11.7.847; pmid: 9106657
28. M. Zhang *et al.*, PTHrP prevents chondrocyte premature hypertrophy by inducing cyclin-D1-dependent Runx2 and Runx3 phosphorylation, ubiquitylation and proteasomal degradation. *J. Cell Sci.* **122**, 1382–1389 (2009). doi: 10.1242/jcs.040709; pmid: 19351720
29. J. Price, S. Allen, Exploring the mechanisms regulating regeneration of deer antlers. *Philos. Trans. R. Soc. London Ser. B* **359**, 809–822 (2004). doi: 10.1098/rstb.2004.1471; pmid: 15293809
30. C. S. Colwell, Linking neural activity and molecular oscillations in the SCN. *Nat. Rev. Neurosci.* **12**, 553–569 (2011). doi: 10.1038/nrn3086; pmid: 21886186
31. H. Ogata *et al.*, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **27**, 29–34 (1999). doi: 10.1093/nar/27.1.29; pmid: 9847135
32. D. Croft *et al.*, Reactome: A database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**, D691–D697 (2011). doi: 10.1093/nar/gkq1018; pmid: 21067998
33. Z. Travnickova-Bendova, N. Cermakian, S. M. Reppert, P. Sassone-Corsi, Bimodal regulation of mPeriod promoters by CREB-dependent signaling and CLOCK/BMAL1 activity. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7728–7733 (2002). doi: 10.1073/pnas.102075599; pmid: 12032351
34. M. Ukai-Tadenuma *et al.*, Delay in feedback repression by cryptochrome 1 is required for circadian clock function. *Cell* **144**, 268–281 (2011). doi: 10.1016/j.cell.2010.12.019; pmid: 21236481
35. J. J. Stubblefield, J. Terrien, C. B. Green, Nocturnin: At the crossroads of clocks and metabolism. *Trends Endocrinol. Metab.* **23**, 326–333 (2012). doi: 10.1016/j.tem.2012.03.007; pmid: 22608110
36. D. A. Golombek, P. V. Agostino, S. A. Plano, G. A. Ferreyra, Signaling in the mammalian circadian clock: The NO/cGMP pathway. *Neurochem. Int.* **45**, 929–936 (2004). doi: 10.1016/j.neuint.2004.03.023; pmid: 15312987
37. T. Hamada *et al.*, The role of inositol trisphosphate-induced Ca^{2+} release from IP3-receptor in the rat suprachiasmatic nucleus on circadian entrainment mechanism. *Neurosci. Lett.* **263**, 125–128 (1999). doi: 10.1016/S0304-3940(99)00111-1; pmid: 10213151
38. D. A. Golombek, R. E. Rosenstein, Physiology of circadian entrainment. *Physiol. Rev.* **90**, 1063–1102 (2010). doi: 10.1152/physrev.00009.2009; pmid: 20664079
39. S. Lingle, W. F. Wilson, Detection and avoidance of predators in white-tailed deer (*Odocoileus virginianus*) and mule deer (*O. hemionus*). *Ethology* **107**, 125–147 (2001). doi: 10.1046/j.1439-0310.2001.00647.x
40. A. S. Wilkins, R. W. Wrangham, W. T. Fitch, The "domestication syndrome" in mammals: A unified explanation based on neural crest cell behavior and genetics. *Genetics* **197**, 795–808 (2014). doi: 10.1534/genetics.114.165423; pmid: 25024034
41. M. J. Montague *et al.*, Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17230–17235 (2014). doi: 10.1073/pnas.1410083111; pmid: 25385592
42. A. L. Pendleton *et al.*, Comparison of village dog and wolf genomes highlights the role of the neural crest in dog domestication. *BMC Biol.* **16**, 64 (2018). doi: 10.1186/s12915-018-0535-2; pmid: 29950181
43. C. Konstantinidou, S. Taraviras, V. Pachnis, Geminin prevents DNA damage in vagal neural crest cells to ensure normal enteric neurogenesis. *BMC Biol.* **14**, 94 (2016). doi: 10.1186/s12915-016-0314-x; pmid: 27776507
44. A. M. Kozlov, A. J. Aberer, A. Stamatakis, ExaML version 3: A tool for phylogenomic analyses on supercomputers. *Bioinformatics* **31**, 2577–2579 (2015). doi: 10.1093/bioinformatics/btv184; pmid: 25819675
45. M. Simões-Costa, M. E. Bronner, Establishing neural crest identity: A gene regulatory recipe. *Development* **142**, 242–257 (2015). doi: 10.1242/dev.105445; pmid: 25564621
46. B. Haase *et al.*, Seven novel *KIT* mutations in horses with white coat colour phenotypes. *Anim. Genet.* **40**, 623–629 (2009). doi: 10.1111/j.1365-2052.2009.01893.x; pmid: 19456317
47. C.-J. Rubin *et al.*, Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19529–19536 (2012). doi: 10.1073/pnas.1217149109; pmid: 23151514
48. H. Hemmer, N. Beckhaus, *Domestication: The Decline of Environmental Appreciation* (Cambridge Univ. Press, 1990).
49. D. M. Bickhart *et al.*, Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. *Nat. Genet.* **49**, 643–650 (2017). doi: 10.1038/ng.3802; pmid: 28263316
50. S. M. Kielbasa, R. Wan, K. Sato, P. Horton, M. C. Frith, Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011). doi: 10.1101/gr.113985.110; pmid: 21209072
51. M. Blanchette *et al.*, Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14**, 708–715 (2004). doi: 10.1101/gr.1933104; pmid: 15060014
52. C. G. Elsik *et al.*, The genome sequence of taurine cattle: A window to ruminant biology and evolution. *Science* **324**, 522–528 (2009). doi: 10.1126/science.1169588; pmid: 19390049
53. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014). doi: 10.1093/bioinformatics/btu033; pmid: 24451623
54. C. Zhang *et al.*, Draft genome of the milu (*Elaphurus davidianus*). *Gigascience* **7**, 1–6 (2018). doi: 10.1093/gigascience/gix130; pmid: 29267854
55. M. Agaba *et al.*, Giraffe genome sequence reveals clues to its unique morphology and physiology. *Nat. Commun.* **7**, 11519 (2016). doi: 10.1038/ncomms11519; pmid: 27187213
56. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007). doi: 10.1093/molbev/msm088; pmid: 17483113
57. C. Xie *et al.*, KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **39** (suppl. 2), W316–W322 (2011). doi: 10.1093/nar/gkr483; pmid: 21715386
58. R. D. Finn *et al.*, The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016). doi: 10.1093/nar/gkv1344; pmid: 26673716
59. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi: 10.1093/bioinformatics/btp324; pmid: 19451168
60. H. Li *et al.*, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009). doi: 10.1093/bioinformatics/btp352; pmid: 19505943
61. M. A. DePristo *et al.*, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011). doi: 10.1038/ng.806; pmid: 21478889

62. J. L. Sussman *et al*., The protein data bank. Bridging the gap between the sequence and 3D structure world. *Genetica* **106**, 149–158 (1999). pmid: 10710721
63. E. F. Pettersen *et al*., UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi: 10.1002/jcc.20084; pmid: 15264254
64. D. G. Gibson *et al*., Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009). doi: 10.1038/nmeth.1318; pmid: 19363495
65. M. Akiyoshi-Shibata *et al*., Further oxidation of hydroxycalcidiol by calcidiol 24-hydroxylase. A study with the mature enzyme expressed in Escherichia coli. *Eur. J. Biochem.* **224**, 335–343 (1994). doi: 10.1111/j.1432-1033.1994.00335.x; pmid: 7925346
66. Y. Zhong, Q. Ye, C. Chen, M. Wang, H. Wang, Ezh2 promotes clock function and hematopoiesis independent of histone methyltransferase activity in zebrafish. *Nucleic Acids Res.* **46**, 3382–3399 (2018). doi: 10.1093/nar/gky101; pmid: 29447387

ACKNOWLEDGMENTS

We thank the reindeer farm of Aoluguya town, Genhe city, Inner Mongolia Autonomous Region, China, for the sample collection in China. **Funding:** This study was supported by the Natural science foundation from Jilin province (20170101158JC), Central Public-interest Scientific Institution Basal Research Fund (Y2019GH13), National Key R&D Program of China (2018YFC1706600), and National Natural Science Foundation of China (31501984) to Z.Li; the NordForsk's Nordic Centre of Excellence (76915) to K.S.K. and K.H.R.; the European Research Council (681396) to M.T.P.G.; the Talents Team Construction Fund of Northwestern Polytechnical University (NWPU) and the Fundamental Research Funds for the Central Universities (3102019JC007) to W.W. and Q.Q.; and the Strategic Priority Research Program of CAS (XDB13000000) and the 1000 Talent Project of Shaanxi Province to W.W. **Author contributions:** Z.Lin, L.C., C.L., B.Y., and K.W. analyzed the genome data; X.C., Y.Z., Y.Y., W.Xia, W.Z., X.L., W.Xiao, and H.W. conducted the in vitro experiment; Y.Y., H.W., K.S.K., K.H.R., and G.L. prepared the samples; Z.Lin, W.Xiao, K.S.K., K.H.R., H.W., R.H., M.T.P.G., Q.Q., W.W., and Z.Li wrote the manuscript; and Z.Li, W.W., and Q.Q. designed the study. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The resequencing data have been deposited in NCBI under accession numbers SRR7630519 to SRR7630521 and SRR8515771 to SRR8515773.

SUPPLEMENTARY MATERIALS

8 October 2018; accepted 16 May 2019
10.1126/science.aav6312

RESEARCH ARTICLE SUMMARY

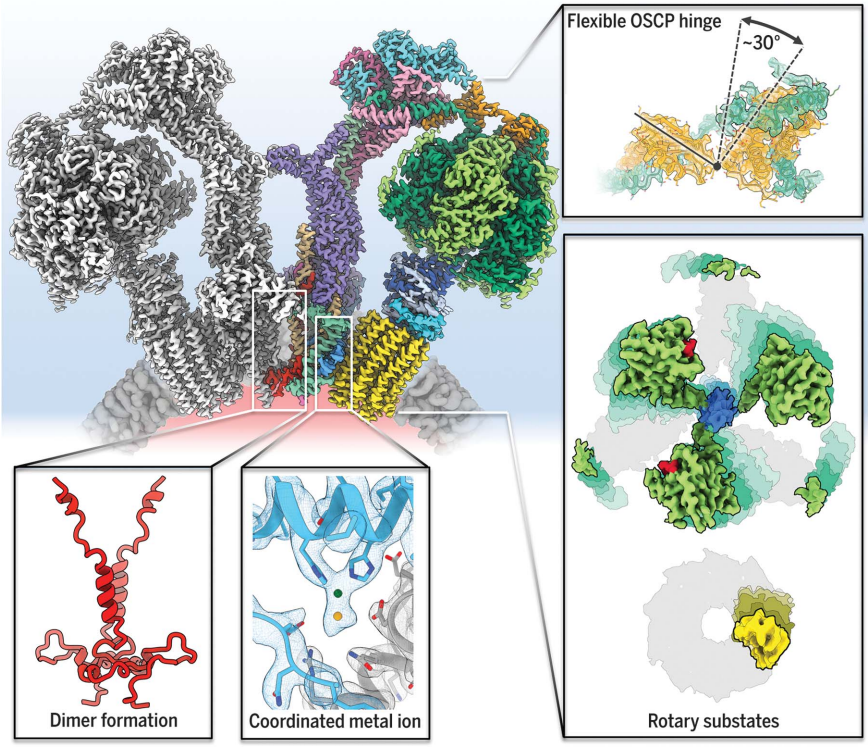
MOLECULAR MACHINES

Rotary substates of mitochondrial ATP synthase reveal the basis of flexible F_1-F_o coupling

Bonnie J. Murphy*, Niklas Klusch*, Julian Langer, Deryck J. Mills, Özkan Yildiz, Werner Kühlbrandt†

INTRODUCTION: Mitochondrial F_1-F_o adenosine triphosphate (ATP) synthases are macromolecular turbines that couple proton translocation across a membrane to ATP synthesis. Protons are translocated through the F_o subcomplex in the lipid bilayer by a rotor composed of a defined number of *c* subunits, each with a proton-binding site, to generate ring rotation. A central stalk is firmly anchored to the *c* ring and conveys rotary motion to the catalytic F_1 subcomplex in the mitochondrial matrix, where ATP is produced by rotary catalysis. A peripheral stalk connects the two subcomplexes to prevent idle rotation of F_1 with F_o.

RATIONALE: Although ATP synthase complexes have been investigated for more than 50 years, several key questions remain. An enduring question is how the stoichiometrically mismatched *c* ring in F_o (composed of 8 to 17 *c* subunits) and the three-fold symmetric F_1 head are efficiently coupled. Another open question is the exact pathway taken by protons through the membrane, which has been the least well characterized part of the mechanism.



Cryo-EM structure of the *Polytomella* ATP synthase dimer. The F_1 head (green) is linked to the c-ring rotor (yellow) by the central stalk and the peripheral stalk. Insets (beginning at top right) show the flexible OSCP hinge (orange); F_1 rotary substates with subunits β (green), γ (blue), and c (yellow); a coordinated metal ion in the proton access channel (light blue); and the dimer-forming subunit ASA10 (red).

RESULTS: We used single-particle cryo–electron microscopy (cryo-EM) to characterize the structure and dynamics of a complete and active dimeric mitochondrial ATP synthase from the chlorophyll-less unicellular alga *Polytomella* sp. Together with data obtained by genome sequencing and mass spectrometry, our 2.7- to 2.8-Å resolution map allowed us to build a full atomic model of the 1.6-MDa complex. The model includes the newly identified subunit ASA10, which interlinks the two ATP synthase monomers on the lumenal side of the membrane. Separation of 13 independent rotary states provides a detailed molecular description of the movements that accompany *c*-ring rotation. We find that the F_1 head rotates together with the central stalk and *c* ring through approximately 30°, or one *c* subunit, at the beginning of each 120° step. Flexible coupling of the F_1 head to the F_o motor is mediated primarily by a hinge at the interdomain link of the oligomycin sensitivity–conferring protein (OSCP) subunit that joins the F_1 head to the peripheral stalk. The extended two-helix bundle of the central stalk γ subunit interacts with the catch-loop region of one β subunit of the F_1 head. The resulting mechanism of flexible coupling is likely to be conserved in other F_1-F_o ATP synthases. Our results provide much-needed context to a wealth of published data indicating that OSCP is a hub of metabolic control in the cell.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaw9128

Our high-resolution map of the proton-translocating F_o complex has revealed a strong density, very likely a metal ion, ligated by two histidine residues. Recent cryo-EM studies of yeast and spinach chloroplast ATP synthase contain unannotated densities at the same position. Mutational experiments in *Escherichia coli* have shown that an equivalent residue is essential to proton translocation. By three-dimensional classification, we separated two different rotational positions of the *c* ring and showed that the coordination environment of the metal ion changes with *c*-ring position. This evidence points toward a role for the metal ion in synchronizing *c*-ring protonation with its rotation.

CONCLUSION: In ATP synthases, the F_1 catalytic head can accompany the rotor through a rotation of ~30° at the beginning of each ~120° step. This movement allows flexible coupling of F_1 and F_o. The interdomain hinge of OSCP facilitates flexible coupling and makes this subunit an apposite point for the regulation of ATP synthesis. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: werner.kuehlbrandt@biophys.mpg.de
Cite this article as B. J. Murphy *et al.*, *Science* **364**, eaaw9128 (2019). DOI: 10.1126/science.aaw9128

RESEARCH ARTICLE

MOLECULAR MACHINES

Rotary substates of mitochondrial ATP synthase reveal the basis of flexible F_1-F_o coupling

Bonnie J. Murphy[1]*, Niklas Klusch[1]*, Julian Langer[2], Deryck J. Mills[1], Özkan Yildiz[1], Werner Kühlbrandt[1]†

F_1F_o–adenosine triphosphate (ATP) synthases make the energy of the proton-motive force available for energy-consuming processes in the cell. We determined the single-particle cryo–electron microscopy structure of active dimeric ATP synthase from mitochondria of *Polytomella* sp. at a resolution of 2.7 to 2.8 angstroms. Separation of 13 well-defined rotary substates by three-dimensional classification provides a detailed picture of the molecular motions that accompany *c*-ring rotation and result in ATP synthesis. Crucially, the F_1 head rotates along with the central stalk and *c*-ring rotor for the first ~30° of each 120° primary rotary step to facilitate flexible coupling of the stoichiometrically mismatched F_1 and F_o subcomplexes. Flexibility is mediated primarily by the interdomain hinge of the conserved OSCP subunit. A conserved metal ion in the proton access channel may synchronize *c*-ring protonation with rotation.

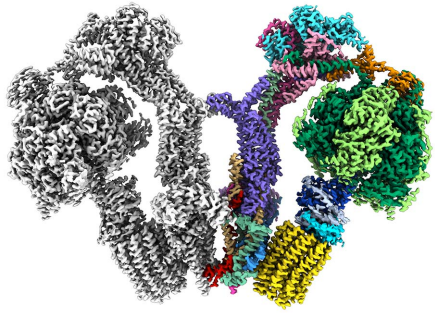
Mitochondria carry out controlled oxidation of reduced substrates to generate an electrochemical gradient across the inner mitochondrial membrane. Movement of protons down the gradient through the F_1F_o-ATP synthase drives the production of the soluble energy carrier ATP by rotary catalysis. The ATP synthases consist of two connected nanomotors: (i) the membrane-embedded F_o subcomplex, where proton translocation generates torque, and (ii) the soluble F_1 subcomplex, where the torque powers adenosine diphosphate (ADP) phosphorylation [see (*1*) for a recent review]. In F_o, a ring composed of 8 to 17 *c* subunits, each containing a carboxylate ion-binding site, rotates against a bearing formed by a near-horizontal helix hairpin of the stationary *a* subunit (*2*). Two aqueous half-channels at the interface of the *c* ring and *a* subunit allow protons to enter and leave the protonation site. Upon protonation from the lumenal channel, a *c*-ring subunit travels by an almost full rotation through the hydrophobic membrane environment before releasing the proton in the matrix channel. In the membrane, the channels are separated by a mere 6 Å (*3*), with a strictly conserved arginine residue facilitating the passage of deprotonated, but not protonated, *c*-ring subunits. The energy from down-gradient proton translocation powers rotation of the *c* ring and the firmly attached central stalk (subunits γ, δ, ε), which conveys rotation generated in F_o to power ATP synthesis. The two-helix bundle of the γ subunit extends deep into the $\alpha_3\beta_3$ hexamer of the F_1 head. Rotation of the central stalk induces conformational changes of subunits α and β that result in ADP phosphorylation. The two-domain oligomycin sensitivity–conferring protein, OSCP, connects the F_1 head to a membrane-anchored peripheral stalk stator, preventing unproductive rotation of F_1.

The dimeric ATP synthase from mitochondria of the chlorophyll-less unicellular alga *Polytomella* sp. contains the signature subunits of mitochondrial ATP synthases, α_3, β_3, γ, δ, ε, c_{10}, *a*, and OSCP. The peripheral stalk and other membrane protein subunits typical of mammalian and fungal ATP synthases (*1*) are replaced in *Polytomella* and related species, including the photosynthetic model organism *Chlamydomonas reinhardtii* (*4*), by proteins known as ATP synthase–associated (ASA) proteins, which bear no homology to other known ATP synthase components (*4*, *5*). These subunits form a bulky peripheral stalk that links the two ATP synthase complexes into a stable dimer.

ATP synthase must couple translocation of 8 to 17 protons (*1*), depending on *c*-ring size, with phosphorylation of three molecules of ADP in the catalytic β-subunit sites of the F_1 head. The stoichiometric mismatch of F_1 and F_o poses a challenge to efficient energy conversion in the complex (*6*, *7*). It has been proposed that the central stalk mediates flexible coupling between F_1 and F_o (*8*, *9*). Some authors suggest that energy is stored by coiling of the γ-subunit two-helix bundle (*10*); others propose that energy is stored in the globular portions of subunits γ and ε, where they are attached to the *c* ring (*9*). Recent cryo–electron microscopy (cryo-EM) studies resolving rotary states of mitochondrial (*11*), bacterial (*12*), and chloroplast (*13*) ATP synthase have all found that the central stalk rotates as a rigid body. Substantial flexibility in the peripheral stalk is observed between rotary states in bacterial (*12*, *14*) and chloroplast ATP synthase (*12*), which suggests that peripheral stalk bending contributes to elastic energy storage in these systems, whereas the peripheral stalk of mitochondrial ATP synthase dimers is bulkier and appears to be less flexible. The robust, rigid peripheral stalks of the *Polytomella* ATP synthase make it an ideal system for examining the dynamics revealed by rotary states and substates by high-resolution cryo-EM.

High-resolution map of a complete F-type ATP synthase dimer

Cryo-EM single-particle analysis (table S1) of purified and active (*2*) (fig. S1) *Polytomella* ATP synthase dimer yielded a map at 2.94 Å overall resolution. Local masking improved the resolution to 2.7 to 2.8 Å in overlapping regions covering the full complex (table S2 and figs. S2 and S3). Symmetry expansion allowed each monomer to be treated independently in image analysis, and three-dimensional (3D) classification of the central stalk and F_1 head enabled separation of three primary rotary states, at resolutions of 2.8 to 2.9 Å, in the F_1 head and rotor (Fig. 1 and Movie 1). The nucleotide-binding sites β_{TP} and β_{DP} contain Mg^{2+}ADP, and β_E is empty (Fig. 1, B to D). The ADP-bound complex was expected as the protein was prepared without added nucleotides. The "arginine finger" (*15*) of αArg^{429} extends toward the nucleotide phosphate in β_{DP} but away from it in β_{TP} (Fig. 1, C and D). The cryo-EM data were complemented by genomic sequencing and mass spectrometry analysis of the purified protein, enabling us to build a full atomic model of the 1.58-MDa complex, with 62 copies of 18 different subunits and a total of 14,240 residues fitted (Fig. 1, fig. S4, and table S3). The model includes the completed polypeptide sequence of subunits ASA2 and ε and the previously unknown subunit ASA10.



Movie 1. Three-dimensional map of the *Polytomella* ATP synthase dimer. One monomer is gray; the other monomer map is colored by subunit, as in Fig. 1.

[1]Department of Structural Biology, Max Planck Institute of Biophysics, Frankfurt 60438, Germany. [2]Department of Molecular Membrane Biology, Max Planck Institute of Biophysics, Frankfurt 60438, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: werner.kuehlbrandt@biophys.mpg.de

Rotary substates reveal concerted rotation of F_1 and central stalk

To describe the conformational space of the F_1F_o-ATP synthases, we sorted the symmetry-expanded dataset into 12 classes. Further subdivision of a mixed class yielded a total of 13 distinct rotary substates (fig. S2), which were refined to 2.8- to 4.2-Å resolution (fig. S3 and table S2). The subclasses can be placed into a meaningful sequence according to the rotational position of the rotor with respect to the stator (fig. S5). Each subclass was assigned to one of three primary rotary states, with six subclasses for state 1 (1A to 1F) (Fig. 2 and figs. S5 and S6), four for state 2 (2A to 2D), and three for state 3

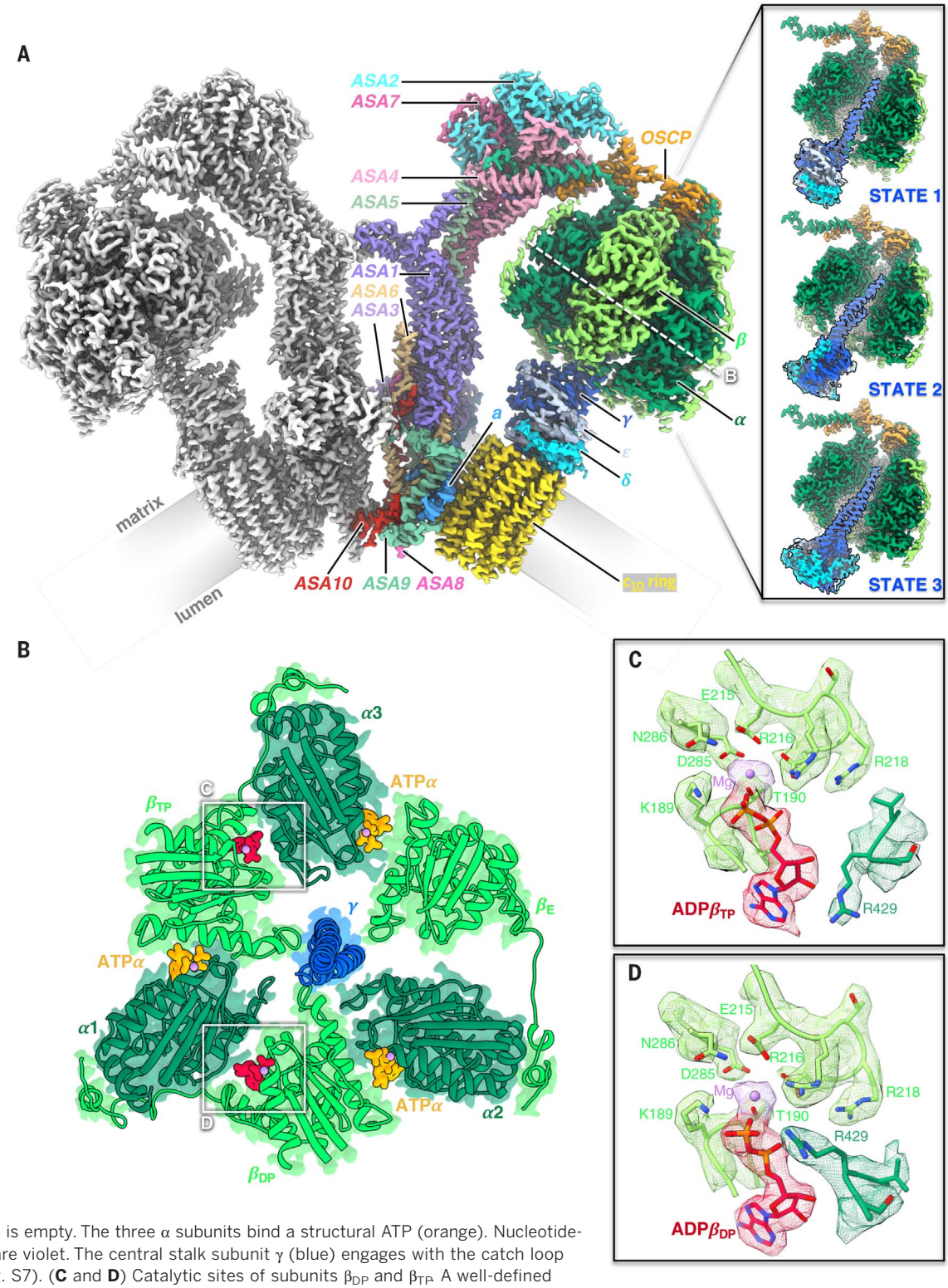


Fig. 1. High-resolution structure of the mitochondrial F_1F_o ATP synthase dimer from *Polytomella* sp. (**A**) Composite cryo-EM map of the 62-subunit, 1.58-MDa *Polytomella* dimer; the right side is colored by subunit. The c_{10} rotor ring and subunit *a* make up the F_o motor complex. Proton translocation through F_o causes rotation of the *c* ring and the attached central stalk subunits γ, δ, and ε. The F_1 head consists of three catalytic β subunits (light green) and three α subunits (dark green). Long C-terminal extensions of β wrap around the α subunits on the outside of F_1 (see Movie 5). The two-domain OSCP subunit links the three α subunits to the peripheral stalk subunits ASA1 to ASA10 (see Fig. 4). Subunit ASA10 connects the two F_1F_o monomers in the membrane (see figs. S9 and S11). The local map resolution is 2.7 Å for the peripheral stalk, F_o, and *c* ring, and 2.8 to 2.9 Å for the F_1 head and central stalk (see fig. S2). Inset: Three primary rotary states (1, 2, and 3) of the F_1F_o monomer are related by ~120° rotation of the central stalk within the $α_3β_3$ assembly. Unless otherwise specified, subunit coloring is consistent throughout all figures. (**B**) Section through F_1 at the level of bound nucleotides. $β_{DP}$ and $β_{TP}$ sites contain ADP (red) and $β_E$ is empty. The three α subunits bind a structural ATP (orange). Nucleotide-coordinating Mg^{2+} ions are violet. The central stalk subunit γ (blue) engages with the catch loop of the $β_E$ subunit (see fig. S7). (**C** and **D**) Catalytic sites of subunits $β_{DP}$ and $β_{TP}$. A well-defined Arg side chain of subunit α (the "arginine finger") extends toward the nucleotide phosphate in the $β_{DP}$ site. ATP in the $β_{TP}$ site has hydrolyzed to ADP during protein isolation. Amino acid abbreviations: D, Asp; E, Glu; K, Lys; N, Asn; R, Arg; T, Thr.

(3A to 3C). The three primary states are separated from one another by ~120° rotation of the central stalk within the F_1 head (Fig. 1A, inset). In contrast, substates of any given rotary state differ by concerted rotation of the F_1 head and rotor within the first 15° to 32° of each ~120° step (Fig. 3, Movies 2 and 3, and fig. S5). Contact between the β_E site and the γ subunit [the "β-catch loop" (*16*)] is maintained between substates (Movie 3 and fig. S7), as is the nucleotide occupancy of β subunits (Fig. 3B and fig. S5). The pivot point of this movement is the single-peptide chain connecting the two OSCP domains (henceforth the "OSCP hinge") (Fig. 4 and Movie 4). The top of the peripheral stalk flexes only slightly. Previous studies have suggested that the interdomain region of OSCP may be flexible (*17*–*19*). In comparing all 13 substates, no coiling of the central stalk is apparent (fig. S6).

In the first ~30° of rotary steps 1 and 2 (in moving from state 1A to 1F and state 2A to 2D), the *c* ring advances by almost one subunit with respect to the *a* subunit (Fig. 3A) while the position of γ with respect to the $\alpha_3\beta_3$ hexamer remains virtually unchanged (Fig. 3B and fig. S5). Between substates of state 3 (in moving from state 3A to 3C), the ring advances by about half a *c* subunit together with the rotor (Fig. 3A). As the rotor moves a further ~90° to the starting position of the next primary rotary state (e.g., from state 1F to 2A), the F_1 head recoils by ~30° to its original position, for a cumulative ~120° power stroke of γ within F_1 for this step. Given that this motion is thermally accessible for the complex at equilibrium, it would almost certainly contribute to flexible coupling of F_1 and F_o also under turnover conditions.

We observe an uneven distribution of particles between the three major rotary states, with 52% in state 1, 23% in state 2, and 21% in state 3, in line with other recent cryo-EM studies of ATP

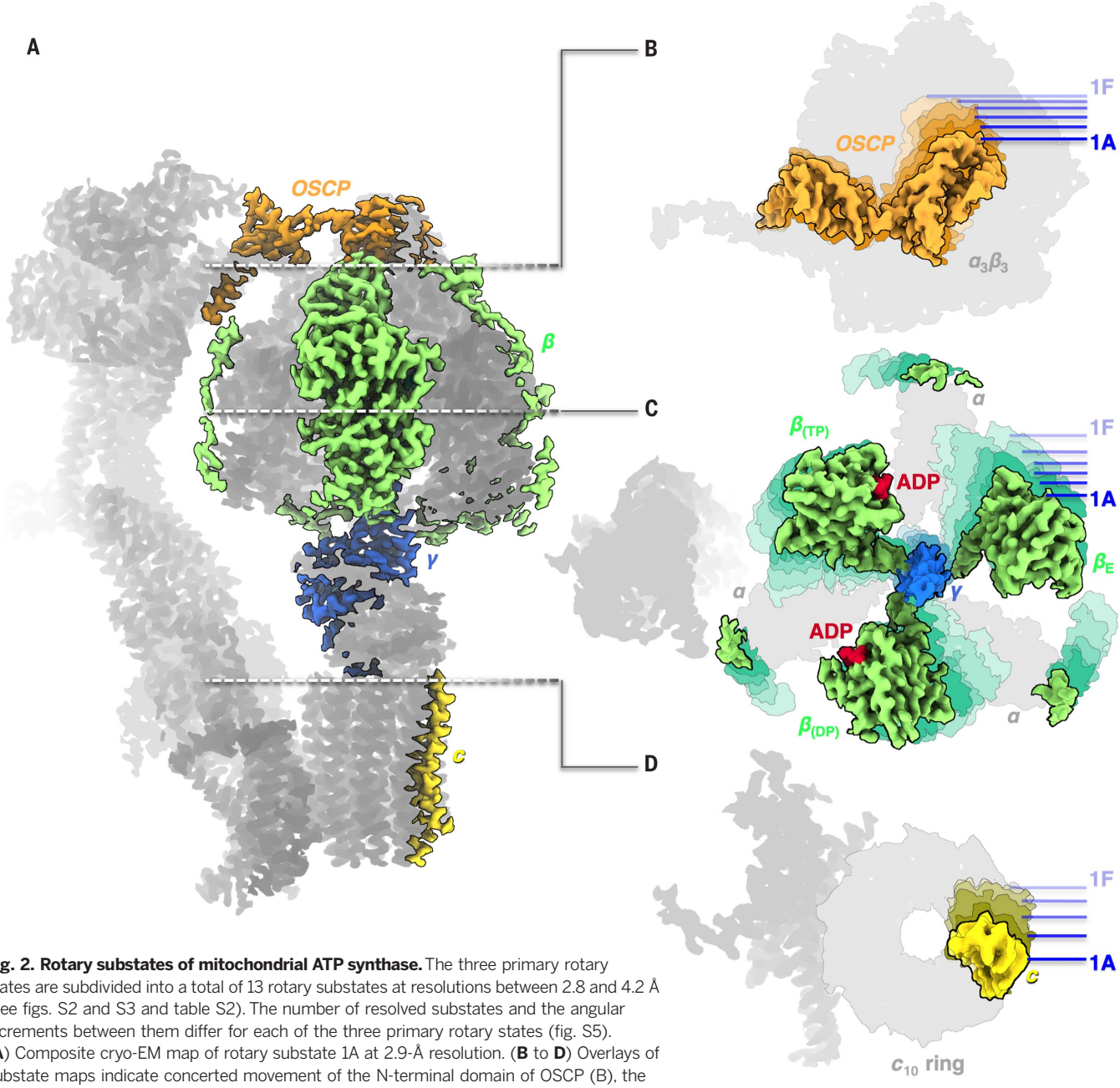


Fig. 2. Rotary substates of mitochondrial ATP synthase. The three primary rotary states are subdivided into a total of 13 rotary substates at resolutions between 2.8 and 4.2 Å (see figs. S2 and S3 and table S2). The number of resolved substates and the angular increments between them differ for each of the three primary rotary states (fig. S5). (**A**) Composite cryo-EM map of rotary substate 1A at 2.9-Å resolution. (**B** to **D**) Overlays of substate maps indicate concerted movement of the N-terminal domain of OSCP (B), the F_1 head with the central γ subunit (C), and the *c* ring (D) from substate 1A to substate 1F. Projected map densities of other subunits are shown in light gray for reference.

synthases (*11*–*13*) and with a recent single-molecule study (*20*) that reported an asymmetry in the time-averaged population of rotary states in the presence of ADP. In comparing substates of a given primary rotary state, the later substates of the *Polytomella* complex (1D, 1E, 1F, 2C, 2D, and 3C) are consistently less populated than the earlier substates (fig. S2 and table S2), which suggests that these are higher-energy states.

The interaction between F_1 and γ appears strongest in the catch loop region (*16*) of β_E, where conserved residues βAsp^{345}, βAsp^{348}, and βThr^{347} (βAsp^{302}, βAsp^{305}, and βThr^{304} in *Escherichia coli*) form ionic and hydrogen-bond interactions with γArg^{296} and γGln^{297} (*E. coli* γArg^{268} and γGln^{269}). The salt bridge between βAsp^{348} and γArg^{296} forms via a resolved water molecule (fig. S7). Throughout the 20° to 30° concerted rotation of F_1 with γ, the interaction between the β_E catch loop and γ is maintained, although subtle loop movements suggest that the hydrogen-bonding partner of βAsp^{345} changes from γGln^{297} to γAsn^{293} between substates (e.g., from state 1A to 1F; fig. S7). Previous disruption of these interactions by mutation in *E. coli* eliminated or strongly reduced ATP hydrolysis activity and the ability to grow on succinate (*16*). A proposed function of such an interaction is that, in ATP synthesis mode, it would stall rotation of γ until the β_E site has bound ADP, to prevent unproductive rotation (*16*). The major impact of catch loop mutations on both synthesis and hydrolysis activity implies that this region plays a broader role in the conformational changes required for catalysis. On the basis of our results, we suggest that this interaction does not stall rotation completely; rather, a proton may be translocated at F_o while the F_1 head waits for substrate binding.

Studies aiming to characterize flexible coupling have focused on the central stalk (*21*), often working with an F_1 and rotor subcomplex. When the whole F_1F_o complex was examined, it was typically attached to substrate and reporter molecules at subunits β and *c* to probe for flexibility along the F_1-F_o axis (*9*, *22*, *23*). A recent single-molecule study of F_1F_o reported forward and backward stepping of the rotor by ~30° (i.e., up to one *c* subunit) (*22*). However, with F_1 anchored to the support by subunit β and a probe attached

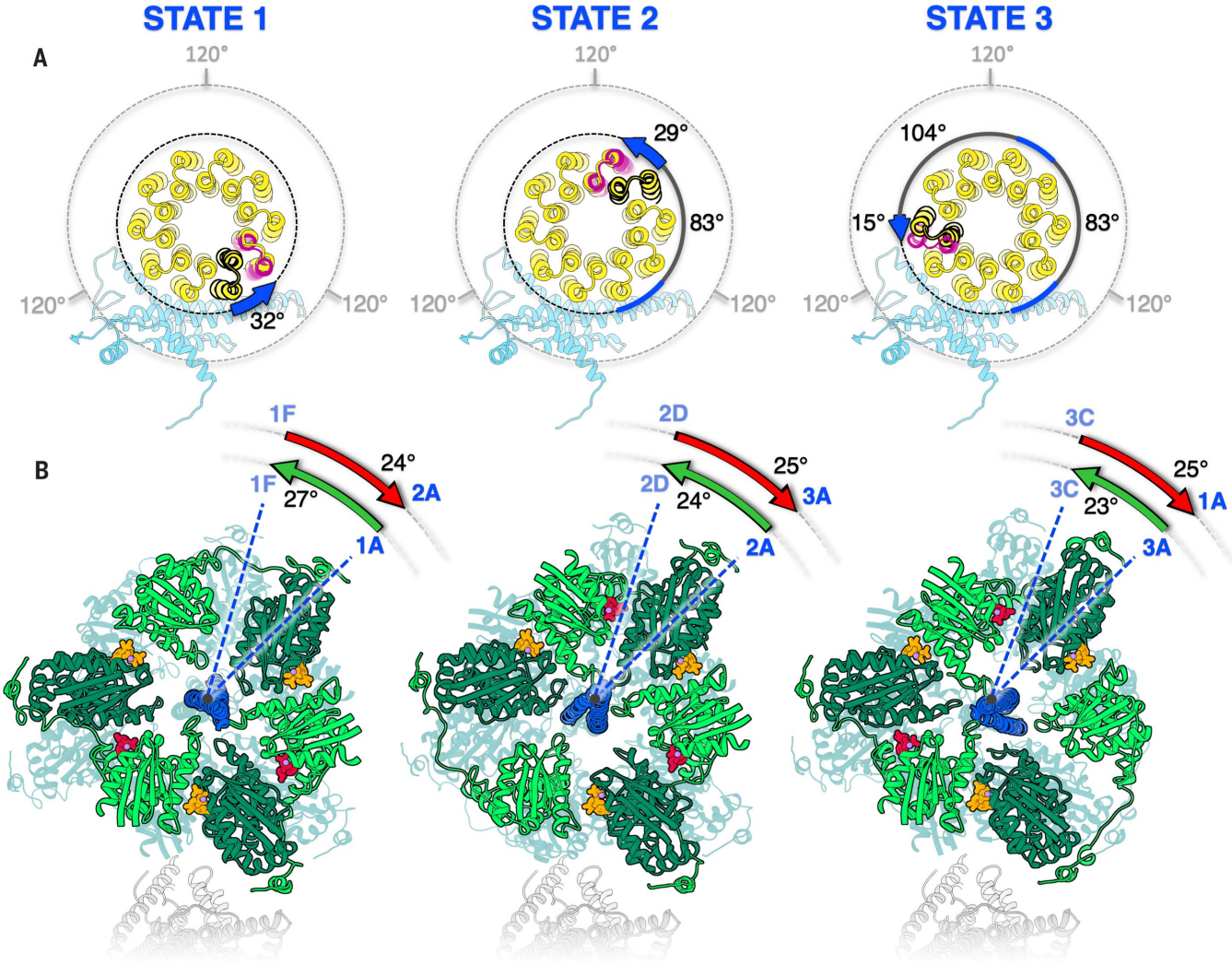


Fig. 3. Concerted rotation of F_1 with central stalk and *c* ring. (**A**) Progressing in the direction of ATP synthesis, the c_{10} ring (yellow) rotates counterclockwise (as seen from F_1) with respect to the *a* subunit (light blue) by up to 32° for substates of the same primary rotary state. The primary rotary states differ by power strokes of ~120°. The position of one *c* subunit in the first (black outline) and last rotary substate (pink) of each primary state is indicated. (**B**) Between substates of a given rotary state, the F_1 head rotates together with the *c* ring and central stalk before recoiling to its original position in the first substate of the subsequent primary rotary state. Subunits α (dark green) and β (light green) with their bound nucleotides (red and orange) are shown for the first substate in each primary state. The position of the last substate in each primary state is indicated in light blue in the background. The peripheral stalk position is shown in gray.

to *c*, flexibility at OSCP would not be detected. The observed ~30° stepping angle appears to reflect regular energy minima in the *c*-ring/*a*-subunit interaction rather than joint rotation of the *c* ring and central stalk with F_1. Single-molecule studies of flexible coupling may need to explore a wider range of anchoring points.

Several lines of evidence support the idea that OSCP-mediated flexible coupling plays a role in other F-type ATP synthases: (i) The structure and polypeptide sequences of β, γ, and OSCP subunits involved in this process are highly conserved. (ii) An OSCP subunit (or δ in chloroplasts and bacteria) is found in all three lineages of ATP synthase known to have independently acquired novel peripheral stalk components (*24*, *25*). (iii) A cryo-EM study of bovine ATP synthase (*11*) identified rotary substates of this complex, albeit at lower resolution, indicating rotation of F_1 relative to the remainder of the complex. Reevaluating these results in light of our data, they do in fact show concerted rotation of F_1 with the central stalk. (iv) Well-known inhibitors and regulators of ATP synthase bind to OSCP, affecting the catalytic rate and Michaelis constant (K_M) (*26*, *27*), suggesting a mechanistic rather than merely structural role for this subunit. The simpler bacterial and chloroplast ATP synthases have a thinner, flexible peripheral stalk relative to *Polytomella*, and previous cryo-EM studies have found substantial flexibility in the peripheral stalk even between primary rotary states (*12*–*14*). In these simpler systems, two or more components may mediate flexible F_1-F_o coupling.

A body of literature suggests that OSCP-mediated flexible coupling (Fig. 4 and Movie 4)

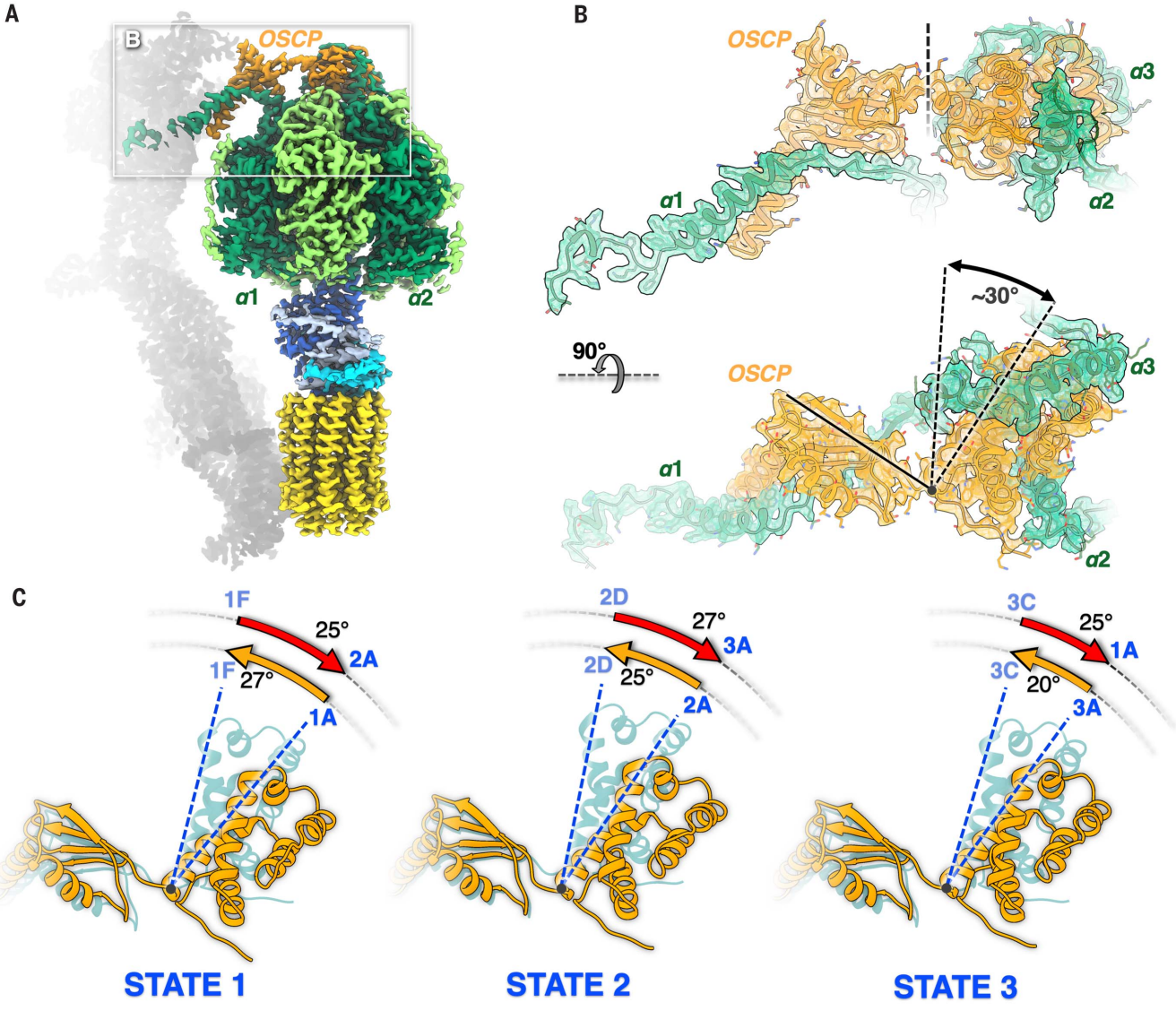


Fig. 4. Subunit OSCP connects the F_1 head and peripheral stalk as a flexible hinge. (**A**) Overview of F_1F_o monomer indicating the position of OSCP (orange) on top of the F_1 head (green) relative to the peripheral stalk (gray). (**B**) The two-domain OSCP interacts with the helical C-terminal extensions of subunits α1, α2, and α3 (green) (see Movie 4). Extensions of α2 and α3 form short helix bundles with helices of the globular N-terminal OSCP domain, whereas α1 binds to the elongated C-terminal OSCP domain that is attached to the peripheral stalk via the α1 extension. The two OSCP domains are connected by a flexible peptide link that facilitates ~30° back-and-forth rotation of the N-terminal domain with F_1 (curved black arrow) relative to the C-terminal OSCP domain and peripheral stalk. (**C**) Between substates of a given primary rotary state, the N-terminal OSCP domain rotates with F_1 by up to 28° (straight dashed lines) in synthesis direction (orange arrows) and then recoils to its starting position in the first substate of the next primary rotary state (red arrows). The OSCP position in the first substate of each primary state is shown in orange; the OSCP position in the last resolved rotary state in each primary state is indicated in light blue in the background.

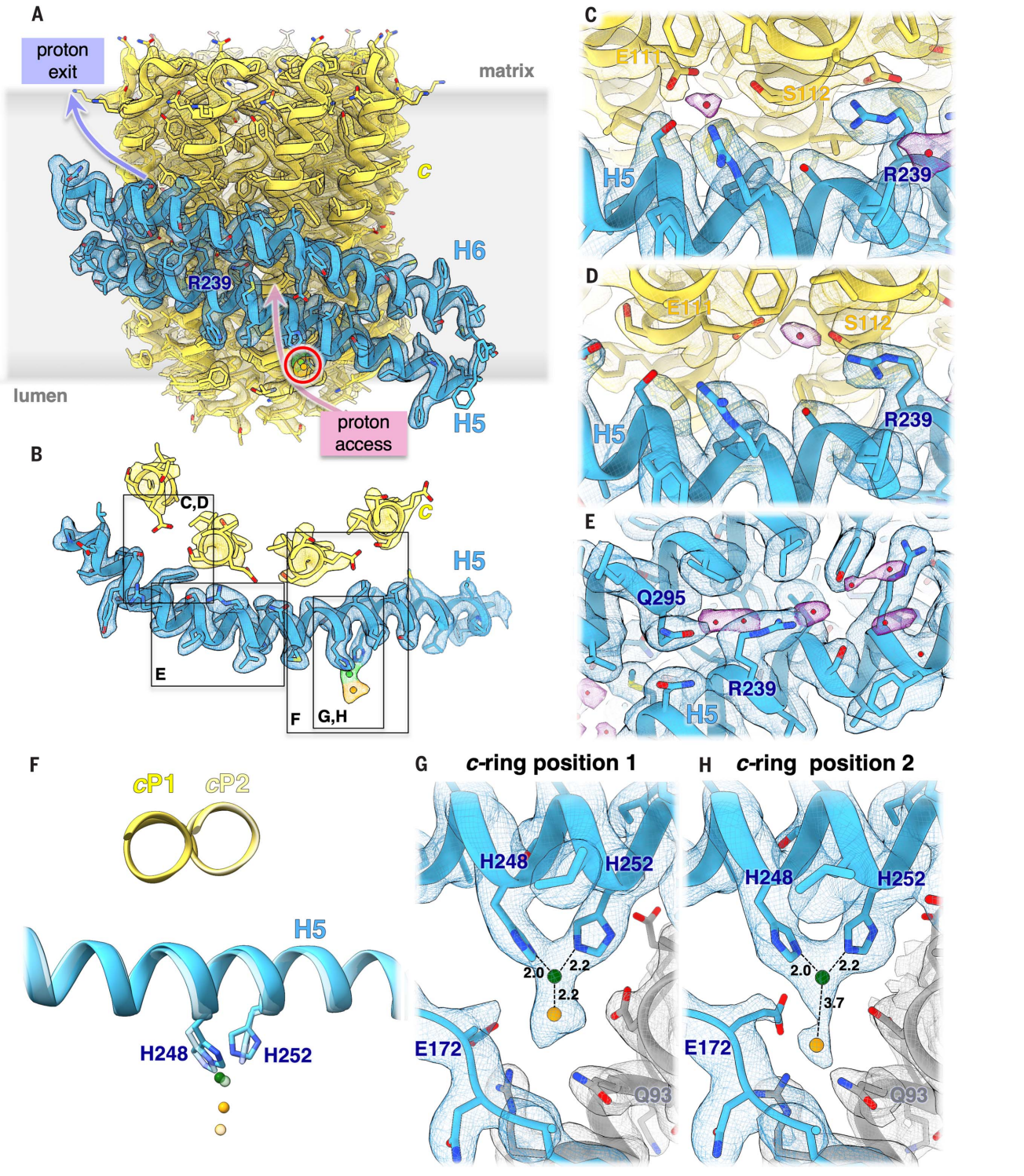


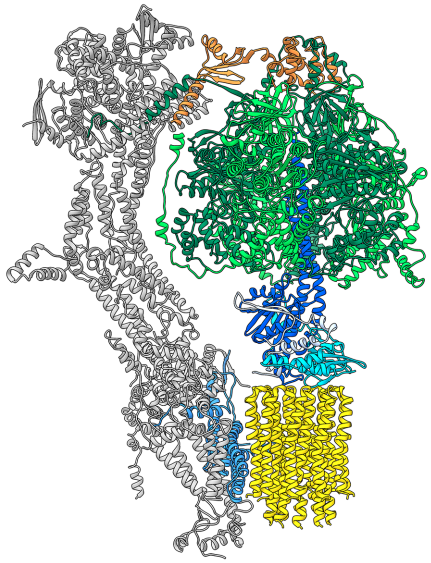
Fig. 5. Metal ion and ordered water molecules in the F_o proton access and release channels. (**A**) The H5/H6 helix hairpin of subunit *a* (light blue) forms the bearing against which the c_{10} ring (yellow) rotates, shown here for *c*-ring position 1. (**B**) Sectional view of H5 and outer-ring *c*-subunit helices with proton-binding *c*Glu[111], *c*-ring position 1. (**C** and **D**) Water molecules (purple density) coordinated by the ion-binding *c*Glu[111] and adjacent *c*Ser[112] in the matrix channel for *c*-ring positions 1 (C) and 2 (D). Maps are displayed at density thresholds of 0.035 for subunit *a*, 0.025 for the *c* ring, and 0.025 (C) or 0.018 (D) for water. (**E**) The strictly conserved *a*Arg[239] and *a*Gln[295] that separate the proton access and release channels in the membrane bind two water molecules, shown in the consensus C2-refined map of the F_o region. (**F** to **H**) Conserved residues *a*His[248] and *a*His[252] in H5 of subunit *a* coordinate a metal ion (green). Bond distances change depending on *c*-ring position. The predominant *c*-ring position [position 1, full color, and (H)] accounts for 58% of particles; a second position, differing by rotation of roughly 13° [position 2, faint yellow, and (G)] accounts for 33% of particles. Except where specified, all maps in a given panel are rendered at the same density threshold. E, Glu; H, His; Q, Gln; R, Arg; S, Ser.

has broad physiological implications, with a number of regulatory mechanisms acting on OSCP. Cyclophilin D, a regulator of mitochondrial permeability transition, binds to OSCP and inhibits ATP synthase activity (*28*), as does the immunomodulator Bz-423, which affects both maximum rate (V_{max}) and K_M for ATP hydrolysis (*27*). Deacetylation of a conserved Lys near the hinge region of OSCP in response to exercise stress affects mitochondrial ATP levels (*29*). The hormone estrogen binds to OSCP (*30*) and inhibits ATP synthase activity (*31*, *32*) with possible relevance to neuroprotection (*33*). The tumor suppressor p53, under nonstress conditions, binds to OSCP and promotes increased O_2 consumption and decreased ROS levels in mitochondria (*34*). These diverse and potent modulations of ATP synthase activity cannot be reconciled with a model of ATP synthase in which OSCP forms a static bridge between the F_1 head and peripheral stalk. Our results show that, on the contrary, OSCP plays a dynamic role in the flexible coupling of F_o and F_1. The concept of a dynamic OSCP will be essential to understand and exploit modulation of ATP synthase activity, with relevance to neurodegeneration and traumatic brain injury, ischemia-reperfusion injury, and immunomodulation, among others.

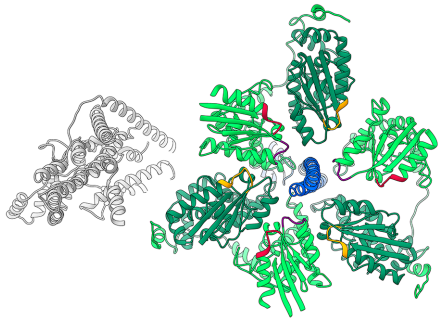
An essential histidine ligates a metal ion that is sensitive to *c*-ring rotation

In respiring mitochondria, the pH of the lumen is 7.2 (*35*), while the pK_a of the *c*-ring carboxylate in an aqueous environment is around 4.5 (*36*); thus, protonation of the *c* ring needs to be assisted by the local environment in the lumenal entrance channel in order to achieve high turnover rates. Mutational studies in *E. coli* (*37*) have established that residues *a*Glu^{219} and *a*His^{245} are essential for both ATP synthesis and proton permeability of F_o in membranes stripped of F_1. The mutations *a*E219H and *a*H245E show low levels of activity, while the double mutant is more active than either single mutant, suggesting that the residues interact to facilitate *c*-ring protonation. Our map shows a strong, nonpeptide density ligated by *a*His^{248} (*a*Glu^{219} in *E. coli*) and *a*His^{252} of *a*H5 (Fig. 5) with bond lengths of 2.0 to 2.2 Å. The strong density and coordination environment favor assignment to a metal ion, provisionally annotated as Zn^{2+} in the deposited models. At high threshold levels, the bound species appears diatomic, perhaps because of a heavy ligand (Fig. 5, B and F to H). Both ligating His residues are present in mammalian ATP synthases (fig. S8). The *a*His^{248} position is occupied by His or Glu in all F-type ATP synthases, with the notable exception of Na^+-translocating complexes (fig. S8A). Recently published high-resolution cryo-EM maps of yeast mitochondrial (*38*) and spinach chloroplast ATP synthase (*13*) both indicate a strong, unattributed nonpeptide density at the same position in the lumenal channel, near the residue equivalent to *a*His^{248} (*a*His^{185} and *a*Glu^{198}, respectively) (fig. S8, C and D). In both instances, a single strong density was observed as compared to the two densities in our map, although this may reflect the lower local resolution (3.6 Å and 3.4 Å) of the yeast and spinach maps.

Masked 3D classification of the *c* ring and *a* subunit yields two distinct *c*-ring rotary positions at 2.7-Å and 3.1-Å resolution, representing 58% and 33% of particles (fig. S2). The positions differ by rotation of the *c* ring by approximately one-third of a *c*-ring subunit (13°) (Fig. 5F). Comparing these maps, it is clear that the configuration of the metal ion in the lumenal channel changes with rotation of the *c* ring, with the distance between the nonpeptide densities changing from 2.2 to 3.7 Å between *c*-ring positions 1 and 2 (Fig. 5, F to H). The coordination of the metal is sensitive to *c*-ring position, suggesting that it may play a role in synchronizing *c*-ring protonation with its rotation. In both positions, distances between the HisεN and metal ion are in the 2.0 to 2.2 Å range, consistent with dative bonds and thus with unprotonated HisεN. Available mutational data strongly suggest that the role of *a*His^{248} is to protonate *a*Glu^{288}, which would require that *a*His^{248} is itself transiently protonated. Protonation of either His or Glu at this position would preclude their ability to coordinate a metal ion. Further work will be needed to clarify the identity, dynamics, and function of this metal ion.



Movie 2. Morph of 13 rotary substates of the *Polytomella* ATP synthase monomer. The *c* ring (yellow) and central stalk (subunits: γ, blue; δ, cyan; ε, pale blue) together rotate in three roughly equal ~120° steps. The F_1 head moves with them for the first ~30° of each step. α, dark green; β, bright green. The two-domain OSCP subunit (orange) works as a hinge between the moving F_1 head and the stationary peripheral stalk (gray).



Movie 3. Section through movie 2 at the level of nucleotide-binding sites in the F_1 head. Catalytic sites of the β subunits (bright green) are red. Binding sites for the structural ATP in the three α subunits (dark green) are orange. The β catch loop is drawn in purple. Central stalk subunit γ, blue; peripheral stalk, gray.

Ordered water in the aqueous half-channels and at the essential Arg^{239}

The membrane-embedded *a* subunit and *c* ring together define the pathway for protons to move across the inner mitochondrial membrane (Fig. 5A). The sequences of both subunits are conserved in all F-type ATP synthases, including those of bacteria and chloroplasts (*1*). The similarity (identity) scores of human and *Polytomella* mitochondrial ATP synthase are 44.9% (27.6%) for subunit *a* and 57.0% (38.7%) for subunit *c*. Previously, we showed (*3*) that the aqueous half-channels that conduct protons to, and away from, the *c*-ring glutamate are separated by a distance of 5 to 7 Å in the center of the membrane. We were able to model ordered water molecules within both channels in the current, high-resolution map (Fig. 5, C to E). In the matrix channel, a water molecule is coordinated by *c*Ser^{112} opposite the ion-binding *c*Glu^{111} of the adjacent *c* subunit, favoring the idea that the *c*-ring glutamate may be directly deprotonated by water in this location. Two water molecules are enclosed by the strictly conserved *a*Arg^{239} and *a*Gln^{295} in a pocket that appears not to be continuous with the aqueous channels (Fig. 5E). The interaction of these conserved residues with each other via a water molecule would reduce the flexibility of the *a*Arg^{239} side chain. In the lumenal proton access channel, the well-ordered water lies between the assigned metal ion density and the strictly conserved *a*Arg^{239}. Examining these waters, it is clear that the channel extends to the *a*/*c* interface close to *a*Arg^{239}, as previously reported (*3*). However, it would appear that these water molecules do not directly protonate the *c*-ring glutamate, because mutation of *a*His^{245} and *a*Glu^{219} in *E. coli* eliminates proton permeability of the F_o complex in stripped membranes (*37*). Instead, the *c* ring is likely protonated as it rotates past *a*Glu^{288} (*a*His^{245} in *E. coli*), immediately before passing into the hydrophobic lipid bilayer. Water in the lumenal channel permeating all the way to *a*Arg^{239} would provide an aqueous environment to prime the *c*-ring glutamate for protonation by inducing an outward-facing conformation (*36*). At the same

time, the small distance between half-channels would generate a substantial field and resulting torque on the *c*-ring glutamate, which may be needed to overcome the energetic barrier of moving the *c*-ring glutamate past the gating Arg^{239} (*1*, *3*).

C-terminal β extensions contact the adjacent α-subunit

In chlorophycean algae, the mitochondrial catalytic β subunit resembles those of the canonical ATP synthases closely, except that the chlorophycean subunit has acquired a ~60-residue C-terminal extension. Our map shows that this extension wraps vertically around the adjacent α subunit (Movie 5) and moves with subunit β as it adopts an open or closed conformation. A ~15-residue N-terminal extension of the α subunits present in chlorophycean algae, but not in other systems, is in a different configuration for each subunit (Fig. 1B, Fig. 4A, and Movie 5). One of these extensions, together with the distal OSCP domain, anchors one α subunit to the peripheral stalk, while the other two form short helix bundles with the proximal globular domain of OSCP, attaching it to F_1. The N- and C-terminal extensions would enhance the stability of the chlorophycean mitochondrial ATP synthase, which may be an advantage under stress conditions, as when *Polytomella* converts to a dormant cyst in nutrient-depleted environments (*39*).

ATP synthase–associated proteins 1 to 10 form a rigid peripheral stalk

In *Polytomella*, dimer formation is mediated by the peripheral stalk consisting of 10 ASA subunits that have no homologs in other known ATP synthases outside this class of green algae. We have modeled all nine previously described ASA subunits unambiguously into the 2.7 Å map (Fig. 1, Movie 1, and fig. S4), as well as an unknown subunit that we term ASA10. Mass spectrometry of the purified complex and DNA sequencing of the ~50-Mb *Polytomella* genome allowed us to identify ASA10 (fig. S9) and to obtain complete sequences for subunits ε and ASA2 (fig. S10). All subunits were built into the cryo-EM map de novo, as there are no homologous structures in the database.

The stability of the *Polytomella* dimer is due to the rigid peripheral stalks and their close interaction. ASA1 is at the core of the soluble stalk region, and forms a bridge between monomers with a helix-turn-helix motif 75 Å above the membrane surface (fig. S11). F_o has six transmembrane helices and four long, membrane-intrinsic helices, which form the two helix hairpins of the conserved *a* subunit (*2*). Of the six transmembrane helices, two belong to ASA6 and one each to ASA5, ASA8, ASA9, and ASA10. The two ASA10 subunits form the primary dimer interface on the lumenal side of the membrane, and their C termini are intertwined (figs. S9 and S11). Little or no other direct interaction is seen within the membrane, with lipids or detergent filling the space between monomers. ASA3 forms the characteristic Armadillo repeat domain above the membrane surface on the matrix side, and ASA2, ASA4, and ASA7 sit atop the peripheral stalk.

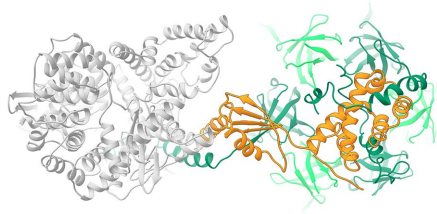
Outlook

We have determined the structures of *Polytomella* ATP synthase in 13 rotary substates, allowing us to visualize most if not all of its thermally accessible conformations. Critically, we have shown that the F_1 head and rotor move together through a 20° to 30° rotation; thus, the *c* ring rotates relative to the *a* subunit while the position of the central stalk within the F_1 head is preserved (fig. S12). This movement ensures flexible coupling of the two symmetry-mismatched nanomotors of ATP synthase, F_o and F_1, for all possible *c*-ring stoichiometries. Flexible coupling appears to be mediated predominantly by OSCP, which is emerging as an important target of cellular and pharmacological control. Whereas most previous structural studies of ATP synthase have characterized inhibited complexes, this work has instead examined the active ADP-bound form. The conformational flexibility we observe in these structures may have been suppressed in the previously determined auto-inhibited and inhibitor-bound states. Further studies will address the energetics of OSCP bending and integrate these results into a full catalytic scheme of ATP synthesis. A metal ion bound at the conserved residue $a His^{248}$ within the lumenal channel is sensitive to the *c*-ring rotary state and is likely to play an important role in protonating the *c*-ring carboxylate. A major unanswered question is how the strictly conserved $a Arg^{239}$ facilitates movement of deprotonated, but not protonated, *c*-ring subunits between the aqueous half-channels. The high-resolution detail provided by our model will be important for answering this question.

Materials and methods

Protein isolation and purification

ATP synthase from *Polytomella* sp. was purified as described (*2*) with slight modifications. Mitochondrial membranes (175 mg) were solubilized for 30 min at 4°C in a volume of 12 ml of buffer containing 30 mM Tris-HCl, pH 7.8, 2 mM $MgCl_2$, 50 mM NaCl and 2.9% (w/v) *n*-dodecyl-β-D-maltoside (DDM) to a final detergent:protein weight ratio of 2:1. After centrifugation (20,000*g*, 15 min, 4°C) to remove unsolubilized material, the filtered supernatant was loaded onto a POROS GoPure HQ column (Thermo Fisher Scientific) equilibrated in buffer A [30 mM Tris-HCl, pH 7.8, 2 mM $MgCl_2$, 50 mM NaCl and 0.015% (w/v) DDM] on an Äkta purifier (GE Healthcare). After washing the column with 100mM NaCl in buffer A, ATP synthase dimers were eluted with a linear 100 mM to 300 mM NaCl gradient in buffer A. For the final purification step, fractions containing ATP synthase dimers were concentrated in Vivaspin 500 columns with 100,000 molecular mass cutoff and then loaded onto a Superose6 Increase 3.2/30 size exclusion column (GE Healthcare) equilibrated in buffer B [30 mM Tris-HCl, pH 7.8, 2 mM $MgCl_2$, 40 mM NaCl, 0.05% (w/v) DDM] on an Ettan purifier (GE Healthcare).

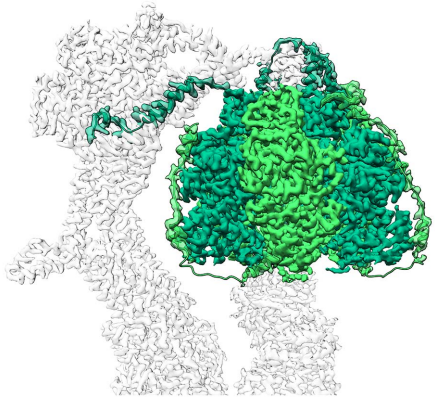


Movie 4. Hinge movement of the two-domain OSCP subunit (orange). The proximal α-helical OSCP domain at the right is attached to the F_1 head by the N-terminal extensions of two of the three α subunits (dark green). The distal β-sheet OSCP domain is attached to the peripheral stalk (gray) by interaction of one OSCP helix and the N-terminal extension of one of the three α subunits.

Fractions of pure and active (*2*) ATP synthase dimers were collected on ice and used directly for cryo-EM specimen preparation. Purity and activity of the sample was analyzed by blue-native polyacrylamide gel electrophoresis (BN-PAGE) and in-gel ATP hydrolysis assay (*40*). For two-dimensional polyacrylamide gel electrophoresis (2D SDS-PAGE) (*41*), subunits of the ATP synthase dimer were first resolved on a 10% acrylamide/8 M urea gel before being separated in the second dimension in a 16% acrylamide gel.

Electron microscopy and image processing

A solution of 3-4 mg/ml purified complex was applied to C-flat 1/1 or 2/1 holey carbon grids (Science Services GmbH) glow-discharged for 45 s at 0.15 mA. Grids were blotted for 4-5 s at blotforce 20 using a Vitrobot, and vitrified in liquid ethane. Electron micrographs were collected on a Titan Krios G2 with Falcon III detector in counting mode, at 300 kV and 75,000× magnification for a calibrated pixel size of 1.053 Å. 81-frame movies were automatically recorded with a dose of 0.4 $e^- Å^{-2} s^{-1}$, using EPU software. Movies were aligned using MotionCor2 within the Relion3-beta wrapper and CTF parameters were calculated using CTFFind4.1.10. Particles were picked using Gautomatch with templates generated by 2D classification of a previous dataset (*3*), low-pass filtering images to 20 Å. The dataset was cleaned using 3D classification in Relion3. Following a refinement of the full dimer with solvent masking, per-particle CTF parameters and beam tilt were refined, followed by Bayesian polishing and a second round of CTF and beam tilt refinement. The resulting 3D reconstruction gave an overall resolution for the dimer of 2.94 Å; masking of the stationary membrane-bound and lower peripheral stalk portions of the complex improved the resolution for this region to 2.69 Å. Symmetry expansion of the dataset using relion_particle_symmetry_expand, followed by c1 refinement of the upper peripheral stalk gave a resolution of 2.75 Å for this portion of the complex. 3D classification of the pre-aligned symmetry-expanded dataset,



Movie 5. Extensions of subunits α and β in *Polytomella* ATP synthase. The polypeptide sequences of subunits α and β in mitochondrial F_1F_o ATP synthases of chlorophycean algae (including *Polytomella* sp.) have characteristic extensions at their N and C termini. The ~15-residue N-terminal extension of subunit α interacts with OSCP (see Fig. 4 and movie 4). The ~60-residue C-terminal extension of the β subunit wraps around the outside of the neighboring subunit α in the F_1 head. Both extensions would render the chlorophycean F_1F_o complex more stable than that of mammalian or fungal mitochondrial ATP synthases.

with T = 20 and no shifts, allowed separation of 12 rotational states, with one class being low resolution and another being a mixture of states; further classification of the latter gave three additional rotary states for a total of 13 classes. Each was refined first with a mask around the selected monomer, and then with a mask around only the F_1 head and rotor portion of the complex. The classes were grouped according to the major rotary state they represent, and these were refined as for the substates. All steps described above were carried out in Relion3. The polished, CTF-refined dataset was imported into cryoSPARC (*42*) and refined using nonuniform refinement with automated masking; this map was used to assist manual building of the less ordered portions of ASA3 and ASA9. For all rotary states and substates, composite maps were generated from the dimer-masked and F_1+rotor-masked refined maps using phenix.combine_focused_maps (*43*).

Map and model analysis

Models of subunits *a*, OSCP, ε, and ASA1 to ASA10 were built manually in Coot (*44*). Models of α, β, γ, δ, and *c* were refined manually in Coot based on homology models calculated using the ModWeb server (*45*) on templates 5DN6 (*46*), 2V7Q (*47*), 3ZIA (*48*), and 2XND (*49*), respectively. Models were refined by real-space refinement with simulated annealing in Phenix (*50*) and checked manually before deposition. Model quality statistics are as given by the program phenix.validation_cryoem and are summarized in table S3. For comparison of rotary states and substates, maps were aligned to a map of F_o and the lower peripheral stalk (EMD-4806). Models of rotary states and substates were aligned using the Matchmaker tool of USCF Chimera (*51*), fitting to the *a* subunit. Figures were made using UCSF Chimera and ChimeraX (*52*).

Polytomella sp. genomic sequence

Genomic DNA was purified following a published protocol (*53*) with modifications. Briefly, *Polytomella* sp. cells were harvested in their logarithmic phase and resuspended in solubilization buffer containing 100 mM Tris-HCl, pH 8.0, 40 mM EDTA, 100 mM NaCl, 2% (w/v) lauryl sarcosine, 1% (w/v) SDS, RNase A (20 μg/ml), and protein kinase K (0.9 mg/ml). After 1 hour at 4°C, DNA was extracted twice with phenol:chloroform: isoamyl alcohol (25:24:1) and the aqueous phase was ethanol-precipitated for 1 hour on ice. Genomic DNA was further purified by 4 M LiCl and 7.5 M ammonium acetate before being ethanol-precipitated again and finally washed twice with water. Genomic sequencing was carried out by Microsynth AG with 50-fold sequence coverage. The genome size was ~50 Mb. A six-frame translation of the sequence was used as a query database for mass spectrometry results.

Mass spectrometry

Tryptic digestion of proteins in solution or excised from 2D gels was performed according to published procedures (*54*). Proteolytic digests were loaded by nano-HPLC (Dionex RSLCnano) on reverse-phase columns (trapping column: Acclaim PepMap c18, particle size 2 μm, L = 20 mm; analytical column: Acclaim PepMap c18, particle size 2 μm, L = 50 cm; Thermo Fisher Scientific) and eluted in organic phase gradients (Buffer A: 95% H_2O, 5% DMSO, 0.1% formic acid; Buffer B: 80% acetonitrile, 15% H_2O, 0.1% formic acid). Typically, gradients were ramped from 4% to 48% buffer B in 80 min at a flow rate of 300 nl/min. Peptides eluting from the column were ionized online using a Nanospray Flex Ion source and analyzed in an Orbitrap Elite or Q Exactive Plus mass spectrometer. Mass spectra were acquired over the 350 to 1600 m/z range at a resolution of 120,000, and sequence information was acquired by computer-controlled, data-dependent automated switching to MS/MS mode in response to collision energies based on mass and charge state of the candidate ions.

Datasets were processed with the Proteome Discoverer software package (version 2.1.0.81). Spectra were internally recalibrated on autoproteolytic trypsin fragments when applicable. Proteins were identified by matching the derived mass lists against a customized *Polytomella* database (combination of NCBI nr *Polytomella*, a full six-frame translation of a dataset acquired by whole-genome shotgun sequencing (Microsynth AG) and a list of common contaminants with the program Sequest (Thermo Fisher Scientific). In general, a mass tolerance of 10 ppm for parent ion spectra and 0.6 Da for fragment ion spectra, two missed cleavages, oxidation of Met (dynamic modification), acetylation of the protein N terminus (dynamic modification) and carbamidomethyl-cysteine (fixed modification) were selected as matching parameters in the search program. Results were evaluated using a percolator node (high-confidence q value, FDR < 0.01) to exclude false positives. Proteome data have been uploaded to the PRIDE online repository (*55*).

Protein sequence analysis

Homologs of the ASA10 sequence were found by BLAST searches of the NCBI (*56*) and Phytozome (*57*) databases. Sequence alignments were carried out using Clustal Omega (*58*) and formatted using JalView (*59*).

REFERENCES AND NOTES

1. W. Kühlbrandt, Structure and mechanisms of F-type ATP synthases. *Annu. Rev. Biochem.* 10.1146/annurev-biochem-013118-110903 (2019). doi: 10.1146/annurev-biochem-013118-110903; pmid: 30901262
2. M. Allegretti *et al.*, Horizontal membrane-intrinsic α-helices in the stator *a*-subunit of an F-type ATP synthase. *Nature* **521**, 237–240 (2015). doi: 10.1038/nature14185; pmid: 25707805
3. N. Klusch, B. J. Murphy, D. J. Mills, Ö. Yildiz, W. Kühlbrandt, Structural basis of proton translocation and force generation in mitochondrial ATP synthase. *eLife* **6**, e33274 (2017). doi: 10.1104/pp.106.094060; pmid: 17468226
4. M. Vázquez-Acevedo *et al.*, The mitochondrial ATP synthase of chlorophycean algae contains eight subunits of unknown origin involved in the formation of an atypical stator-stalk and in the dimerization of the complex. *J. Bioenerg. Biomembr.* **38**, 271–282 (2006). doi: 10.1007/s10863-006-9046-x; pmid: 17160464
5. R. van Lis, G. Mendoza-Hernández, G. Groth, A. Atteia, New insights into the unique structure of the F_0F_1-ATP synthase from the chlamydomonad algae Polytomella sp. and Chlamydomonas reinhardtii. *Plant Physiol.* **144**, 1190–1199 (2007). doi: 10.1104/pp.106.094060; pmid: 17468226
6. D. A. Cherepanov, A. Y. Mulkidjanian, W. Junge, Transient accumulation of elastic energy in proton translocating ATP synthase. *FEBS Lett.* **449**, 1–6 (1999). doi: 10.1016/S0014-5793(99)00386-5; pmid: 10225416
7. N. Soga, K. Kimura, K. Kinosita Jr., M. Yoshida, T. Suzuki, Perfect chemomechanical coupling of F_oF_1-ATP synthase. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4960–4965 (2017). doi: 10.1073/pnas.1700801114; pmid: 28442567
8. A. Wächter *et al.*, Two rotary motors in F-ATP synthase are elastically coupled by a flexible rotor and a stiff stator stalk. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3924–3929 (2011). doi: 10.1073/pnas.1011581108; pmid: 21368147
9. H. Sielaff *et al.*, Domain compliance and elastic power transmission in rotary F_oF_1-ATPase. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17760–17765 (2008). doi: 10.1073/pnas.0807683105; pmid: 19001275
10. J. L. Martin, R. Ishmukhametov, D. Spetzler, T. Hornung, W. D. Frasch, Elastic coupling power stroke mechanism of the F_1-ATPase molecular motor. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 5750–5755 (2018). doi: 10.1073/pnas.1803147115; pmid: 29760063
11. A. Zhou *et al.*, Structure and conformational states of the bovine mitochondrial ATP synthase by cryo-EM. *eLife* **4**, e10180 (2015). doi: 10.1104/pp.106.094060; pmid: 17468226
12. M. Sobti *et al.*, Cryo-EM structures of the autoinhibited E. coli ATP synthase in three rotational states. *eLife* **5**, e21598 (2016). doi: 10.1104/pp.106.094060; pmid: 17468226
13. A. Hahn, J. Vonck, D. J. Mills, T. Meier, W. Kühlbrandt, Structure, mechanism, and regulation of the chloroplast ATP synthase. *Science* **360**, eaat4318 (2018). doi: 10.1126/science.aat4318; pmid: 29748256
14. M. Sobti *et al.*, Cryo-EM reveals distinct conformations of E. coli ATP synthase on exposure to ATP. *eLife* **8**, e43864 (2019). doi: 10.1104/pp.106.094060; pmid: 17468226
15. Y. Komoriya *et al.*, Principal role of the arginine finger in rotary catalysis of F_1-ATPase. *J. Biol. Chem.* **287**, 15134–15142 (2012). doi: 10.1074/jbc.M111.328153; pmid: 22403407
16. M. D. Greene, W. D. Frasch, Interactions among γR268, γQ269, and the β subunit catch loop of Escherichia coli F_1-ATPase are important for catalytic activity. *J. Biol. Chem.* **278**, 51594–51598 (2003). doi: 10.1074/jbc.M309948200; pmid: 14532272

17. K. R. Vinothkumar, M. G. Montgomery, S. Liu, J. E. Walker, Structure of the mitochondrial ATP synthase from *Pichia angusta* determined by electron cryo-microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12709–12714 (2016). doi: 10.1073/pnas.1615902113; pmid: 27791192
18. D. M. Rees, A. G. W. Leslie, J. E. Walker, The structure of the membrane extrinsic region of bovine ATP synthase. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21597–21601 (2009). doi: 10.1073/pnas.0910365106; pmid: 19995987
19. V. Giorgio, F. Fogolari, G. Lippe, P. Bernardi, OSCP subunit of mitochondrial ATP synthase: Role in regulation of enzyme function and of its transition to a pore. *Br. J. Pharmacol.* 10.1111/bph.14513 (2018). doi: 10.1111/bph.14513; pmid: 30291799
20. H. Sielaff, S. Yanagisawa, W. D. Frasch, W. Junge, M. Börsch, Structural Asymmetry and Kinetic Limping of Single Rotary F-ATP Synthases. *Molecules* **24**, 504 (2019). doi: 10.3390/molecules24030504; pmid: 30704145
21. K. Okazaki, G. Hummer, Elasticity, friction, and pathway of γ-subunit rotation in F_oF_1-ATP synthase. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10720–10725 (2015). doi: 10.1073/pnas.1500691112; pmid: 26261344
22. S. Yanagisawa, W. D. Frasch, Protonation-dependent stepped rotation of the F-type ATP synthase c-ring observed by single-molecule measurements. *J. Biol. Chem.* **292**, 17093–17100 (2017). doi: 10.1074/jbc.M117.799940; pmid: 28842481
23. R. Ishmukhametov, T. Hornung, D. Spetzler, W. D. Frasch, Direct observation of stepped proteolipid ring rotation in E. coli F_oF_1-ATP synthase. *EMBO J.* **29**, 3911–3923 (2010). doi: 10.1038/emboj.2010.259; pmid: 21037553
24. P. Balabaskaran Nina *et al*., ATP synthase complex of Plasmodium falciparum: Dimeric assembly in mitochondrial membranes and resistance to genetic disruption. *J. Biol. Chem.* **286**, 41312–41322 (2011). doi: 10.1074/jbc.M111.290973; pmid: 21984828
25. E. Perez *et al*., The mitochondrial respiratory chain of the secondary green alga *Euglena gracilis* shares many additional subunits with parasitic Trypanosomatidae. *Mitochondrion* **19B**, 338–349 (2014). doi: 10.3389/fphys.2018.00803; pmid: 30022951
26. I. Starke, G. D. Glick, M. Börsch, Visualizing Mitochondrial F_oF_1-ATP Synthase as the Target of the Immunomodulatory Drug Bz-423. *Front. Physiol.* **9**, 803 (2018). doi: 10.3389/fphys.2018.00803; pmid: 30022951
27. K. M. Johnson *et al*., Identification and validation of the mitochondrial F_1F_0-ATPase as the molecular target of the immunomodulatory benzodiazepine Bz-423. *Chem. Biol.* **12**, 485–496 (2005). doi: 10.1016/j.chembiol.2005.02.012; pmid: 15850986
28. V. Giorgio *et al*., Cyclophilin D modulates mitochondrial F_oF_1-ATP synthase by interacting with the lateral stalk of the complex. *J. Biol. Chem.* **284**, 33982–33988 (2009). doi: 10.1074/jbc.M109.020115; pmid: 19801635
29. A. Vassilopoulos *et al*., SIRT3 deacetylates ATP synthase F_1 complex proteins in response to nutrient- and exercise-induced stress. *Antioxid. Redox Signal.* **21**, 551–564 (2014). doi: 10.1089/ars.2013.5420; pmid: 24252090
30. J. Zheng, V. D. Ramirez, Purification and identification of an estrogen binding protein from rat brain: Oligomycin sensitivity-conferring protein (OSCP), a subunit of mitochondrial F_oF_1-ATP synthase/ATPase. *J. Steroid Biochem. Mol. Biol.* **68**, 65–75 (1999). doi: 10.1016/S0960-0760(98)00161-7; pmid: 10215039
31. J. Zheng, V. D. Ramirez, Rapid inhibition of rat brain mitochondrial proton F_oF_1-ATPase activity by estrogens: Comparison with Na^+,K^+-ATPase of porcine cortex. *Eur. J. Pharmacol.* **368**, 95–102 (1999). doi: 10.1016/S0014-2999(99)00012-6; pmid: 10096774
32. F. Massart, S. Paolini, E. Piscitelli, M. L. Brandi, G. Solaini, Dose-dependent inhibition of mitochondrial ATP synthase by 17β-estradiol. *Gynecol. Endocrinol.* **16**, 373–377 (2002). doi: 10.1080/gye.16.5.373.377; pmid: 12587531
33. J. W. Simpkins, K. D. Yi, S.-H. Yang, J. A. Dykens, Mitochondrial mechanisms of estrogen neuroprotection. *Biochim. Biophys. Acta* **1800**, 1113–1120 (2010). doi: 10.1016/j.bbagen.2009.11.013; pmid: 19931595
34. M. Bergeaud *et al*., Mitochondrial p53 mediates a transcription-independent regulation of cell respiration and interacts with the mitochondrial F_1F_0-ATP synthase. *Cell Cycle* **12**, 2781–2793 (2013). doi: 10.4161/cc.25870; pmid: 23966169
35. B. Rieger, W. Junge, K. B. Busch, Lateral pH gradient between OXPHOS complex IV and F_oF_1 ATP-synthase in folded mitochondrial membranes. *Nat. Commun.* **5**, 3103 (2014). doi: 10.1038/ncomms4103; pmid: 24476986
36. D. Pogoryelov *et al*., Microscopic rotary mechanism of ion translocation in the F_o complex of ATP synthases. *Nat. Chem. Biol.* **6**, 891–899 (2010). doi: 10.1038/nchembio.457; pmid: 20972431
37. B. D. Cain, R. D. Simoni, Interaction between Glu-219 and His-245 within the a subunit of F_1F_0-ATPase in Escherichia coli. *J. Biol. Chem.* **263**, 6606–6612 (1988). pmid: 2896197
38. H. Guo, S. A. Bueler, J. L. Rubinstein, Atomic model for the dimeric F_O region of mitochondrial ATP synthase. *Science* **358**, 936–940 (2017). doi: 10.1126/science.aao4815; pmid: 29074581
39. E. Lewis, C. Munger, R. Watson, D. Wise, Life Cycle of *Polytomella caeca* (Phytomonadida, Polyblepharidae). *J. Protozool.* **21**, 647–649 (1974). doi: 10.1111/j.1550-7408.1974.tb03720.x
40. E. Zerbetto, L. Vergani, F. Dabbeni-Sala, Quantification of muscle mitochondrial oxidative phosphorylation enzymes via histochemical staining of blue native polyacrylamide gels. *Electrophoresis* **18**, 2059–2064 (1997). doi: 10.1002/elps.1150181131; pmid: 9420170
41. I. Rais, M. Karas, H. Schägger, Two-dimensional electrophoresis for the isolation of integral membrane proteins and mass spectrometric identification. *Proteomics* **4**, 2567–2571 (2004). doi: 10.1002/pmic.200400829; pmid: 15352231
42. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). doi: 10.1038/nmeth.4169; pmid: 28165473
43. P. V. Afonine *et al*., New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D* **74**, 814–840 (2018). doi: 10.1107/S2059798318009324; pmid: 30198894
44. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010). doi: 10.1107/S0907444910007493; pmid: 20383002
45. N. Eswar *et al*., Tools for comparative protein structure modeling and analysis. *Nucleic Acids Res.* **31**, 3375–3380 (2003). doi: 10.1093/nar/gkg543; pmid: 12824331
46. E. Morales-Rios, M. G. Montgomery, A. G. W. Leslie, J. E. Walker, Structure of ATP synthase from Paracoccus denitrificans determined by X-ray crystallography at 4.0 Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13231–13236 (2015). doi: 10.1073/pnas.1517542112; pmid: 26460036
47. J. R. Gledhill, M. G. Montgomery, A. G. W. Leslie, J. E. Walker, How the regulatory protein, IF_1, inhibits F_1-ATPase from bovine mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15671–15676 (2007). doi: 10.1073/pnas.0707326104; pmid: 17895376
48. G. C. Robinson *et al*., The structure of F_1-ATPase from Saccharomyces cerevisiae inhibited by its regulatory protein IF_1. *Open Biol.* **3**, 120164 (2013). doi: 10.1098/rsob.120164; pmid: 23407639
49. I. N. Watt, M. G. Montgomery, M. J. Runswick, A. G. W. Leslie, J. E. Walker, Bioenergetic cost of making an adenosine triphosphate molecule in animal mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16823–16827 (2010). doi: 10.1073/pnas.1011099107; pmid: 20847295
50. P. V. Afonine *et al*., Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018). doi: 10.1107/S2059798318006551; pmid: 29872004
51. E. F. Pettersen *et al*., UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi: 10.1002/jcc.20084; pmid: 15264254
52. T. D. Goddard *et al*., UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018). doi: 10.1002/pro.3235; pmid: 28710774
53. J. Fan, R. W. Lee, Mitochondrial genome of the colorless green alga Polytomella parva: Two linear DNA molecules with homologous inverted repeat Termini. *Mol. Biol. Evol.* **19**, 999–1007 (2002). doi: 10.1093/oxfordjournals.molbev.a004180; pmid: 12082120
54. S. Buschmann *et al*., The structure of cbb_3 cytochrome oxidase provides insights into proton pumping. *Science* **329**, 327–330 (2010). doi: 10.1126/science.1187303; pmid: 20576851
55. J. A. Vizcaíno *et al*., 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2016). doi: 10.1093/nar/gkv1145; pmid: 26527722
56. G. M. Boratyn *et al*., Domain enhanced lookup time accelerated BLAST. *Biol. Direct* **7**, 12 (2012). doi: 10.1186/1745-6150-7-12; pmid: 22510480
57. D. M. Goodstein *et al*., Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186 (2012). doi: 10.1093/nar/gkr944; pmid: 22110026
58. F. Sievers, D. G. Higgins, Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* **27**, 135–145 (2018). doi: 10.1002/pro.3290; pmid: 28884485
59. A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009). doi: 10.1093/bioinformatics/btp033; pmid: 19151095

ACKNOWLEDGMENTS

Funding: Supported by the Max Planck Society, Deutsche Forschungsgemeinschaft grant SFB 807 (W.K.), and EMBO long-term fellowship EMBO LTF 702-2016 (B.J.M.). We thank I. Wöllenweber and F. Rupprecht for preparative and technical support. **Author contributions:** W.K. initiated and supervised the study; N.K. grew *Polytomella* cultures and isolated, purified, and analyzed ATP synthase dimers biochemically; B.J.M. and N.K. prepared cryo-EM specimens and collected data; B.J.M. performed image analysis and identified and analyzed rotary substates; J.L. performed mass spectrometry; D.J.M. maintained cryo-EM instrumentation and EM alignment; B.J.M., N.K., J.L., Ö.Y., and W.K. analyzed data; and B.J.M., N.K., and W.K. wrote the manuscript. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** Maps and models are publicly available through the Electron Microscopy Data Bank and Protein Data Bank databases (tables S2 and S3). Mass spectrometry data are publicly available through the PRIDE database (accession number PXD013530) (*55*). All other data discussed are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/eaaw9128/suppl/DC1
Figs. S1 to S12
Tables S1 to S3
References (*60*–*64*)

5 February 2019; accepted 26 April 2019
10.1126/science.aaw9128

RESEARCH ARTICLE

CANCER

The glycan CA19-9 promotes pancreatitis and pancreatic cancer in mice

Dannielle D. Engle[1,2]*, Hervé Tiriac[1,2]†, Keith D. Rivera[1], Arnaud Pommier[1,2]‡, Sean Whalen[3], Tobiloba E. Oni[1,2], Brinda Alagesan[1,2], Eun Jung Lee[1,2], Melissa A. Yao[1,2], Matthew S. Lucito[1,2], Benjamin Spielman[1,2], Brandon Da Silva[1,2], Christina Schoepfer[1,2], Kevin Wright[1,2], Brianna Creighton[1,2], Lauren Afinowicz[1,2], Kenneth H. Yu[4,5], Robert Grützmann[6], Daniela Aust[7], Phyllis A. Gimotty[8], Katherine S. Pollard[3,9], Ralph H. Hruban[10], Michael G. Goggins[10,11], Christian Pilarsky[6], Youngkyu Park[1,2], Darryl J. Pappin[1], Michael A. Hollingsworth[12], David A. Tuveson[1,2]§

Glycosylation alterations are indicative of tissue inflammation and neoplasia, but whether these alterations contribute to disease pathogenesis is largely unknown. To study the role of glycan changes in pancreatic disease, we inducibly expressed human fucosyltransferase 3 and β1,3-galactosyltransferase 5 in mice, reconstituting the glycan sialyl-Lewis^a, also known as carbohydrate antigen 19-9 (CA19-9). Notably, CA19-9 expression in mice resulted in rapid and severe pancreatitis with hyperactivation of epidermal growth factor receptor (EGFR) signaling. Mechanistically, CA19-9 modification of the matricellular protein fibulin-3 increased its interaction with EGFR, and blockade of fibulin-3, EGFR ligands, or CA19-9 prevented EGFR hyperactivation in organoids. CA19-9–mediated pancreatitis was reversible and could be suppressed with CA19-9 antibodies. CA19-9 also cooperated with the *Kras^{G12D}* oncogene to produce aggressive pancreatic cancer. These findings implicate CA19-9 in the etiology of pancreatitis and pancreatic cancer and nominate CA19-9 as a therapeutic target.

Pancreatitis, or inflammation of the pancreas, is a painful, recurrent, and occasionally lethal medical disorder with limited treatment options. The incidence of acute and chronic pancreatitis is rising (*1*). Pancreatitis accounts for more than 275,000 hospitalizations in the United States per year, and the number of hospital admissions has increased by 20% over the past decade (*2*). The causes of pancreatitis include blockage of the pancreatic duct by gallstones, alcohol and certain drugs that cause acinar cell damage, medical procedures or trauma that damage pancreatic tissue, and autoimmune diseases (*2*). In approximately one-third of cases, the underlying etiology of the pancreatitis is unknown (idiopathic) (*3*, *4*). Most acute pancreatitis cases will resolve with supportive care; however, up to 20% of patients will develop severe tissue damage and will either succumb to multiorgan system failure or suffer from bouts of recurrent disease with markedly diminished quality of life (*2*–*4*). Individuals with hereditary acute pancreatitis progress to chronic pancreatitis with a much higher penetrance and have a 40 to 55% lifetime risk of developing pancreatic cancer (*1*, *5*). Indeed, chronic pancreatitis promotes mutant Kras-mediated development of pancreatic cancer in mice (*6*).

The glycan carbohydrate antigen 19-9 (CA19-9) is found in the serum of 10 to 30% of pancreatitis patients and in 75% of pancreatic cancer patients, as well as in patients with other gastrointestinal diseases (*7*–*16*). CA19-9 elevation is also detected in pancreatic intraepithelial neoplasms (PanINs), which are precursors to pancreatic ductal adenocarcinoma (PDAC) (*17*). CA19-9 [sialyl-Lewis^a (sLe^a)] is generated by the stepwise addition of sugar moieties to type I precursor chains present on proteins and other molecules, culminating in the α1,4 linkage of fucose to *N*-acetylglucosamine (fig. S1A). Fucosyltransferase 3 (FUT3) is the only enzyme with the ability to add fucose moieties through an α1,4 linkage and generate CA19-9. Mice lack this enzyme because Fut3 is a pseudogene in rodents (*18*, *19*).

To facilitate the discovery of PDAC biomarker candidates, we sought to create a mouse model of PDAC that recapitulated the elevation of CA19-9 observed in human patients. This model would enable prioritization of biomarkers that outperform CA19-9. Furthermore, changes to glycosylation often result in functional consequences. In this study, we investigated the role of CA19-9 elevation in mouse and organoid models of pancreatic disease.

Recapitulation of CA19-9 elevation and regulation in cultured mouse PDAC cells

To express CA19-9 in mouse cells, we transduced mouse PDAC cells with human FUT3. Expression of FUT3 alone was insufficient for CA19-9 production but did lead to increased levels of Lewis^x antigens after removal of terminal galactose moieties present in rodents but not humans (Fig. 1A). The generation of the related Lewis^x epitopes suggested that reprogramming of the precursor substrates would be necessary for the production of CA19-9 in pancreatic ductal cells. β1,3-galactosyltransferase 5 (β3GALT5) is required for the production of type I chain precursors (*20*), which serve as the precursors for the Lewis^a modification (fig. S1A). Accordingly, expression of both FUT3 and β3GALT5 in mouse PDAC cells led to the cell surface expression of CA19-9 at levels equivalent to those observed in human cancer cell lines (COLO205, SUIT2) (Fig. 1B and fig. S1B). Comparable CA19-9 levels were observed in the blood of mice after orthotopic transplantation of the CA19-9–expressing mouse and human cells (fig. S1C).

To determine whether the murine PDAC cell proteins harboring CA19-9 modification are similar to those in human PDAC cells, CA19-9 protein carriers were immunoprecipitated (IP) and identified by mass spectrometry (MS) (fig. S2A and table S1). These analyses identified known CA19-9 protein carriers in the FUT3- and β3GALT5-expressing mouse cells (n = 3; FC1199, FC1242, FC1245), including CD44, LGALS3BP, MUC1, and MUC5AC (*21*–*24*). The human PDAC cell line, MiaPaCa-2, is CA19-9 negative, and therefore, we used it as a control to identify human CA19-9 core proteins in the CA19-9–positive cell lines, Capan-2, SUIT2, and hM1-2D (fig. S2B and table

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. [2]Lustgarten Foundation Pancreatic Cancer Research Laboratory, Cold Spring Harbor, NY 11724, USA. [3]Gladstone Institutes, San Francisco, CA 94158, USA. [4]David M. Rubenstein Center for Pancreatic Cancer Research, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [5]Joan and Sanford I. Weill Medical College, Cornell University, New York, NY 10065, USA. [6]Department of Surgery, Universitätsklinikum Erlangen, 91054 Erlangen, Germany. [7]Institute for Pathology, Universitätsklinikum Dresden, 01307 Dresden, Germany. [8]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA 19104, USA. [9]Department of Epidemiology and Biostatistics, Institute for Human Genetics, Quantitative Biology Institute, Institute for Computational Health Sciences, and Chan Zuckerberg Biohub, University of California, San Francisco, San Francisco, CA 94158, USA. [10]Sidney Kimmel Cancer Center, The Sol Goldman Pancreatic Cancer Research Center, and Department of Pathology, School of Medicine, Johns Hopkins University, Baltimore, MD 21231, USA. [11]Departments of Medicine and Oncology, School of Medicine, Johns Hopkins University, Baltimore, MD 21231, USA. [12]Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Omaha, NE 68198, USA.
*Present address: The Salk Institute for Biological Studies, La Jolla, CA 92037, USA. †Present address: Department of Surgery, University of California, San Diego, La Jolla, CA 92093, USA. ‡Present address: Department of Immunology, Cochin Institute, 75014 Paris, France.
§Corresponding author. Email: dtuveson@cshl.edu

S2). We compared mouse and human CA19-9 protein carriers and found that an average of 72.3% (95% confidence interval 60.3 to 84.3%; $n = 3$) of the CA19-9–modified proteins identified in all three human PDAC cell lines were also found in the engineered murine PDAC cell lines. Thus, expression of the human FUT3 and β3GALT5 genes in mouse cells largely recapitulates the human CA19-9 carrier profile (Fig. 1C).

CA19-9 elevation in mice leads to acute and chronic pancreatitis

We next generated a mouse model with inducible CA19-9 expression to study the effects of this glycan on pancreatic disease pathogenesis (fig. S2, C to E). Using PDX1-Cre (C), we restricted expression of the transgenes to the pancreas, duodenum, and bile duct. Cre-mediated excision of a LoxP-flanked stop (LSL) cassette enables ubiquitous expression of the reverse tetracycline-controlled transactivator 3 and mKate2 from the CAGS promoter within the ROSA26 locus (R^{LSL}) (*25*). Addition of doxycycline (Dox) to the diet activates the reverse tetracycline-controlled transactivator, enabling FUT3, β3GALT5, and enhanced green fluorescent protein (eGFP) expression from the tetracycline response element promoter in the COLA1 locus (F) (*26*). In addition, we excised the LSL cassette in the R allele to generate animals with Dox-inducible, whole-body CA19-9 expression (R;F). Both C;R^{LSL};F and R;F Dox-treated mice exhibited histologic signs of pancreatitis, including interstitial edema, lymphocyte infiltration, and collagen deposition (Fig. 2A and fig. S3, A and B). All other nonpancreatic tissues appeared histologically normal in both Dox-treated and untreated models and their littermate controls (figs. S4 and S5). The mice progressed from acute to chronic pancreatitis after 28 days of Dox treatment as demonstrated by the clinical histopathological hallmarks of acinar atrophy, accumulation of metaplastic ductal lesions, and persistent fibroinflammatory disease (figs. S3, A and B, and S6, A and B) (*27*). Pancreatitis was highly penetrant in both models (Fig. 2B and fig. S6, C and D).

Both mouse models demonstrated elevated serum levels of the pancreatic enzymes amylase and lipase within 24 hours of Dox treatment, but no changes were observed in untreated mice and littermate controls (Fig. 2C and fig. S6, C and D). Patients with acute pancreatitis have elevated serum levels of amylase and lipase, but in the chronic phase of pancreatitis, the levels of these enzymes either normalize or decrease even further. We found in our mouse models that the levels of amylase and lipase began to normalize after 4 weeks of Dox treatment, but both C;R^{LSL};F and R;F mice still exhibited histologic signs of chronic pancreatitis. Expression of eGFP and CA19-9 was detected in the pancreas after Dox treatment of C;R^{LSL};F mice and R;F mice, but they were not detected in untreated or control littermates (Fig. 2A and figs. S3C, S7, and S8). In both mouse models, CA19-9 was predominantly expressed in intralobular, intercalated, and metaplastic pancreatic ducts as well as in islet cells (Fig. 2A and figs. S3C, S7, and S8). Both models exhibited elevated CA19-9 levels in the circulation (Fig. 2D and fig. S6, C and D). Secreted CA19-9 was also observed coating eGFP-negative endothelial cells as well as fibroblasts (figs. S7B and S8). E-selectin is an endogenous receptor of CA19-9 expressed by endothelial cells (*28*, *29*) and may explain the accumulation of CA19-9 within the vasculature. Despite recombination in the acinar compartment, acinar cells were rarely observed to be CA19-9 positive. R;F mice exhibited Dox-dependent expression of CA19-9 and eGFP in all tissues examined, whereas in C;R^{LSL};F mice, expression was limited to the pancreas, duodenum, and gall bladder (Fig. 2A and figs. S3, C to G, and S7 to S11).

We next sought to determine the prevalence of CA19-9 elevation in human pancreatic disease and to compare the CA19-9 tissue expression pattern observed in human patients and the CA19-9 genetically engineered mouse models. We therefore evaluated CA19-9 levels and expression patterns in patients, including specimens of pancreatic cancer and adjacent normal tissue ($n = 72$) as well as surgically resected chronic pancreatitis samples ($n = 44$) (figs. S12 and S13). CA19-9 was expressed at low levels in normal homeostatic pancreatic ducts and at elevated levels in adjacent reactive ducts that consist of atypical, benign pancreatic ducts that occur even in the absence of substantial inflammation (fig. S12A), preinvasive carcinomas (PanIN-1A and PanIN-1B), and invasive PDAC specimens (fig. S12, B to D) (*17*). In chronic pancreatitis specimens, CA19-9 was expressed at high levels in the reactive and metaplastic ducts with more sporadic expression in the centroacinar and acinar compartments (fig. S13, A and B, and table S3). Elevated levels of CA19-9 were detected in the serum of 20% of the chronic pancreatitis patients we examined, whereas more than 93% of these patients exhibited local elevation of CA19-9 in the resected pancreatic specimens by immunohistochemistry (IHC) ($n = 44$) (fig. S13, C and D). Together, these data indicate that CA19-9 elevation is a common feature of chronic pancreatitis and that the CA19-9 expression pattern is similar between the CA19-9 mouse models and human patients.

Patients with severe acute or chronic pancreatitis often also exhibit weight loss (*30*, *31*). Therefore, we assessed the health and weight of both mouse models. Whereas most C;R^{LSL};F mice maintained their overall health status after Dox treatment, Dox-treated R;F mice exhibited significant body weight loss (fig. S14A). In severe cases, the weight loss exceeded 20%, and euthanasia was required. The weight loss was not due to pathology of the stomach, duodenum, or colon and was not associated with an autoimmune reaction, induction of endoplasmic reticulum stress in the pancreas, or loss of glycemic control (*32*) (fig. S14, B to D). Pancreatic exocrine

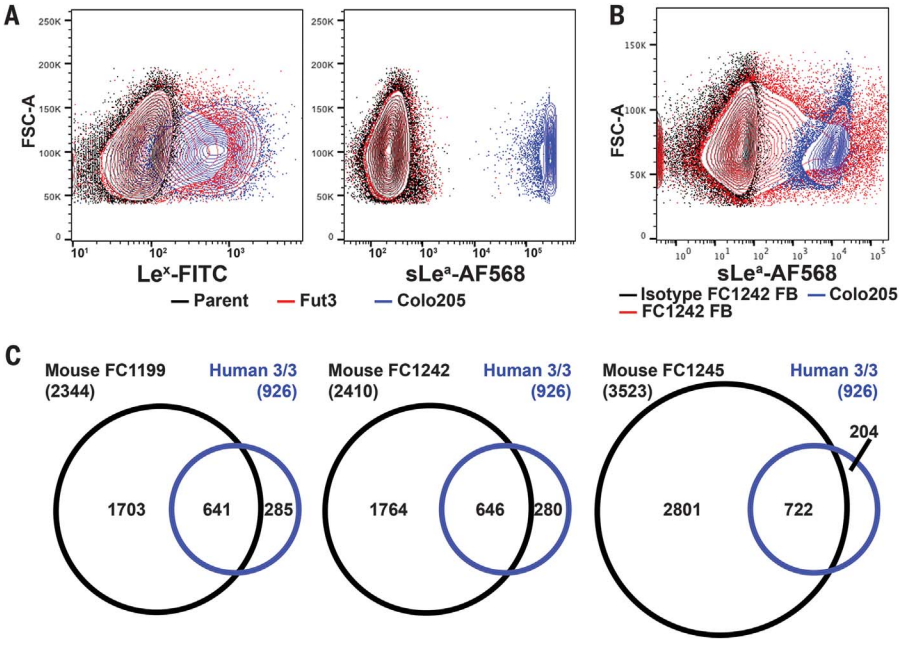


Fig. 1. FUT3 with β3GALT5 expression enables CA19-9 production in engineered mouse pancreatic cancer cells. (**A**) Ectopic FUT3 induces Lewis^x (Le^x) but not CA19-9–sLe^a expression in mouse PDAC cells by flow cytometric analysis. The Le^x- and CA19-9–sLe^a—positive human cell line COLO205 and Le^x- and CA19-9–sLe^a—negative parental KPC cell lines are shown. FITC, fluorescein isothiocyanate; FSC-A, forward scatter–area; AF568, Alexafluor 568. (**B**) CA19-9 flow cytometry of mouse PDAC cells stably and constitutively expressing FUT3 with β3GALT5 (FB) compared with the isotype control antibody. (**C**) Overlap between CA19-9 protein carriers identified in three out of three human PDAC cell lines ($n = 926$) with three independent mouse PDAC cell lines expressing FUT3 and β3GALT5.

insufficiency is often observed in cases of chronic pancreatitis and can cause severe weight loss and destruction of the acinar compartment in patients (*31*). R;F mice exhibited a significant Dox-dependent decrease in pancreatic mass, whereas no change was observed in C;R^{LSL};F mice and littermate controls after Dox treatment (fig. S14E). In addition, we observed a significant and sustained increase in steatorrhea in R;F mice, whereas C;R^{LSL};F mice exhibited a transient increase in fecal triglycerides that normalized within 4 weeks of Dox treatment (fig. S14F). The pancreatic atrophy and steatorrhea in R;F mice was accompanied by a significant reduction of fecal elastase in R;F mice within the first week of Dox treatment (fig. S14G). Together, these data indicate that R;F mice suffer from pancreatic atrophy and exocrine insufficiency. Both the increased penetrance of CA19-9 expression in the Cre-independent R;F model and expression of CA19-9 in other tissues may contribute to the severe pancreatitis observed in this model.

Human pancreatitis is associated with an influx of macrophages to the pancreas during the acute phase of the disease; after progression to a chronic condition, T and B cells also infiltrate the pancreas (*33*, *34*). To determine whether this aspect of human pancreatitis is also found in mouse models of pancreatitis, we examined the immune infiltrate in C;R^{LSL};F mice and the established cerulein model of acute pancreatitis. Flow cytometry revealed an influx of immune cells into the pancreas, including recruitment of inflammatory monocytes and macrophages, after induction of pancreatitis (Fig. 2E and figs. S15 and S16A). This result was confirmed by IHC (fig. S16B). Neither model showed significant recruitment of T or B cells at the acute pancreatitis time points examined by flow cytometry (1 to 3 days) (fig. S15). At the chronic phase of pancreatitis, both R;F and C;R^{LSL};F models contained intrapancreatic T and B cells as indicated by IHC (fig. S17). Therefore, both the CA19-9 and cerulein mouse models of pancreatitis exhibited immune infiltrates similar to those found in patients with pancreatitis.

Pancreatic cell proliferation is a common feature of human pancreatitis (*35*). Cerulein-treated mice exhibited an elevation in Ki67-positive nuclei in the infiltrating immune cells as well as in the ductal and acinar compartments. Similarly, both R;F and C;R^{LSL};F mice exhibited increased Ki67-positive nuclei in the immune, acinar, and ductal compartments in the acute phases of pancreatitis, which gradually diminished as the mice progressed to chronic pancreatitis (fig. S18, A to E). Increased ethynyl deoxyuridine (EdU) incorporation was also observed in the pancreas, including elevation in the immune, ductal, and

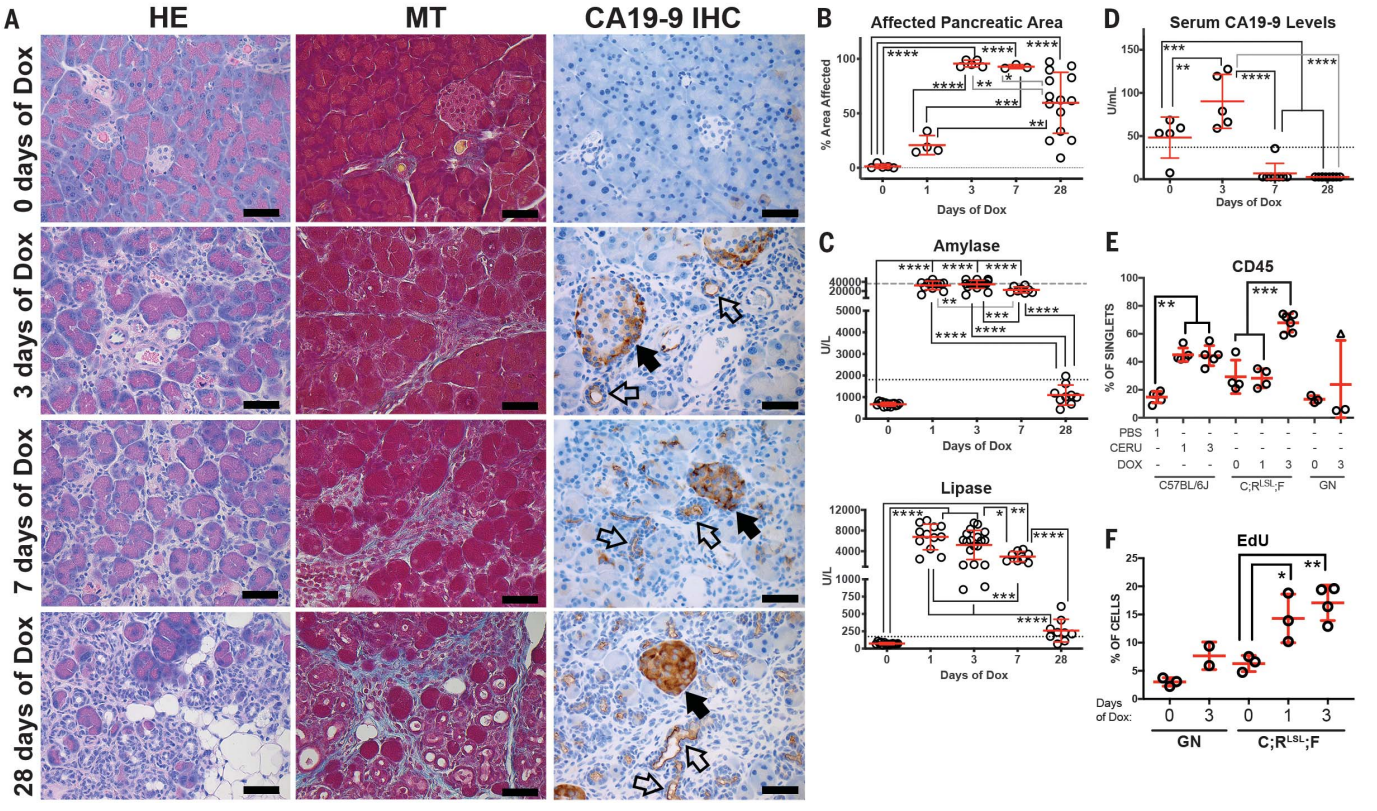


Fig. 2. CA19-9 expression promotes pancreatitis in mice. (**A**) Histologic evaluation of C;R^{LSL};F mice by hematoxylin and eosin (HE) staining, Masson's trichrome staining (MT) (blue indicates collagen deposition), and CA19-9 expression (open arrow, CA19-9^+ duct; closed arrow, CA19-9^+ islet) by IHC after treatment with Dox. Scale bars = 50 μm. (**B**) Quantification of the percentage of pancreatic area exhibiting histologic signs of pancreatitis after treatment of C;R^{LSL};F mice with Dox. (**C**) Circulating levels (U/L) of the pancreatic enzymes amylase and lipase in C;R^{LSL};F mice after treatment with Dox (days). The dotted line indicates the threshold elevation required for the diagnosis of pancreatitis. The dashed line indicates the maximum level of detection possible for amylase. (**D**) The circulating level of CA19-9 (U/ml) after treatment of C;R^{LSL};F mice with Dox. Values that exceed 37 U/ml (dotted line) are elevated. (**E**) Immune cell infiltration evaluated by flow cytometry in mice treated with phosphate-buffered saline (PBS) or cerulein (Ceru) for 2 days followed by a 1- or 3-day recovery period (C57Bl/6j; n = 5, 5, and 5, respectively) and C;R^{LSL};F mice treated with Dox (n = 4, 4, and 6, respectively) compared with "genetically negative" controls (GN) (n = 3 and 3, respectively). Outlier analysis using Grubb's method identified one data point (triangle symbol, GN, 3 days of Dox) as an outlier. (**F**) EdU incorporation in the pancreas was evaluated by flow cytometry in C;R^{LSL};F mice (n = 3, 3, and 4, respectively) and littermate GNs (n = 3 and 2, respectively) after treatment with Dox. Middle horizontal red lines represent the mean, and error bars represent the standard deviation; each data point represents a measurement from an individual mouse. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001 for multiple comparisons using Holm-Sidak's procedure following a one-way analysis of variance.

fibroblast compartments (Fig. 2F and fig. S18F). Together, these data demonstrate that CA19-9–mediated pancreatitis in mice bears similarity to the human disease.

CA19-9 expression results in hyperactivation of epidermal growth factor receptor signaling

To identify the molecular mechanisms underlying the pancreatitis phenotype observed in mice, we focused our studies on the C;R^{LSL};F model. Other researchers have established that epidermal growth factor receptor (EGFR) signaling is necessary and sufficient for the induction of pancreatitis and is important for the initiation of pancreatic cancer in mice (*36*, *37*). Accordingly, we found that the levels of tyrosine phosphorylated and activated EGFR were prominently elevated in the pancreatic ducts of C;R^{LSL};F mice after Dox treatment (Fig. 3A). To identify the alterations in cell signaling pathways that occur in direct response to CA19-9 expression, we isolated pancreatic ductal organoids from C;R^{LSL};F mice (*38*). RNA-sequencing analyses of the C;R^{LSL};F organoids after induction of CA19-9 expression revealed the expected changes in transgene expression (fig. S19, A to F, and table S4). FUT3 and β3GALT5 expression levels were comparable between mouse and human ductal organoids (fig. S19D) (*39*).

Gene set enrichment analysis identified significant changes to 21 pathways in the hallmarks of cancer gene sets and 64 pathways annotated in the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway gene sets (Fig. 3B and table S4). CA19-9 expression triggered a reduction in expression of genes associated with the unfolded protein response and an increase in expression of genes associated with chemokine, hedgehog, Janus kinase–signal transducers and activators of transcription, and transforming growth factor–β signaling. The latter genes may participate in the activation of the fibroblast compartment and recruitment of the immune infiltrate (e.g., interleukin-1a, colony-stimulating factor 1, tumor necrosis factor). We further explored enrichment of the extracellular matrix (ECM)–receptor interaction, ERBB, phosphatidylinositol 3-kinase (PI3K), and AKT signaling pathways (Fig. 3B). CA19-9 expression in C;R^{LSL};F organoids resulted in elevated EGFR phosphorylation (Y1068 and Y1148) and decreased total EGFR, indicating greater flux through this pathway (*40*) (Fig. 3C and fig. S20, A, B, and E to G). The increase in phosphorylated EGFR was accompanied by increased phosphorylation of focal adhesion kinase (FAK) (Y397), AKT, and extracellular signal–regulated kinase 1 and 2 (ERK1/2) (Fig. 3C and fig. S20F). Coincident FAK and EGFR activation through integrin-mediated interaction with the ECM has been previously reported in three-dimensional but not monolayer culture models (*41*, *42*). EGFR stimulation by EGF was more rapid and robust in organoids that were previously cultured in EGF-free conditions, but EGF stimulation did not induce FAK phosphorylation (fig. S20, C and D). Although the alterations of phosphorylated EGFR, total EGFR, and their ratio were consistent among biological replicates, no consistent changes to total or phosphorylated S6, HER2, and p65 were observed (fig. S20, A and B). These findings were independent of the inclusion of murine EGF in the media of the cultures (fig. S20E), and the level of phospho- and total EGFR and downstream effector induction exceeded the increase observed after treatment with additional EGF (fig. S20, B to D).

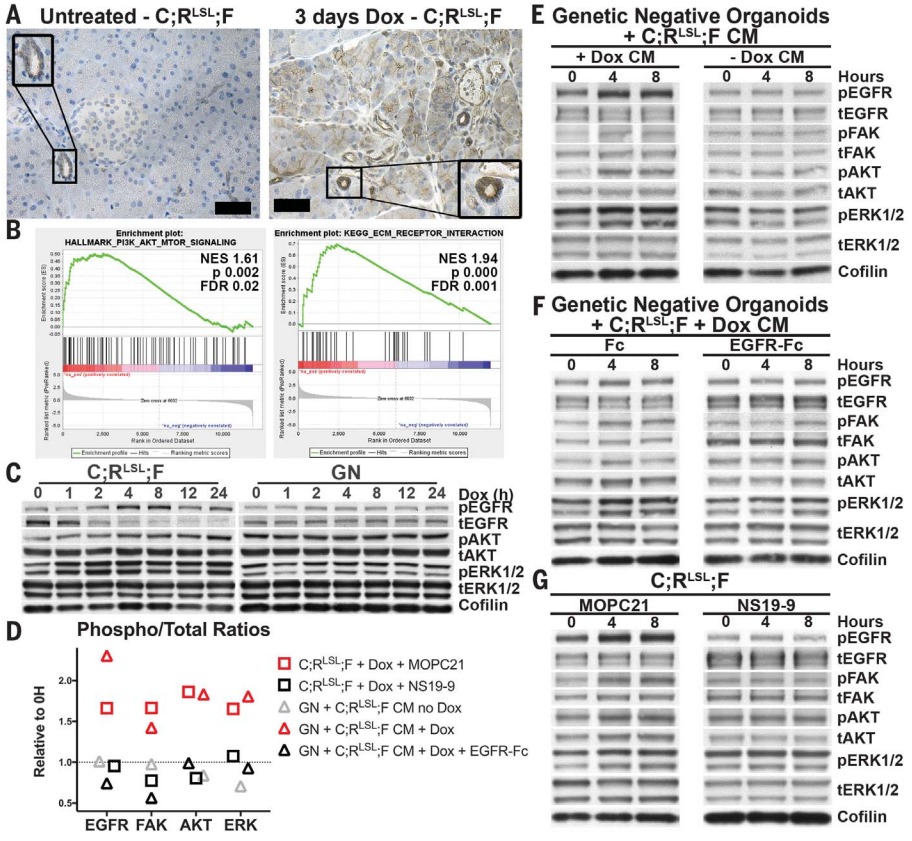


Fig. 3. CA19-9 expression activates EGFR signaling. (**A**) Representative phosphorylated EGFR IHC in C;R^{LSL};F mice that were treated for 0 (n = 4) or 3 (n = 5) days with Dox. Insets are higher magnification of pancreatic ducts. Scale bars = 50 μm. (**B**) Gene Set Enrichment Analysis of C;R^{LSL};F organoids identified enrichment in PI3K/AKT/mTOR and ECM-receptor signaling and downstream effector pathways. FDR, false discovery rate; NES, normalized enrichment score. (**C**) C;R^{LSL};F organoids (n = 3 biological replicates) and GN organoids (n = 2 biological replicates) were evaluated by immunoblot for the activation of signaling pathways after treatment with Dox. Blots are representative of three technical replicates, three biological replicates of C;R^{LSL};F organoids, and two biological replicates of GN organoids. p, phosphorylated; t, total. (**D**) Quantification of changes to the ratio of phosphorylated to total protein is shown for Fig. 3, E to G. (**E** and **F**) Organoids from GN littermates were incubated with conditioned media (CM) from C;R^{LSL};F organoids with or without previous Dox treatment (24 hours) in the presence or absence of Fc control or EGFR-conjugated Fc and evaluated by immunoblot. (**G**) C;R^{LSL};F organoids (n = 3 biological replicates) were treated with Dox (hours) in the presence of isotype control (MOPC-21) or CA19-9–blocking monoclonal antibody (NS19-9) and evaluated by immunoblot. Blots are representative of two technical and three biological replicates.

The glycosylation state of EGFR has been reported to affect its ability to activate and to respond to targeted kinase inhibitors (*43*, *44*). Accordingly, we examined whether EGFR was modified by CA19-9 in mouse organoids, but this was not detected (fig. S20H). To investigate whether the CA19-9–dependent activation of EGFR was due to a soluble ligand, we evaluated conditioned medium from CA19-9–expressing organoids and found that it stimulated EGFR phosphorylation in control murine ductal organoids (Fig. 3, D to F). This activity was attenuated by the addition of EGFR-Fc (EGF trap), further supporting the presence of one or more EGFR ligands (Fig. 3, D and F). Addition of a monoclonal antibody to CA19-9 (NS19-9) was also sufficient to block EGFR phosphorylation in C;R^{LSL};F organoids (Fig. 3, D and G). These findings suggested the presence of a CA19-9–modified, secreted EGFR ligand.

EGFR complexes from C;R^{LSL};F organoid lysates were identified by IP-MS (Fig. 4A and table S5). We observed elevated levels of endogenous fibulin-3 [Efemp1 (FBLN3)] by IP-MS in C;R^{LSL};F, but not control organoids, after Dox treatment (Fig. 4A). FBLN3, a secreted matricellular glycoprotein with five EGF-like domains,

has been proposed to be a ligand for EGFR (*45*). Furthermore, we performed CA19-9 IP-MS from the conditioned media using two different CA19-9 monoclonal antibody clones and identified FBLN3 as a secreted, CA19-9–modified protein (Fig. 4B, fig. S21A, and table S6).

To confirm the glycosylation status of FBLN3, we coexpressed FLAG epitope–tagged FBLN3 in the presence or absence of FUT3 and β3GALT5 in mouse PDAC cells and detected CA19-9 modification of secreted FBLN3 (fig. S21B). The mRNA expression of fibulin family members and EGFR ligands did not change after Dox treatment of C;R^{LSL};F organoids (fig. S21C). Furthermore, the total protein levels of FBLN3 were unchanged after CA19-9 expression in organoid-conditioned media and in mouse plasma in vivo (fig. S21D). Secreted FLAG-FBLN3 isolated by immunoprecipitation coprecipitated with EGFR in PDAC cell lysates, consistent with an association of FBLN3 with EGFR (Fig. 4C). Because the association of ectopic FLAG-FBLN3 with EGFR in cell lysates occurs irrespective of its CA19-9 glycosylation status, the finding that the glycosylation status of endogenous FBLN3 increases association with EGFR may reflect the interaction of sLe^a–modified FBLN3 with additional extracellular proteins. To determine whether FBLN3 is necessary for the activation of the EGFR pathway after CA19-9 expression, multiple FBLN3-blocking antibodies and short-hairpin RNA (shRNA) vectors were utilized and shown to prevent EGFR phosphorylation and the activation of downstream effector pathways in C;R^{LSL};F organoids (Fig. 4, D and E, and fig. S21, E and F).

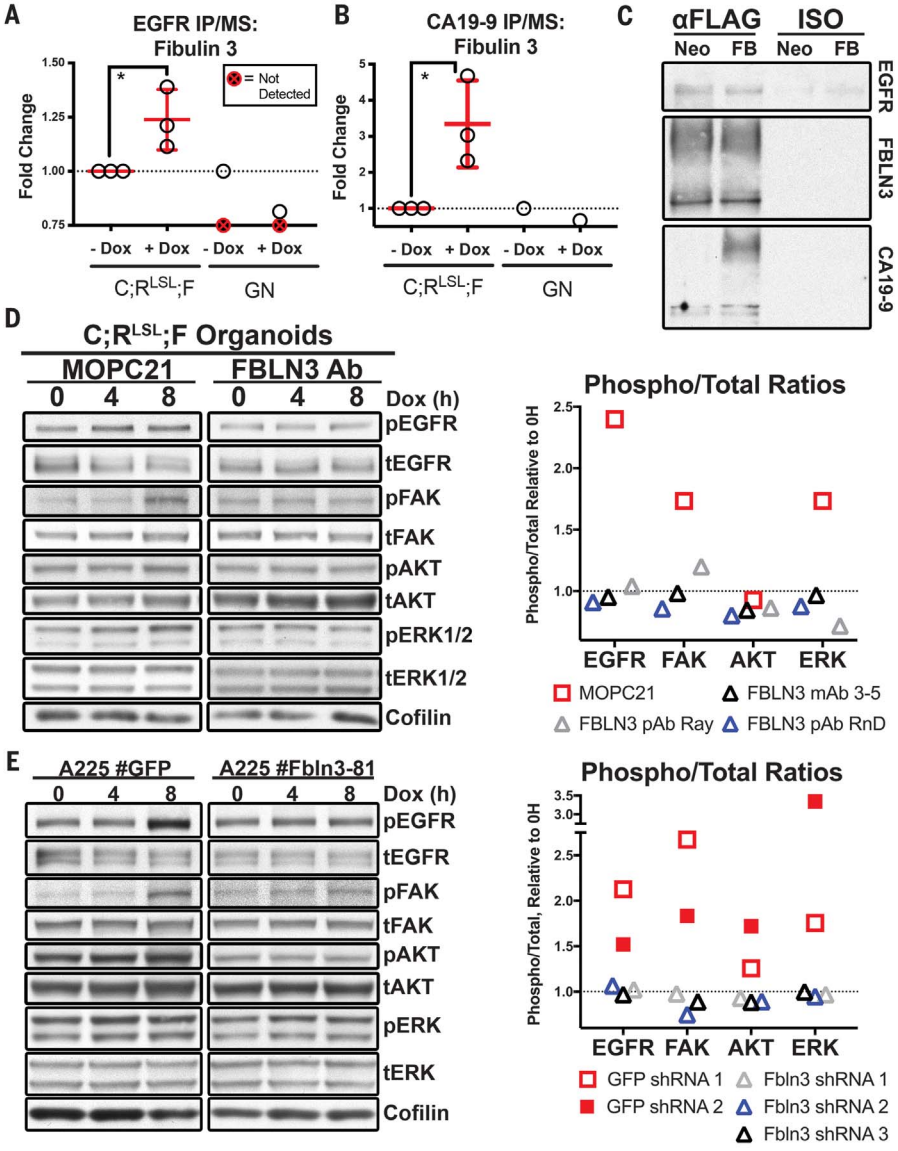


Fig. 4. CA19-9–modified fibulin-3 activates EGFR. (**A**) IP-MS of EGFR complexes from C;R^{LSL};F (n = 3) and GN organoid (n = 2) whole-cell lysates after Dox administration (8 hours). (**B**) CA19-9 protein carrier IP-MS analysis of fibulin-3 from conditioned media from C;R^{LSL};F organoids treated with Dox for 0 to 8 hours (n = 3) relative to untreated and treated GN organoids (n = 1) using 5B1 CA19-9 antibody clone. (**C**) Immunoblot of IP fibulin-3 (FLAG) after incubation with lysates from pancreatic cancer cells lacking FLAG-FBLN3 expression. ISO, isotype control antibody. (**D**) C;R^{LSL};F organoids (n = 3) treated with Dox in the presence of MOPC-21 or three independent fibulin-3 antibodies (FBLN3 Ab) and evaluated by immunoblot. Changes to the ratio of phosphorylated to total protein levels are included. mAb, monoclonal antibody; pAb, polyclonal antibody. (**E**) Organoids transduced with hairpins to GFP or to FBLN3 were immunoblotted after Dox treatment. Changes to the ratio of phosphorylated to total protein levels are included. Middle horizontal red lines represent the mean, and error bars represent the standard deviation. *P < 0.05 by unpaired, two-tailed t test.

CA19-9 as a therapeutic target in pancreatitis

We designed a Dox pulse-chase approach to determine if CA19-9 expression is required to maintain pancreatitis. CA19-9–mediated pancreatitis was completely reversed in C;R^{LSL};F mice after a 3-day Dox pulse and 4-day recovery period (fig. S22, A to C). Partial resolution of chronic pancreatitis in the C;R^{LSL};F model (28-day Dox treatment) was also observed after a 14-day recovery period (fig. S22, D and E). In a preventive setting of acute pancreatitis, two antibodies directed against CA19-9 both reduced immune infiltration, ductal metaplasia, and fibrosis in vivo (Fig. 5A and fig. S22, F and G). In addition, decreased release of amylase and lipase into the circulation and reduced hyperactivation of EGFR were observed (Fig. 5, B and C, and fig. S22, H and I). In an intervention setting of existing acute pancreatitis, we found that two forms of 5B1 significantly reduced the secretion of amylase into the circulation with modest normalization of the pancreatic histology (fig. S22, J to L). CA19-9 antibody treatment of existing acute pancreatitis also reduced the levels of phosphorylated EGFR in both the ductal and acinar compartments and decreased recruitment of macrophages (fig. S22, M and N). These data suggest that CA19-9 plays a role in disease pathogenesis and maintenance and that CA19-9–targeted therapy may warrant further therapeutic exploration.

Inhibition of EGFR in vivo by using erlotinib was not as effective as CA19-9 antibody blockade to mitigate pancreatitis induction in mice (fig. S23, A to F). Indeed, erlotinib treatment of CA19-9–expressing mice caused severe weight loss, necessitating euthanasia. These effects were unrelated to erlotinib toxicity in control mice. Although serum levels of amylase and lipase decreased in erlotinib-treated animals, pancreatic atrophy and acinar cell vacuolization were observed, suggesting that the weight loss were due to increased pancreatitis severity and resulting exocrine insufficiency (fig. S23, B and C). Erlotinib treatment also incompletely blocked phospho-EGFR levels in vivo (fig. S22D). Acinar-ductal metaplasia (ADM) can be detected by SOX9 IHC (*36*) and occurred after 7 days of Dox treatment in the C;R^{LSL};F model (fig. S22E). Although ADM was less apparent in the erlotinib-treated

C;R^{LSL};F mice (fig. S22F), lymphocyte infiltration and fibrosis remained unaffected. CA19-9 sequestration is unlikely to interfere with ADM survival mechanisms in the acinar compartment, given that these cells are largely CA19-9 negative (*46*). Therefore, EGFR kinase inhibition alone cannot substitute for CA19-9 blockade and may be a harmful therapy in the setting of pancreatitis.

Pancreatic tumorigenesis is accelerated by CA19-9–mediated pancreatitis

To determine whether CA19-9 expression promoted PDAC, we intercrossed the C;R^{LSL};F alleles with the conditional Kras$^{LSL\text{-}G12D}$ allele (K;C;R^{LSL};F) (fig. S24A) (*47*, *48*). CA19-9 expression significantly accelerated pancreatic cancer lethality relative to untreated mice and control littermates (Fig. 6A). When treated with Dox, K;C;R^{LSL};F mice rapidly succumbed to primary and metastatic pancreatic cancer with a median survival of 202 days relative to 460 days in the K;C control cohort and 420 days in the untreated K;C;R^{LSL};F cohort. The primary tumors were anaplastic with glandular features (Fig. 6B). Widespread metastases were observed in the peritoneum, diaphragm, liver, and lung in multiple Dox-treated K;C;R^{LSL};F mice. To better

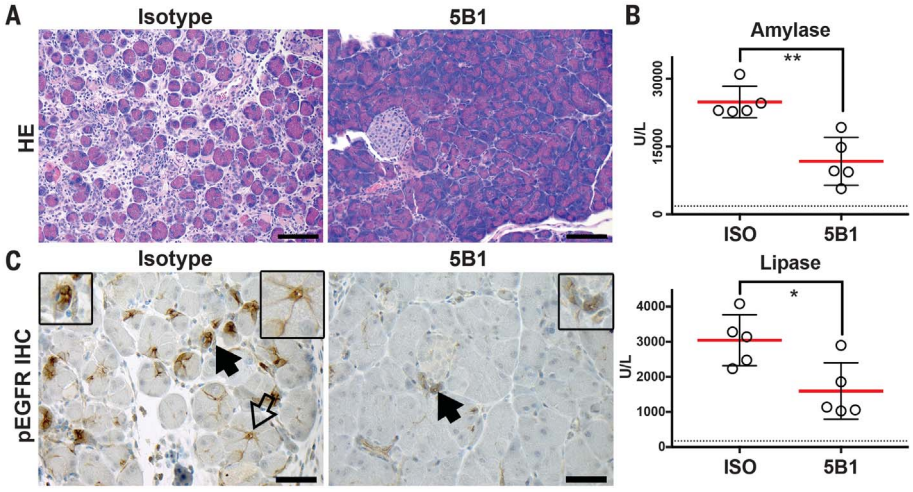


Fig. 5. CA19-9 as a therapeutic target for the treatment of pancreatitis. (**A**) HE staining of representative areas from C;R^{LSL};F mice after treatment with human isotype control (hIgG1; $n = 5$) or the CA19-9 antibody clone 5B1 ($n = 5$) for 8 days and Dox for the last 7 days. Scale bars = 100 μm. (**B**) Serum amylase and lipase levels in C;R^{LSL};F mice treated with 5B1 or isotype control. (**C**) Phospho-EGFR IHC on representative mice treated with either isotype or 5B1 as described in Fig. 5A. Scale bars = 50 μm. Middle horizontal red lines represent the mean, and error bars represent the standard deviation; each data point represents a measurement from an individual mouse. *P* value was determined by using an unpaired, parametric, two-tailed *t* test. $^{*}P < 0.05$, $^{**}P < 0.01$.

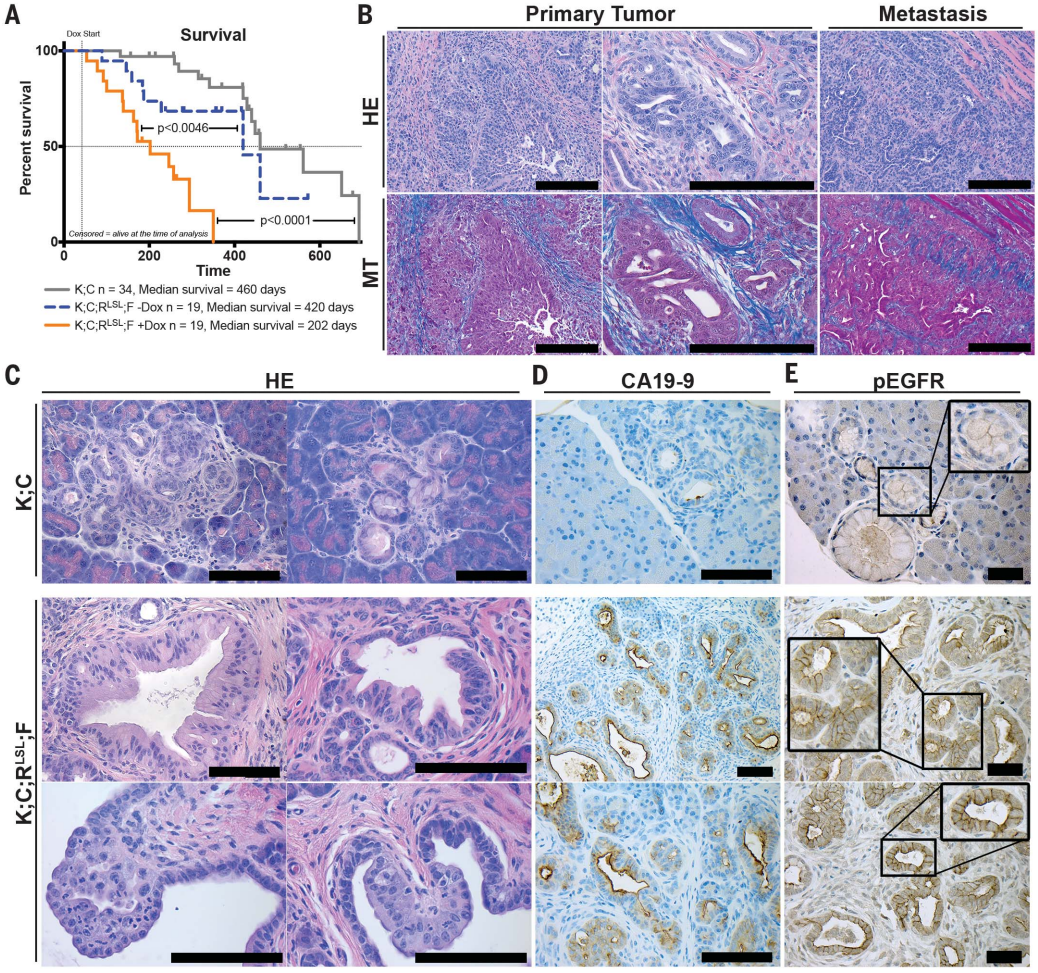


Fig. 6. CA19-9 promotes rapid and aggressive pancreatic tumorigenesis. (**A**) Survival curve for untreated (eight deaths out of 19 mice) and Dox-treated K;C;R^{LSL};F mice (14 deaths out of 19 mice) and K;C GNs (13 deaths out of 34 mice). The *P* value was determined by a log-rank Mantel-Cox test. (**B**) Representative histology of pancreatic tumors and metastatic lesions from Dox-treated K;C;R^{LSL};F mice. Scale bars = 200 μm. (**C**) Representative histology (scale bars = 100 μm), (**D**) CA19-9 IHC (scale bars = 100 μm), and (**E**) pEGFR IHC (scale bars = 50 μm) of the pancreata from K;C and K;C;R^{LSL};F mice after 2 to 4 weeks of Dox treatment.

understand the role of CA19-9 in PDAC initiation, we examined the effect of short-term CA19-9 expression on pancreatic transformation. Whereas littermate controls exhibited the expected low burden of mPanIN-1A lesions, CA19-9–expressing animals harbored a high penetrance of cystic and fibroinflammatory disease with abundant mPanIN-1B and occasional mPanIN-2 lesions after 2 weeks of Dox (Fig. 6C) and increased macrophage infiltration relative to the K;C control cohort (fig. S24, B and C). After 4 weeks of Dox, cystic papillary neoplasia and invasive carcinoma could be detected (Fig. 6C). CA19-9 expression was elevated in normal, benign-reactive, and metaplastic ducts as well as in mPanIN and PDAC lesions in K;C;R^{LSL};F mice, similar to the human expression pattern (Fig. 6D and fig. S12). Equivalent levels of CA19-9 were also observed in human PDAC and mouse K;C;R^{LSL};F organoids (fig. S24D). Phosphorylated EGFR (Fig. 6E) was detected at high levels in Dox-treated K;C;R^{LSL};F mice, whereas CA19-9–negative K;C mice exhibited low to negative levels of phospho-EGFR.

Discussion

CA19-9 is expressed at low levels in the normal ducts of the pancreas but becomes elevated in benign-reactive, metaplastic, and malignant ducts in humans. It is conceivable that the degree of CA19-9 elevation could impart a way to control the degree of fibroinflammatory response in both pancreatitis and PDAC. Mice retain the ability to respond to elevations in CA19-9 despite the fact that they do not normally express this glycan. The ability of mice to respond to CA19-9 elevation may be due to the existence of similar regulatory mechanisms for related Lewis antigens (e.g., sialyl-$Lewis^x$) that are involved in similar processes in other organs and have also been shown to substitute for CA19-9 in individuals lacking this glycan (*49*).

Patients with pancreatitis have an elevated risk (2.7- to 16.5-fold) of developing PDAC, and individuals with hereditary pancreatitis have a 40 to 55% lifetime risk of developing pancreatic cancer (*1*, *5*). Therapeutic options for pancreatitis patients are now focused on treating the symptoms, and little can be done to facilitate the resolution of idiopathic pancreatitis or to prevent its recurrence, highlighting the need for new treatments. Prophylactic intervention could also be beneficial in the setting of recurrent or hereditary pancreatitis and after certain routine procedures for which pancreatitis is a common outcome. For example, 3.5% of the >700,000 patients undergoing endoscopic retrograde cholangiopancreatography each year in the United States will develop pancreatitis (*50*, *51*); several risk factors can be used to identify patients at elevated risk for endoscopic retrograde cholangiopancreatography–associated pancreatitis (40%) (*52*). Fully human CA19-9 antibodies have passed phase 1A clinical trials for positron emission tomography imaging of pancreatic cancer (*53*). This could potentially facilitate rapid translation of CA19-9–targeted therapy to the clinic for the treatment of pancreatitis.

REFERENCES AND NOTES

1. D. Yadav, A. B. Lowenfels, *Gastroenterology* **144**, 1252–1261 (2013).
2. C. E. Forsmark, S. S. Vege, C. M. Wilcox, *N. Engl. J. Med.* **375**, 1972–1981 (2016).
3. J.-L. Frossard, M. L. Steer, C. M. Pastor, *Lancet* **371**, 143–152 (2008).
4. S. E. van Brummelen, N. G. Venneman, K. J. van Erpecum, G. P. VanBerge-Henegouwen, *Scand. J. Gastroenterol. Suppl.* **38**, 117–122 (2003).
5. A. B. Lowenfels *et al.*, *N. Engl. J. Med.* **328**, 1433–1437 (1993).
6. C. Guerra *et al.*, *Cancer Cell* **11**, 291–302 (2007).
7. Y. Zhang *et al.*, *Int. J. Clin. Exp. Med.* **8**, 11683–11691 (2015).
8. C. Haglund, J. Lindgren, P. J. Roberts, S. Nordling, *Br. J. Cancer* **53**, 189–195 (1986).
9. J. Makovitzky, *Virchows Arch. B* **51**, 535–544 (1986).
10. S. H. Itzkowitz *et al.*, *Cancer Res.* **48**, 3834–3842 (1988).
11. K. Iwase *et al.*, *Gastroenterology* **91**, 576–580 (1986).
12. A. Malesci *et al.*, *Gastroenterology* **92**, 60–67 (1987).
13. F. Safi *et al.*, *Pancreas* **2**, 398–403 (1987).
14. W. M. Steinberg *et al.*, *Gastroenterology* **90**, 343–349 (1986).
15. S. S. Pinho, C. A. Reis, *Nat. Rev. Cancer* **15**, 540–555 (2015).
16. M. Thomsen *et al.*, *Br. J. Cancer* **118**, 1609–1616 (2018).
17. C. Shi, N. Merchant, G. Newsome, D. M. Goldenberg, D. V. Gold, *Arch. Pathol. Lab. Med.* **138**, 220–228 (2014).
18. B. Xia, C. L. Feasley, G. P. Sachdev, D. F. Smith, R. D. Cummings, *Anal. Biochem.* **387**, 162–170 (2009).
19. H. Narimatsu, in *Handbook of Glycosyltransferases and Related Genes*, N. Taniguchi, K. Honke, M. Fukuda, Eds. (Springer Japan, 2002), pp. 218–225.
20. N. Ishijima *et al.*, *J. Biol. Chem.* **286**, 25256–25264 (2011).
21. M. Sperandio, C. A. Gleissner, K. Ley, *Immunol. Rev.* **230**, 97–113 (2009).
22. T. Yue *et al.*, *PLOS ONE* **6**, e29180 (2011).
23. T. Yue *et al.*, *Proteomics* **11**, 3665–3674 (2011).
24. Y. Hirao, S. Ogasawara, A. Togayachi, *J. Mol. Biomark. Diagn.* **2**, 124 (2012).
25. L. E. Dow *et al.*, *PLOS ONE* **9**, e95236 (2014).
26. C. Beard, K. Hochedlinger, K. Plath, A. Wutz, R. Jaenisch, *Genesis* **44**, 23–28 (2006).
27. D. C. Whitcomb *et al.*, *Pancreatology* **16**, 218–224 (2016).
28. A. Takada *et al.*, *Biochem. Biophys. Res. Commun.* **179**, 713–719 (1991).
29. A. Takada *et al.*, *Cancer Res.* **53**, 354–361 (1993).
30. S. Majumder, S. T. Chari, *Lancet* **387**, 1957–1966 (2016).
31. M. R. Struyvenberg, C. R. Martin, S. D. Freedman, *BMC Med.* **15**, 29 (2017).
32. P. G. Falk, L. Bry, J. Holgersson, J. I. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1515–1519 (1995).
33. A. Habtezion, *Curr. Opin. Gastroenterol.* **31**, 395–399 (2015).
34. K. S. Inman, A. A. Francis, N. R. Murray, *World J. Gastroenterol.* **20**, 11160–11181 (2014).
35. G. Klöppel, N. V. Adsay, *Arch. Pathol. Lab. Med.* **133**, 382–387 (2009).
36. C. M. Ardito *et al.*, *Cancer Cell* **22**, 304–317 (2012).
37. C. Navas *et al.*, *Cancer Cell* **22**, 318–330 (2012).
38. S. F. Boj *et al.*, *Cell* **160**, 324–338 (2015).
39. H. Tiriac *et al.*, *Cancer Discov.* **8**, 1112–1129 (2018).
40. A. V. Vieira, C. Lamaze, S. L. Schmid, *Science* **274**, 2086–2089 (1996).
41. H. M. Bill *et al.*, *Mol. Cell. Biol.* **24**, 8586–8599 (2004).
42. F. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14821–14826 (1998).
43. Y. C. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11332–11337 (2011).
44. K. Matsumoto *et al.*, *Cancer Sci.* **99**, 1611–1617 (2008).
45. P. Camaj *et al.*, *Biol. Chem.* **390**, 1293–1302 (2009).
46. E. P. Sandgren, N. C. Luetteke, R. D. Palmiter, R. L. Brinster, D. C. Lee, *Cell* **61**, 1121–1135 (1990).
47. E. L. Jackson *et al.*, *Genes Dev.* **15**, 3243–3248 (2001).
48. S. R. Hingorani *et al.*, *Cancer Cell* **4**, 437–450 (2003).
49. H. Tang *et al.*, *Mol. Cell. Proteomics* **14**, 1323–1333 (2015).
50. B. J. Elmunzer, Prevention of ERCP-induced pancreatitis. *Pancreapedia: The Exocrine Pancreas Knowledge Base* (2015).
51. P. J. Parekh, R. Majithia, S. K. Sikka, T. H. Baron, *Mayo Clin. Proc.* **92**, 434–448 (2017).
52. M. L. Freeman *et al.*, *N. Engl. J. Med.* **335**, 909–918 (1996).
53. R. Sawada *et al.*, *Clin. Cancer Res.* **17**, 1024–1032 (2011).
54. J. A. Vizcaíno *et al.*, *Nucleic Acids Res.* **44**, 11033 (2016).
55. R. Edgar, M. Domrachev, A. E. Lash, *Nucleic Acids Res.* **30**, 207–210 (2002).

ACKNOWLEDGMENTS

We thank P. W. Maffuid and J. S. Lewis for the 5B1 antibodies. We acknowledge the assistance of the Cold Spring Harbor Laboratory (CSHL) Animal and Genetic Engineering, Animal and Tissue Imaging, Microscopy, Flow Cytometry, Antibody and Phage Display, and Mass Spectrometry Shared Resources; the New York University (NYU) Experimental Pathology Immunohistochemistry Core Laboratory and Rodent Genetic Engineering Core; the Center for Comparative Medicine and Pathology at Memorial Sloan Kettering Cancer Center; and the University of Cincinnati Mouse Metabolic Phenotyping Center. The TROMA III (CK19) antibody was obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa. **Funding:** This work was supported by the Lustgarten Foundation and Pancreatic Cancer UK. D.A.T. is also supported by the CSHL Association, the CSHL and Northwell Health Affiliation, and the National Institutes of Health (5P30CA45508-29, 5P50CA101955-07, P20CA192996-03, 1U10CA180944-04, 1R01CA188134-01, and 1R01CA190092-04 for D.A.T.; U01CA210240-01A1 for D.A.T. and M.A.H.; PO1CA217798 and P50CA127297 for M.A.H.; 5T32CA148056 and 5K99CA204725 for D.D.E.; and R50CA211506 for Y.P.). In addition, we are grateful for the following support: SWOG ITSC (5U10CA180944-04 for D.A.T. and H.T.). The CSHL Shared Resources are funded by the NIH Cancer Center Support Grant (5P30CA045508). The NYU Experimental Pathology Immunohistochemistry Core Laboratory and Rodent Genetic Engineering Core are supported in part by the Laura and Isaac Perlmutter Cancer Center Support Grant; NIH-NCI (P30CA016087) and the NIH S10 grants; and NIH-ORIP (S10OD01058 and S10OD018338). **Author contributions:** D.D.E. generated and characterized the CA19-9–expressing cells, organoids, and mice; analyzed human tissues; and cowrote the manuscript. H.T. assisted with organoid and immunoblotting experiments and immunohistochemistry. A.P. assisted with EdU and immune cell flow cytometry experiments. K.D.R. and D.J.P. performed LC-MS/MS and analyzed proteomic results. T.E.O. performed immunization experiments. B.A. assisted with KC survival studies. C.S., M.A.Y., B.S., B.D.S., B.C., E.J.L., L.A., K.W., and M.S.L. assisted with animal experiments and organoid generation. P.A.G. assisted with statistical analyses. K.S.P. and S.W. analyzed RNA-sequencing data. K.H.Y. provided human pancreatitis and pancreatic cancer specimens, M.G.G. provided the normal-adjacent pancreas and tumor TMA, and R.H.H. provided pathology diagnoses. C.P., R.G., and D.A. collected human pancreatitis serum and developed the pancreatitis TMA with associated clinical data. Y.P. assisted in the generation of the CA19-9 mouse models and provided oversight on studies involving mice. M.A.H. and D.A.T. conceived the project, directed the experiments, and cowrote the manuscript. **Competing interests:** A provisional patent application, with D.D.E. and D.A.T. as named inventors, that covers the use of anti–CA19-9 antibodies for the treatment and prevention of pancreatitis was filed on behalf of CSHL under a license from CSHL to MabVax Therapeutics; this license has subsequently been transferred to BioNTech Research and Development, Inc. D.A.T. serves on the scientific advisory boards of Leap Therapeutics, Surface Oncology, and Bethyl Laboratory, which are not related to the subject matter of this manuscript. D.A.T. also serves on the Board of Scientific Advisors for the NCI, the Scientific Advisory Board of AACR, the Scientific Advisory Council of Stand Up To Cancer, and the Scientific Advisory Committee of the Georg-Speyer-Haus Institute for Tumor Biology and Experimental Therapy. D.A.T. is a distinguished scholar of the Lustgarten Foundation and director of the Lustgarten Foundation–designated Laboratory of Pancreatic Cancer Research. **Data and materials availability:** The models created and used herein will be available for distribution upon execution of a materials transfer agreement (D.A.T.). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (*54*) partner repository with the dataset identifier PXD008564 (www.ebi.ac.uk/pride/archive/). The RNA-sequencing data have been deposited in NCBI's Gene Expression Omnibus (accession number GSE130854) (*55*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1156/suppl/DC1
Materials and Methods
Figs. S1 to S24
Tables S1 to S6
References (*56*–*70*)

8 December 2018; resubmitted 25 March 2019
Accepted 14 May 2019
10.1126/science.aaw3145

REPORT

PHYSICS

Quantum amplification of mechanical oscillator motion

S. C. Burd[1,2]*, R. Srinivas[1,2], J. J. Bollinger[1], A. C. Wilson[1], D. J. Wineland[1,2,3], D. Leibfried[1], D. H. Slichter[1], D. T. C. Allcock[1,2,3]

Detection of the weakest forces in nature is aided by increasingly sensitive measurements of the motion of mechanical oscillators. However, the attainable knowledge of an oscillator's motion is limited by quantum fluctuations that exist even if the oscillator is in its lowest possible energy state. We demonstrate a technique for amplifying coherent displacements of a mechanical oscillator with initial magnitudes well below these zero-point fluctuations. When applying two orthogonal squeezing interactions, one before and one after a small displacement, the displacement is amplified, ideally with no added quantum noise. We implemented this protocol with a trapped-ion mechanical oscillator and determined an increase by a factor of up to 7.3 (±0.3) in sensitivity to small displacements.

Mechanical oscillators are essential components in an increasing variety of precision sensing applications, including gravitational wave detection (*1*), atomic force microscopy (*2*), cavity optomechanics (*3*), and measurement of weak electric fields (*4*). Quantum mechanically, any harmonic oscillator can be described by a pair of noncommuting observables; for a mechanical oscillator, these are typically position and momentum. The precision of measurement of these observables is limited by unavoidable quantum fluctuations that are present even if the oscillator is in its ground state. Using the method of "squeezing," these zero-point fluctuations can be manipulated while preserving their product as dictated by the Heisenberg uncertainty relation. This squeezing allows for improved measurement precision for one observable at the expense of increased fluctuations in the other (*5*).

Although squeezed states have been created in a variety of physical systems, including electromagnetic fields (*6*), spin systems (*7*), micromechanical oscillators (*8*–*10*), and the motional modes of single trapped ions (*11*, *12*), exploiting squeezing for enhanced metrology has been challenging. In particular, noise added during the detection process will limit the metrological enhancement unless it is smaller than the squeezed noise. The requirement of low-noise detection can be overcome by increasing the magnitude of the signal to be measured. In optical interferometry (*13*) and in spin systems (*14*), it has been shown that reversal of squeezing interactions can magnify small phase shifts, thereby relaxing detection requirements (*15*). Photon field displacements in microwave cavities have also been amplified using similar phase-sensitive amplification schemes (*16*, *17*). However, in mechanical oscillator systems, technical challenges in implementing reversible squeezing interactions have prevented prior use of such methods.

We present a protocol, based on reversible squeezing, for ideally noiseless phase-sensitive amplification (*5*) of mechanical oscillator displacements. This amplification method (Fig. 1) is applicable to any harmonic oscillator where reversible squeezing can be implemented faster than system decoherence. By first squeezing the motional ground state, quantum fluctuations along a particular phase space quadrature are suppressed. A small initial displacement α_i (to be amplified) is then applied along the squeezed axis. At this stage, although the signal-to-noise ratio (SNR) for measuring α_i has been improved by squeezing, resolution below the zero-point fluctuations would require a detection method with yet lower noise. Finally, by reversing the squeezing interaction, the oscillator returns to a minimum-uncertainty coherent state with a larger amplitude $\alpha_f = G\alpha_i$, where G is the gain. Ideally, this process adds no noise in either quadrature. For an oscillator described using creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$, the amplification is given by the identity

$$\hat{D}(\alpha_f) = \hat{S}^\dagger(\xi)\hat{D}(\alpha_i)\hat{S}(\xi) \quad (1)$$

(*18*), where $\hat{D}(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$ is the displacement operator, and $\hat{S}(\xi) = \exp[(\xi^*\hat{a}^2 - \xi\hat{a}^{\dagger 2})/2]$ is the squeezing operator with complex squeezing parameter $\xi(r, \theta) = r\exp(i\theta)$. For arbitrary orientations of the displacement α_i with respect to the initial squeezing axis, $\alpha_f = \alpha_i\cosh(r) + \alpha_i^*\exp(i\theta)\sinh(r)$. Maximum amplification is achieved if the displacement is along the squeezed axis where $\arg(\alpha_i) = \theta/2$, giving $G = \exp(r)$.

We demonstrate this technique using a single trapped $^{25}Mg^+$ ion as the mechanical oscillator (*19*). The ion is held ~30 μm above a linear surface-electrode radio-frequency trap (*20*, *21*), which is cryogenically cooled to 18 K. Experiments are performed on a radial motional mode of the ion with frequency $\omega_r \approx 2\pi \times 6.3$ MHz, energy eigenstates denoted by $|n\rangle$, and zero-point wave function extent of ~5.7 nm (*19*). To analyze the motional state, we use qubit states $|\downarrow\rangle \equiv |F = 3, m_F = 1\rangle$ and $|\uparrow\rangle \equiv |F = 2, m_F = 1\rangle$ within the $^2S_{1/2}$ electronic ground-state hyperfine manifold, where F is the total angular momentum and m_F is its projection along the direction of the quantization magnetic field of approximately 21.3 mT. The qubit transition frequency $\omega_0 \approx 2\pi \times 1.686$ GHz is first-order insensitive to magnetic field fluctuations, giving a qubit coherence time longer than 200 ms (*21*). The qubit state can be manipulated with resonant microwave carrier pulses. In each experiment, the ion is initialized in the electronic and motional ground state $|\downarrow\rangle|0\rangle$ with optical pumping, resolved-sideband laser cooling (*22*), and microwave pulses. Qubit readout is accomplished by applying a laser resonant with the $^2S_{1/2} \leftrightarrow {}^2P_{3/2}$ cycling transition and detecting state-dependent ion fluorescence. We analyze the motional state of the ion by applying sideband interactions to map it onto the qubit states (*11*, *12*). Applying a sideband interaction for various durations results in qubit Rabi oscillations with multiple frequency components whose amplitudes depend on the Fock state populations. We generate these

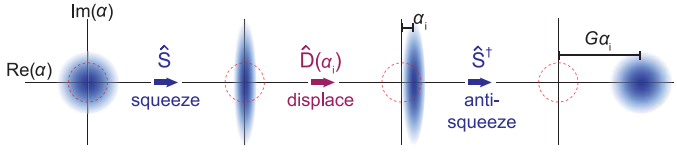


Fig. 1. Conceptual illustration of the amplification protocol. Each panel shows a Wigner function phase space distribution (not to scale) in a frame rotating at the oscillator frequency. A displacement α_i of an initially squeezed ground state is amplified by subsequent reversed squeezing ("anti-squeezing"), resulting in a final coherent state with amplitude $G\alpha_i$ with no added noise. Dashed red circles indicate the characteristic extent of the initial ground-state fluctuations.

[1]Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO 80305, USA. [2]Department of Physics, University of Colorado, Boulder, CO 80309, USA. [3]Department of Physics, University of Oregon, Eugene, OR 97403, USA.
*Corresponding author. Email: shaun.burd@colorado.edu

interactions using oscillating magnetic field gradients (*21*). The blue sideband (BSB) interaction induces transitions between the states $|\downarrow\rangle|n\rangle$ and $|\uparrow\rangle|n+1\rangle$ with Rabi frequencies proportional to $\sqrt{n+1}$. The red sideband (RSB) interaction drives transitions between $|\downarrow\rangle|n\rangle$ and $|\uparrow\rangle|n-1\rangle$ with Rabi frequencies proportional to $\sqrt{n}$ and will not cause a qubit transition if the ion is in $|\downarrow\rangle|0\rangle$.

Squeezing of the motional state is accomplished by applying an oscillating potential at twice the motional frequency ($2\omega_r$) to the radio-frequency electrodes of the trap (*23*). This modulates the confining potential for the ion, yielding the interaction picture Hamiltonian

$$\hat{H} = i\hbar\frac{g}{2}\left[\hat{a}^2\exp(-i\theta) - \hat{a}^{\dagger 2}\exp(i\theta)\right] \quad (2)$$

(*19*), where $\hbar$ is the Planck constant divided by 2π, g is the parametric coupling strength, and θ is the phase of the parametric modulation. Applying this Hamiltonian for duration t implements the unitary squeezing operator $\hat{S}(\xi)$ with $r = gt$. Although electronic parametric modulation has been used with single ions to squeeze a thermal state of motion (*24*) and for phase-sensitive parametric amplification of highly displaced thermal states (*25*), it has not previously been implemented on pure quantum states. Optical forces can also be used for parametric modulation (*11*), but decoherence due to photon scattering and higher-order nonlinearities in the optical field have limited the achievable squeezing (*11*). Squeezed mechanical oscillator states can also be prepared using dissipative reservoir engineering (*8*, *12*). However, this is not a unitary squeezing operation, as is required for the amplification method described here.

We characterize our squeezing process using motional sideband analysis to extract Fock state populations (*19*) (Fig. 2). To characterize the unitarity of our squeezing operations, we measure the ground-state population after squeezing and anti-squeezing, $\langle 0|\hat{S}^{\dagger}\hat{S}|0\rangle$. For $r < 2$, this population is ~0.98, which is consistent with the measured value without squeezing and anti-squeezing ($r = 0$). The population in $n = 0$ remains above 0.93 for $r < 2.37$ (±0.03), or 20.6 (±0.3) dB of squeezing (*19*). The calibrated parametric coupling strength is $g = 2\pi \times 50.2$ (±0.2) kHz, equivalent to a squeezing rate of 2.75 (±0.02) dB/μs.

We use this unitary squeezing interaction to demonstrate amplification of harmonic oscillator displacements (see Fig. 1). Displacements are implemented by applying an oscillating potential resonant with the motional mode (at frequency ω_r) to an electrode of the ion trap (*19*). All control fields are digitally synthesized with the same reference clock, enabling stable and deterministic control of the relative phases between the displacement, squeezing, sideband, and carrier interactions. At each stage of the amplification process, we verify the Fock state composition of the ion's motional state using sideband analysis (Fig. 2, A to C). The measured gain for various values of the squeezing parameter closely follows the theoretically expected exponential growth of the coherent state amplitude (Fig. 2E).

Using this amplification technique, we achieve increased sensitivity when measuring displacements much smaller than the zero-point fluctuations. To map the final displacements α_f onto the qubit states, we use a phase-sensitive red sideband (PSRSB) method (*26*) (Fig. 3A), which reaches the standard quantum limit for $|\alpha_f| \ll 1$ (*19*). Here, a displacement of the motional ground state results in a probability of measuring $|\downarrow\rangle$ of $P_{\downarrow} = \frac{1}{2}[1 - C(|\alpha_f|)\cos\phi]$, where $C(|\alpha_f|)$ is the signal contrast and ϕ is the phase of a carrier $\pi/2$ pulse, which follows the RSB π mapping pulse. For $|\alpha_f| \ll 1$, $C(\alpha_f) \approx 2|\alpha_f|$. In comparison, simply measuring the qubit directly after the RSB π pulse gives a signal $P_{\downarrow} \propto |\alpha_f|^2$. Without amplification, $\alpha_f = \alpha_i$ and the PSRSB contrast is $C(|\alpha_i|)$. With amplification, the initial displacement amplitude α_i is ideally increased by a factor of G and the PSRSB contrast becomes $C(|G\alpha_i|)$. This increase in contrast is shown in Fig. 3B, where the presence of oscillations for the state after amplification indicates that it has a well-defined motional phase. The carrier phase dependence in this figure is a feature of the PSRSB method, not of the amplification protocol. Figure 3C highlights the phase-sensitive nature of the amplification protocol by plotting the contrast C of the PSRSB fringe against the squeezing phase θ for a fixed displacement. Maximum amplification is achieved when the displacement is oriented along the squeezed axis of the initial squeezed state in motional phase space (see Fig. 1). Figure 3D shows the measured signal contrast as a function of $|\alpha_i|$ for various parametric drive durations. For each displacement, the contrast is defined as $C \equiv P_{\downarrow,\max} - P_{\downarrow,\min}$, where $P_{\downarrow,\max}$ and $P_{\downarrow,\min}$ are the maximum and minimum, respectively, of the fringes shown in Fig. 3B. The uncertainty in measuring the contrast is $\sigma(C) = \sqrt{\sigma(P_{\downarrow,\max})^2 + \sigma(P_{\downarrow,\min})^2}$, where $\sigma(P_{\downarrow,\max(\min)})^2$ is the variance of the projection noise associated with measuring $P_{\downarrow,\max(\min)}$. The SNR for a displacement measurement is then $s(G) = C(G\alpha_i)/\sigma[C(G\alpha_i)]$. For a given number of experiments, amplification allows the SNR for a displacement measurement to be improved in comparison to the ideal PSRSB measurement with no squeezing (Fig. 3D, black solid line), giving a measurement sensitivity enhancement of $s(G)/s(G = 1)$. For measurements where $C \lesssim 0.25$, the contrast varies linearly with $|\alpha_i|$, and the gain in contrast $C(G\alpha_i)/C(\alpha_i)$ sets a lower bound (which becomes exact as $|\alpha_i| \to 0$) on the measurement sensitivity enhancement, because the projection noise decreases monotonically with increasing contrast. Increasing the squeezing results in increased contrast for $|\alpha_f| \ll 1$, up to a squeezing time of approximately 8 μs [corresponding to $r = 2.54$

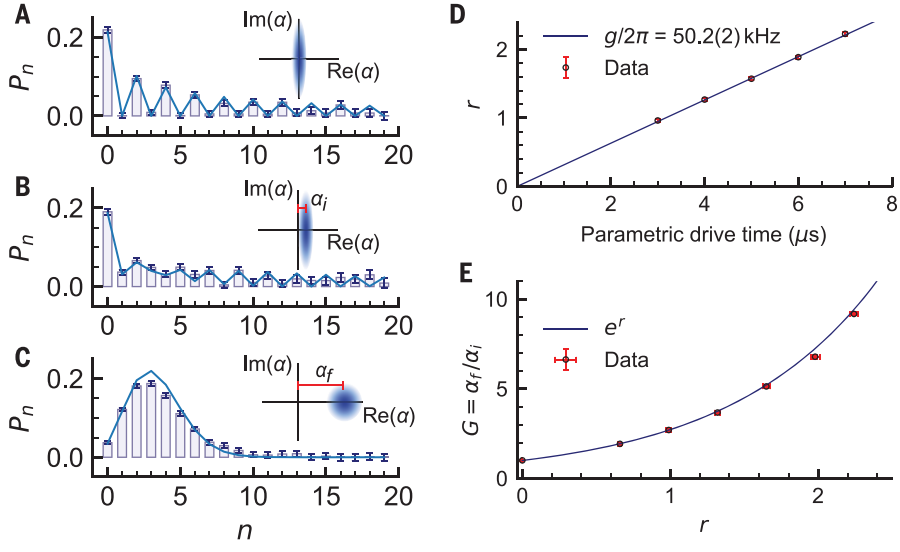


Fig. 2. Fock state population analysis. (**A** to **C**) Histograms of Fock state populations extracted by fitting to BSB Rabi oscillations. Vertical bars are derived by fitting to an unconstrained population distribution. Solid blue lines are fits assuming parameterized functional forms of the ideal Fock state populations, yielding values of r, α_i, and α_f (*19*). Insets show Wigner function illustrations of the corresponding motional states. (A) Initial squeezed motional ground state with $r = 2.26$ (±0.02). (B) After displacing this state by $\alpha_i = 0.200$ (±0.002). (C) Final coherent state with amplitude $\alpha_f = 1.83$ (±0.01), following the reversed squeezing operation. The initial displacement is amplified by $G = \alpha_f/\alpha_i = 9.17$ (±0.09). (**D**) Squeezing parameter r (black circles) as a function of the parametric drive duration. The solid line is a linear fit whose slope gives the parametric coupling strength g. (**E**) Measured gain (black circles) as a function of the squeezing parameter r. The solid line is the theoretical gain $G = \exp(r)$. Error bars denote SD.

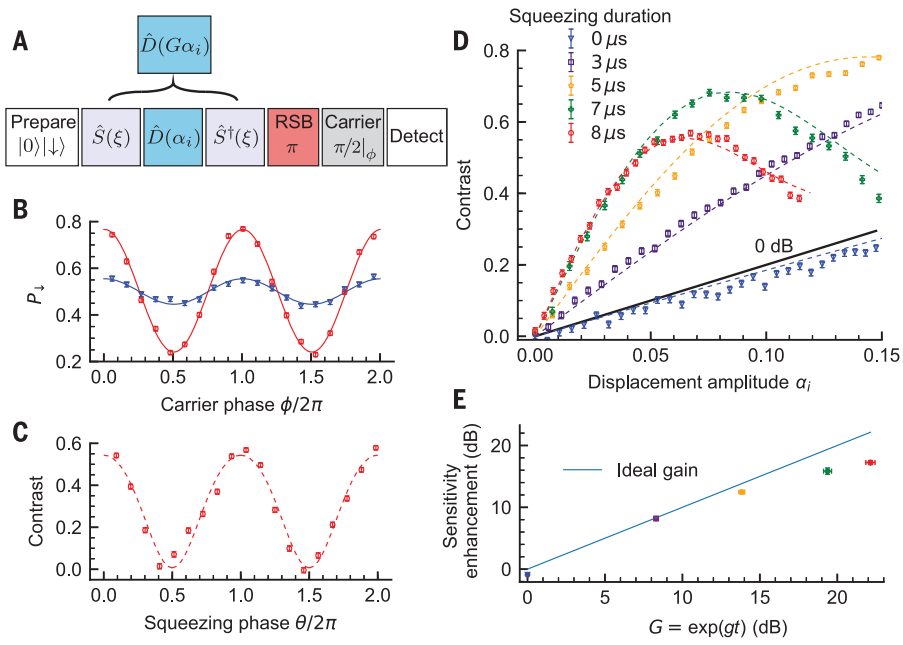


Fig. 3. Measurement sensitivity enhancement. (**A**) Pulse sequence for displacement sensing protocol with PSRSB detection. (**B**) Population in $|\downarrow\rangle$ as a function of the carrier $\pi/2$ pulse phase. Blue inverted triangles, data with no squeezing; red circles, data with amplification. Solid lines show sinusoidal fits to the data. (**C**) Contrast of the carrier phase scan, as shown in (B), as a function of the squeezing phase θ for a fixed displacement. (**D**) Contrast as a function of the displacement amplitude $|\alpha_i|$ for different initial squeezing pulse durations. Each data point is calculated from ~10^4 experiments. The data shown in (B) and (C) have initial $|\alpha_i| = 0.0578$ (±0.0006) and a squeezing duration of $t = 8$ μs [nominally $r = 2.54$ (±0.03)]. The solid black line in (D) is the maximum theoretical contrast without squeezing. Dashed lines in (C) and (D) are derived from a numerical model that includes motional decoherence. (**E**) Measurement sensitivity enhancement in the linear small-displacement regime as a function of the ideal gain $G = \exp(gt)$. For each squeezing duration, the enhancement is determined by dividing the slope of the contrast for $C \lesssim 0.25$ [obtained by fitting a straight line to data points in (D) with $C \lesssim 0.25$] by the slope of the 0 dB black line, which represents the standard quantum limit (*19*). Error bars denote SD.

(±0.03), and ideally 22.0 (±0.3) dB of squeezing]. Here, we achieved a contrast gain of 7.3 (±0.3), corresponding to a factor of 53 (±4) reduction in the number of measurements required to achieve a given SNR, equivalent to a 17.2 (±0.3) dB enhancement in measurement sensitivity. For larger squeezing durations, degradation of the contrast due to background motional heating and dephasing in our trap prevents a further increase in gain. This is not a limitation of the amplification process or our squeezing method. We note that with amplification, we can achieve a SNR of 1 for measuring a displacement of one Bohr radius (~0.0529 nm, corresponding to $\alpha \approx 0.00467$), less than the extent of the ground-state vacuum fluctuations ($\alpha = 0.5$) by a factor of 107, in ~200 experiments.

We have implemented a fast unitary squeezing interaction in a simple mechanical oscillator and used it to amplify and detect coherent motional displacements that are significantly smaller than the quantum zero-point fluctuations. This amplification technique can enhance measurement sensitivity in protocols that use phase-stable displacements, such as photon-recoil spectroscopy (*26*, *27*), where the phase of momentum kicks from photon absorption can be controlled by modulating the photon source. Our method can be extended to amplify displacements of unknown frequency or phase, following the recent proposal in (*28*). The parametric modulation used for squeezing can also be combined with a spin-dependent force to enhance phonon-mediated spin-spin interactions (*28*, *29*), which are used to create entanglement in quantum simulation and quantum information processing experiments. Our methods are also applicable to the generation of exotic nonclassical motional states and to continuous-variable quantum information processing (*30*). Finally, we note that the squeezing, displacement, spin-motion coupling, and qubit control interactions used in this work are all generated without lasers, thereby eliminating spontaneous emission, simplifying control of relative phases, and enabling use with other charged particles lacking optical transitions such as electrons, positrons, and (anti-)protons (*23*).

REFERENCES AND NOTES

1. LIGO Scientific Collaboration and Virgo Collaboration, *Phys. Rev. Lett.* **116**, 061102 (2016).
2. H.-J. Butt, B. Cappella, M. Kappl, *Surf. Sci. Rep.* **59**, 1–152 (2005).
3. M. Aspelmeyer, T. J. Kippenberg, F. Marquardt, *Rev. Mod. Phys.* **86**, 1391–1452 (2014).
4. M. J. Biercuk, H. Uys, J. W. Britton, A. P. VanDevender, J. J. Bollinger, *Nat. Nanotechnol.* **5**, 646–650 (2010).
5. C. M. Caves, *Phys. Rev. D* **23**, 1693–1708 (1981).
6. R. E. Slusher, L. W. Hollberg, B. Yurke, J. C. Mertz, J. F. Valley, *Phys. Rev. Lett.* **55**, 2409–2412 (1985).
7. O. Hosten, N. J. Engelsen, R. Krishnakumar, M. A. Kasevich, *Nature* **529**, 505–508 (2016).
8. E. E. Wollman *et al.*, *Science* **349**, 952–955 (2015).
9. F. Lecocq, J. B. Clark, R. W. Simmonds, J. Aumentado, J. D. Teufel, *Phys. Rev. X* **5**, 041037 (2015).
10. J.-M. Pirkkalainen, E. Damskägg, M. Brandt, F. Massel, M. A. Sillanpää, *Phys. Rev. Lett.* **115**, 243601 (2015).
11. D. M. Meekhof, C. Monroe, B. E. King, W. M. Itano, D. J. Wineland, *Phys. Rev. Lett.* **76**, 1796–1799 (1996).
12. D. Kienzler *et al.*, *Science* **347**, 53–56 (2015).
13. B. Yurke, S. L. McCall, J. R. Klauder, *Phys. Rev. A* **33**, 4033–4054 (1986).
14. E. Davis, G. Bentsen, M. Schleier-Smith, *Phys. Rev. Lett.* **116**, 053601 (2016).
15. D. Linnemann *et al.*, *Phys. Rev. Lett.* **117**, 013001 (2016).
16. M. Malnou *et al.*, *Phys. Rev. X* **9**, 021023 (2019).
17. A. Eddins *et al.*, *Phys. Rev. Lett.* **120**, 040505 (2018).
18. M. M. Nieto, D. R. Truax, *Fortschr. Phys.* **45**, 145–156 (1997).
19. See supplementary materials.
20. S. Seidelin *et al.*, *Phys. Rev. Lett.* **96**, 253003 (2006).
21. C. Ospelkaus *et al.*, *Nature* **476**, 181–184 (2011).
22. C. Monroe *et al.*, *Phys. Rev. Lett.* **75**, 4011–4014 (1995).
23. D. J. Heinzen, D. J. Wineland, *Phys. Rev. A* **42**, 2977–2994 (1990).
24. V. Natarajan, F. DiFilippo, D. E. Pritchard, *Phys. Rev. Lett.* **74**, 2855–2858 (1995).
25. N. Yu, H. Dehmelt, W. Nagourney, *J. Appl. Phys.* **73**, 8650–8652 (1993).
26. C. Hempel *et al.*, *Nat. Photonics* **7**, 630–633 (2013).
27. Y. Wan *et al.*, *Nat. Commun.* **5**, 3096 (2014).
28. C. Arenz, D. I. Bondar, D. Burgarth, C. Cormick, H. Rabitz, arXiv:1806.00444 [quant-ph] (1 June 2018).
29. W. Ge *et al.*, *Phys. Rev. Lett.* **122**, 030501 (2019).
30. C. Flühmann *et al.*, *Nature* **566**, 513–517 (2019).
31. S. C. Burd *et al.*, Data for "Quantum amplification of mechanical oscillator motion". NIST Public Data Repository (2019).

ACKNOWLEDGMENTS

We thank W. Ge, D. Kienzler, and D. M. Lucas for stimulating discussions, and S. M. Brewer and S. S. Kotler for helpful comments on the manuscript. These experiments were performed using the ARTIQ control system. This paper is a contribution of NIST and is not subject to U.S. copyright. **Funding:** S.C.B., R.S., and D.T.C.A. are Associates in the Professional Research Experience Program (PREP) operated jointly by NIST and the University of Colorado, Boulder, under award 70NANB18H006 from the U.S. Department of Commerce, National Institute of Standards and Technology. This work was supported by ARO, ONR, and the NIST Quantum Information Program. **Author contributions:** S.C.B. and D.T.C.A. conceived and carried out the experiments with assistance from R.S. and D.H.S.; the squeezing and amplification protocols were developed by S.C.B. and D.T.C.A. in collaboration with J.J.B. and D.J.W.; D.T.C.A., D.H.S., R.S., and S.C.B. designed and built the apparatus; S.C.B. analyzed the data and wrote the manuscript with input from all authors; and D.T.C.A. and D.H.S. supervised all work, with support from D.L., A.C.W., and D.J.W. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** The data from the main text and supplementary materials are available from the NIST Public Data Repository (*31*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1163/suppl/DC1
Supplementary Text
Figs. S1 to S4
References (*32*–*36*)

17 December 2018; accepted 11 April 2019
10.1126/science.aaw2884

BIOCATALYSIS

Photoexcitation of flavoenzymes enables a stereoselective radical cyclization

Kyle F. Biegasiewicz[1]*, Simon J. Cooper[1]*, Xin Gao[1]*, Daniel G. Oblinsky[1], Ji Hye Kim[1], Samuel E. Garfinkle[1], Leo A. Joyce[2], Braddock A. Sandoval[1], Gregory D. Scholes[1], Todd K. Hyster[1]†

Photoexcitation is a common strategy for initiating radical reactions in chemical synthesis. We found that photoexcitation of flavin-dependent "ene"-reductases changes their catalytic function, enabling these enzymes to promote an asymmetric radical cyclization. This reactivity enables the construction of five-, six-, seven-, and eight-membered lactams with stereochemical preference conferred by the enzyme active site. After formation of a prochiral radical, the enzyme guides the delivery of a hydrogen atom from flavin—a challenging feat for small-molecule chemical reagents. The initial electron transfer occurs through direct excitation of an electron donor-acceptor complex that forms between the substrate and the reduced flavin cofactor within the enzyme active site. Photoexcitation of promiscuous flavoenzymes has thus furnished a previously unknown biocatalytic reaction.

Radical enzymes catalyze radical-mediated reactions with exquisite chemo-, regio-, and enantioselectivity (*1*). These catalysts, however, exploit strategies for radical formation that do not translate well outside of the natural setting and thus do not lend themselves readily to be harnessed for robust synthetic organic chemistry, in which operational simplicity, generality, and ease of reaction execution are desirable (*2*). To achieve these features, it would be attractive to develop biocatalytic reactions that use mechanisms of radical formation commonly used in chemical synthesis. If these mechanisms were to occur within active sites of established enzymatic platforms known to be substrate promiscuous, general catalysts for asymmetric radical reactions could be accessed. Moreover, by repurposing known enzymes, desired biocatalytic functions can be achieved while retaining characteristics such as ease of handling, substrate promiscuity, and evolvability already associated with these catalysts (*3*).

Photoexcitation is widely used in organic synthesis to facilitate radical reactions (*4*) but is less common in enzyme catalysis. Light has been used to activate enzymes through conformational changes (*5*), drive charge separation in multiprotein systems (*6*), and facilitate cofactor turnover or substrate epimerization (*7*). Direct use of photonic energy to drive native, biological reactions is much more limited but has been documented for protochlorophyllide reductase (*8*), fatty acid photodecarboxylase (*9*), and DNA photolyase (*10*). We recently found that nicotinamide-dependent ketoreductases (KREDs) and double-bond reductases (DBRs) can use photoinduced electron transfer to effect asymmetric radical hydrodehalogenation and hydrodeacetoxylation reactions (*11*, *12*). These examples illustrate that irradiation of common oxidoreductases can draw out a non-natural reaction mode for catalysis, a feature we anticipate can be exploited to address challenges in chemical synthesis.

The stereoselective coupling of electrophilic radicals and unactivated alkenes enables the

[1]Department of Chemistry, Princeton University, Princeton, NJ 08544, USA. [2]Department of Process Research and Development, Merck, Rahway, NJ 07065, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: thyster@princeton.edu

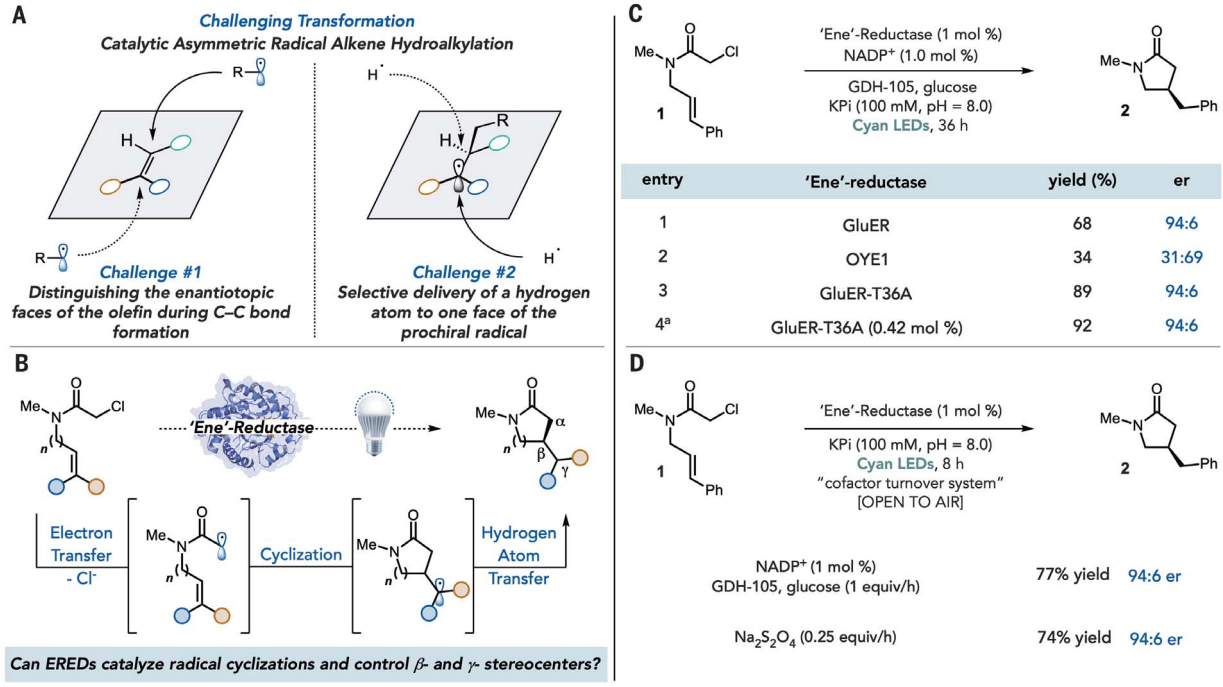


Fig. 1. Strategy for achieving a biocatalytic radical-mediated C–C bond formation. (**A**) Challenges associated with achieving a catalytic asymmetric radical hydroalkylation. (**B**) Proposed catalytic cycle for the biocatalytic radical cyclization to set β- and γ- positions. (**C**) Enzyme screen and optimization by means of mutagenesis. [a]Reaction run by using cell free lysate on 1-g scale. (**D**) Reaction conditions and results for running this reaction open to air.

preparation of structural motifs found in agrochemical and pharmaceutical agents (*13*). Catalytic strategies for rendering this type of transformation stereoselective remain elusive (*14*). This is primarily because of the challenge of maintaining association between radical species and chiral catalysts throughout the stereoselectivity-determining step and the lack of reagents capable of supplying a hydrogen atom to one face of a prochiral radical (Fig. 1A) (*15*). Given these challenges, we imagined that enzymes would be an ideal catalyst for this transformation because of their ability to precisely control the conformational landscape of reactive species. As a model for this family of reactivity, we targeted the development of a biocatalytic radical cyclization of α-chloroamides to afford β-stereogenic lactams (Fig. 1B). The lactam motif is prevalent in medicinally valuable molecules (*16*), and the proposed synthesis would be distinct from existing biocatalytic approaches for generating *N*-heterocycles (*17*). Although this cyclization is well known in the radical literature, it is plagued by preferential formation of the hydrodehalogenated and oligomerized product, and there are no known catalytic asymmetric variants (*18*, *19*). New enzymatic Csp^3–Csp^3 bond–forming reactions are desired in biocatalysis (*20*, *21*).

We looked to flavin-dependent "ene"-reductases (EREDs) as a potential catalyst family for the proposed cyclization of α-chloroamides. Because EREDs have large, substrate-promiscuous active sites, we hypothesized that these enzymes would be able to bind the substrate in a conformation to enable cyclization (*22*, *23*). Moreover, our recent studies demonstrated that these enzymes are able to control the stereochemical outcome of radical hydrodehalogenation reactions (*24*). Unfortunately, flavin hydroquinone (FMN_{hq}) is a modest single-electron reductant [$E_{1/2}$ = −0.45 V versus saturated calomel electrode (SCE)], making electron transfer to α-chloroamides ($E_{p/2}^{red}$ = −1.65 V versus SCE) thermodynamically challenging (fig. S23). The excited state of the flavin hydroquinone (FMN_{hq}*) ($E_{1/2}$* = −2.26 V versus SCE), however, should be capable of accomplishing this initial electron transfer (*25*). Previous work has investigated the fundamental photophysics of the FMN_{hq}* in flavin-dependent enzymes (*26*, *27*). If this excited state could be used for the envisioned chemistry, it would transform flavin-dependent enzymes into chiral photo the envisioned chemistry, substantially expanding their synthetic utility.

We began by testing the cyclization of α-chloroacetamide **1** to afford γ-lactam **2** under visible light irradiation (table S1). An ERED from *Gluconobacter oxydans* (GluER) was effective when irradiated with near–ultraviolet (UV) light (390 nm), furnishing the desired cyclization product with modest yield and enantioselectivity (table S2). Optimization of the lighting source revealed cyan light (497 nm) to provide product with improved yield and enantioselectivity while producing less than 2% of hydrodehalogenation product (Fig. 1C, entry 1). The opposite enantiomer is favored by Old Yellow Enzyme 1 (OYE1) (Fig. 1C, entry 2). Control experiments confirm that light, ERED, nicotinamide adenine dinucleotide

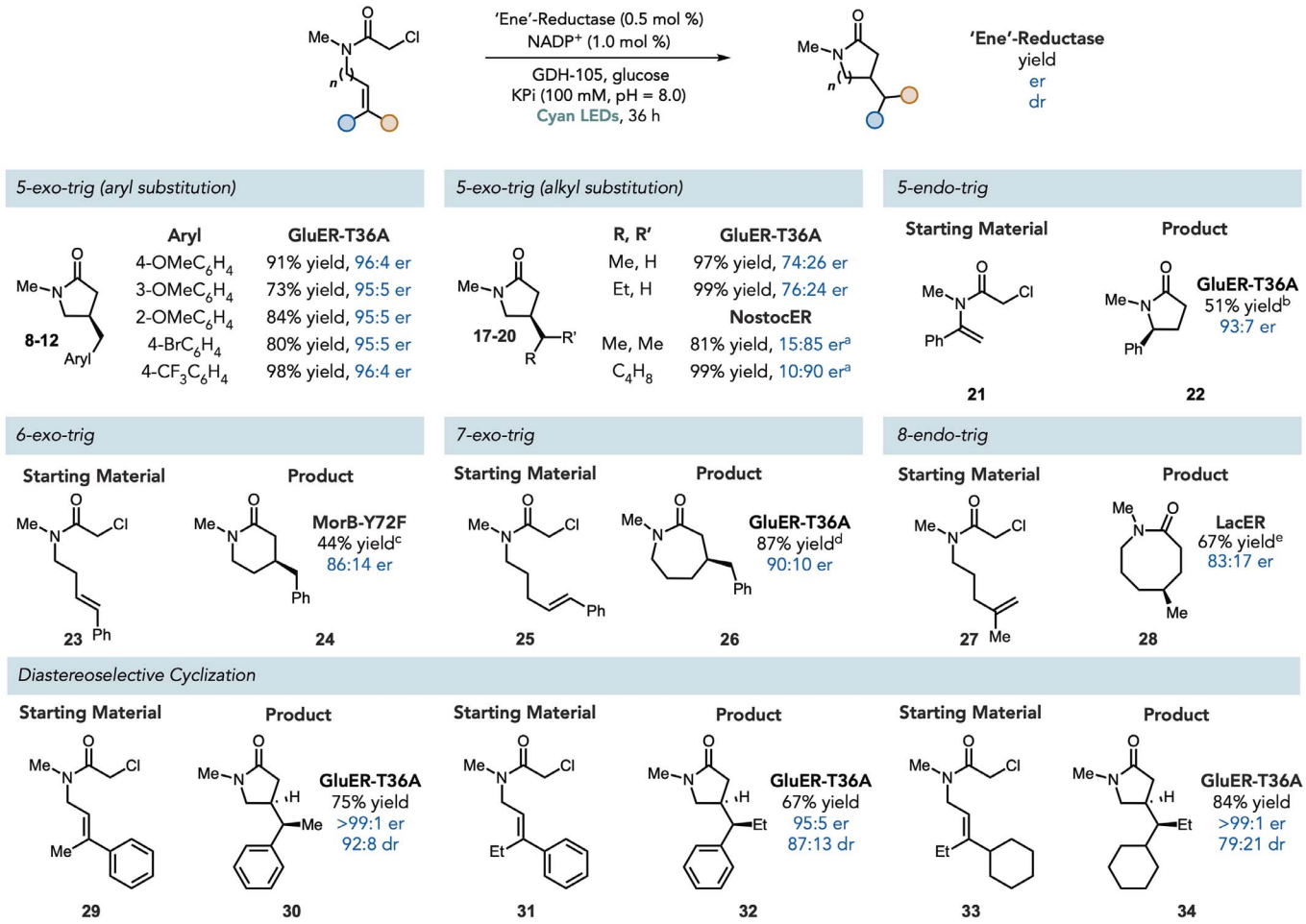


Fig. 2. Scope of the biocatalytic radical cyclization. Reaction conditions: GluER-T36A (0.5 mol %), $NADP^+$ (1 mol %), GDH-105 (0.2 mg/mg starting material), glucose (6 equiv.), KPi (pH = 8.0, 100 mM, with 10% glycerol), and substrate (1 equivalent). 50 W Cyan LED. Average temperature, 35°C. [a]Results obtained by using NostocER and favors the opposite product enantiomer. [b]12% formation of the hydrodehalogenated product. [c]1 mol % enzyme loading (18 hours). [d]2 mol % enzyme loading (18 hours). [e]4 mol % enzyme loading with 8 mol % NAD^+ (72 hours).

phosphate ($NADP^+$), glucose, and glucose dehydrogenase (GDH-105) are required for reactivity (table S1). To increase the activity of GluER for this non-natural function, we conducted mutagenesis and found that mutation of T36, a residue located at the surface of the protein, to alanine (T36A) furnished improved yields of product and only trace quantities of the hydrodehalogenated product (Fig. 1C, entry 3). We solved crystal structures of GluER and GluER-T36A and found no differences in the structure that would explain the improved yield (backbone root mean square deviation of 0.53 Å) (figs. S48 to S52). This mutation does not have a detrimental effect on the native function of this enzyme (fig. S45) and allows the catalyst loading to be decreased to 0.5 mole % (mol %) without compromising yield or enantioselectivity (fig. S46). Moreover, this reaction can be conducted with lyophilized cell-free lysate. This feature enabled the model cyclization to be run on gram scale, providing product with no change in yield or enantioselectivity (Fig. 1C, entry 4). The reaction could be run open to air with slow addition of glucose over 8 hours, providing comparable yields and enantioselectivities to reactions run under inert atmosphere (Fig. 1D). Alternatively, GDH-105, $NADP^+$, and glucose could be replaced with periodic addition of sodium dithionite and degassed buffer; when run open to air, there was no observable change in the reaction outcome (Fig. 1D).

With optimized conditions in hand, we explored the scope and limitations of this reaction (Fig. 2). Aromatic substituted alkenes are tolerated for 5-exo-trig cyclizations, affording product in high yield and enantioselectivity with substituents at the para-, meta-, and ortho positions (Fig. 2, **8-12**). The electronic characteristics of these substituents had only a modest effect on reaction efficiency. Alkyl substituents on the olefin are also tolerated, affording product in nearly quantitative yield in all cases, albeit with diminished enantioselectivities (Fig. 2, **17-20**). Because this mode of reactivity should be accessible across the entire ERED family, we suspected that other members might provide improved levels of enantioselectivity. An ERED from *Nostoc* sp. (NostocER) accordingly provided improved enantioselectivities for the more sterically demanding alkyl-substituted substrates (Fig. 2, **19, 20**).

We next shifted our attention to other cyclization modes. GluER-T36A catalyzes a 5-endo-trig cyclization, in which the stereogenic center is created by means of hydrogen atom transfer (HAT) to yield the 5-substituted γ-lactam (Fig. 2, **22**). The ERED is thus capable of controlling the delivery of a hydrogen atom to sites distal to the carbonyl, a feat largely unknown in the small molecule literature (*10*, *28*). A 6-exo-trig cyclization to furnish δ-lactams occurs with GluER-T36C in modest yield and enantioselectivity. Improved levels of enantioselectivity and yield can be achieved by using a variant of Morphinone reductase (MorB-Y72F) (Fig. 2, **24**). GluER-T36A can also catalyze a 7-exo-trig cyclization (Fig. 2, **26**), albeit with increased catalyst loadings. Last, an 8-endo-trig cyclization is also possible by using GluER-T36A, affording product in high yields but with essentially no enantioselectivity. An ERED from *Lactobacillus casei* (LacER) provides significantly improved levels of enantioselectivity but slightly diminished yields, with the remaining mass balance being unreacted starting material (Fig. 2, **28**). The latter examples are particularly exciting because they are underrepresented cyclization modes in the small molecule literature (*29*).

Inspired by the selectivity of the HAT event, we considered the possibility of ERED-controlled hydrogen atom delivery to exocyclic prochiral radicals. Because these radicals are conformationally flexible, controlling HAT is challenging. We began by testing substrates that possess trisubstituted alkenes for the 5-exo-trig cyclization mode. Substrates that contain phenyl/methyl and phenyl/ethyl substituents at the terminal position of the olefin afforded lactams with excellent enantio- and diastereoselectivities (Fig. 2, **30** and **32**). Because these substrates provide an equimolar mixture of diastereomers when prepared with photoredox catalysts, there does not appear to be a thermodynamic preference for one diastereomer, demonstrating that the enzyme is responsible for controlling the delivery of the hydrogen atom. The cyclohexyl/Et substituted alkene substrate **33**, which lacks an aromatic group, was also a substrate for this reaction, with diastereoselectivity only slightly diminished (Fig. 2, **34**). Substrates that contain substituents with steric bulk may thus be ideal for achieving high levels of diastereoselectivity.

We conducted a series of initial rate experiments to better understand how the kinetic profile of the reaction compares with native alkene reduction. We found that the model reaction follows Michaelis-Menten kinetics with a Michaelis constant (K_m) = 9.7 mM, an apparent unimolecular rate constant (k_{cat}) = 0.012 s^{-1}, and a k_{cat}/K_m of 0.0012 mM^{-1} s^{-1} (fig. S27). On the

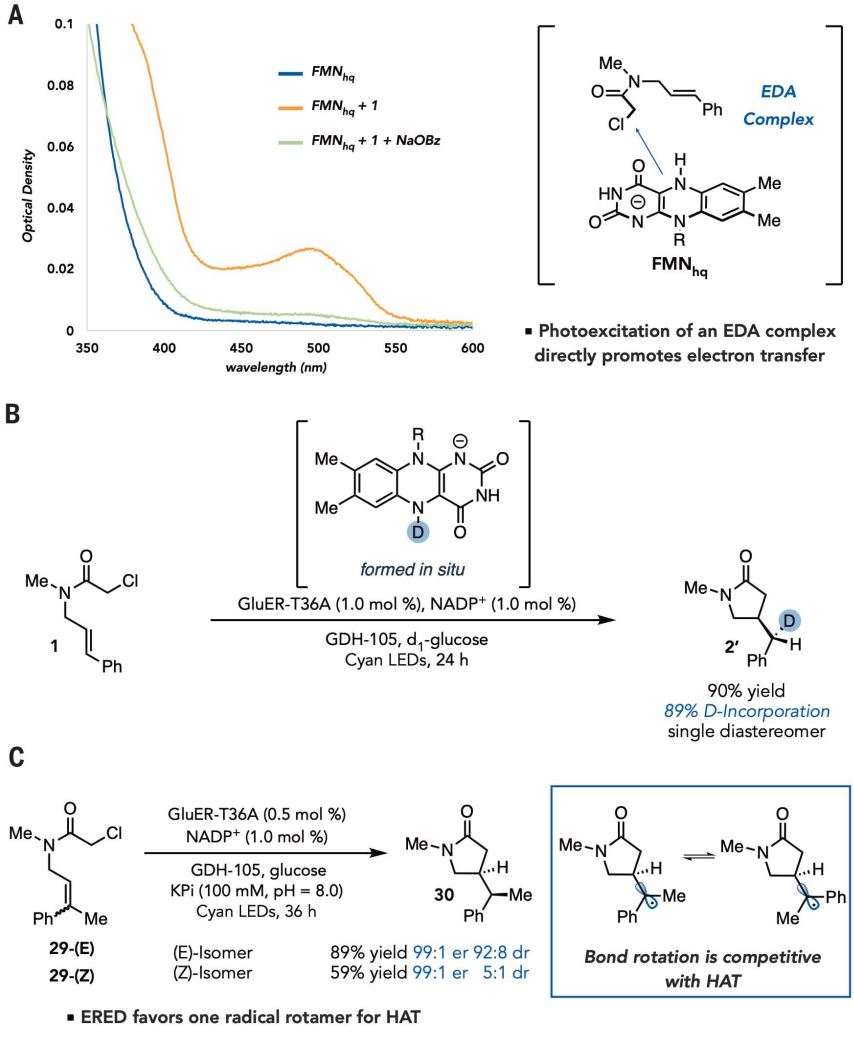


Fig. 3. Mechanistic experiments. (**A**) UV-visible spectrum of reduced GluER-T36A (FMN_{hq}) in the presence of substrate (FMN_{hq} + **1**) and in the presence of substrate and sodium benzoate (FMN_{hq} + **1** + NaOBz). (**B**) Isotope labeling. (**C**) Conformation-selective HAT.

basis of our observation that increased light intensities afford increased rates and higher yields, we suspect that the decreased k_{cat} is due to the reaction being photon limited (fig. S28). We observed a quantum yield of 7.8%.

We conducted a series of mechanistic experiments to better understand the nuances of this reaction, including the superior reactivity with cyan light-emitting diodes (LEDs) [peak maximum (λ_{max}) = 497 nm]. Reduction of GluER to the FMN_{hq} oxidation state with dithionite revealed formation of the diagnostic FMN_{hq} signature with minimal absorption at 497 nm, suggesting that direct excitation of the cofactor is not responsible for the observed wavelength preference (Fig. 3A, FMN_{hq}). Addition of 100 equivalents of chloroamide **1** resulted in formation of a broad absorption band at λ_{max} = 500 nm (Fig. 3A, FMN_{hq} + **1**). Because this absorption feature is not lost upon addition of sodium dithionite, we do not attribute it to FMN_{sq}. It is lost, however, upon addition of 150 equivalents of sodium benzoate (Fig. 3A, FMN_{hq} + **1** + NaOBz). These data are consistent with the formation of an electron donor-acceptor complex between the substrate and FMN_{hq} within the enzyme active site. We propose that excitation of this charge transfer band promotes the initial electron transfer from FMN_{hq} to **1**.

Next, we sought evidence for the existence of radical intermediates by preparing a substrate containing a cyclopropyl ring. GluER-T36A produced only products of cyclopropane ring opening, supporting the existence of a radical intermediate (fig. S9). To determine the terminal hydrogen atom source, we used labeled $FMND_{hq}$ generated in situ from D-glucose-1-d_1 (Fig. 3A). Deuterium was incorporated exclusively at the benzylic carbon, with a high level of diastereocontrol. These experiments support a reaction mechanism in which substrate is initially reduced by one electron after irradiation of the electron donor-acceptor complex formed between the substrate and FMN_{hq} within the enzyme active site. The α-acyl radical can react with the pendent olefin to afford an exocyclic radical that is terminated through hydrogen atom transfer from neutral flavin semiquinone (FMN_{sq}) to afford the product and oxidized flavin (fig. S47).

On the basis of this mechanistic hypothesis, we reasoned that the configuration of the alkene may be responsible for the observed levels of diastereoselectivity if HAT is faster than rotation of the exocyclic C–C bond. We thus performed the reaction on starting material **29** with (Z)-alkene geometry rather than the (E)-isomer. Both alkene geometries favor the same diastereomer and produce no change in the enantioselectivity (Fig. 3C). The enzyme thus favors HAT from one rotamer of the prochiral radical at rates that are competitive with bond rotation. The diminished levels of diastereoselectivity observed with the (Z)-alkene isomer are presumably due to a small degree of hydrogen atom transfer before bond rotation.

Transient absorption spectroscopy with olefin isomers of **29** provides additional support for this mechanistic hypothesis. Initial excitation at 370 nm followed by a broad-band probe pulse reveals underlying dynamics of the flavin redox states. We used a sequential kinetic scheme to fit the data through global analysis. We found three time constants that describe the temporal evolution of the various flavin species through evolutionary associated difference spectra (EADS). When **(E)-29** is used as the substrate, a signal corresponding to the charge transfer state decays in 10 ps to the FMN_{sq}. This time scale is consistent with previously reported rates of decomposition for the α-chloroacetophenone ketyl radical anion (*30*). The neutral FMN_{sq} decays on the time scale of 250 ps to the flavin quinone (FMN_{ox}) (figs. S42 and S43). This decay likely corresponds to the rate of cyclization and termination of the radical through hydrogen atom transfer from FMN_{sq}. When the same experiment is conducted on **(Z)-29**, the FMN_{sq} lifetime increases to 700 ps (figs. S39 and S40). These data are consistent with post-cyclization rotation of the exo-cyclic radical to a conformation in which HAT from the FMN_{sq} to the substrate-centered radical is kinetically facile.

We anticipate that the light-promoted reactivity demonstrated here will be replicable widely in EREDs, KREDs, and other classes of oxidoreductases known in biocatalysis for their promiscuity and adaptability. These flavoenzymes may serve as stereoselective catalysts for unexpected radical transformations beyond those demonstrated here, depending on the active site organization, provided starting material, and properties of the excited state flavin. By more widely investigating and exploiting photoexcited states of cofactors, it should be possible to photoinitiate radical-based reactions within enzymes that are otherwise inaccessible in the ground state.

REFERENCES AND NOTES

1. C. M. Jäger, A. K. Croft, *Chem. Bio. Eng.* **5**, 143–162 (2018).
2. A. Studer, D. P. Curran, *Angew. Chem. Int. Ed.* **55**, 58–102 (2016).
3. U. T. Bornscheuer, R. J. Kazlauskas, *Angew. Chem. Int. Ed.* **43**, 6032–6040 (2004).
4. N. A. Romero, D. A. Nicewicz, *Chem. Rev.* **116**, 10075–10166 (2016).
5. S. Seifert, S. Brakmann, *ACS Chem. Biol.* **13**, 1914–1920 (2018).
6. J. Gao, H. Wang, Q. Yuan, Y. Feng, *Front. Plant Sci.* **9**, 357 (2018).
7. L. Schmermund *et al*., *ACS Catal.* **9**, 4115–4144 (2019).
8. M. Gabruk, B. Mysliwa-Kurdziel, *Biochemistry* **54**, 5255–5262 (2015).
9. D. Sorigué *et al*., *Science* **357**, 903–907 (2017).
10. A. Sancar, *Chem. Rev.* **103**, 2203–2237 (2003).
11. M. A. Emmanuel, N. R. Greenberg, D. G. Oblinsky, T. K. Hyster, *Nature* **540**, 414–417 (2016).
12. K. F. Biegasiewicz, S. J. Cooper, M. A. Emmanuel, D. C. Miller, T. K. Hyster, *Nat. Chem.* **10**, 770–775 (2018).
13. K. Hung, X. Hu, T. J. Maimone, *Nat. Prod. Rep.* **35**, 174–202 (2018).
14. M. P. Sibi, S. Manyem, J. Zimmerman, *Chem. Rev.* **103**, 3263–3296 (2003).
15. Y. Cai, B. P. Roberts, D. A. Tocher, *J. Chem. Soc., Perkin Trans. 1* (11): 1376–1386 (2002).
16. E. Vitaku, D. T. Smith, J. T. Njardarson, *J. Med. Chem.* **57**, 10257–10274 (2014).
17. S. P. France *et al*., *ACS Catal.* **6**, 3753–3759 (2016).
18. D. P. Curran, J. Tamine, *J. Org. Chem.* **56**, 2746–2750 (1991).
19. K. Hiroi, M. Ishii, *Tetrahedron Lett.* **41**, 7071–7074 (2000).
20. N. G. Schmidt, E. Eger, W. Kroutil, *ACS Catal.* **6**, 4286–4311 (2016).
21. N. J. Turner, E. O'Reilly, *Nat. Chem. Biol.* **9**, 285–288 (2013).
22. H. S. Toogood, N. S. Scrutton, *ACS Catal.* **8**, 3532–3549 (2018).
23. K. Heckenbichler *et al*., *Angew. Chem. Int. Ed.* **57**, 7240–7244 (2018).
24. B. A. Sandoval, A. J. Meichan, T. K. Hyster, *J. Am. Chem. Soc.* **139**, 11313–11316 (2017).
25. J. J. Warren, M. E. Ener, A. Vlček Jr., J. R. Winkler, H. B. Gray, *Coord. Chem. Rev.* **256**, 2478–2487 (2012).
26. S. Ghisla, V. Massey, J.-M. Lhoste, S. G. Mayhew, *Biochemistry* **13**, 589–597 (1974).
27. V. Massey, P. Hemmerich, W. R. Knappe, H. J. Duchstein, H. Fenner, *Biochemistry* **17**, 9–16 (1978).
28. Z. G. Brill, H. K. Grover, T. J. Maimone, *Science* **352**, 1078–1082 (2016).
29. L. Yet, *Tetrahedron* **55**, 9349–9403 (1999).
30. D. D. Tanner, J. J. Chen, L. Chen, C. Luelo, *J. Am. Chem. Soc.* **113**, 8074–8081 (1991).

ACKNOWLEDGMENTS

We thank M. Souza for preparing glassware for spectroscopy studies, C. Kraml and L. Wilson at Lotus Separations for compound purification, P. Jeffrey for assistance with x-ray structure determination, K. Conover for assistance in photoNMR data acquisition, and the staff of NSLS-II beamline FMX (17-ID-2) for help with data collection. **Funding:** Research reported in this publication was supported by the National Institutes of Health (NIH) National Institute of General Medical Sciences (NIGMS) (R01 GM127703), the Searle Scholars Award (SSP-2017-1741), Sloan Research Fellowship, the Princeton Catalysis Initiative, and Princeton University. D.G.O. acknowledges support from the Postgraduate Scholarships Doctoral Program of the Natural Sciences and Engineering Research Council of Canada. D.G.O. and G.D.S. acknowledge support from the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences of the U.S. Department of Energy (DOE) through grant DE-SC0019370. The AMX (17-ID) beamline of The Life Science Biomedical Technology Research (LSBR) resource is primarily supported by NIH, NIGMS through a Biomedical Technology Research Resource P41 grant (P41GM111244), and by the DOE Office of Biological and Environmental Research (KP1605010). As a National Synchrotron Light Source II facility resource at Brookhaven National Laboratory, work performed at LSBR is supported in part by the DOE Office of Science, Office of Basic Energy Sciences Program under contract DE-SC0012704 (KC0401040). **Author contributions:** T.K.H. conceived and directed the project. T.K.H., K.F.B., S.J.C., X.G., and J.H.K. designed the experiments. K.F.B., S.J.C., X.G., and J.H.K. performed and analyzed results. B.A.S. cloned the improved variant, and S.E.G. obtained x-ray quality crystals and solved the crystal structures. L.A.J. determined the absolute configuration. D.G.O. designed and conducted the spectroscopic studies, and D.G.O. and G.D.S. analyzed the data. **Competing interests:** The authors declare no conflicts of interest. **Data and materials availability:** All data are available in the main text or the supplementary materials. Crystallographic models and structure factors have been deposited in the Protein Data Bank with accession nos. 6O08 and 6MYW for GluER and GluER-T36A, respectively. X-ray crystallographic data for lactam **2** have been deposited in the Cambridge Crystallographic Data Centre (1878605).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1166/suppl/DC1
Materials and Methods
Figs. S1 to S52
Tables S1 to S3
References (*31*–*56*)

7 December 2018; accepted 29 May 2019
10.1126/science.aaw1143

RADIOTRACER CHEMISTRY

Direct arene C–H fluorination with $^{18}F^-$ via organic photoredox catalysis

Wei Chen[1]*, Zeng Huang[1]*, Nicholas E. S. Tay[2]*, Benjamin Giglio[1], Mengzhe Wang[1], Hui Wang[1], Zhanhong Wu[1], David A. Nicewicz[2]†, Zibo Li[1]†

Positron emission tomography (PET) plays key roles in drug discovery and development, as well as medical imaging. However, there is a dearth of efficient and simple radiolabeling methods for aromatic C–H bonds, which limits advancements in PET radiotracer development. Here, we disclose a mild method for the fluorine-18 (^{18}F)–fluorination of aromatic C–H bonds by an [^{18}F]F^- salt via organic photoredox catalysis under blue light illumination. This strategy was applied to the synthesis of a wide range of ^{18}F-labeled arenes and heteroaromatics, including pharmaceutical compounds. These products can serve as diagnostic agents or provide key information about the in vivo fate of the labeled substrates, as showcased in preliminary tracer studies in mice.

Fluorine-18 (^{18}F) is one of the most important radioisotopes in the radiopharmaceutical industry because it has a relatively long half-life ($t_{1/2}$ = 110 min) and decays with high efficiency by positron emission (97%) (*1*). A primary application of this radioisotope is in the form of 2-[^{18}F]fluorodeoxyglucose ([^{18}F]FDG), which is used for oncological diagnoses, neuroimaging, and studying glucose metabolism (*2*). Uptake of [^{18}F]FDG and other ^{18}F-labeled agents is monitored by positron emission tomography (PET) imaging, which quantifies spatial distributions and metabolic perturbations, as well as site-specific chemical reactivity and ensuing in vivo biological processes (*3*).

Many small-molecule pharmaceuticals and therapeutics contain aromatic or heteroaromatic systems within their framework, thus presenting a common organic subunit for the installation of radioisotopes to yield radiotracers with imaging utility. In particular, the direct conversion of arene C–H into C–^{18}F bonds is ideal owing to the prevalence of aromatic C–H bonds and the increasing importance of C(sp^2)–F bonds in small-molecule therapeutics and probes (*4*, *5*). Direct ^{18}F-fluorination of aromatics currently requires the use of electrophilic fluorine sources, the simplest of which is [^{18}F]F_2; however, this gaseous reagent is incompatible with many common organic functional groups and suffers from low molar activity (the measured radioactivity per mole of compound) as a result of its production methods (*6*–*8*). Most modern methods for C–H to C–F bond conversion require electrophilic fluorine in the form of relatively expensive but bench-stable reagents such as *N*-fluorobenzenesulfonimide (NFSI) and 1-chloromethyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octane-bis(tetra-fluoroborate) (Selectfluor). However, their utility for ^{18}F radiolabeling via electrophilic fluorination is diminished because [^{18}F]NFSI and [^{18}F]Selectfluor are prepared from [^{18}F]F_2 (*9*), which results in even lower molar activities of the ^{18}F-labeled tracers. Therefore, a method for direct conversion of a C–H to a C–^{18}F bond via a nucleophilic arene fluorination strategy is highly attractive because the synthesis of ^{18}F-radiolabeled pharmaceutical compounds can be accomplished with high molar activity fluoride ([^{18}F]F^-). Furthermore, direct C–H to C–^{18}F conversion should alleviate synthetic burdens associated with the synthesis of radiotracer precursors (*10*) and allow for the recovery of precious, unreacted starting material. However, the primary hurdle for this approach is the lack of reactivity for a majority of simple aromatics with [^{18}F]F^-.

Nucleophilic aromatic substitution of electron-deficient arenes has been the standard method for [^{18}F]F^- incorporation (*1*), but prefunctionalization of the aromatic subunit with electron-withdrawing groups is required for this strategy (*3*). Thus, modern radiofluorination methods have sought to generalize the arene scope for this transformation. Current strategies include ^{18}F-deoxyfluorination of phenols via uronium (*11*) and *N*-arylsydnone (*12*) intermediates; displacement of sulfonium salts (*13*); fluorodemetalation of preformed palladium or nickel arene complexes from the requisite aryl halides or boronic acids (*14*, *15*); and copper-mediated cross-coupling of preformed or in situ–generated aryliodoniums (*16*, *17*), aryl boronic acids (*18*), esters (*19*), and arylstannanes (*20*) (Fig. 1A). Despite the advances in radiofluorination, these approaches can be technically challenging for radiochemists. Moreover, the use of metal reagents, especially in stoichiometric or super-stoichiometric amounts, can complicate the quality control process for translational studies because additional analysis on residual metal levels is required.

With these challenges in mind, we sought to develop an arene C–H fluorination method compatible with [^{18}F]F^- (Fig. 1B). The Nicewicz lab has developed a research program using organic single-electron photooxidants to catalytically generate arene radical cations as reactive intermediates for arene C–H functionalization reactions. Thus, we applied this strategy for direct C–H to C–^{18}F bond conversion. Considering the low molar amount of no-carrier-added [^{18}F]F^-, radiation exposure, and the potentially high cost of [^{18}F]F^- production, we decided to first use ^{19}F-fluoride for the development of a photoredox-catalyzed method for arene C–H fluorination. Using diphenyl ether as the aromatic substrate, we were able to obtain 17% of fluorinated adducts in a 13:1 para:ortho ratio using an acridinium-based photooxidant (**1**), cesium fluoride (CsF), the phase-transfer reagent tetrabutylammonium bisulfate (TBA-HSO_4) and 2,2,6,6-tetramethyl-1-piperidine 1-oxyl (TEMPO) as a redox co-mediator under aerobic conditions for 24 hours (Fig. 2 and table S1). We attribute the low yields to the relative recalcitrance of fluoride, which is a Brønsted base and can form strong hydrogen

[1]Biomedical Research Imaging Center, Department of Radiology, and UNC Lineberger Comprehensive Cancer Center, University of North Carolina–Chapel Hill, Chapel Hill, NC 27514, USA. [2]Department of Chemistry, University of North Carolina–Chapel Hill, Chapel Hill, NC 27599, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: nicewicz@unc.edu (D.A.N.); ziboli@med.unc.edu (Z.L.)

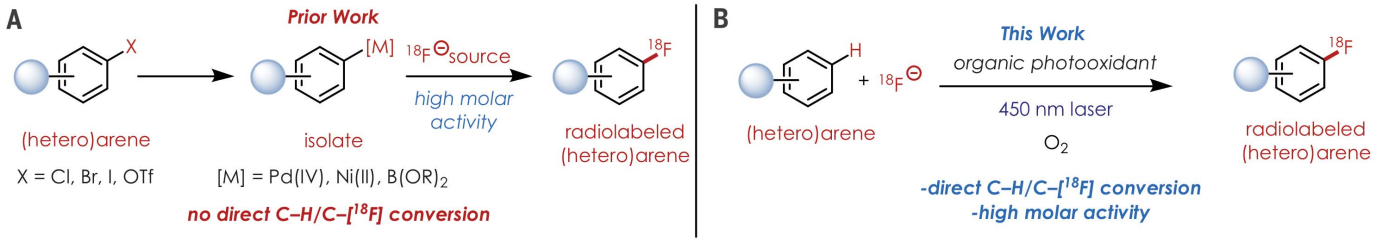


Fig. 1. An organic photoredox approach to ^{18}F labeling of arenes for PET studies. (**A**) Prior work relies on the isolation of aryl organometallic species. (**B**) This study obviates the need for metalation by using organic photoredox catalysis.

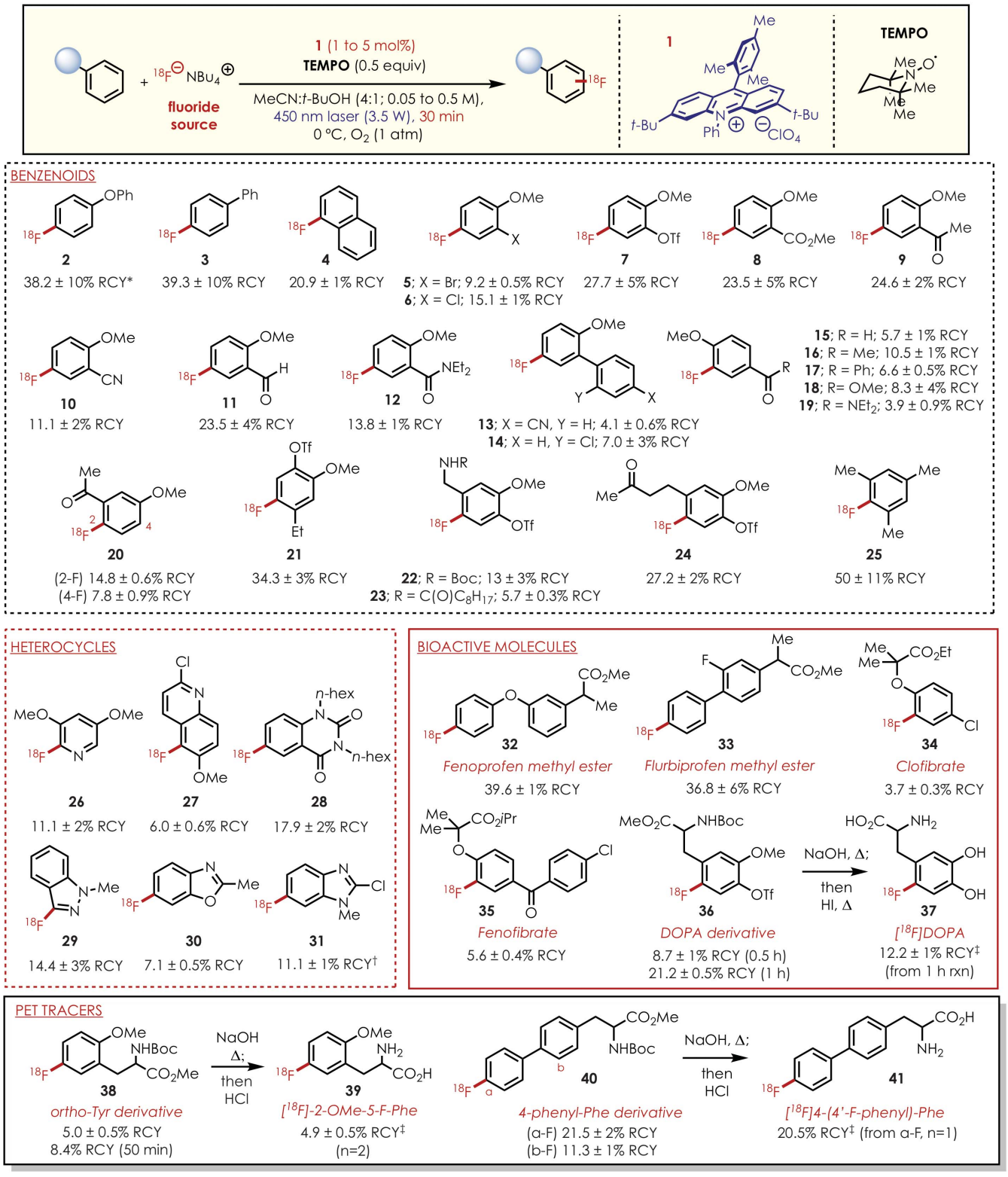


Fig. 2. Reaction scope of ^{18}F-fluorination of aromatics. All RCYs are decay-corrected and averaged over three experiments unless otherwise noted. Asterisk indicates yield averaged over five experiments. Single-dagger symbol indicates that 24.8% RCY of the 2-fluoro dechlorinated product is formed. Double-dagger symbol indicates that the RCYs listed are based on the product deprotection yields (see SM5.6). Bu, butyl; MeCN, acetonitrile; hex, hexyl; iPr, isopropyl; *t*-Bu, *tert*-butyl; Me, methyl; Ph, phenyl; OTf, trifluoromethanesulfonate; Et, ethyl; Boc, butoxycarbonyl.

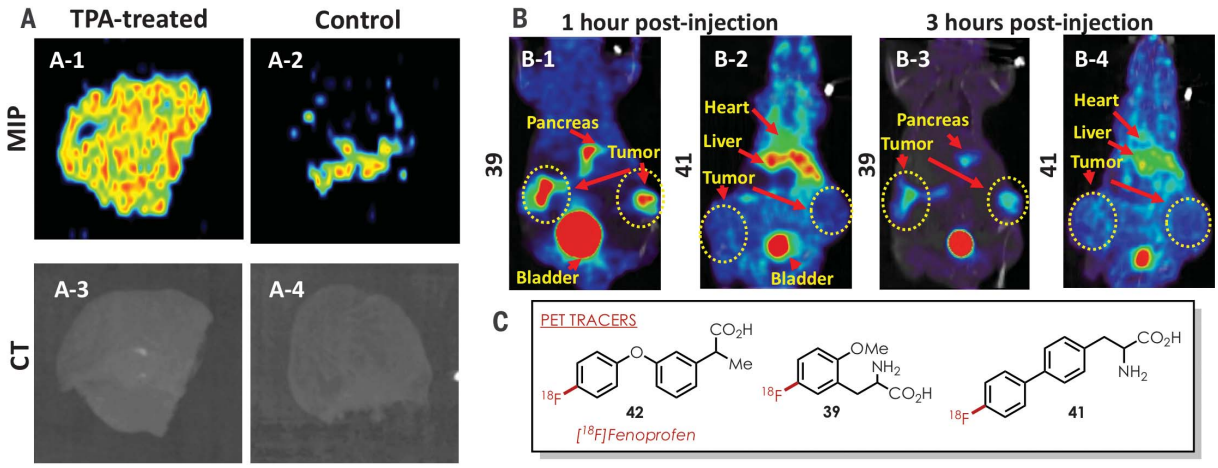


Fig. 3. Examples of PET tracers synthesized via arene C–H radiofluorination. (**A**) Maximum intensity projection (MIP) PET images of [^{18}F]-fenoprofen (**42**) demonstrate higher uptake in TPA-treated mouse ear (A-1) compared with control (A-2) mouse ear. (**B**) PET/CT images demonstrate preferential tumor (MCF-7) accumulation of **39**, compared with longer blood circulation and higher nonspecific binding of **41** at 1 hour after injection. (**C**) Structures of the tracers used in the preceding panels are shown.

bonds in aqueous environments (*21*). Despite lower yields for the aryl C–H fluorination with $^{19}F^-$, C–H fluorination was feasible and thus we decided to extend our method to radiofluorination.

In this new system, [^{18}F]F$^-$ is employed as the limiting reagent, and we envisioned that the relatively higher concentration of arene radical cation would efficiently capture [^{18}F]F$^-$. Transitioning from F$^-$ to [^{18}F]F$^-$ necessitated a reexamination of fluoride sources and irradiation method while aiming to achieve reasonable ^{18}F-labeling yields on 30 min to 1 hour time scales, given the fleeting half-life of the fluorine radioisotope. High molar activity aqueous [^{18}F]F$^-$ is prepared via proton bombardment of [^{18}O] water and subsequent elution of ^{18}F-fluoride with tetrabutylammonium bicarbonate to yield [^{18}F]TBAF, or potassium carbonate complexed with the aminopolyether cryptand Kryptofix 2.2.2 to form [^{18}F]KF-K$_{222}$ (*1*). In both of these systems, excess tetrabutylammonium bicarbonate and potassium carbonate are present. Although the initial radiochemical yields (RCYs) were relatively low (0.57%), [^{18}F]TBAF was the most effective ^{18}F-fluorinating agent in the synthesis of [^{18}F]**2** [see section 5.5 of the supplementary materials and methods (SM5.5)]. Initial RCYs using 455-nm light-emitting diodes were relatively low (0.57%), which prompted us to reevaluate the light flux for the transformation. Using top-down irradiation with a 3.5-W laser (450 nm) (*22*) for 30 min (cooled to 0°C to prevent solvent loss from laser-generated heat) under an aerobic atmosphere afforded a marked increase in the yields of the ^{18}F-fluorinated adducts of diphenyl ether (25.8 and 2.0% for the para and ortho adducts, respectively). After extensive optimization (SM5.5), we identified a system that afforded the para and ortho fluorinated adducts (37.1 and 2.0% yield, respectively) but observed a drop in the molar activity of [^{18}F]-TBAF. We attributed this inconsistency to the exchange of [^{18}F]F$^-$ with fluoride from the BF_4^- counterion of the acridinium catalyst, which in turn lowers the net molar activity of [^{18}F]-TBAF. To circumvent this problem, we synthesized an acridinium containing a perchlorate (ClO_4^-) counterion and found that ^{18}F-labeled diphenyl ether (**2**) adducts could be isolated with a molar activity of 1.37 Ci/μmol (compared with 0.39 Ci/μmol with BF_4^- counterion) and comparable RCYs of 38.2 and 1.8% for the para and ortho products, respectively.

Having identified the optimal catalytic conditions for the transformation, we turned our attention to evaluating the scope of this method with a range of aromatic and heteroaromatic substrates. Biphenyl (**3**) and naphthalene (**4**) were fluorinated at the 4- and 1-positions, respectively, in good RCY. 2-Substituted methoxyarenes (**5** to **11**) were also efficiently fluorinated at the C–H site para to the methoxy group, following a similar trend to the amination (*23*) and cyanation (*24*) methods from our laboratory. Halogenated and pseudohalogenated methoxyarenes 2-bromoanisole (**5**), 2-chloroanisole (**6**), and 2-OTf-anisole (**7**) were all fluorinated at the 4-position in good RCY, thus demonstrating the compatibility of halogenated and pseudohalogenated arenes with our system, as these functional groups are typically susceptible to oxidative addition in transition metal–catalyzed methods. A range of carbonyl-containing functional groups, such as esters (**8**), ketones (**9**), nitriles (**10**), aldehydes (**11**), and amides (**12**), were all compatible with this method, affording single regioisomers of the ^{18}F adducts. Methoxy-substituted biphenyl systems (**13** and **14**) were also competent substrates for fluorination, affording the single regioisomers in moderate RCY. We did not observe significant ortho fluorination for 2-substituted anisoles, which we attribute to a greater gain in positive charge density on the para position of the arene cation radical (*25*) and potential steric hindrance. Thus, we were curious to explore the amenability of our radiofluorination protocol to methoxyarenes bearing substitution at the para position. We found that these 4-substituted congeners (**15** to **19**) were compatible fluorination partners, although the RCYs observed were significantly lower than their 2-substituted counterparts. Nevertheless, the isolated yields for these 4-substituted methoxyarenes are acceptable for PET imaging applications. Preliminary computational studies (figs. S97 and S98) suggest that there is little correlation between the calculated electrophilicities of the cation radical for the 2- and 4-substituted arenes and the observed yields, thus suggesting a greater contribution of steric effects in determining site selectivity. When 3-methoxyacetophenone was subjected to the radiofluorination conditions, a 2:1 ratio of ^{18}F-labeled adducts was obtained with a preference for fluorination para to the methoxy group (**20**). Computational studies suggest that the 6-position gains the most electrophilic character relative to the 2- and 4-positions (fig. S99) upon single electron oxidation; however, only the latter products are observed. Taken together, these results demonstrate a strong preference for the para functionalization of methoxyarenes except in cases where the para position is occupied, and they suggest that steric effects may play a role in dictating site selectivity for C–H fluorination.

Arenes containing guaiacol motifs are found in numerous plant and animal metabolites; thus, the application of our C–H radiofluorination strategy could enable the facile synthesis of ^{18}F radiotracers from renewable sources. We found that ethylguaiacol (**21**), vanillylamine (**22**), nonivamide (**23**), and zingerone (**24**) derivatives were all successfully fluorinated to provide single regioisomers of the ^{18}F analogs. Furthermore, this reaction is not limited to methoxyarenes, as demonstrated by mesitylene undergoing fluorination (<9%) under aerobic conditions to give **25**. This isolated RCY was improved to 50% by employing a modified anaerobic system (using TEMPO as a

net oxidant in MeCN with N_2 bubbling). Fluorinated heterocycles are privileged and highly desirable motifs in pharmaceutical and agrochemical research, and substantial resources are directed toward measuring the pharmacokinetics of these compounds (*26*). Our radiofluorination protocol was applied to several heterocyclic classes: We found that 2,5-dimethoxypyridine (**26**), 2-chloro-6-methoxyquinoline (**27**), and *N*,*N*-dihexylquinazolinedione (**28**) were all successfully fluorinated in good to moderate RCY. Benzazoles common to many therapeutics, such as *N*-methylindazole (**29**), benzoxazole (**30**), and benzimidazole (**31**), all underwent fluorination at the most electrophilic positions of their respective cation radicals. Selective late-stage arene C–H fluorination is an attractive synthetic strategy, as it circumvents the need for complicated labeling precursors and enables the straightforward conversion of bioactive molecules and drugs into PET agents for in vitro companion diagnosis or for pharmacodynamic and pharmacokinetic studies.

We chose to apply our ^{18}F radiofluorination method to several nonsteroidal anti-inflammatory drugs (NSAIDs), which are an important class of pharmaceuticals that alleviate pain and inflammation by inhibiting the activity of cyclooxygenase enzymes (COX-1 and COX-2). Although there have been recent advances in the radiolabeling of COX-1 and COX-2 inhibitors, many examples use ^{11}C as the radionuclide, which has the disadvantage of a shorter half-life ($t_{1/2}$ = 20.2 min) than ^{18}F (*27*). Existing ^{18}F-labeled COX inhibitors typically incorporate ^{18}F in the radioprobe as part of a phenol-appended fluorinated alkyl chain (*28*). However, these functional groups are prone to metabolic degradation and thus may be less-effective radiotracers (*27*, *29*). Fluorination of the aromatic ring is a strategy typically used to study drug metabolism, as fluorine is a hydrogen bioisostere and its substitution slows the metabolic degradation of drug molecules by cytochrome P450 (*29*). Fluorinated aromatics can also act as metabolic tracers because hydroxylated fluoroarene metabolites undergo a 1,2-fluoride shift (NIH shift), thus allowing for the detection and quantification of metabolic byproducts (*30*, *31*). Furthermore, the introduction of fluorine into the aromatic system can improve the potency and cell permeability of drug molecules through noncovalent interactions (*5*). The development of the NSAID celecoxib is an instructive example, in which the substitution of various aryl C–H bonds for aryl C–F bonds was used to bias in vitro COX-2 selectivity (*32*). However, routes for the analogous synthesis of C(sp^2)–^{18}F bonds in COX inhibitors with aromatic moieties are underexplored because of difficulties with designing late-stage precursors for ^{18}F radiolabeling (*33*). Thus, we envisioned that our method would enable the introduction of ^{18}F into known COX inhibitors. The NSAID derivatives fenoprofen methyl ester (**32**), flurbiprofen methyl ester (**33**), and *O*-methyl methyl salicylate (**8**) were all fluorinated in good to moderate RCYs. Given the ubiquitous use of these commercial NSAIDs (*34*), the synthesis of their radiotracer counterparts could provide researchers with a method for visualizing their immediate in vivo metabolic fates that is complementary to the longitudinal metabolism studies enabled by ^{3}H- and ^{14}C-labeling strategies (*35*).

The hypolipidemic agents clofibrate (**34**) and fenofibrate (**35**), as well as a derivative of the biological neurotransmitter precursor *DL*-DOPA (**36**), were selectively fluorinated in moderate RCY after 30 min. We found that the fluorination protocol was influenced by reaction times, as extending the runtime to 1 hour increased the RCY of the fluorinated DOPA derivative to 21.2%. This result is especially noteworthy because [^{18}F]DOPA is an important radioprobe for the PET imaging of CNS disorders (*36*), but published routes to it typically require extensive and sensitive synthesis with ^{18}F precursors (*37*) or the fluorination of prefunctionalized DOPA analogs (*16*, *19*). This fluorinated DOPA derivative was then subjected to facile global deprotection to yield [^{18}F]-DOPA (**37**) in 12.3% RCY. Other aromatic amino acids, such as the protected variants of *O*-Me-*ortho*-tyrosine and 4-phenyl-phenylalanine, were also successfully radiofluorinated (**38** and **40**, respectively), and their deprotected forms (**39** and **41**, respectively) were accessed with relative ease.

A major objective of our synthetic methodology was to develop clinically relevant PET tracers from readily available bioactive molecules without first requiring arene prefunctionalization. To test this concept, we examined the feasibility of converting the NSAID fenoprofen into a PET agent by direct C–H fluorination. Fenoprofen has notable anti-inflammatory activity (*38*), but there have been minimal studies with its fluorinated analogs. Given the widely exploited hydrogen bioisosterism of fluorine in medicinal chemistry (*5*), we were interested in examining the viability of the radiofluorinated analog for PET studies. [^{18}F]-Fenoprofen (**42**) was readily accessed from **32**, which was then used to detect inflammation induced by 12-*o*-tetradecanoylphorbol-13-acetate (TPA) in mouse ears (Fig. 3A and fig. S94) (see supplementary materials and methods for more details) (*39*). Preliminary ex vivo PET studies (Fig. 3A) show significantly higher uptake of [^{18}F]-fenoprofen in the ear inflammation model relative to the control 30 min after intravenous introduction of the radioprobe. These data suggest that **42** is a potential PET agent that demonstrates preferential accumulation in inflamed tissue. Additional biological evaluations for **42** are needed but are currently beyond the scope of this Report.

Another application of our C–H radiofluorination method is rapid radioligand screening in drug discovery and development. We chose to highlight synthetic aromatic amino acids as a class of bioactive metabolites owing to their applicability for oncological PET imaging (*40*–*42*). We are especially interested in the tyrosine scaffold, as fluorination on the aromatic ring is an important functionalization mode for developing PET probes. We selected *O*-Me-*ortho*-tyrosine and 4-phenyl-phenylalanine as the working examples, and our method provides facile access to radiofluorinated **39** and **41**. In vivo PET studies of mice containing MCF7 (breast cancer) and U87MG (glioblastoma) tumor xenografts demonstrate prominent uptake of **39** with minimal accumulation in other organs except the pancreas and bladder (Fig. 3B and fig. S95) (see supplementary materials and methods for more details). Conversely, **41** displayed low uptake in similar tumor models with significantly higher retention in the mouse circulatory system. On the basis of these preliminary studies, **39** shows promise as a selective amino acid radioprobe for tumor detection, and further studies will be needed to examine its biological activity and pharmacology. These results further demonstrate the potential of our radiofluorination method for the discovery of new PET agents that circumvents the need for prefunctionalized (hetero)arenes.

^{18}F radiolabeling is an important tool for noninvasive studies of biological systems, and we anticipate that the applicability of our radiofluorination method to commercial pharmaceuticals and metabolites will enable direct access to new classes of translationally relevant ^{18}F radiotracers, either as diagnostic agents or as target probes for elucidating the in vivo fate of metabolites or pharmaceuticals.

REFERENCES AND NOTES

1. S. Preshlock, M. Tredwell, V. Gouverneur, *Chem. Rev.* **116**, 719–766 (2016).
2. G. J. Kelloff *et al.*, *Clin. Cancer Res.* **11**, 2785–2808 (2005).
3. H. H. Coenen *et al.*, *Nucl. Med. Biol.* **37**, 727–740 (2010).
4. E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly, N. A. Meanwell, *J. Med. Chem.* **58**, 8315–8359 (2015).
5. N. A. Meanwell, *J. Med. Chem.* **61**, 5822–5880 (2018).
6. R. J. Nickles, M. E. Daube, T. J. Ruth, *Int. J. Appl. Radiat. Isot.* **35**, 117–122 (1984).
7. A. Krzyczmonik, T. Keller, A. K. Kirjavainen, S. Forsback, O. Solin, *J. Labelled Comp. Radiopharm.* **60**, 186–193 (2017).
8. H. H. Coenen *et al.*, *Nucl. Med. Biol.* **55**, v–xi (2017).
9. H. Teare *et al.*, *Angew. Chem. Int. Ed.* **49**, 6821–6824 (2010).
10. N. J. Taylor *et al.*, *J. Am. Chem. Soc.* **139**, 8267–8276 (2017).
11. C. N. Neumann, J. M. Hooker, T. Ritter, *Nature* **534**, 369–373 (2016).
12. M. K. Narayanam, G. Ma, P. A. Champagne, K. N. Houk, J. M. Murphy, *Angew. Chem. Int. Ed.* **56**, 13006–13010 (2017).
13. T. Gendron *et al.*, *J. Am. Chem. Soc.* **140**, 11125–11132 (2018).
14. E. Lee *et al.*, *Science* **334**, 639–642 (2011).
15. E. Lee, J. M. Hooker, T. Ritter, *J. Am. Chem. Soc.* **134**, 17456–17458 (2012).
16. N. Ichiishi *et al.*, *Org. Lett.* **16**, 3224–3227 (2014).
17. M. S. McCammant *et al.*, *Org. Lett.* **19**, 3939–3942 (2017).
18. A. V. Mossine *et al.*, *Org. Lett.* **17**, 5780–5783 (2015).
19. M. Tredwell *et al.*, *Angew. Chem. Int. Ed.* **53**, 7751–7755 (2014).
20. K. J. Makaravage, A. F. Brooks, A. V. Mossine, M. S. Sanford, P. J. H. Scott, *Org. Lett.* **18**, 5440–5443 (2016).
21. Y. Sasson *et al.*, *ACS Symp. Ser.* **659**, 148–162 (1997).
22. K. C. Harper, E. G. Moschetta, S. V. Bordawekar, S. J. Wittenberger, *ACS Cent. Sci.* **5**, 109–115 (2019).
23. N. A. Romero, K. A. Margrey, N. E. Tay, D. A. Nicewicz, *Science* **349**, 1326–1330 (2015).
24. J. B. McManus, D. A. Nicewicz, *J. Am. Chem. Soc.* **139**, 2880–2883 (2017).
25. K. A. Margrey, J. B. McManus, S. Bonazzi, F. Zecri, D. A. Nicewicz, *J. Am. Chem. Soc.* **139**, 11288–11299 (2017).

26. E. Vitaku, D. T. Smith, J. T. Njardarson, *J. Med. Chem.* **57**, 10257–10274 (2014).
27. M. Laube, T. Kniess, J. Pietzsch, *Molecules* **18**, 6311–6355 (2013).
28. X. Deng *et al.*, *Angew. Chem. Int. Ed.* **58**, 2580–2605 (2019).
29. B. K. Park, N. R. Kitteringham, P. M. O'Neill, *Annu. Rev. Pharmacol. Toxicol.* **41**, 443–470 (2001).
30. J. Koerts, A. E. M. F. Soffers, J. Vervoort, A. De Jager, I. M. C. M. Rietjens, *Chem. Res. Toxicol.* **11**, 503–512 (1998).
31. P. Shah, A. D. Westwell, *J. Enzyme Inhib. Med. Chem.* **22**, 527–540 (2007).
32. T. D. Penning *et al.*, *J. Med. Chem.* **40**, 1347–1365 (1997).
33. A. Lebedev *et al.*, *PLOS ONE* **12**, e0176606 (2017).
34. B. Cryer, M. Feldman, *Am. J. Med.* **104**, 413–421 (1998).
35. E. M. Isin, C. S. Elmore, G. N. Nilsson, R. A. Thompson, L. Weidolf, *Chem. Res. Toxicol.* **25**, 532–542 (2012).
36. M. Pretze *et al.*, *Nucl. Med. Biol.* **45**, 35–42 (2017).
37. M. Pretze, C. Wängler, B. Wängler, *BioMed Res. Int.* **2014**, 674063 (2014).
38. R. N. Brogden, R. M. Finder, T. M. Speight, G. S. Avery, *Drugs* **13**, 241–265 (1977).
39. M. Madsen *et al.*, *BMC Dermatol.* **16**, 9 (2016).
40. A. Zhu, H. Shim, *Nucl. Med. Mol. Imaging* **45**, 1–14 (2011).
41. Y. Qi, X. Liu, J. Li, H. Yao, S. Yuan, *Oncotarget* **8**, 60581–60588 (2017).
42. A. Sun, X. Liu, G. Tang, *Front Chem.* **5**, 124 (2018).
43. W. Chen, Z. Huang, N. Tay, B. Giglio, M. Wang, H. Wang, Z. Wu, D. Nicewicz, Z. Li, #39 in MCF-7 and U87MG tumor models at different time points, Zenodo (2019); https://doi.org/10.5281/zenodo.3228508.
44. W. Chen, Z. Huang, N. Tay, B. Giglio, M. Wang, H. Wang, Z. Wu, D. Nicewicz, Z. Li, #41 in MCF-7 tumor model at different time points, Zenodo (2019); https://doi.org/10.5281/zenodo.3228551.
45. W. Chen, Z. Huang, N. Tay, B. Giglio, M. Wang, H. Wang, Z. Wu, D. Nicewicz, Z. Li, #42 in ear inflammation model, Zenodo (2019); https://doi.org/10.5281/zenodo.3228546.
46. W. Chen, Z. Huang, N. Tay, B. Giglio, M. Wang, H. Wang, Z. Wu, D. Nicewicz, Z. Li, #42 dynamic PET scan, Zenodo (2019); https://doi.org/10.5281/zenodo.3228534.
47. W. Chen, Z. Huang, N. Tay, B. Giglio, M. Wang, H. Wang, Z. Wu, D. Nicewicz, Z. Li, File index, Zenodo (2019); https://doi.org/10.5281/zenodo.3228554.

ACKNOWLEDGMENTS

Funding: Financial support was provided in part by the National Institutes of Health (NIGMS) awards R01GM120186 (D.A.N.) and 5R01EB014354 (Z.L.) and by the UNC Department of Radiology, Biomedical Research Imaging Center, and UNC Lineberger Comprehensive Cancer Center (start-up fund to Z.L.). N.E.S.T. is grateful for an NSF Graduate Research Fellowship. PET instrumentation was supported via an NIH High-End Instrumentation Grant (1S10OD023611-01). **Author contributions:** W.C. established the final labeling conditions and conducted the majority of the aromatic labeling experiments. Z.H. established the initial labeling conditions and ran a portion of the aromatic labeling scope. N.E.S.T. identified the ^{19}F-fluorination conditions and the catalyst system, synthesized ^{19}F-aromatic standards, and co-wrote the manuscript. B.G. contributed to the initial labeling design and discussion. M.W. performed initial labeling and PET imaging experiments. H.W. established the inflammation and tumor models and performed PET imaging analysis. Z.W. contributed to the initial labeling design and discussion. D.A.N. and Z.L. conceived and supervised the project and experiments and also co-wrote the manuscript. **Competing interests:** N.E.S.T. and D.A.N. are inventors on a patent filed by UNC currently pending (U.S. Patent Application No. 15/826,092). **Data and materials availability:** Experimental procedures, additional data, and analysis are included in the supplementary materials. Data for PET and PET/CT images in Fig. 3 and figs. S94 and S95 have been submitted to Zenodo (*43*–*47*).

SUPPLEMENTARY MATERIALS

11 October 2018; resubmitted 6 February 2019
Accepted 29 May 2019
10.1126/science.aav7019

COLLOIDAL MATERIALS

Particle analogs of electrons in colloidal crystals

Martin Girard[1,2,3]*, Shunzhi Wang[3,4]*, Jingshan S. Du[1,3]*, Anindita Das[3,4]*, Ziyin Huang[1,3], Vinayak P. Dravid[1,3], Byeongdu Lee[5], Chad A. Mirkin[1,3,4]†, Monica Olvera de la Cruz[1,2,3,4]†

A versatile method for the design of colloidal crystals involves the use of DNA as a particle-directing ligand. With such systems, DNA-nanoparticle conjugates are considered programmable atom equivalents (PAEs), and design rules have been devised to engineer crystallization outcomes. This work shows that when reduced in size and DNA grafting density, PAEs behave as electron equivalents (EEs), roaming through and stabilizing the lattices defined by larger PAEs, as electrons do in metals in the classical picture. This discovery defines a new property of colloidal crystals—metallicity—that is characterized by the extent of EE delocalization and diffusion. As the number of strands increases or the temperature decreases, the EEs localize, which is structurally reminiscent of a metal-insulator transition. Colloidal crystal metallicity, therefore, provides new routes to metallic, intermetallic, and compound phases.

The interactions among electrons and atoms to form molecules and materials are foundational in physics and chemistry. However, in the science of colloidal crystals, in which particles are often analogized with atoms, a particle analog to electrons has not been invoked, despite the synthesis of hundreds of colloidal crystals (*1*–*24*) and the development of certain approaches into elaborate forms of crystal engineering (*9*–*21*). In particular, colloidal crystal engineering with DNA has led to the design of structures with diverse symmetries, lattice parameters, and crystal habits (*12*–*22*). However, to date, the particles modified with DNA that define such structures behave as programmable atom equivalents (PAEs) and have fixed particle positions at set stoichiometric ratios. These systems are governed by a set of design rules and the complementary contact model (CCM) (*14*), the premise that they organize themselves to maximize contacts that lead to hybridization and structures such as ionic compounds.

We report on an electron-atom duality analog in colloidal crystal engineering with DNA, in which the resulting colloidal assemblies are better classified as "metallic" structures. In such structures, small DNA-functionalized NPs become mobile and "electron-like" [or electron equivalents (EEs)] and are essential for maintaining the positions of the larger PAE "atom" components. Mixtures of complementary DNA-functionalized nanoparticles (NPs) that vary in size and DNA surface density were assembled and characterized by means of electron microscopy, synchrotron small-angle x-ray scattering (SAXS), and scale-accurate molecular dynamics (MD) simulations with explicit hybridization (*17*, *25*, *26*). Through a combination of theory, simulations, and experiments, we show that small particles grafted with low numbers of DNA strands (for example, <6), when mixed with complementary functionalized NPs (Fig. 1A), form crystals but do not occupy specific lattice sites and diffuse through the crystal in a manner reminiscent of classical electrons in metals, as described by the original Drude model. The PAEs alone will not form crystals because they are almost solely repulsive. The delocalized EEs that move freely through the lattice are responsible for stabilizing it, a type of bonding more reminiscent of metals than ionic compounds (Fig. 1B).

Furthermore, when the interactions are tailored by increasing the number of potential DNA bonds or lowering the temperature, these EEs condense into specific locations, yielding a transition akin to a metal-insulator transition. Last, by taking advantage of this duality and the structural features of the DNA-modified particles that govern it, we realized three polymorphic crystal phases—body-centered cubic (bcc), face-centered cubic (fcc), and Frank-Kasper A15—and analyzed the distribution and diffusion of the particles (EEs and PAEs) within them as a function of temperature and number of linkers per EE.

In a typical set of experiments, 10-nm-diameter Au NPs were densely modified with single-stranded propylthiolated DNA to yield conjugates with ~160 strands per NP. These modified NPs were hybridized with a complementary strand to form a rigid duplex region (18 bases) with a six-base single-stranded overhang (fig. S1). NPs with average diameters of 10, 5, 2, and 1.4 nm, respectively, were modified in a similar manner but with a second type of complementary DNA overhang (Fig. 1A). The average number of DNA overhanging strands available for bonding (number of linkers per EE) is a function of input linker concentrations in the solution (Fig. 1F). For the first three combinations of NPs (10 + 10, 10 + 5, and 10 + 2 nm), all formed the expected CsCl lattice (space group $Pm\bar{3}m$) (*27*) based on the conventional CCM model and the description of them as ionic compound analogs (*14*).

However, with the 10 + 1.4 nm combination, the 10-nm NPs assumed a bcc lattice (space group $Im\bar{3}m$), but the 1.4-nm NPs were invisible to SAXS (Fig. 1C). The lattice assignments were all verified by means of electron microscopy with low-angle annular dark field (LAADF) imaging after the structures were encased in silica (*28*), whereas the 1.4-nm NPs in the 10 + 1.4 nm combination did not appear at specific lattice sites (Fig. 1D). For the first three combinations, the NP positions were determined by the length of the DNA bonding elements that define them (Fig. 1E). However, for the 10 + 1.4 nm combination, there was a marked decrease in the interparticle distance compared with the expected value based on CCM prediction (*14*).

To visualize the positions of the 1.4-nm NPs, we performed cryogenic transmission electron microscopy (cryo-TEM) on the as-synthesized bcc lattice formed from the 10 + 1.4 nm NPs and then stacked the repeating "unit cell" images and EE locations along the [111] zone axis (Fig. 2A). Cryo-TEM showed that the large NPs (PAEs) assumed a bcc lattice, and the small NPs (EEs) were randomized throughout that lattice (Fig. 2B), which is in agreement with the MD simulations (Fig. 2D). In the simulations, crystalline structures were obtained for mixtures of complementary DNA–functionalized Au NPs at a fixed size ratio (10- to 2-nm diameter) but with variable EE:PAE ratios from 4:1 to 12:1 and number of linkers per EE from four to eight (Fig. 2, D and E). Because the MD simulations were performed in the isobaric-isothermal ensemble (NPT) with pressure near zero (supplementary materials), these complementary EEs, which are delocalized from specific lattice sites (Fig. 2D), were responsible for the attraction that holds the large PAEs in crystalline positions in these metal-like assemblies.

We determined the degree of delocalization of EEs in the MD simulations by discretizing the unit cell volumes V_{cell} into $(128)^3$ voxels of equal volume a^3 so that $V_{cell} = (128a)^3$ and then counting the EE visitation frequency in each voxel. This frequency gives a probability distribution f_k in each voxel, k. A quantitative measure of clustering tendency, S_{cl}, is defined as

$$S_{cl} = -\sum_k f_k \ln(f_k) + \ln(V_{cell}/a_0^3) \qquad (1)$$

where a_0 is the average of a over all simulations and used as a constant value to normalize the volume. The quantity S_{cl} can be associated with

[1]Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA. [2]Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA. [3]International Institute for Nanotechnology, Northwestern University, Evanston, IL 60208, USA. [4]Department of Chemistry, Northwestern University, Evanston, IL 60208, USA. [5]X-ray Science Division, Advanced Photon Source, Argonne National Laboratory, Lemont, IL 60439, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: m-olvera@northwestern.edu (M.O.d.l.C.); chadnano@northwestern.edu (C.A.M.)

an information entropy (*29*). This quantity is minimized if the EEs are localized and maximized if they are delocalized (when they have a uniform distribution). The resulting S_{cl} shows a strong correlation with both EE:PAE ratio and number of linkers per EE (Fig. 2F): Either an increase in the number of EEs in the lattice or a decrease in the number of duplexed *DNA* linkers on the EEs resulted in a more randomized density distribution of EEs in the lattice.

The EEs in the PAE-EE assemblies are classical particles and as such follow a Boltzmann distribution. Thus, the movement of EEs in time is related to free-energy barriers that allow for an exponential relationship between their trapping time and temperature (Fig. 2G), which is directly related to the degree of EE delocalization. To visualize the spatial distribution of EEs from MD simulations, we calculated the cumulative density of EEs by integrating f_k from its maxima (where the density of EEs is the highest) and drew isosurfaces that separate the unit cell into equally probable accessible volumes for the EEs, termed Boltzmann volumes (supplementary materials). The Boltzmann volumes for an assembly with a low number of DNA linkers per EE were widely dispersed across the volume of the crystal (Fig. 2H). Increasing the number of linkers per EE resulted in the localization of EEs into a group of locations near the B sites in AB_6-type binary compounds (Fig. 2I and Table 1). This localized state is similar to the electron charge-density distributions in semiconductors and insulators (*30*). Furthermore, the Boltzmann volumes became more dispersed as the EE:PAE ratio increased (Fig. 2J and fig. S33). Such a response was also experimentally observed by comparing the projected EE locations determined with cryo-TEM on samples of bcc assemblies with varying input EE:PAE ratios and EE linker concentrations in the solution. The EE local density in the crystal with a low EE:PAE ratio and high linker concentration (Fig. 2C) around the predicted localization sites was substantially higher than the uniform distribution baseline and than the local densities in the assemblies with higher EE:PAE ratios and lower linker concentrations (Fig. 2B and fig. S17).

Compared with PAEs that reside on relatively fixed lattice sites, EEs in a PAE-EE assembly are macroscopically mobile beyond local vibrations. Time-series MD simulation snapshots showed that the EEs could diffuse between unit cells (Fig. 3A, fig. S39, and movie S4). To further probe the diffusion of EEs in experimentally realized assemblies, 10-nm PAEs and Cy5-DNA–labeled 1.4-nm EEs were assembled in the bcc structure. Subsequently, these crystals were incubated with a solution of Cy3-DNA–labeled EEs. To track any exchange of EEs between the crystalline assembly and EEs in solution, the ultraviolet-visible (UV-vis) extinction spectra of the supernatant were measured over time (Fig. 3B, top). A simultaneous increase of Cy5 signal and reduction of Cy3 signal suggested that the EEs were highly mobile and could diffuse macroscopically be-

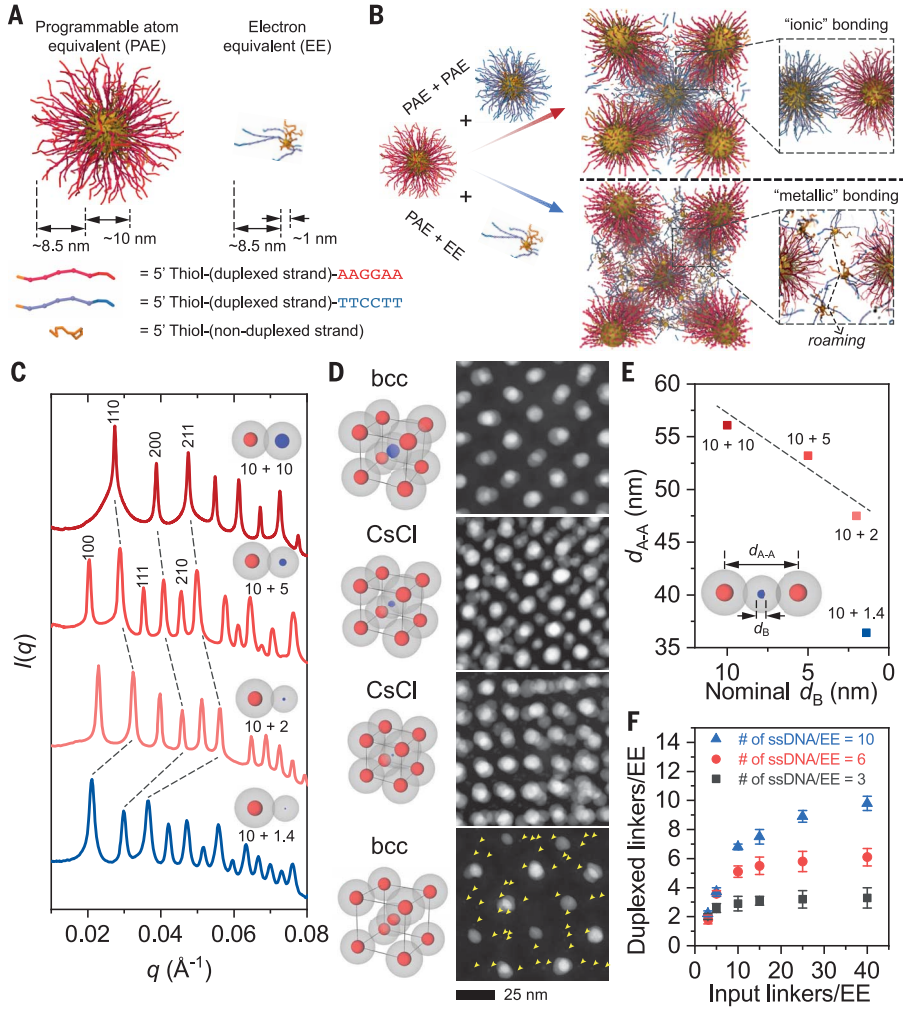


Fig. 1. Transition from PAE-PAE systems to PAE-EE systems. (**A**) Illustrations of DNA-functionalized Au NPs behaving as programmable atom equivalents (PAEs) or electron equivalents (EEs) used in the MD simulation. (**B**) Snapshots from the MD simulation depicting "ionic" bonding behavior shown by PAE + PAE assemblies, and "metallic" bonding behavior shown by PAE + EE assemblies, where roaming EEs hold the crystal of repulsive PAEs together. (**C** and **D**) Four crystalline lattices assembled from 10-nm PAEs and complementary DNA–functionalized Au NPs (nominal core diameters of 10, 5, 2, and 1.4 nm, respectively). Shown are (C) SAXS spectra, (D) models, and cross-sectional LAADF images of silica-encapsulated samples. Scale bar, 25 nm. The 1.4-nm Au NPs in (D) (yellow arrows indicate visually identified ones) are dispersed randomly in the lattice and do not occupy specific lattice sites. (**E**) SAXS-determined distance between bonding 10-nm PAE pairs (same DNA type, defined in the inset according to CCM assumptions). (**F**) Quantification of linker DNA strands duplexed on 1.4-nm Au NPs (EEs) as a function of the input number of linkers per EE in the solution.

Table 1. Symmetry of PAE-EE assemblies in the fully localized state. Wyckoff positions in square brackets have higher energies than the ground-state configurations.

PAE lattice	Space group	Localized EE Wyckoff positions
bcc	$Im\bar{3}m$	12d
fcc	$Fm\bar{3}m$	8c, 32f
A15	$Pm\bar{3}n$	6c, 16i, 24k, [12f], [24j]

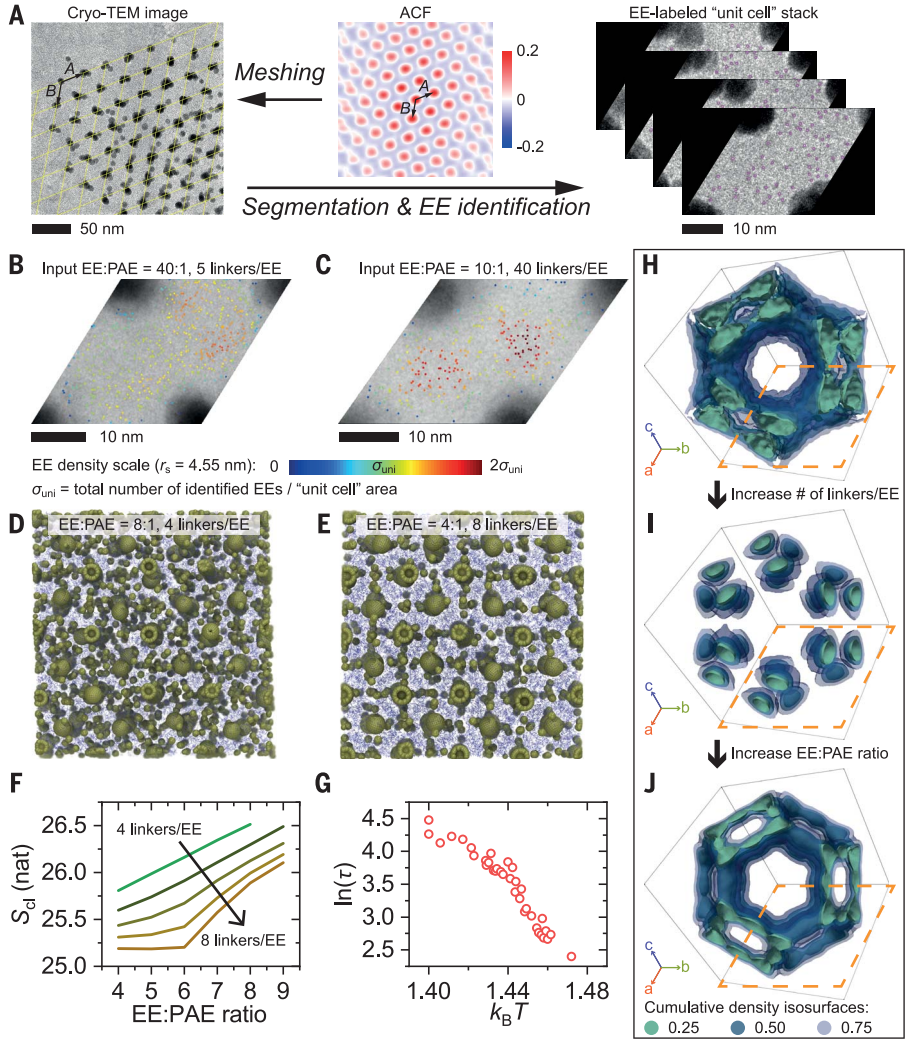


Fig. 2. Spatial probability distribution of EEs in PAE-EE assemblies in the bcc structure. (**A**) Workflow for obtaining EE location-labeled "unit cell" images from cryo-TEM by using image segmentation. ACF, auto-correlation function. (**B** and **C**) Overlay of averaged-intensity TEM images and identified EE locations in "unit cells" along the [111] direction. The input parameters refer to the ratio of each substance added to the solution, not within a crystal. (**D** and **E**) MD simulation snapshots of PAE-EE assemblies. (**F**) Measure of clustering tendency (S_{cl}) of EEs in MD simulations. (**G**) Temperature-dependent trapping time (τ) of EEs in MD simulations. k_B, Boltzmann constant. (**H** to **J**) Simulated Boltzmann volumes of EEs viewed along the [111] direction with EE:PAE = 4:1 and (H) 4 or (I) 8 linkers per EE, or with (J) EE:PAE = 9:1 and 8 linkers per EE. Orange dashes approximate a repeating "unit cell" used in cryo-TEM image analysis.

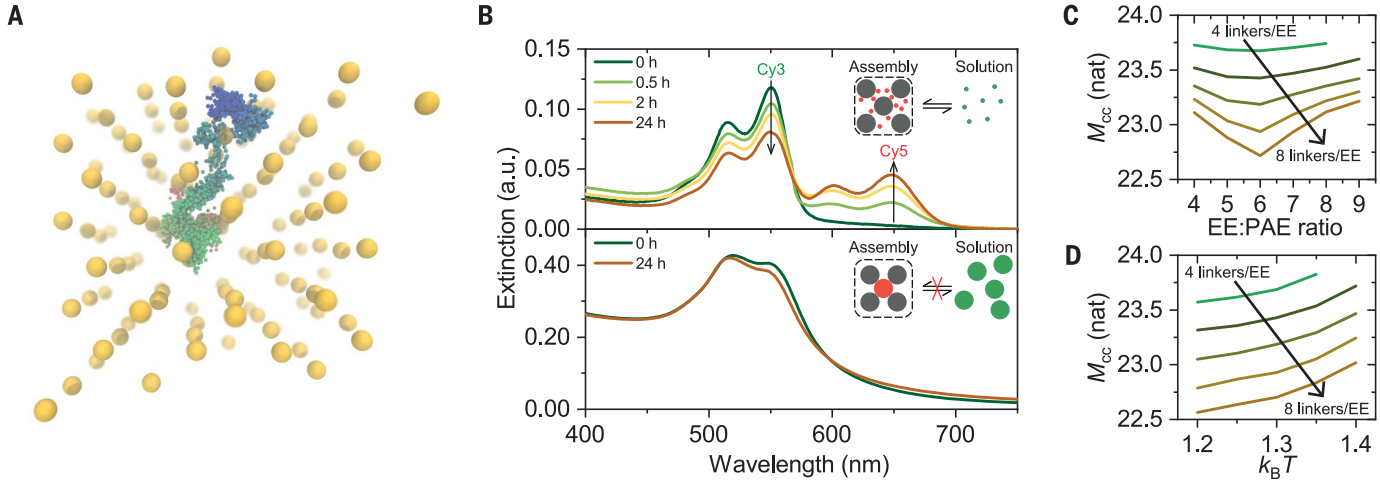


Fig. 3. Diffusion of EEs in PAE-EE assemblies in the bcc structure. (**A**) Trajectory of one EE over time in a lattice of PAEs (yellow) from the MD simulation. The color of the EE positions (red to green to blue) represents different time points. (**B**) The exchange of dye-labeled particles between crystalline lattices and solution was monitored by the change of light extinction in solution by means of UV-vis spectroscopy. Cy5-DNA–labeled EEs were exchanged from "metallic" PAE-EE assemblies (10 + 1.4 nm, bcc) by Cy3-DNA–labeled EEs in the supernatant over 24 hours (top), whereas no appreciable exchange was observed for PAE-PAE assemblies (10 + 10 nm, bcc) (bottom). (**C** and **D**) Predicted colloidal crystal metallicity (M_{cc}) in bcc assemblies with different number of linkers per EE. The M_{cc} value has a minimum against the EE:PAE ratio at 6:1 (C) but is monotonic against temperature (D).

tween the colloidal assemblies and solution. In comparison, no appreciable exchange events were observed for the bcc assemblies formed by 10 + 10 nm PAEs modified with identical dye-labeled DNA (Fig. 3B, bottom), suggesting a much weaker diffusion of PAEs as compared with that of EEs. In both cases, the crystallinity of the assemblies was preserved after the exchange (fig. S18).

The tunable spatial density distribution of EEs (based on number, temperature, and linker density) and their macroscale diffusion inside a crystalline PAE framework establish the concept of colloidal crystal metallicity, M_{cc}, as

$$M_{cc} = S_{cl} - \ln(N_{cell}) \quad (2)$$

where N_{cell} is the number of EEs per unit cell and is used to make the metallicity independent of crystal unit cell and to reduce to the ideal entropy of a gas in the limit of noninteracting particles. In Eq. 1, the EEs are considered as a group (Fig. 2F), but in Eq. 2, M_{cc} is a quantitative measure of the average degree of delocalization per EE, which can decrease when the EE:PAE ratio increases (Fig. 3C) even if S_{cl} increases or remains nearly the same (Fig. 2F). A metallicity minimum was attained at EE:PAE = 6:1, suggesting that the metal-like bonds in the colloidal system become saturated. Increasing the number of duplexed strands on EEs led to crystals with lower M_{cc} values (Fig. 3C) because the EEs were more localized or trapped as the frequency of DNA binding events increased. When the system temperature increased, M_{cc} also increased (Fig. 3D) because of the decrease in the number of hybridized sticky ends. These results show that the EEs can undergo a classical (nonquantized) process analogous to a metal-insulator transition observed in atomic solids (*31*) because the degree of EE delocalization and diffusion drastically changed when either the temperature, the number of linkers per EE, or the EE:PAE ratio was varied.

Colloidal crystal metallicity depends on the number of particles, the number of DNA strands that can engage in bonding, and the strength of the bonds formed. Because these parameters are difficult if not impossible to control in purely electrostatic systems owing to electroneutrality requirements, metallicity has been neither observed nor defined in conventional ionic colloidal crystals (crystals formed from oppositely charged particles) (*4*, *9*). The complementary DNA design on NPs ensures that a crystal will not form from PAEs alone (and vice versa from EEs alone), which distinguishes this system from a description of small particles doped in a crystalline solid formed from larger particles. Moreover, although this concept of metallicity superficially resembles "sublattice melting" in superionic conductors (*32*), in which one sublattice of an ionic compound loses long-range order while the other is fixed (for example, AgI), ion diffusion in such systems is mediated by Frenkel defects because of the constraint of charge balance. Thus, ionic systems have limited tunability and more strict stoichiometry requirements compared with those of the colloidal crystals formed through DNA-directed assembly events.

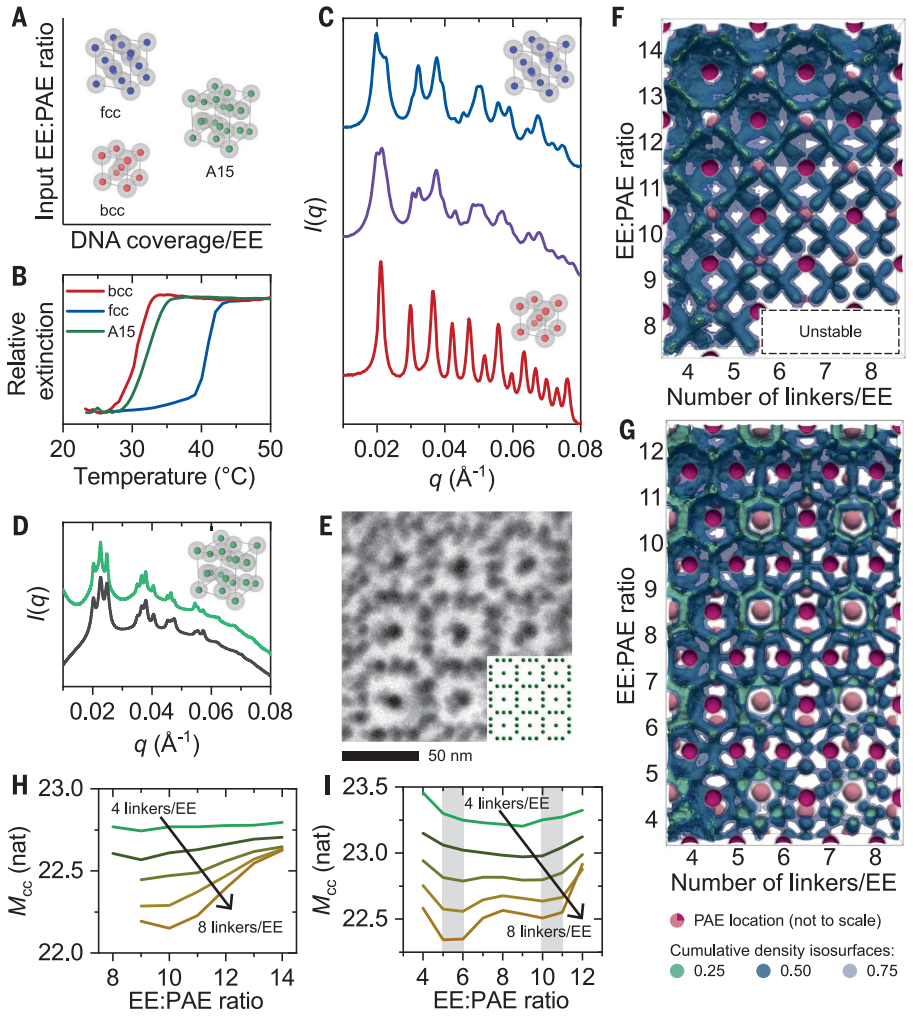


Fig. 4. Equilibrium phases realized by PAE-EE assemblies. (**A**) Schematic representation of the equilibrium conditions of bcc, fcc, and A15 phases and (**B**) their corresponding thermal melting transitions. (**C**) SAXS spectra showing the equilibrium phase transition from bcc (red, input EE:PAE = 10:1) to a bcc/fcc mixture (purple, input EE:PAE = 20:1), and then to a majority fcc phase (blue, input EE:PAE = 40:1). (**D**) Experimental (green) and simulated (black) SAXS spectra of A15 assemblies. (**E**) Cryo-TEM image of an A15 assembly. (Inset) A15 lattice model along the [001] direction. (**F** and **G**) Simulated Boltzmann volumes of fcc (F) and A15 (G) phases as a function of the number of linkers per EE and EE:PAE ratio at a constant temperature (k_BT = 1.30). A whole graph is reconstructed by 1/8 of the unit cells from each combination. (**H** and **I**) Predicted colloidal crystal metallicity (M_{cc}) in (H) fcc and (I) A15 assemblies as a function of the number of linkers per EE and EE:PAE ratio. Two low-metallicity configurations (gray shades) are present in (I).

New phases were accessed by adjusting the input EE:PAE ratio in solution and the total DNA coverage on EEs. For example, as the input EE:PAE ratio was progressively increased, an fcc lattice (space group $Fm\bar{3}m$) emerged (Fig. 4, A and C). The increase in EE:PAE ratio resulted in stronger cumulative bonding interactions (there are more DNA bonding connections under such circumstances), which was reflected by the increase in the crystal melting temperature, T_m, from 31° (bcc) to 41°C (fcc) (Fig. 4B). In addition, if both the total DNA coverage on EEs (characterized by the total number of duplexed and nonduplexed strands) and the input EE:PAE ratio in solution increase, the Frank-Kasper A15 phase (space group $Pm\bar{3}n$) emerged (Fig. 4, D and E). The formation of the A15 phase may be associated with its tendency to minimize the contact area between repulsive PAEs, which is similar to the trend observed in dendrimer assemblies (fig. S27) (*33*) but different from the Cr_3Si structure in binary superlattices (*14*, *34*). These phases were also observed in the simulations (table S9). The three phases realized in

this system mimic the metal tungsten, in which bcc, fcc, and A15 structures have all been realized either in the bulk or in thin films (*35*) (table S6).

Critically, the MD simulations allowed us to identify the positions that the EEs reside in at low M_{cc} values. For example, when the number of DNA linkers per EE was maximized and the EE:PAE ratio was decreased, the EEs settled into distinct locations in the fcc and A15 assemblies, as shown by the Boltzmann volumes (Fig. 4, F and G) and the corresponding M_{cc} values (Fig. 4, H and I). The localized lattice for the fcc structure resembled a fully filled high-temperature Cu_2Se lattice (fig. S34, A_4B_{40}), whereas the A15 structure shows two possible configurations, either clathrate type I (fig. S35A, A_8B_{46}) (*36*) or an unreported lattice (fig. S35B, A_8B_{82}). The latter configuration, in which M_{cc} reached a local minimum, contained all sites in the clathrate structure that were fully occupied, and the additional EEs that could not occupy the lowest energy positions began to fill the higher-energy 12f and 24j positions (Table 1).

Taken together, this work makes the case for describing certain classes of colloidal crystals in a fundamentally new way, in which, in the case of mobile particles (EEs), the concept of metallicity becomes important. By understanding the factors that govern EE diffusion and delocalization, we have a better understanding of the structures and phases that can be accessed through colloidal crystals, potentially including metals, intermetallics, and complex metal alloys. It also challenges the colloidal science community to identify exotic new properties that arise from the PAE-to-EE transition and structures that exhibit high degrees of metallicity as well as to develop theoretical models that capture the effects that lead to metallicity.

REFERENCES AND NOTES

1. C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff, *Nature* **382**, 607–609 (1996).
2. C. A. Batista, R. G. Larson, N. A. Kotov, *Science* **350**, 1242477 (2015).
3. U. Gasser, E. R. Weeks, A. Schofield, P. N. Pusey, D. A. Weitz, *Science* **292**, 258–262 (2001).
4. M. E. Leunissen *et al*., *Nature* **437**, 235–240 (2005).
5. Y. Wang *et al*., *Nature* **491**, 51–55 (2012).
6. M. A. Boles, M. Engel, D. V. Talapin, *Chem. Rev.* **116**, 11220–11289 (2016).
7. F. A. Aldaye, A. L. Palmer, H. F. Sleiman, *Science* **321**, 1795–1799 (2008).
8. S.-J. Park, A. A. Lazarides, J. J. Storhoff, L. Pesce, C. A. Mirkin, *J. Phys. Chem. B* **108**, 12375–12380 (2004).
9. E. V. Shevchenko, D. V. Talapin, N. A. Kotov, S. O'Brien, C. B. Murray, *Nature* **439**, 55–59 (2006).
10. Q. Chen, S. C. Bae, S. Granick, *Nature* **469**, 381–384 (2011).
11. A. M. Kalsin *et al*., *Science* **312**, 420–424 (2006).
12. S. Y. Park *et al*., *Nature* **451**, 553–556 (2008).
13. D. Nykypanchuk, M. M. Maye, D. van der Lelie, O. Gang, *Nature* **451**, 549–552 (2008).
14. R. J. Macfarlane *et al*., *Science* **334**, 204–208 (2011).
15. E. Auyeung *et al*., *Nature* **505**, 73–77 (2014).
16. J. D. Brodin, E. Auyeung, C. A. Mirkin, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4564–4569 (2015).
17. M. N. O'Brien *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10485–10490 (2016).
18. W. Liu *et al*., *Science* **351**, 582–586 (2016).
19. H. Lin *et al*., *Science* **355**, 931–935 (2017).
20. J. R. McMillan *et al*., *J. Am. Chem. Soc.* **139**, 1754–1757 (2017).
21. S. E. Seo, M. Girard, M. Olvera de la Cruz, C. A. Mirkin, *Nat. Commun.* **9**, 4558 (2018).
22. S. Angioletti-Uberti, B. M. Mognetti, D. Frenkel, *Phys. Chem. Chem. Phys.* **18**, 6373–6393 (2016).
23. P. L. Biancaniello, A. J. Kim, J. C. Crocker, *Phys. Rev. Lett.* **94**, 058302 (2005).
24. Y. Wang *et al*., *Nat. Commun.* **6**, 7253 (2015).
25. C. Knorowski, S. Burleigh, A. Travesset, *Phys. Rev. Lett.* **106**, 215501 (2011).
26. T. I. N. G. Li, R. Sknepnek, R. J. Macfarlane, C. A. Mirkin, M. O. de la Cruz, *Nano Lett.* **12**, 2509–2514 (2012).
27. When characterized by means of SAXS and electron microscopy, the 10 + 10 nm assembly shows a bcc lattice because the complementary DNA-functionalized Au NPs are indistinguishable.
28. E. Auyeung, R. J. Macfarlane, C. H. J. Choi, J. I. Cutler, C. A. Mirkin, *Adv. Mater.* **24**, 5181–5186 (2012).
29. C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379–423 (1948).
30. J. M. Zuo, M. Kim, M. O'Keeffe, J. C. H. Spence, *Nature* **401**, 49–52 (1999).
31. M. Imada, A. Fujimori, Y. Tokura, *Rev. Mod. Phys.* **70**, 1039–1263 (1998).
32. S. Hull, *Rep. Prog. Phys.* **67**, 1233–1314 (2004).
33. P. Ziherl, R. D. Kamien, *J. Phys. Chem. B* **105**, 10147–10158 (2001).
34. A. Travesset, *Phys. Rev. Lett.* **119**, 115701 (2017).
35. E. Lassner, W.-D. Schubert, *Tungsten: Properties, Chemistry, Technology of the Element, Alloys, and Chemical Compounds* (Springer, 1999), pp. 1–59.
36. S. Yamanaka, E. Enishi, H. Fukuoka, M. Yasukawa, *Inorg. Chem.* **39**, 56–58 (2000).

ACKNOWLEDGMENTS

The authors thank H. Lopez-Rios [Northwestern University (NU)] and M. G. Kanatzidis (NU) for helpful discussions, A. M. Geller (NU) for rendering the Boltzmann volume data, E. W. Roth (NU) for ultramicrotomy, and J. Remis (NU) for cryo-TEM tomography. **Funding:** This material is based on work supported by the Center for Bio-Inspired Energy Science (CBES), an Energy Frontier Research Center funded by the U.S. Department of Energy (DOE) Office of Basic Energy Sciences (DE-SC0000989, for computational studies), the Air Force Office of Scientific Research (FA9550-17-1-0348, for synthesis, spectroscopy, and electron microscopy), the Vannevar Bush Faculty Fellowship program sponsored by the Basic Research Office of the Assistant Secretary of Defense for Research and Engineering and funded by the Office of Naval Research (N00014-15-1-0043), the Sherman Fairchild Foundation (for electron microscopy and computational support), and the Biotechnology Training Program of NU (for cryo-TEM). This work made use of facilities at the NUANCE Center at NU (NSF ECCS-1542205 and NSF DMR-1720139), the Structural Biology Facility at NU (NCI CCSG P30 CA060553), and the DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) of the Advanced Photon Source (APS) Sector 5 (DOE DE-AC02-06CH11357). **Author contributions:** C.A.M. and M.O.d.l.C. directed the research. M.G. performed simulations. S.W. and A.D. performed synthesis and x-ray scattering experiments. J.S.D., S.W., and Z.H. performed electron microscopy studies. All authors contributed to data analysis and manuscript preparation. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in this manuscript are present in the main text or the supplementary materials. Additional data or codes are available upon request to the corresponding authors.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1174/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S39
Tables S1 to S11
References (*37*–*52*)
Movies S1 to S6

30 January 2019; accepted 28 May 2019
10.1126/science.aaw8237

MUCOSAL IMMUNITY

Akkermansia muciniphila induces intestinal adaptive immune responses during homeostasis

Eduard Ansaldo[1], Leianna C. Slayden[1], Krystal L. Ching[1]*, Meghan A. Koch[1]†, Natalie K. Wolf[1], Damian R. Plichta[2], Eric M. Brown[2], Daniel B. Graham[2,3,4,5], Ramnik J. Xavier[2,3,4,5], James J. Moon[6], Gregory M. Barton[1]‡

Intestinal adaptive immune responses influence host health, yet only a few intestinal bacteria species that induce cognate adaptive immune responses during homeostasis have been identified. Here, we show that *Akkermansia muciniphila*, an intestinal bacterium associated with systemic effects on host metabolism and PD-1 checkpoint immunotherapy, induces immunoglobulin G1 (IgG1) antibodies and antigen-specific T cell responses in mice. Unlike previously characterized mucosal responses, T cell responses to *A. muciniphila* are limited to T follicular helper cells in a gnotobiotic setting, without appreciable induction of other T helper fates or migration to the lamina propria. However, *A. muciniphila*–specific responses are context dependent and adopt other fates in conventional mice. These findings suggest that, during homeostasis, contextual signals influence T cell responses to the microbiota and modulate host immune function.

Adaptive immune cells are critical contributors to tissue homeostasis in the intestine, with B cell–derived immunoglobulin A (IgA) playing a major role in establishing barrier function (*1*). T cell–independent (TI) IgA recognizes a broad fraction of intestinal microbes. By contrast, only a few species that induce T cell–dependent (TD) IgA have been identified (*2*–*4*). Until recently, intestinal IgG antibody responses were only thought to occur in the context of mucosal barrier disruption (*5*) or in response to enteric pathogens (*6*) and certain pathobionts that breach the intestinal barrier (*7*). However, recent work has revealed that mice also generate broadly reactive TI anticommensal IgG2b and IgG3 in a largely Toll-like receptor 2 (TLR2)- and TLR4-dependent manner (*8*).

Despite the abundance of foreign commensal antigens and the high frequency of effector and memory T cells within the intestine, only a small number of intestinal bacteria species that induce antigen-specific T cell responses during homeostasis have been identified (*1*, *9*, *10*). Inappropriate T cell responses against the microbiota are believed to contribute to the pathogenesis of inflammatory bowel disease (*11*, *12*). However, anticommensal effector T cell responses have also been hypothesized to provide bystander protection in the context of enteric infections (*13*). Thus, understanding which commensal species induce cognate T cell responses, the signals that mediate their induction and regulation, and their effects on host physiology remain important goals. Here, we describe an anticommensal TD IgG1 response and use it to identify and characterize a commensal species that induces T cell responses during homeostasis.

We examined serum antibody binding to intact commensal bacteria by staining fecal samples with paired mouse sera and isotype-specific secondary antibodies as previously described (*8*). Surprisingly, comparing wild-type (WT) and T cell–deficient (*Tcrb*$^{-/-}$) mice revealed that mice mount a microbiota-reactive IgG1 antibody response that is entirely dependent on αβ T cells (Fig. 1, A and B, and fig. S1, A to C), whereas anticommensal IgG2b, IgG3, and IgA are induced in a TI manner, as previously reported (fig. S1A) (*4*, *8*, *14*). The fraction of commensals

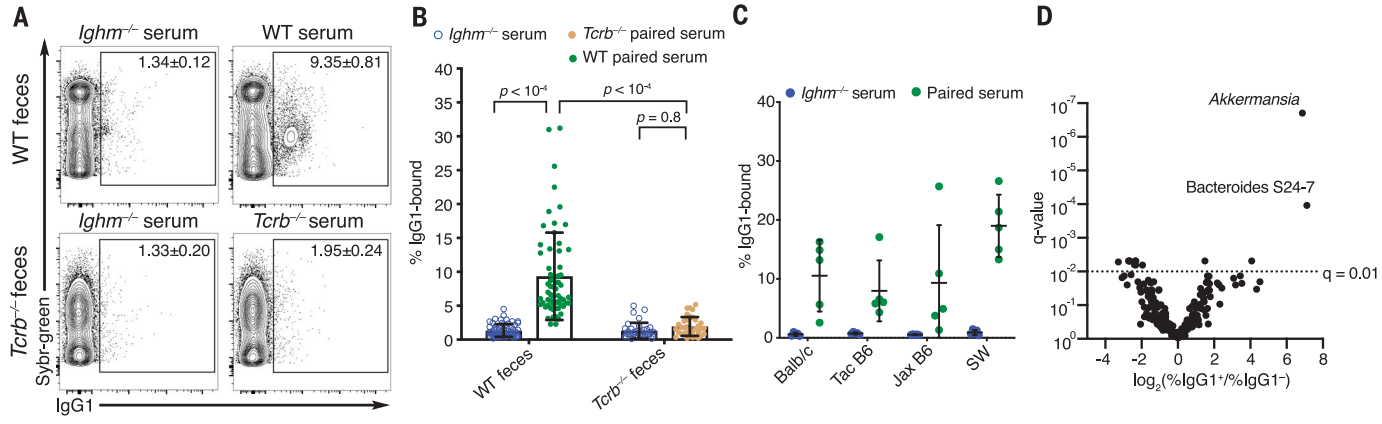


Fig. 1. Mice generate anticommensal IgG1 antibodies during homeostasis. (**A**) Representative IgG1 flow cytometric analysis of fecal microbiota with sera from WT and T cell–deficient (*Tcrb*$^{-/-}$) mice. Feces and sera originated from the same mouse (paired serum) except when using antibody-deficient (*Ighm*$^{-/-}$) serum as a negative staining control. SYBR Green labels a fraction of the microbiota, ensuring that SYBR^{hi} events are bacteria, whereas some of the SYBR^{lo} events are also commensals that are less permeable to the dye (*8*). (**B**) IgG1 microbiota flow cytometric analysis compiled from eight independent experiments. All mice were housed at UC Berkeley. WT, n = 63; *Tcrb*$^{-/-}$, n = 35. (**C**) IgG1 microbiota flow cytometric analysis with paired feces and sera from mice of the indicated genetic backgrounds and vivaria. Balb/c mice were from The Jackson Laboratory. Jax B6 and Tac B6 C57BL/6 mice were from the Jackson Laboratory and Taconic Biosciences, respectively. Swiss Webster (SW) mice were from Taconic Biosciences (n = 5 mice per group). Data are representative of two independent experiments. (**D**) Results from sorting and 16S rDNA sequencing of IgG1-bound and -unbound fractions (n = 12 mice). Graph depicts the average log2 ratio of abundances between both fractions for each individual OTU and the corresponding q value. Data are representative of two independent experiments. Each symbol represents a mouse [(B) and (C)] or an OTU (D). Error bars represent mean ± SD. Gates on flow cytometry plots show mean ± SEM. P values were calculated by Kruskal-Wallis test followed by Dunn's multiple-comparisons test (B) or by paired-ratio Student's t test followed by the Benjamini, Krieger, and Yekutieli two-stage false discovery rate (FDR) correction for multiple comparisons, with an FDR of 0.01 (D).

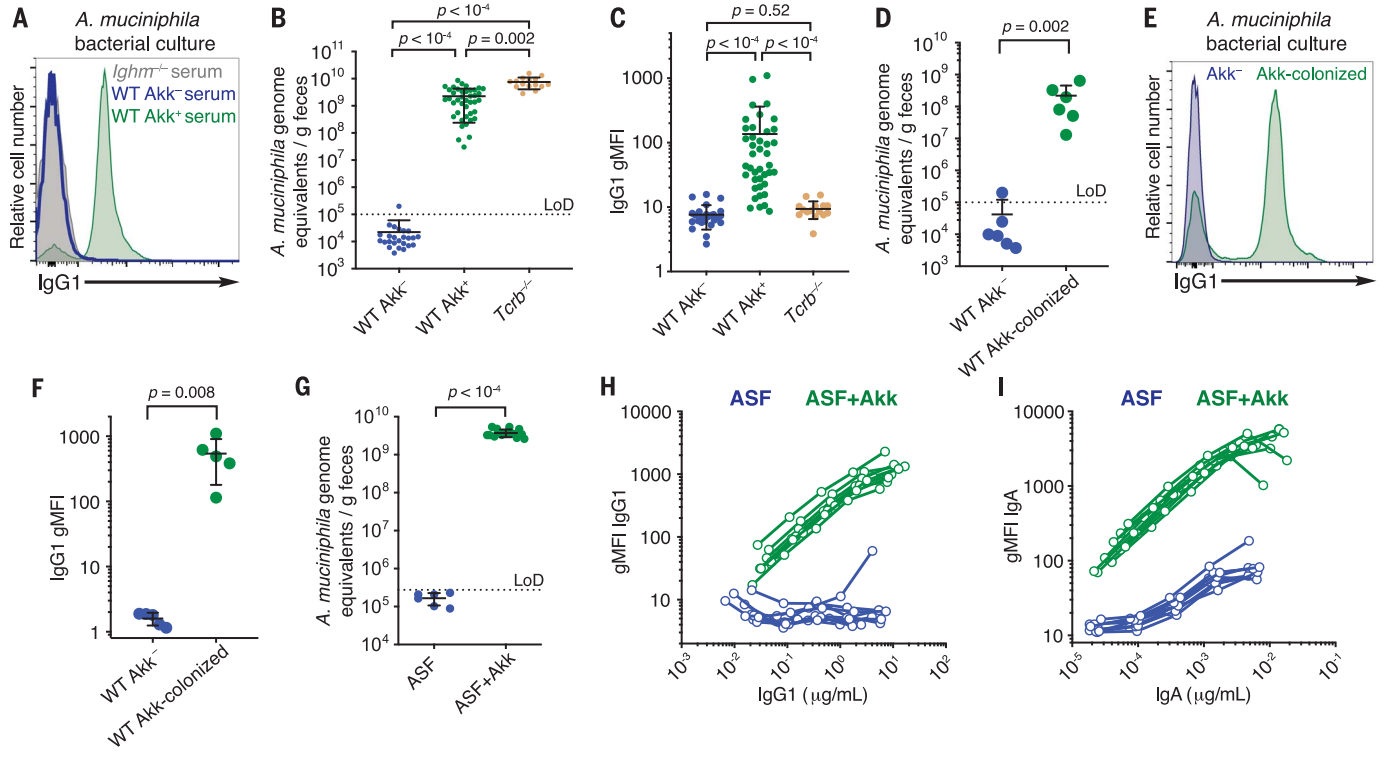


Fig. 2. *A. muciniphila* is necessary and sufficient to induce cognate *A. muciniphila*–specific IgG1 antibody responses. (**A**) Representative IgG1 bacterial flow cytometric analysis of *A. muciniphila* incubated with the indicated mouse sera. Geometric mean fluorescence intensity (gMFI) of the assay was quantified in (C) for multiple mice. (**B**) Quantification of *A. muciniphila* colonization by fecal 16S quantitative polymerase chain reaction (qPCR) for mice described in (C). LoD, limit of detection. (**C**) *A. muciniphila* IgG1 bacterial flow cytometric analysis for mice of the indicated genotypes and indicated *A. muciniphila* colonization status. WT Akk$^-$, n = 25; WT Akk$^+$, n = 41; *Tcrb*$^{-/-}$, n = 15. Data are compiled from seven independent experiments. (**D**) Quantification of *A. muciniphila* colonization by fecal 16S qPCR before (WT Akk$^-$) and 5 weeks after (WT Akk-colonized) a single *A. muciniphila* oral gavage of 10^9 colony-forming units. n = 6 mice. Data are representative of three independent experiments. (**E** and **F**) Representative plot (E) and quantification (F) of *A. muciniphila* IgG1 bacterial flow cytometric analysis using sera from mice before (Akk$^-$) and 5 weeks after (Akk-colonized) colonization. n = 5 mice. Data are representative of three independent experiments. (**G**) Quantification of *A. muciniphila* colonization by fecal 16S qPCR. ASF, n = 6 mice; ASF+Akk, n = 16 mice. Data are representative of two independent experiments. (**H** and **I**) *A. muciniphila* bacterial flow cytometric analysis with serial dilution of serum in ASF and ASF+Akk mice. Each line represents one mouse. The x-axis denotes total serum IgG1 (H) or serum IgA (I) concentration in the assay. n = 9 mice per group. Data are representative of two independent experiments. Each symbol represents a mouse; error bars represent mean ± SD. P values were calculated with a Kruskal-Wallis test followed by Dunn's multiple-comparisons test (B and C) or Mann-Whitney test (D, F, and G).

bound by TD IgG1 was ~10% (Fig. 1B), far less than the percentage recognized by IgA, IgG2b, and IgG3 (fig. S1A). Thus, mice produced IgG1 that appeared to recognize only a subset of microbes in the intestine, in contrast to the polyreactive TI antibodies that bind diverse bacteria (*8*, *15*). T cell–sufficient (*Tcrb*$^{+/-}$) and *Tcrb*$^{-/-}$ cohoused littermates were compared with control for the composition of the microbiota, and an anticommensal IgG1 response was found only in *Tcrb*$^{+/-}$ mice (fig. S1B). Moreover, analysis of mice from multiple vendors and of different genetic backgrounds demonstrated that this IgG1 response was a general feature of healthy mice (Fig. 1C).

To identify the commensal species targeted by this humoral response, we sorted IgG1-bound and -unbound populations from fecal samples stained with sera from corresponding mice (fig. S2, A and B) and performed 16*S* ribosomal DNA (rDNA) sequencing on the resulting fractions. Analysis of operational taxonomic units (OTUs) that were significantly enriched in the IgG1-positive fractions compared with the IgG1-negative fractions revealed two major IgG1 targets (Fig. 1D and fig. S2, C and D). These OTUs correspond to the *Akkermansia* genus (OTU2) and the Bacteroides S24-7 family (OTU63). Bacteroides S24-7 comprises a poorly characterized family of mouse intestinal microbes (*16*). *Akkermansia* is a genus of commensals in the Verrucomicrobia phylum, which until recently, only contained one member, *Akkermansia muciniphila* (*17*). *A. muciniphila* can degrade mucin (*17*) and is an abundant member of the human intestinal microbiota (*18*). Colonization with *A. muciniphila* has been reported to have protective effects in diet-induced obesity (*19*, *20*), to promote mucosal wound healing (*21*), and to increase antitumor responses during anti-PD-1 immunotherapy (*22*). The mechanisms by which *A. muciniphila* mediates these diverse effects remain poorly understood, as little is known about host sensing of this bacterium. Therefore, based on the ability of *A. muciniphila* to induce TD IgG1 responses and its reported effects on host physiology, we sought to characterize the immune response to this bacterium.

We isolated *A. muciniphila* from mice in our colony by plating feces on selective media (*17*). We then used bacterial flow cytometric analysis

[1]Division of Immunology and Pathogenesis, Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. [2]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [3]Center for Computational and Integrative Biology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [4]Department of Molecular Biology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [5]Center for Microbiome Informatics and Therapeutics, MIT, Cambridge, MA, USA. [6]Center for Immunology and Inflammatory Diseases and Division of Pulmonary and Critical Care Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA.
*Present address: Sackler Institute of Graduate Biomedical Sciences, NYU School of Medicine, New York, NY, USA.
†Present address: Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.
‡Corresponding author. Email: barton@berkeley.edu

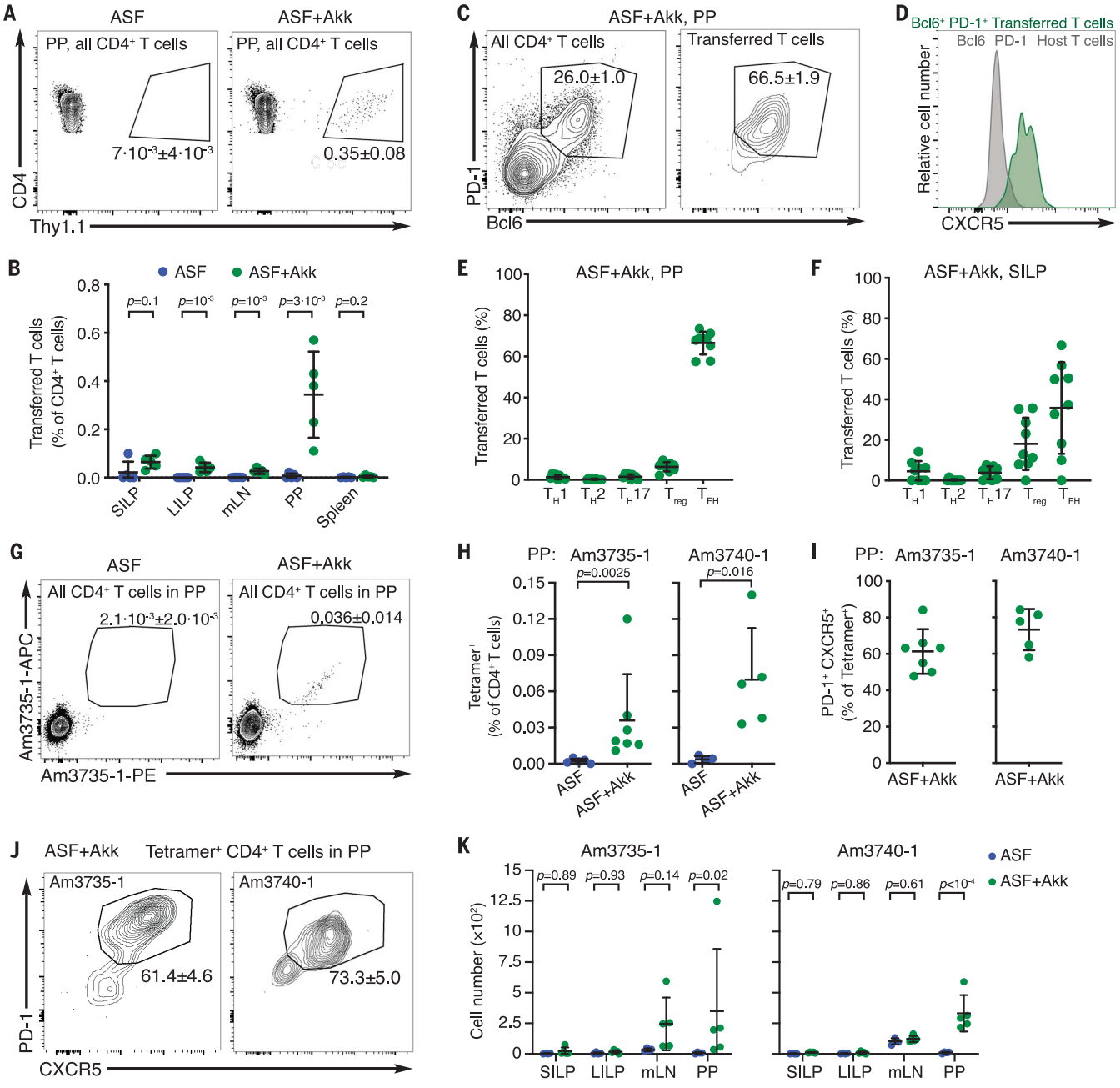


Fig. 3. *A. muciniphila* induces antigen-specific T_{FH} cell responses during homeostasis. (**A**) Representative flow cytometric analysis depicting transferred T cells (Thy1.1$^+$) as a percentage of all CD4$^+$ T cells in the PPs of ASF and ASF+Akk mice 12 days after low-frequency adoptive transfer of Amuc124 (TCR-transgenic) T cells. (**B**) Frequencies of transferred T cells in intestinal tissues of ASF and ASF+Akk mice. n = 5 mice per group; data are representative of six independent experiments. (**C** and **D**) Representative flow cytometric analysis of expression of T_{FH} markers (PD-1, Bcl6, and CXCR5) by endogenous and transferred T cells in the PPs of ASF+Akk mice. (**E** and **F**) Expression of T_H1 (T-bet$^+$ FOXP3$^-$), T_H2 (GATA3$^+$ FOXP3$^-$), T_H17 (RORγt$^+$ FOXP3$^-$), T_{reg} (FOXP3$^+$), or T_{FH} (Bcl6$^+$ PD-1$^+$) markers by transferred T cells in the PPs (E) and SILP (F) of ASF+Akk mice. n = 9 mice. Data are representative of six independent experiments. (**G**) Representative Am3735-1 tetramer flow cytometric analysis of PPs from ASF and ASF+Akk mice. (**H**) Frequencies of Am3735-1- or Am3740-1 tetramer–positive endogenous T cells in ASF and ASF+Akk mice as a percentage of total CD4$^+$ T cells. ASF, n = 4 mice; ASF+Akk, n = 5 to 7 mice. Data are representative of three (Am3740-1) or six (Am3735-1) independent experiments. (**I**) Frequencies of Am3735-1- or Am3740-1 tetramer–positive cells expressing T_{FH} markers [PD-1 and CXCR5, as shown in (J)] in the PPs of ASF+Akk mice. n = 5 to 7 mice; data are representative of three (Am3740-1) or six (Am3735-1) independent experiments. (**J**) Representative flow cytometric analysis of expression of T_{FH} markers (PD-1 and CXCR5) by endogenous Am3735-1- or Am3740-1 tetramer–positive cells. (**K**) Numbers of Am3735-1- and Am3740-1 tetramer–positive cells in all intestinal tissues in ASF and ASF+Akk mice. ASF, n = 4 mice; ASF +Akk, n = 5 mice. Data are representative of two (Am3740-1) or three (Am3740-1) independent experiments. Each symbol represents a mouse. Error bars represent mean ± SD. Gates on flow cytometry plots show mean ± SEM. P values were calculated with unpaired Student's t test (B and K) or Mann-Whitney test (H).

to confirm the presence of *A. muciniphila*–specific IgG1 antibodies in the sera of mice that harbored *A. muciniphila* at steady state (Fig. 2A). These IgG1 antibody responses to *A. muciniphila* could consist of preexisting natural polyreactive specificities or could be antigen specific. To discriminate between these possibilities, we identified C57BL/6 mice lacking *A. muciniphila* in their microbiota (Fig. 2B). Comparing IgG1 antibody responses between *A. muciniphila*–negative and *A. muciniphila*–positive mice confirmed that the induction of *A. muciniphila*–specific serum IgG1 responses required colonization with *A. muciniphila* as well as T cells (Fig. 2, A to C, and fig. S3A). *A. muciniphila*–positive mice also mounted serum *A. muciniphila*–specific TD IgA responses (fig. S3B). Moreover, de novo colonization of *A. muciniphila*–negative mice by oral gavage was sufficient to induce *A. muciniphila*–specific IgG1 antibodies (Fig. 2, D to F, and fig. S3C). Thus, IgG1 responses to *A. muciniphila* are not derived from preexisting cross-reactive specificities. Rather, mice mount an antigen-specific TD IgG1 antibody response upon *A. muciniphila* colonization.

We noted that titers of serum IgG1 responses against *A. muciniphila* were variable across *A. muciniphila*–positive mice. A small number of mice lacked *A. muciniphila*–specific IgG1 altogether despite similar colonization (Fig. 2, B and C, and fig. S3A). One explanation for this variability is that variation within intestinal microbial communities may alter the response to *A. muciniphila*. Indeed, previous studies have shown that intestinal infection or inflammation can lead to altered bystander responses against commensal microbes (*23*). To overcome such complications, we established a defined gnotobiotic system to determine whether direct engagement of the mucosal immune system by *A. muciniphila* underlies the TD IgG1 response. To this end, we introduced *A. muciniphila* into gnotobiotic C57BL/6 mice colonized with altered Schaedler flora (ASF) (*24*, *25*) to generate two mouse colonies with identical microbiota except for the presence of *A. muciniphila* in the ASF+Akk colony. *A. muciniphila* colonized ASF+Akk mice to high levels and was vertically transmitted (Fig. 2G and fig. S3D). We restricted all of our analyses to progeny (or progeny of progeny) of ASF+Akk mice that acquired *A. muciniphila* via vertical transmission. Mice colonized with the ASF+Akk flora, but not the ASF flora alone, mounted IgG1 and IgA antibody responses specific for *A. muciniphila*, with very consistent titers between mice (Fig. 2, H and I). Thus, *A. muciniphila* directly engages the immune system to induce TD IgG1 and IgA.

To explore T cell responses to *A. muciniphila*, we expanded *A. muciniphila*–specific T cell lines from intestinal tissues of *A. muciniphila*–colonized mice and generated T cell hybridomas by fusion with BWZ.36 cells (*26*) (fig. S4, A and B). We identified the antigens recognized by two of these hybridomas (124-2 and 168-H10) by screening an *A. muciniphila* genomic expression library for clones that stimulated each T cell hybridoma (fig. S4, C to E). The 124-2 hybridoma recognized a peptide derived from Amuc_RS03735, an outer membrane autotransporter domain–containing protein. Next, we generated *A. muciniphila*–specific, T cell receptor (TCR)–transgenic mice (Amuc124) using the TCRα and TCRβ chains from the 124-2 T cell hybridoma (figs. S4A and S5, A and B). Importantly, we maintained Amuc124 mice free of *A. muciniphila* colonization for all subsequent experiments.

We sought to track the T cell response to *A. muciniphila* by performing low-frequency adoptive transfers of naïve, congenically marked (Thy1.1) Amuc124 CD4$^+$ T cells into ASF and ASF+Akk mice (fig. S6A). *A. muciniphila*–specific T cells expanded in ASF+Akk mice but were undetectable in ASF mice, indicating that *A. muciniphila* antigens are presented under homeostatic conditions (Fig. 3, A and B; fig. S6, A and B; and fig. S3A). Amuc124 T cells were localized to the Peyer's patches (PPs) and to the mesenteric lymph nodes (mLNs), but very few cells were found in the large intestine or small intestine lamina propria (LILP or SILP, respectively) (Fig. 3, A and B, and fig. S6B). The majority of transferred T cells expressed the T follicular helper (T_{FH}) cell markers PD-1, Bcl6, and CXCR5 (Fig. 3,

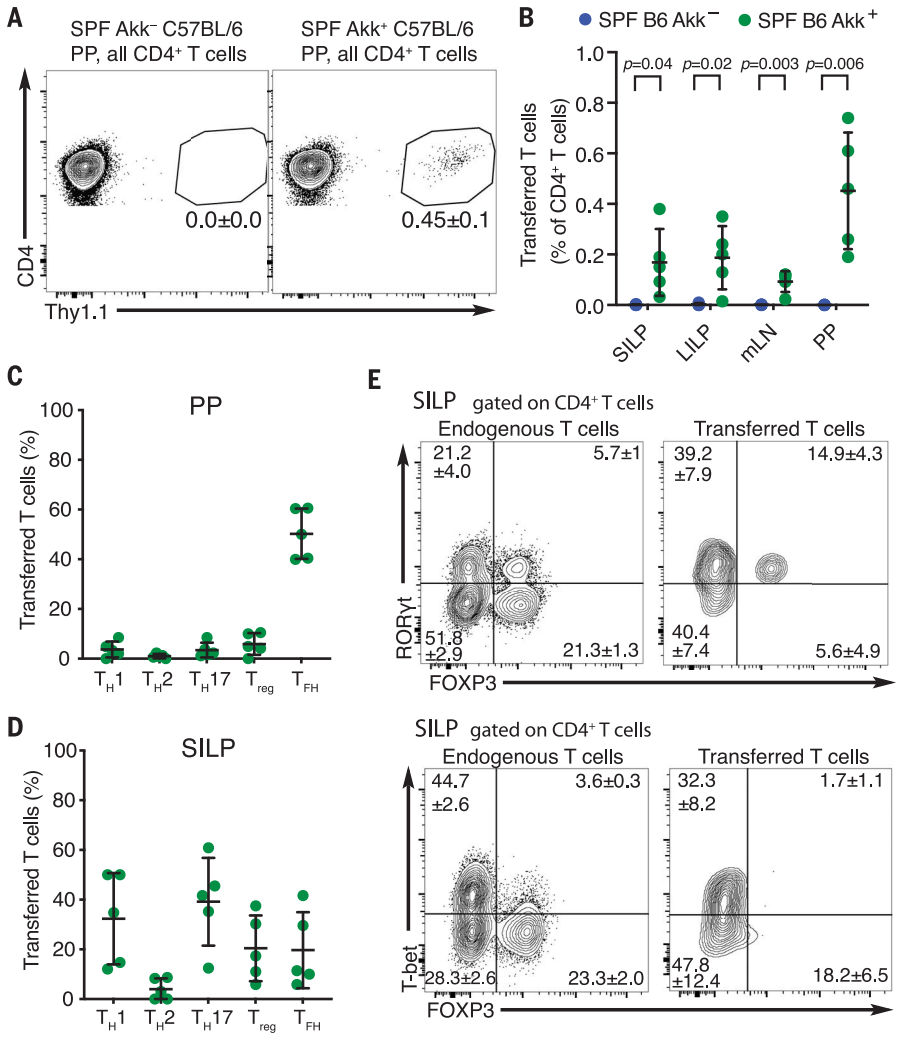


Fig. 4. *A. muciniphila*–specific T cells also adopt other fates in the context of a complex microbiota. (**A**) Representative flow cytometric analysis depicting transferred Amuc124 T cells (Thy1.1$^+$) as a percentage of all CD4$^+$ T cells in the PPs of SPF Akk$^-$ and SPF Akk$^+$ mice. (**B**) Frequencies of transferred T cells as a percentage of all CD4$^+$ T cells in intestinal tissues of conventional SPF *A. muciniphila*–negative (n = 4) and SPF *A. muciniphila*–positive (n = 5) mice. Data are representative of three independent experiments. (**C** and **D**) Expression of T_H1 (T-bet$^+$ FOXP3$^-$), T_H2 (GATA3$^+$ FOXP3$^-$), T_H17 (RORγt$^+$ FOXP3$^-$), T_{reg} (FOXP3$^+$), or T_{FH} (Bcl6$^+$ PD-1$^+$) markers by transferred T cells in the PPs (C) or SILP (D) of SPF *A. muciniphila*–positive mice. n = 5 mice. Data are representative of three independent experiments. (**E**) Representative flow cytometric analysis of expression of T_H1 and T_H17 markers by endogenous total CD4$^+$ T cells and *A. muciniphila*–specific (transferred) T cells in the SILP of SPF *A. muciniphila*–positive mice. Each symbol represents a mouse. Error bars represent mean ± SD. Gates on flow cytometry plots show mean ± SEM. *P* values were calculated with unpaired Student's *t* test (B).

C to E, and fig. S6, C and D), with a small percentage adopting the regulatory T cell (T_{reg}) marker FOXP3. The small number of T cells detectable in the LP were also skewed toward a T_{FH} cell phenotype, most likely indicating that secondary or tertiary lymphoid tissues were not completely excluded from these preparations (Fig. 3F). Transferred *Rag1*$^{+/+}$ or a *Rag1*$^{-/-}$ Amuc124 T cells yielded similar results (fig. S7, A to D). Surprisingly, transferred T cells expressing markers for T helper 1 (T_H1), T_H2, or T_H17 cells (T-bet, GATA3, or RORγt, respectively) were not detected in appreciable numbers (fig. S6D).

To track endogenous *A. muciniphila*–specific T cells, we generated I–A^b tetramers loaded with the peptide TLYIGSGAILS (Thr-Leu-Tyr-Ile-Gly-Ser-Gly-Ala-Ile-Leu-Ser) from the outer membrane protein Amuc_RS03735 (Am3735-1) (fig. S4, C to F). Endogenous, tetramer-positive, *A. muciniphila*–specific T cells were identified in the PPs of ASF+Akk but not ASF mice (Fig. 3, G and H), consistent with the results obtained with Amuc124 T cells after adoptive transfer. Moreover, the endogenous *A. muciniphila*–specific T cells identified by tetramer staining also expressed T_{FH} cell markers (Fig. 3, I and J). An independent tetramer with a different *Akkermansia* epitope from Amuc_RS03740 (Am3740-1) yielded very similar results (Fig. 3, H to J). Finally, we probed all gut-associated lymphoid tissue (LILP, PPs, and SILP) and mLNs with both tetramers. *A. muciniphila*–specific T cells were detected in the PPs but not in the LP (Fig. 3K), which confirmed results obtained with Amuc124 T cell transfers. Thus, *A. muciniphila* induces antigen-specific T cell responses in the intestine that manifest primarily as T_{FH} cells in the PPs.

Finally, we returned to specific pathogen–free (SPF) mice, in which we had observed greater variability in the *A. muciniphila* IgG1 response, and characterized the T cell response to *A. muciniphila* in the context of a conventional microbiota. Similar to our findings for the ASF system, transferred Amuc124 T cells expanded and localized to the PPs in *A. muciniphila*–positive mice but were undetectable in *A. muciniphila*–negative mice (Fig. 4, A and B, and fig. S8A). The majority of transferred T cells in the PPs also adopted T_{FH} cell markers (Fig. 4C and fig. S8B). Therefore, *A. muciniphila* also induces a T_{FH} cell response in the context of a complex microbiota. However, unlike the ASF+Akk system, greater numbers of transferred T cells were detected in the intestinal LP of *A. muciniphila*–positive mice (Fig. 4B), some of which adopted markers consistent with proinflammatory T cell fates (Fig. 4, D and E, and fig. S8, C to E). Consistent with the variable *A. muciniphila*–specific IgG1 titers observed in SPF mice (Fig. 2C), some cohorts of SPF Akk$^+$ mice lacked detectable T cell activation and proliferation after transfer despite the presence of *A. muciniphila*. Thus, T cell responses to *A. muciniphila* appear to be context dependent, resulting in the induction of other CD4 T effector fates in addition to T_{FH} cells in certain conditions. Interestingly, and in contrast to what has been described for segmented filamentous bacteria (SFB) and *Helicobacter* spp. (*9*, *10*, *27*), T cell responses to *A. muciniphila* in SPF mice were mixed between different CD4 T cell fates (fig. S8E).

Commensal-specific TD antibodies were previously thought to be restricted to IgA responses specific to a small subset of commensal species (*2*). Our work reveals that such TD antibodies also include IgG1, which recognizes a small subset of commensals, including *A. muciniphila*. These bacteria may share certain features, such as proximity to the intestinal epithelium, which may increase their potential to cause disease during barrier disruption. Indeed, systemic antibodies specific for commensal bacteria have been reported to protect against gut-derived septicemia (*7*, *28*), and *A. muciniphila* can promote disease in certain immunodeficient settings (*29*).

Very few commensal antigens have been identified to date that lead to antigen-specific T cell responses in the gut during homeostasis. SFB and *Helicobacter* spp. both elicit very defined responses. SFB induce RORγt$^+$ T_H17 cells both in conventional and monocolonized mice (*27*), whereas *Helicobacter* spp. induce RORγt$^+$ FOXP3$^+$ iT_{reg} cells in the LILP (*9*, *10*). T_{FH} cells made up a small proportion of SFB- and *Helicobacter* spp.–specific T cells, but these were in the context of T_H17 cell- or T_{reg}-dominated responses, respectively (*10*). T_{FH} cells in the intestine have been suggested to differentiate from either T_H17 cells or T_{regs} (*30*–*32*). By contrast, the ASF+Akk system produced commensal-specific T cell responses dominated by the induction of T_{FH} cells, with very few T_{reg}, T_H1, T_H2, or T_H17 cells. Thus, commensal-specific T_{FH} responses can occur in the absence of a primary CD4 T cell response of a different fate and may differentiate from naïve commensal-specific T cells in the mesenteric lymph nodes or PPs. Consequently, *A. muciniphila* appears to engage the mucosal immune system in a manner distinct from previously described T cell–activating commensal bacteria. This commensal-specific T_{FH} response appears in conjunction with robust anticommensal TD IgG1 and IgA.

Surprisingly, in the context of a conventional microbiota, and in addition to the induction of T_{FH} cells, *A. muciniphila* can induce CD4 T cells of other fates that home to the LP. Together, these results support the hypothesis that T cell responses against commensals can be context dependent, not just in the setting of acute gastrointestinal infection or inflammation (*23*), but also during homeostasis. Interactions with certain microbes may change the localization or function of *A. muciniphila*, or signals provided by other microbes may shape the immune response against this commensal bacterium.

Prior work has established that *A. muciniphila* mediates effects on host metabolism (*19*, *20*) and can influence the efficacy of anti-PD-1–based immunotherapy against cancer (*22*). The mechanisms for these effects remain poorly understood, but both appear to be immune mediated and correlated with type 1 immunity. In particular, responses to anti-PD-1–based immunotherapy in humans were correlated with interferon gamma production by peripheral T cells incubated with *A. muciniphila* antigens in vitro (*22*). Interestingly, not all patients generated type 1 responses against *A. muciniphila*. Our results provide a potential explanation for this varied response and suggest that differential skewing of *A. muciniphila*–specific T cell responses in individuals due to differences in microbiota composition or other environmental signals may have profound systemic effects. Defining the mechanisms by which these commensal-specific T cell responses can be skewed toward different fates is an important goal with clear clinical implications.

REFERENCES AND NOTES

1. Y. Belkaid, O. J. Harrison, *Immunity* **46**, 562–576 (2017).
2. J. J. Bunker, A. Bendelac, *Immunity* **49**, 211–224 (2018).
3. N. W. Palm *et al.*, *Cell* **158**, 1000–1010 (2014).
4. J. J. Bunker *et al.*, *Immunity* **43**, 541–553 (2015).
5. E. Slack *et al.*, *Science* **325**, 617–620 (2009).
6. N. Kamada *et al.*, *Cell Host Microbe* **17**, 617–627 (2015).
7. M. Y. Zeng *et al.*, *Immunity* **44**, 647–658 (2016).
8. M. A. Koch *et al.*, *Cell* **165**, 827–841 (2016).
9. J. N. Chai *et al.*, *Sci. Immunol.* **2**, eaal5068 (2017).
10. M. Xu *et al.*, *Nature* **554**, 373–377 (2018).
11. R. B. Sartor, *Nat. Clin. Pract. Gastroenterol. Hepatol.* **3**, 390–407 (2006).
12. Y. Belkaid, T. W. Hand, *Cell* **157**, 121–141 (2014).
13. Y. Belkaid, N. Bouladoux, T. W. Hand, *Trends Immunol.* **34**, 299–306 (2013).
14. A. J. Macpherson *et al.*, *Science* **288**, 2222–2226 (2000).
15. J. J. Bunker *et al.*, *Science* **358**, eaan6619 (2017).
16. N. H. Salzman *et al.*, *Microbiology* **148**, 3651–3660 (2002).
17. M. Derrien, E. E. Vaughan, C. M. Plugge, W. M. de Vos, *Int. J. Syst. Evol. Microbiol.* **54**, 1469–1476 (2004).
18. M. Derrien, M. C. Collado, K. Ben-Amor, S. Salminen, W. M. de Vos, *Appl. Environ. Microbiol.* **74**, 1646–1648 (2008).
19. A. Everard *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9066–9071 (2013).
20. R. L. Greer *et al.*, *Nat. Commun.* **7**, 13329 (2016).
21. A. Alam *et al.*, *Nat. Microbiol.* **1**, 15021 (2016).
22. B. Routy *et al.*, *Science* **359**, 91–97 (2018).
23. T. W. Hand *et al.*, *Science* **337**, 1553–1556 (2012).
24. M. Wymore Brand *et al.*, *ILAR J.* **56**, 169–178 (2015).
25. M. B. Geuking *et al.*, *Immunity* **34**, 794–806 (2011).
26. S. Sanderson, D. J. Campbell, N. Shastri, *J. Exp. Med.* **182**, 1751–1757 (1995).
27. Y. Yang *et al.*, *Nature* **510**, 152–156 (2014).
28. J. R. Wilmore *et al.*, *Cell Host Microbe* **23**, 302–311.e3 (2018).
29. S. S. Seregin *et al.*, *Cell Reports* **19**, 733–745 (2017).
30. M. Tsuji *et al.*, *Science* **323**, 1488–1492 (2009).
31. Y. Cong, T. Feng, K. Fujihashi, T. R. Schoeb, C. O. Elson, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 19256–19261 (2009).
32. K. Hirota *et al.*, *Nat. Immunol.* **14**, 372–379 (2013).

ACKNOWLEDGMENTS

We thank H. Nolla and A. Valeros for assistance with cell sorting, M. Manoharan for initial help with the peptide screen, F. Gonzalez and N. Shastri for help and advice regarding T cell hybridomas, E. Robey and E. Pamer for providing the pTα and pTβ cassettes, E. Wu and B. Russell for technical assistance, and Life Sciences Addition vivarium staff for mouse colony maintenance. We also thank R. Vance and members of the Barton and Vance labs for comments on the manuscript and helpful discussions. **Funding:** This work was supported by the NIH (P01AI063302 and R01AI142926 to G.M.B.; R21AI124143 and P30 DK043351 to J.J.M), a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease award (G.M.B.), and a John P. Stock Faculty Fellow award (G.M.B.). E.A. was supported by a fellowship from "La Caixa" foundation. M.A.K. was supported by postdoctoral fellowship no. 252507 from the Crohn's and Colitis Foundation of America. D.G.B and R.J.X. were supported by the Center of Microbiome Informatics and Translation (MIT). **Author**

contributions: E.A. and G.M.B. designed experiments. E.A. performed most of the experiments and data analysis. L.C.S. and K.L.C. helped perform the experiments shown in Fig. 3. L.C.S. also contributed to the experiments shown in Fig. 4 and figs. S1 and S3. N.K.W. performed the peptide screen with the T cell hybridomas. M.A.K. contributed to the experiments shown in Fig. 1. D.R.P., E.M.B., D.B.G., and R.J.X. performed in silico prediction and validation of the peptide for Am3740.1 (D.R.P.: computational analysis; E.M.B.: experimental validation). J.J.M. generated tetramers and provided guidance in tetramer staining and data analysis. E.A. wrote the manuscript. G.M.B. revised and edited the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** 16S rDNA sequencing raw data can be accessed at ncbi.nlm.nih.gov/ with BioProject accession number PRJNA514379 and BioSample accession number SAMN10724098. All other data needed to evaluate the conclusions in this paper are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1179/suppl/DC1
Materials and Methods
Figs. S1 to S8
Table S1
References (*33*–*43*)

21 January 2019; accepted 29 May 2019
10.1126/science.aaw7479

STRUCTURAL BIOLOGY

Structural identification of a hotspot on CFTR for potentiation

Fangyu Liu[1,2]*, Zhe Zhang[1]*, Anat Levit[3], Jesper Levring[1], Kouki K. Touhara[4]†, Brian K. Shoichet[3], Jue Chen[1,5]‡

Cystic fibrosis is a fatal disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR). Two main categories of drugs are being developed: correctors that improve folding of CFTR and potentiators that recover the function of CFTR. Here, we report two cryo–electron microscopy structures of human CFTR in complex with potentiators: one with the U.S. Food and Drug Administration (FDA)–approved drug ivacaftor at 3.3-angstrom resolution and the other with an investigational drug, GLPG1837, at 3.2-angstrom resolution. These two drugs, although chemically dissimilar, bind to the same site within the transmembrane region. Mutagenesis suggests that in both cases, hydrogen bonds provided by the protein are important for drug recognition. The molecular details of how ivacaftor and GLPG1837 interact with CFTR may facilitate structure-based optimization of therapeutic compounds.

The cystic fibrosis transmembrane conductance regulator (CFTR) is an anion channel widely expressed on epithelial surfaces of different organs, including the lung and intestine (*1*). It belongs to the family of the ATP-binding cassette (ABC) transporters, but functions as an anion channel. CFTR consists of two transmembrane domains (TMDs) that form the pore, two cytoplasmic nucleotide-binding domains (NBDs) that bind and hydrolyze adenosine 5′-triphosphate (ATP), and a regulatory (R) domain that must be phosphorylated to allow the channel to open (*2*). More than 300 mutations have been identified to cause cystic fibrosis (CF); details on the variants are given at the CFTR2 website (*3*). The most prevalent mutation is the deletion of a single amino acid, F508, which makes CFTR prone to degradation before reaching the cell's plasma membrane (*4*). Other mutants, such as R117H and G551D, are expressed on the cell membrane but do not gate properly (*5*).

Over the past eight decades, medical advances have improved the treatment of cystic fibrosis. The average survival age of patients has been lengthened from early infancy in the 1930s to around 47 years at present. Most treatments offer symptomatic relief, including pancreatic enzyme supplements to aid digestion, antibiotics to prevent and treat infection, mucus-thinning drugs to clear the airway, and lung transplants. Recently, therapies have been developed to target the CFTR protein. Small-molecule CFTR modulators include correctors that increase the abundance of CFTR at the cell surface and potentiators that increase the ion flux of mutant CFTR (*6*–*8*). Currently, two correctors (lumacaftor and tezacaftor) and one potentiator (ivacaftor), all developed by Vertex Pharmaceuticals, are available to patients (*9*–*11*). In addition, many other candidates to enhance the function of CFTR are in the drug discovery pipeline (*7*, *12*, *13*). All of these CFTR modulators were discovered through intensive high-throughput screening and iterative medicinal chemistry optimization. Rational drug discovery has not been feasible, owing to the lack of structural information. To address this issue, we report here cryo–electron microscopy (EM) structures of the human CFTR in complex with two different potentiators: the Vertex drug ivacaftor (*6*) and GLPG1837, an investigational drug developed by Galapagos (*7*).

[1]Laboratory of Membrane Biophysics and Biology, The Rockefeller University, New York, NY 10065, USA. [2]Tri-Institutional Training Program in Chemical Biology, The Rockefeller University, New York, NY 10065, USA. [3]Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94158, USA. [4]Laboratory of Molecular Neurobiology and Biophysics, The Rockefeller University, New York, NY 10065, USA. [5]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.
*These authors contributed equally to the work. †Present address: Department of Physiology, University of California, San Francisco, CA 94143, USA.
‡Corresponding author. Email: juechen@rockefeller.edu

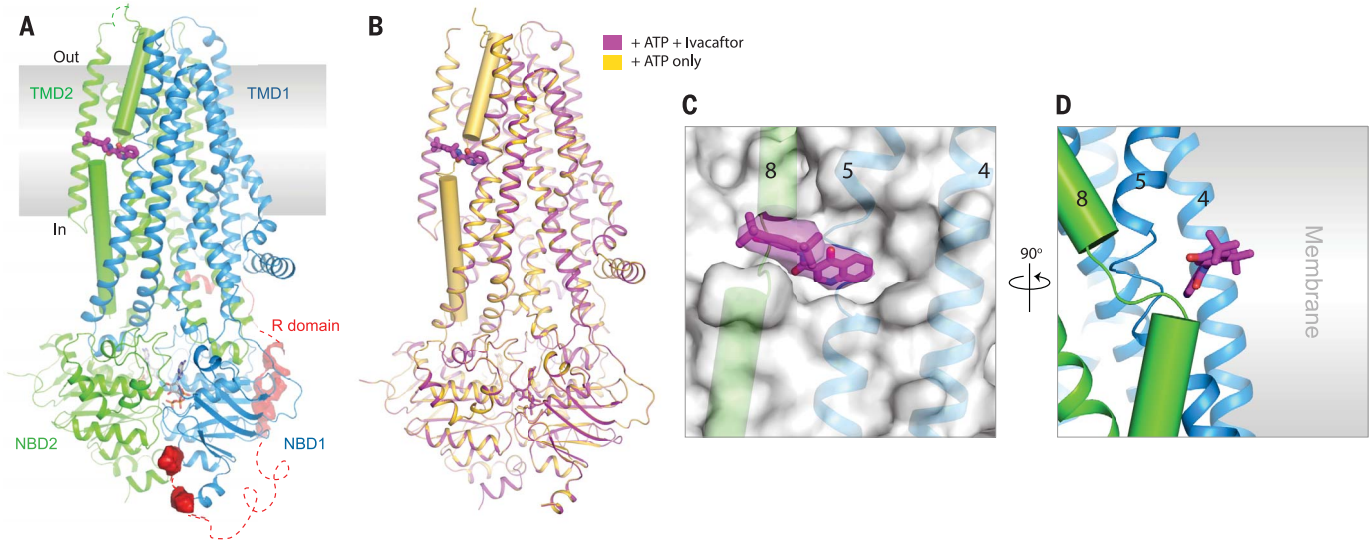


Fig. 1. Ivacaftor binds CFTR inside the membrane. (**A**) Overall structure of the phosphorylated, ATP-bound human CFTR in complex with ivacaftor (shown in magenta). TMD1 and NBD1 are shown in blue, TMD2 and NBD2 in green, and R domain in red. TM8 is highlighted in cylinder representation. Regions not resolved in the structure are shown as dashed lines. (**B**) Superposition of phosphorylated, ATP-bound CFTR in the absence (yellow) and presence of ivacaftor (magenta). (**C**) A magnified view of the ivacaftor-binding site. CFTR is shown as a transparent surface model with TM 4, 5, and 8 indicated. Ivacaftor is shown as a stick model together with the corresponding EM density. (**D**) Ivacaftor binds at the protein-lipid interface, exposing half of its surface to the lipid bilayer.

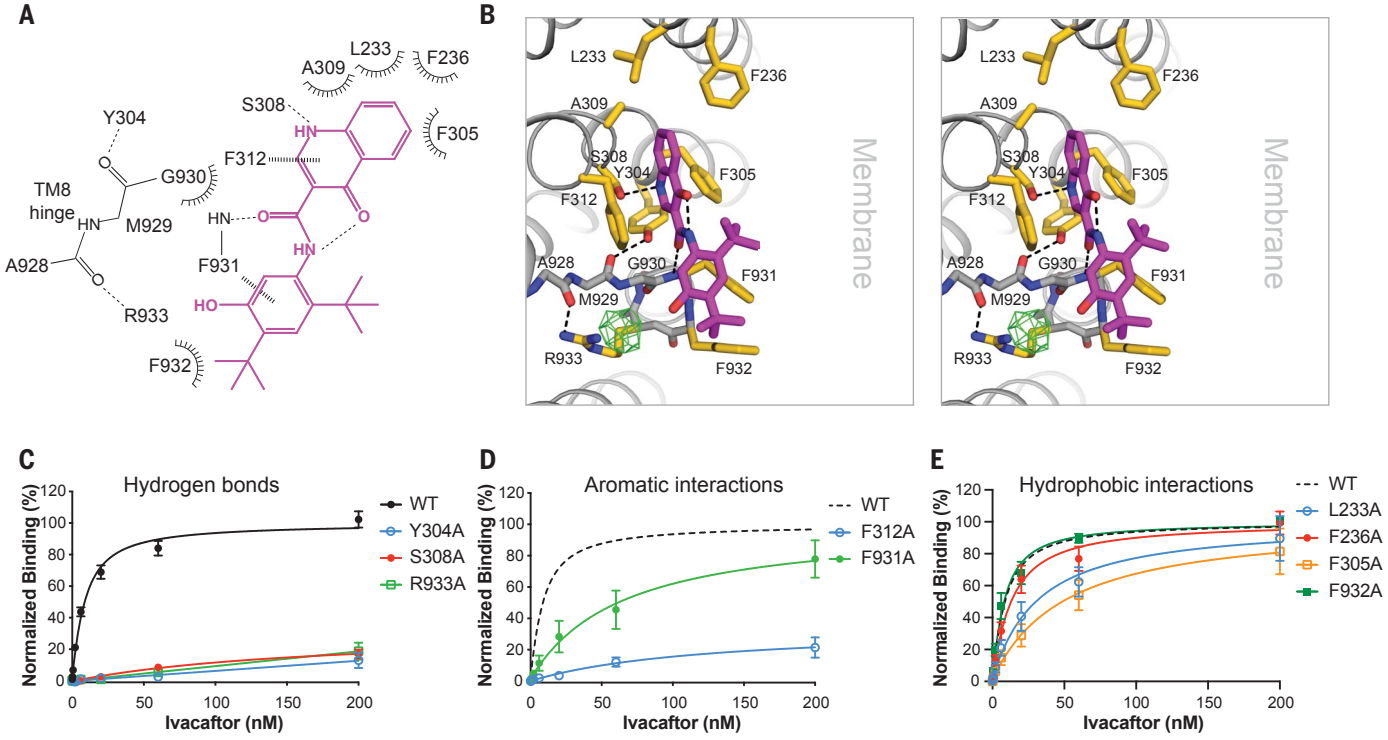


Fig. 2. Contribution of individual residues to ivacaftor binding. (**A**) Schematic drawing of interactions formed between ivacaftor (magenta) and the CFTR-binding site. Residues within van der Waals distances (<4.5 Å) are shown. Representations: black dashed lines, hydrogen bonds; blue vertical lines, aromatic interactions; spokes, hydrophobic interactions. The hinge region in TM8 is shown as black sticks and labeled. (**B**) Stereo view of the ivacaftor-binding site. Residues within van der Waals distances are shown in yellow, and hydrogen bonds are depicted as black dashed lines. An unknown density between R933 and ivacaftor is shown as green mesh. (**C** to **E**) Binding affinities of WT CFTR and mutants replacing residues making (C) hydrogen bonds, (D) aromatic interactions, and (E) hydrophobic interactions with ivacaftor. Data points represent the means and SEMs of at least three measurements. The calculated K_d values are listed in Table S2. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; L, Leu; M, Met; Q, Gln R, Arg; S, Ser; and Y, Tyr.

Ivacaftor was discovered by screening compounds that increase anion flux in G551D-CFTR–expressing cells (*6*). Subsequent studies have shown that ivacaftor increases the open probability (P_o) of both wild-type (WT) and mutant CFTRs in membrane patches, proteoliposomes, and planar lipid bilayers (*14*–*16*). The potentiation by ivacaftor requires phosphorylation of CFTR by protein kinase A (PKA), but is independent of ATP (*15*). These results suggest that ivacaftor acts directly on CFTR, rather than functioning through other regulatory mechanisms.

To describe the specific molecular interactions between ivacaftor and CFTR, we determined a cryo-EM structure of ivacaftor in complex with phosphorylated E1371Q CFTR in the presence of saturating ATP-Mg^{2+} (10 mM) (Fig. 1A, figs. S1 to S3, and table S1). The final map has an overall resolution of 3.3 Å, showing well-defined density throughout the protein, except for the R domain. Density for both ATP-Mg^{2+} molecules are visible at the NBD dimer interface (fig. S3). An additional strong density is observed on the outer surface of the TMDs in the center of the lipid bilayer (Fig. 1C and fig S1C). This density, not observed in any of the previous CFTR structures (*17*–*20*), has a shape and size consistent with the chemical structure of ivacaftor (Fig. 1C and fig S3). Within this density, we built a model of ivacaftor and examined multiple orientations of it in the site, using molecular docking (*21*, *22*) followed by energy minimization of the protein-ligand complex. Complexes were prioritized by their energetic complementarity, ability to make favorable polar interactions, and subsequent refinement to the electron density maps.

We previously reported the structure of the phosphorylated E1371Q construct in the presence of ATP-Mg^{2+} but in the absence of ivacaftor (*20*). The ivacaftor-bound E1371Q exhibits the same protein conformation in which the ATP-bound NBDs form a closed dimer, and the two TMDs pack closely together to form an ion conduction pathway open to the cytoplasmic solution. The R domain, largely unstructured, is located along the peripheral surface of NBD1 and the cytoplasmic region of the TMDs (Fig. 1A). No significant protein conformational changes were observed upon binding of ivacaftor, and the overall root mean square deviation between the two structures is 0.14 Å (Fig. 1B).

Ivacaftor binds CFTR at the protein-lipid interface, docking into a cleft formed by transmembrane (TM) helices 4, 5, and 8 (Fig. 1, C and D). The binding site coincides with a hinge region in TM8, a structural feature of CFTR not found in other ABC transporters (*19*, *20*). The extracellular segment of TM 8 rotates around this hinge upon ATP binding (*17*–*19*); stabilizing this rotation may explain the drug's efficacy. Whereas about 40% of the molecular surface of ivacaftor is buried against CFTR, the remaining 60% is exposed to the hydrophobic region of the membrane (Fig. 1, C and D).

The interactions between ivacaftor and CFTR include two hydrogen bonds, two aromatic interactions, and six hydrophobic interactions (Fig. 2, A and B). To evaluate how each residue contributes to ivacaftor binding, we developed a scintillation proximity assay (SPA) to measure the apparent affinity of ivacaftor for CFTR (Fig. 2, C to E). Every residue in the binding site was individually substituted by alanine. Every mutant eluted from the size-exclusion column as a monomeric peak, similar to the WT CFTR, indicating that these mutations did not alter CFTR folding (fig. S4). Specific binding of ivacaftor to the WT CFTR increased as a function of ivacaftor concentration (Fig. 2C). Nonlinear regression analysis shows that the data fit well to a single-site binding model with an equilibrium dissociation constant (K_d) of 6.6 ± 1.2 nM. In comparison, all but one mutation resulted in a reduced binding

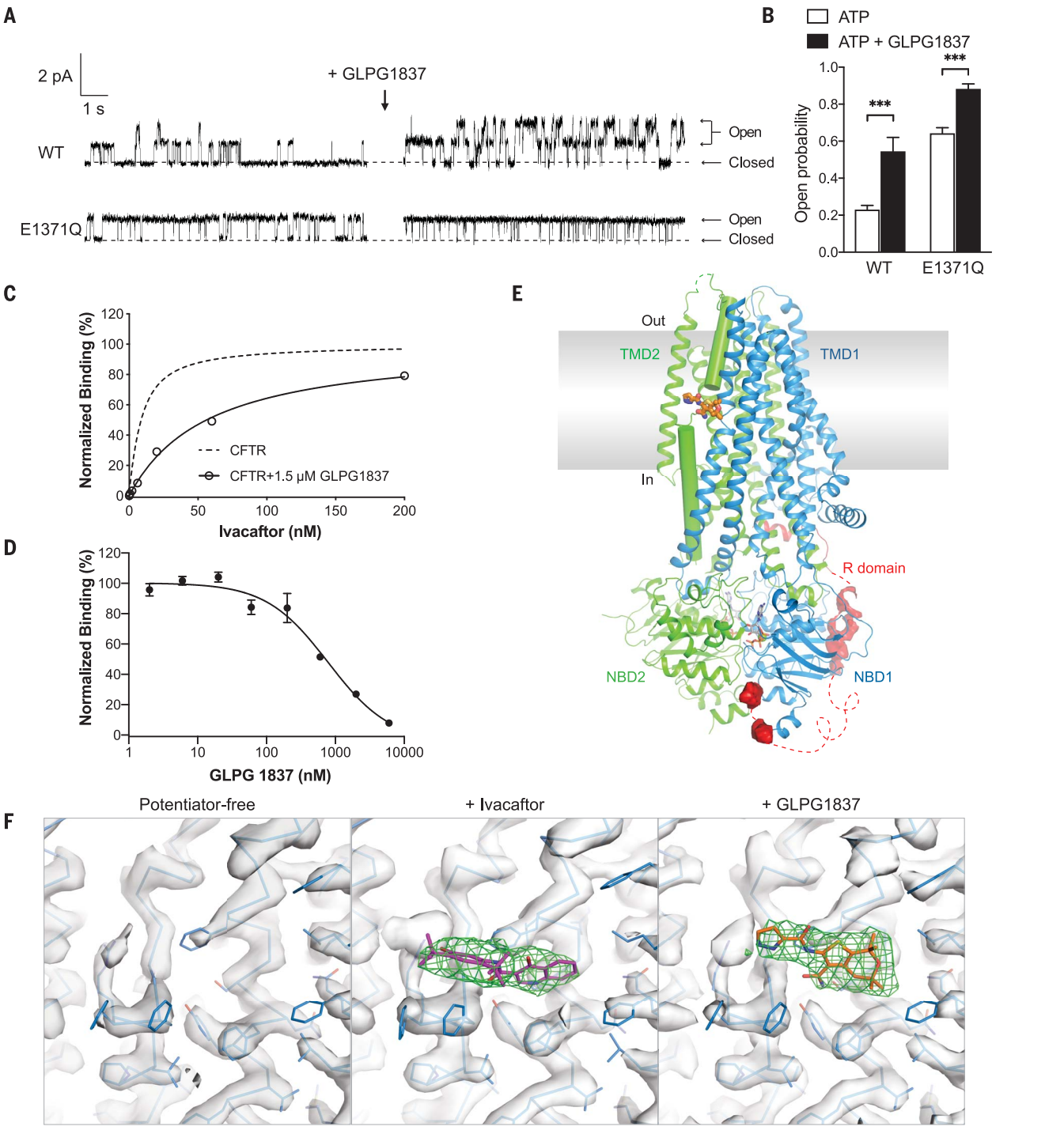


Fig. 3. GLPG1837 binds to the same site as ivacaftor. (**A**) Representative recordings of WT (upper trace) and E1371Q (lower trace) CFTR reconstituted in synthetic lipid bilayers. CFTR was phosphorylated with PKA prior to fusion with bilayers. Recordings were performed on individual membranes with 2 mM ATP before (left) and after (right) addition of 10 μM GLPG1837. (**B**) Open probabilities of WT and E1371Q CFTR before (open bar) and after (filled bar) addition of 10 μM GLPG1837. Data points represent the means and SEMs of three to nine membranes. P_o = 0.23 ± 0.02, for WT; P_o = 0.54 ± 0.08, for WT + GLPG1837; P_o = 0.64 ± 0.03, for E1371Q; P_o = 0.88 ± 0.03, for E1371Q + GLPG1837. ***$p < 0.001$ by Student's t test. (**C**) The presence of 1.5 μM GLPG1837 shifts the apparent K_d of ivacaftor from 6.6 ± 1.2 nM (dashed line) to 54 ± 4 nM (solid line, n = 9). (**D**) Competition binding assay. Ivacaftor was kept at a constant concentration of 8 nM, and its binding to WT CFTR is plotted as a function of GLPG1837 concentration. K_i = 0.30 ± 0.08 μM. (**E**) Ribbon diagram of the phosphorylated, ATP-bound CFTR in complex with GLPG1837 (stick representation, orange). (**F**) EM density at the potentiator-binding site in the potentiator-free (left), ivacaftor-bound (center), and GLPG1837-bound (right) reconstructions. Densities corresponding to CFTR are shown in gray, whereas densities only observed in the potentiator-bound maps are shown as green meshes. All maps are contoured at 9σ.

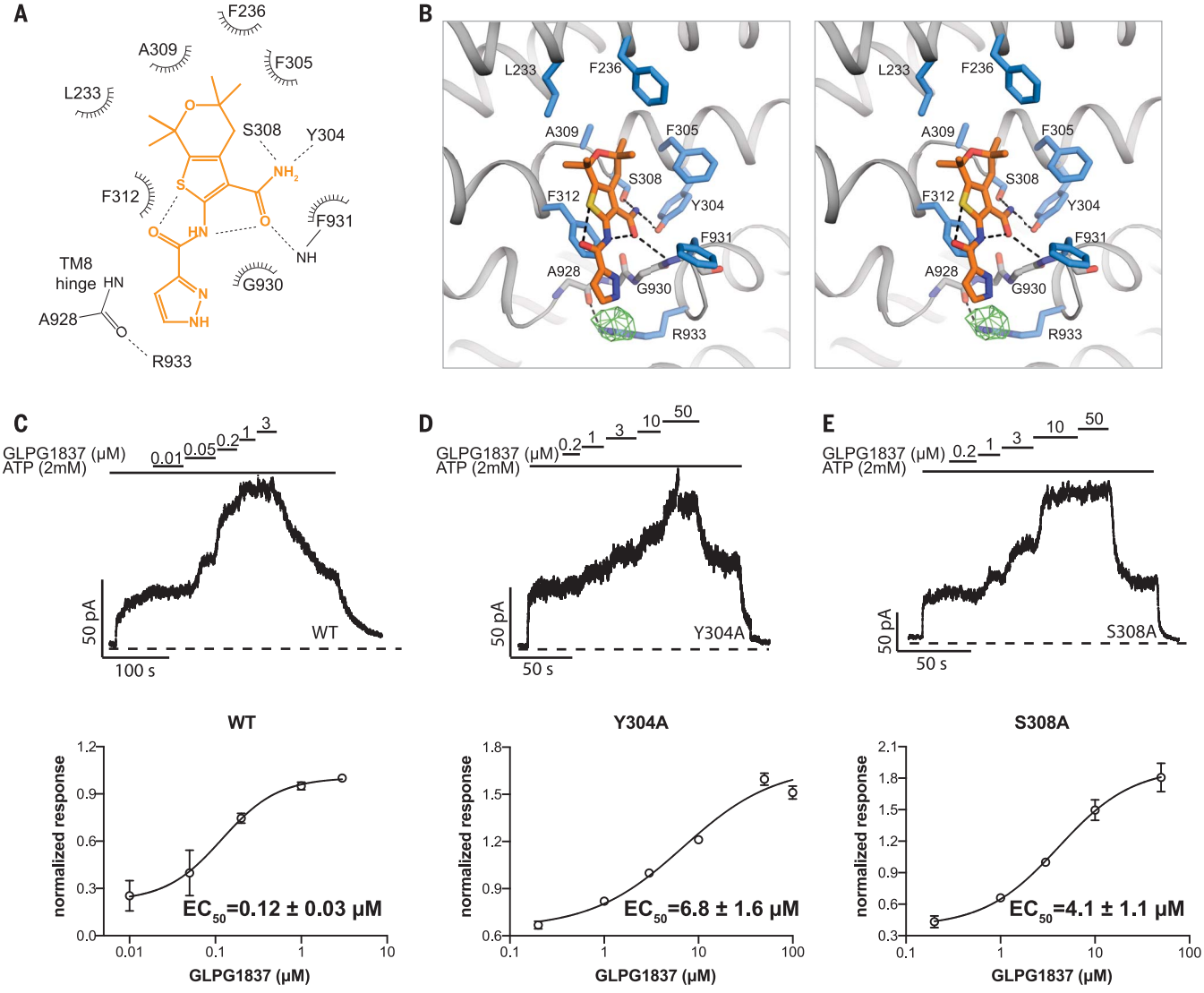


Fig. 4. Molecular details of GLPG1837 binding. (**A**) Schematic drawing of the interactions between GLPG1837 (orange) and CFTR. Hydrogen bonds are represented by black dashed lines, and hydrophobic interactions are shown by the spokes. (**B**) Stereo view of the GLPG1837 binding site. Residues within van der Waals distances (<4.5 Å) are shown as blue sticks, and hydrogen bonds are depicted as black dashed lines. An unknown density between R933 and GLPG1837 is shown in green mesh. (**C** to **E**) Upper panel: Representative macroscopic current traces of WT and mutant CFTR in response to GLPG1837 perfusion. Different concentrations of GLPG1837 applied after channel activation are marked above the trace. Lower panel: Dose-response curves with estimated EC_{50} values. CFTR-containing membrane patches were fully phosphorylated by PKA in the presence of saturating amount of ATP before GLPG1837 titration. The total current at 3 μM GLPG1837 was used to normalize the current potentiated by different concentrations of GLPG1837. The dose responses were fitted with the Hill equation. Each data point represents values determined from five to nine patches.

affinity for ivacaftor (Fig. 2, C to E, and table S2). Four residues appear to be most important—their alanine substitutions nearly abolishing ivacaftor binding (Fig. 2, C and D). Among them, S308 and F312 directly coordinate the oxoquinoline moiety through a hydrogen bond and a π-π stacking interaction, respectively (Fig. 2, A and B). The other two residues, R933 and Y304A, hydrogen bond to main-chain carbonyls in the TM 8 hinge, thus stabilizing the overall structure of the binding site (Fig. 2, A and B). Another important residue, F931, forms an edge-to-face interaction with the phenol ring of ivacaftor. Mutation of F931 to alanine decreased drug affinity by about 10-fold (Fig. 2D and table S2). Mutating F305, L233, and F236, the three residues within van der Waals distance from the oxoquinoline of the drug, also decreased its affinity (Fig. 2E and table S2). By contrast, substitution of F932, which interacts with one of the lipid-exposed tert-butyl groups, had no effect (Fig. 2E and table S2). These mutational effects are consistent with the structure of the CFTR–ivacaftor complex, indicating that hydrogen bonds are critical for ligand recognition in the low-dielectric environment of the membrane.

The structure of CFTR–ivacaftor complex largely explains the structure activity relationship (SAR) for the series of analogs that led to this drug (*6*). From the published SAR, 48 ivacaftor analogs, ranging from the initial high-throughput screening hit to optimized leads, may be readily docked into the ivacaftor site, making favorable interactions (fig. S5). The docked poses superpose with the ivafactor structure, recapitulating more and more of the the drug's interactions with CFTR as the molecules are optimized. Key interactions common to most of the docked complexes include the internal hydrogen bond between the conserved side-chain amide and the ubiquitous oxoquinoline oxygen, the hydrogen bond between the main-chain nitrogen of

F931 and that same amide, and the interaction between the oxoquinoline nitrogen and S308. Similarly, the stacking observed between F312 and ivacaftor's oxoquinoline ring is conserved among the analogs. The phenolic hydroxyl, which appears late in the affinity maturation and is retained in ivacaftor, docks to interact with R933, as observed in the ivacaftor complex; addition of this group substantially increases affinity, at least partly reflecting its new interaction with the arginine (fig. S5; a more detailed analysis is presented in the supplementary text). These results are consistent with the ivafactor structure determined here and the SAR observed in its development (*6*).

Recently, a new potentiator, GLPG1837, has been discovered to have higher efficacy than that of ivacaftor (*7*). We first studied the effects of GLPG1837 in a planar bilayer system, where detergent-purified CFTR channels were reconstituted into liposomes then fused with a bilayer lipid membrane. At saturating ATP concentration, the P_o of the phosphorylated WT CFTR increased from 0.23 to 0.54 upon addition of 10 μM GLPG1837 (Fig. 3, A and B). The P_o of E1371Q was also increased by GLPG1837, from 0.64 to 0.88 (Fig. 3, A and B). The P_o values are lower than those measured in cellular membranes (*23*, *24*), possibly owing to differences in their lipid compositions. A competitive binding assay shows that GLGP1837 reduced the apparent affinity of ivacaftor (Fig. 3C); the inhibition constant (K_i) was determined to be 0.30 ± 0.08 μM (Fig. 3D).

Next, we determined the cryo-EM structure of phosphorylated E1371Q CFTR in complex with ATP-Mg^{2+} and GLPG1837 (Fig. 3E, figs. S6 to S8, and table S3). The overall structure, at 3.2 Å resolution, is essentially indistinguishable from those of drug-free and ivacaftor-bound forms. Although GLPG1837 clearly binds in the same pocket, its orientation and shape differ from that of the ivacaftor density (Fig. 3F). Here too, a model of the CFTR–GLPG1837 complex was built by molecular docking, followed by energy minimization and refinement of the ligand-receptor complex. In the final model, the drug fits well into the electron density, with favorable polar and nonpolar interactions, and a relatively unstrained ligand geometry (Fig. 3F).

Although the chemical structure of GLPG1837 differs from that of ivacaftor, residues interacting with GLPG1837 largely overlap with those engaging ivacaftor (Fig. 4, A and B). Specifically, S308 and Y304 form hydrogen bonds; and L233, F236, F305, A309, F312 form hydrophobic interactions with the drug (Fig. 4, A and B). Four polar groups on GLPG1837 are engaged in intramolecular interactions (Fig. 4, A and B), also observed in the crystal structure of the compound itself (*7*), that shield charges, permitting this relatively polar molecule to permeate the membrane.

To evaluate the contribution of the intermolecular hydrogen bonds, we measured the half-maximal effective concentration (EC_{50}) values of GLPG1837 for the WT, Y304A, and S308A CFTR (Fig. 4, C to E). The EC_{50} value of the WT CFTR was determined to be 0.12 ± 0.03 μM (Fig. 4C), similar to the reported value of 0.23 ± 0.12 μM (*24*). Replacing Y304 or S308 with an alanine increased the EC_{50} values by 57- and 34-fold, respectively (Fig. 4, D and E), underscoring the importance of the structurally observed hydrogen bonds in GLPG1837 recognition.

Here, we have described the structures of CFTR in complex with two separate potentiators. These structures allow us to reach several conclusions. First, although these two potentiators are chemically dissimilar, they both bind to the same site within the transmembrane region of CFTR. This explains why ivacaftor and GLPG1837 are competitive in electrophysiological and binding assays (Fig. 3, C and D) (*24*). Second, because the drug binding site coincides with a hinge involved in gating (*19*, *20*), we propose that the presence of a drug in the pocket stabilizes the open configuration of the pore relative to the closed. In electrophysiological experiments, this stabilization is manifested as an increased opening rate and a decreased closing rate (Fig. 3A) (*14*, *24*). The absence of observable protein structural differences between the drug-bound and drug-free conformations is not surprising given that an open probability increase from 0.64 to 0.88 (Fig. 3A) corresponds to a drug-induced energy change in the closed-open equilibrium (i.e., $\Delta\Delta G$) of just over 1 k_BT (where k_B is the Boltzmann constant and T is temperature). Third, the drug-binding pocket identified here is likely a hotspot for the action of CFTR potentiators. It is now possible to use these structures of CFTR bound to two different potentiator molecules, together with computation, to identify new potentiators.

REFERENCES AND NOTES

1. J. S. Elborn, *Lancet* **388**, 2519–2531 (2016).
2. S. H. Cheng *et al*., *Cell* **66**, 1027–1036 (1991).
3. https://www.cftr2.org/mutations_history.
4. S. H. Cheng *et al*., *Cell* **63**, 827–834 (1990).
5. M. J. Welsh, A. E. Smith, *Cell* **73**, 1251–1254 (1993).
6. S. Hadida *et al*., *J. Med. Chem.* **57**, 9776–9795 (2014).
7. S. E. Van der Plas *et al*., *J. Med. Chem.* **61**, 1425–1435 (2018).
8. F. Van Goor *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18843–18848 (2011).
9. J. L. Taylor-Cousar *et al*., *N. Engl. J. Med.* **377**, 2013–2023 (2017).
10. C. E. Wainwright *et al*., *N. Engl. J. Med.* **373**, 1783–1784 (2015).
11. B. W. Ramsey *et al*., *N. Engl. J. Med.* **365**, 1663–1672 (2011).
12. D. Keating *et al*., *N. Engl. J. Med.* **379**, 1612–1620 (2018).
13. J. C. Davies *et al*., *N. Engl. J. Med.* **379**, 1599–1611 (2018).
14. K. Y. Jih, T. C. Hwang, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4404–4409 (2013).
15. P. D. Eckford, C. Li, M. Ramjeesingh, C. E. Bear, *J. Biol. Chem.* **287**, 36639–36649 (2012).
16. F. Van Goor *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18825–18830 (2009).
17. F. Liu, Z. Zhang, L. Csanády, D. C. Gadsby, J. Chen, *Cell* **169**, 85–95.e8 (2017).
18. Z. Zhang, J. Chen, *Cell* **167**, 1586–1597.e9 (2016).
19. Z. Zhang, F. Liu, J. Chen, *Cell* **170**, 483–491.e8 (2017).
20. Z. Zhang, F. Liu, J. Chen, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12757–12762 (2018).
21. R. G. Coleman, M. Carchia, T. Sterling, J. J. Irwin, B. K. Shoichet, *PLOS ONE* **8**, e75992 (2013).
22. T. E. Balius *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6839–E6846 (2017).
23. Y. C. Yu, Y. Sohma, T. C. Hwang, *J. Physiol.* **594**, 3227–3244 (2016).
24. H. I. Yeh, Y. Sohma, K. Conrath, T. C. Hwang, *J. Gen. Physiol.* **149**, 1105–1118 (2017).

ACKNOWLEDGMENTS

We thank M. Ebrahim and J. Sotiris at Rockefeller's Evelyn Gruss Lipper Cryo-Electron Microscopy Resource Center for assistance in data collection and F. Glickman of the Rockefeller High-throughput and Spectroscopy Resource Center for help with the SPA experiments. We also thank P. Olinares and B. Chait for their mass spectrometry efforts to identify the density associated with R933, and D. Gadsby, R. MacKinnon, and T.-C. Hwang for advice on the electrophysiology experiments. Lastly, we thank C.-H. Lee and M. Lin for helpful discussions. **Funding**: This work is supported by the Howard Hughes Medical Institute (to J.C), the Cystic Fibrosis Foundation Therapeutics (to J.C), the Charles H. Revson fellowship in Biomedical Science (to Z.Z), and R35GM122481 (to B.K.S). **Author contributions**: F.L. and Z.Z. determined the cryo-EM structures of the CFTR–ivacaftor and CFTR–GLPG1837 complexes, respectively. F.L. performed the SPA assays. J.L. performed the bilayer experiments. F.L., Z.Z., and K.K.T. performed the patch-clamp experiments. A.L and B.K.S refined the structures of ivacaftor and GLPG1837 and performed the SAR analysis. J.C. conceptualized the study and analyzed the structures. J.C and F.L. wrote the manuscript with input from all authors. **Competing interests**: The authors have no competing interests. **Data and materials availability:** Cryo-EM density maps of the CFTR–drug complexes have been deposited in the Electron Microscopy Data Bank under the accession codes EMD-0611 (CFTR–Ivacaftor) and EMD-0606 (CFTR–GLPG1837). Atomic coordinates have been deposited in the Protein Data Bank under accession codes 6O2P (CFTR–Ivacaftor) and 6O1V (CFTR–GLPG1837). Raw data for Figs. 2 to 4 are available in the supplementary materials. This work is dedicated to the memory of David C. Gadsby.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1184/suppl/DC1
Supplementary Text
Materials and Methods
Figs. S1 to S8
Tables S1 to S3
References (*25*–*50*)

23 January 2019; accepted 31 May 2019
10.1126/science.aaw7611

CORAL REEFS

Demographic dynamics of the smallest marine vertebrates fuel coral reef ecosystem functioning

Simon J. Brandl[1,2]*, Luke Tornabene[3], Christopher H. R. Goatley[4], Jordan M. Casey[5,6,7], Renato A. Morais[8,9], Isabelle M. Côté[1], Carole C. Baldwin[10], Valeriano Parravicini[5,6], Nina M. D. Schiettekatte[5,6], David R. Bellwood[8,9]

How coral reefs survive as oases of life in low-productivity oceans has puzzled scientists for centuries. The answer may lie in internal nutrient cycling and/or input from the pelagic zone. Integrating meta-analysis, field data, and population modeling, we show that the ocean's smallest vertebrates, cryptobenthic reef fishes, promote internal reef fish biomass production through extensive larval supply from the pelagic environment. Specifically, cryptobenthics account for two-thirds of reef fish larvae in the near-reef pelagic zone despite limited adult reproductive outputs. This overwhelming abundance of cryptobenthic larvae fuels reef trophodynamics via rapid growth and extreme mortality, producing almost 60% of consumed reef fish biomass. Although cryptobenthics are often overlooked, their distinctive demographic dynamics may make them a cornerstone of ecosystem functioning on modern coral reefs.

How coral reefs maintain high diversity and productivity in oligotrophic tropical oceans—often termed "Darwin's paradox"—remains poorly understood (*1*–*3*). Both pelagic subsidies (*2*) and internal energy and nutrient cycling (*3*) have been put forward as drivers of productivity on reefs, but their relative importance has not been resolved.

Fishes form the largest reservoir of consumer biomass on reefs and are involved in most internal energy and nutrient fluxes (*4*), but they also bridge the pelagic zone–reef interface during their larval development (*5*) and may therefore subsidize reefs with pelagic productivity (*6*). Over the past two decades, reconstructions of larval dispersal pathways (*7*, *8*) have revealed the importance of larval dynamics for reef fish ecology, evolution, and conservation (*9*–*11*). However, the role of larval stages in coral reef ecosystem functioning remains virtually unknown.

[1]Department of Biological Sciences, Simon Fraser University, Burnaby, BC V5A 1S6, Canada. [2]Tennenbaum Marine Observatories Network, Smithsonian Institution, Edgewater, MD 21037, USA. [3]School of Aquatic and Fishery Sciences and the Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98105, USA. [4]Function, Evolution and Anatomy Research (FEAR) Lab and Palaeoscience Research Centre, School of Environmental and Rural Science, University of New England, Armidale 2351, Australia. [5]PSL Université Paris: EPHE-UPVD-CNRS, USR 3278 CRIOBE, Université de Perpignan, 66860 Perpignan, France. [6]Laboratoire d'Excellence "CORAIL," Perpignan, France. [7]Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20650, USA. [8]ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville 4811, Australia. [9]College of Science and Engineering, James Cook University, Townsville 4811, Australia. [10]Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20650, USA.
***Corresponding author. Email: simonjbrandl@gmail.com**

Integrating published surveys of coral reef fish larvae, new data on adult reef fish communities, and a theoretical population model (*12*), we demonstrate that cryptobenthic reef fishes [a group of 17 reef-associated fish families characterized by species <50 mm in length (*12*, *13*)] play a previously unrecognized but critically important role in coral reef ecosystem functioning because of their larval dynamics, rapid growth, and extreme mortality.

Cryptobenthic larvae greatly outnumber large reef fish larvae near coral reefs (Fig. 1) despite limited adult reproductive outputs. As more gametes produce more larvae, family-specific larval abundance predictably increases with gamete output for both cryptobenthic and large reef fishes [calculated from adult densities, fecundity, and spawning frequency (*12*, *14*, *15*)]. However, the respective slopes differ drastically, with the relationship for cryptobenthics being more than an order of magnitude steeper than that for large reef fishes (Fig. 2 and table S1), and this difference is consistent across the locations studied (Australia, Belize, and French Polynesia). We found no statistical evidence for effects of pelagic larval duration (i.e., the time that larvae spend in the plankton) or broadcast (i.e., eggs released into the water column) versus benthic clutch spawning (i.e., distinct clutches of guarded eggs) on larval supply (table S2) despite the high prevalence of benthic clutch spawning in cryptobenthics. Furthermore, the difference in slopes was not affected by the uncertainty in the input data, including increases in spawning frequency for cryptobenthics (fig. S2 and tables S3 and S4) and uncertainties in the body mass–fecundity relationship, adult densities, and larval composition (figs. S3 to S5). Thus, cryptobenthics are disproportionally more successful at converting adult gametes into larval supply, which likely underpins their extreme abundance in the nearshore ichthyoplankton.

The abundance of cryptobenthic larvae, despite limited gamete output, may provide the continuous (*16*) and copious inflow of larvae that is assumed necessary for adult population maintenance in small, short-lived reef fishes (*17*). Our results show that this, in turn, forms the basis of a critical energy and nutrient pump that operates across the pelagic zone–reef interface and may help to explain the enigmatic productivity of coral reef ecosystems (Fig. 3). Using a demographic model that simulates the daily arrival, growth, and death of larval recruits from all reef fish families over 1 year (fig. S6), we estimate that juvenile and adult cryptobenthics provide most (57.5 ± 0.1% SE) of the consumed fish biomass on reefs. However, because of extreme mortality rates and small sizes, cryptobenthics appear to make negligible contributions to net productivity and standing fish biomass (Fig. 3C), which mirrors empirical evidence (*18*). Standing biomass is the most commonly quantified metric of ecosystem functioning on reefs (*19*); however, it does not capture the rapid turnover in cryptobenthic populations (693.1 ± 2.7% SE annually) that is enabled by their extraordinary demographic dynamics. Thus, cryptobenthics represent the "dark productivity" of coral reefs, which fuels reef fish biomass production but is rarely perceived because it is consumed almost as quickly as it is produced.

This role of cryptobenthics is empirically reflected in their extreme mortality [up to 70% per week (*20*, *21*)] and their consumption by virtually any predator capable of eating them (*21*, *22*). Although the community-wide representation of cryptobenthics in fish diets frequently appears lower than the ~58% identified herein (*23*), their true contribution to coral reef trophodynamics may be obscured by (i) rapid digestion, precluding reliable visual identification (*24*); (ii) predation on cryptobenthics by invertebrates (*22*), which are fed on by larger fishes; and (iii) predation on cryptobenthics by juvenile predatory fishes [e.g., cryptobenthics comprise up to 88.6% of fish prey for juvenile groupers (*25*, *26*)], which are rarely included in community-wide dietary analyses.

The key to the unique demographic dynamics of cryptobenthics and their productivity might be a shift away from long-range dispersal toward retention of larvae in the immediate vicinity of natal (home) reefs. Four lines of evidence, along with our findings, support this hypothesis. First, larval dispersal models show that larval supply easily maintains adult populations in large-bodied, long-lived reef fishes (*7*). Conversely, small-bodied, short-lived taxa appear unable to sustain local populations even when active swimming by larvae is considered (*7*), yet cryptobenthic populations persist. Near-complete retention of cryptobenthic larvae close to natal reefs may solve this paradox. Second, driven by olfactory and auditory cues, cryptobenthic larvae show stronger natal homing than similarly sized large reef fishes (*27*) and can have very short dispersal distances (*28*). Third,

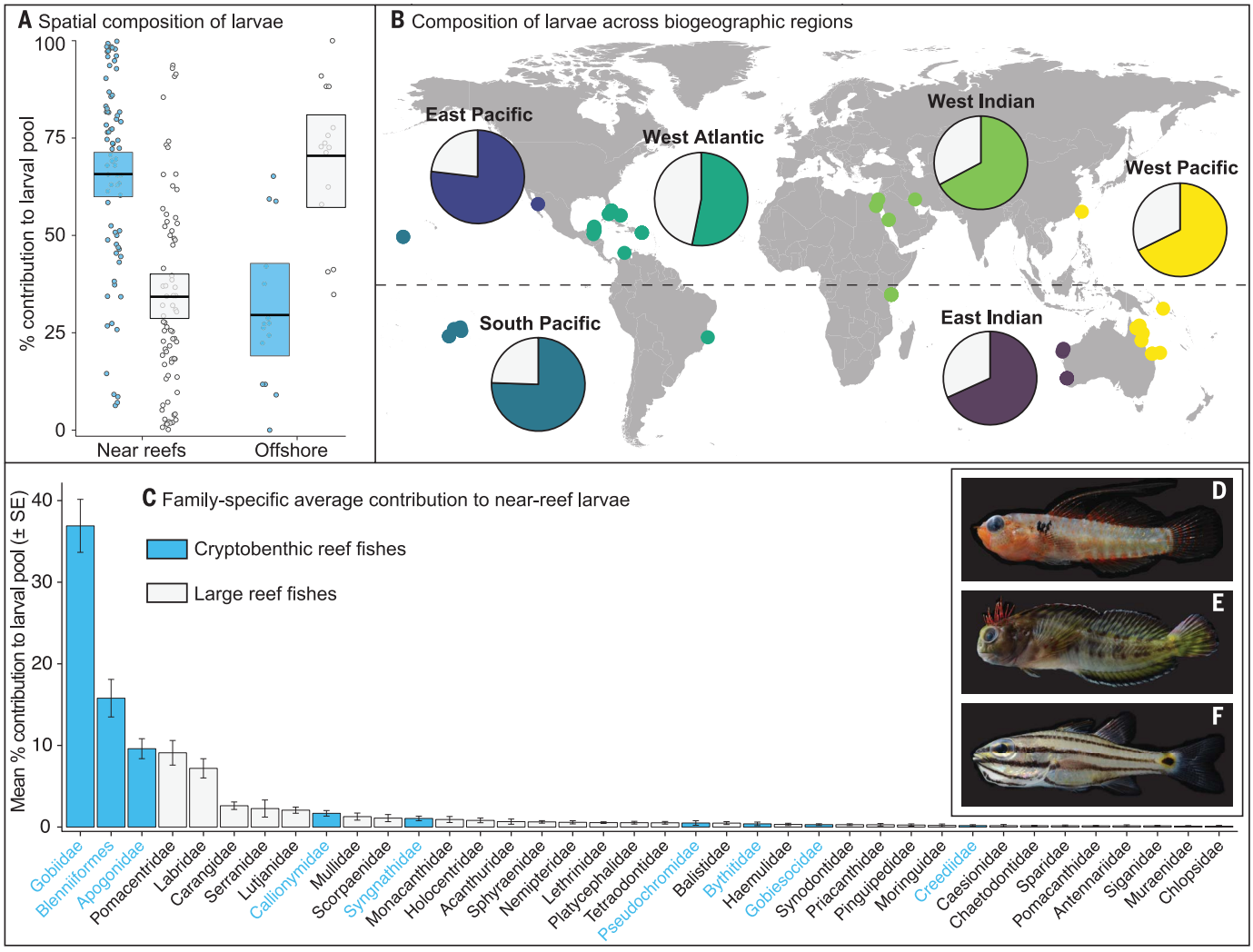
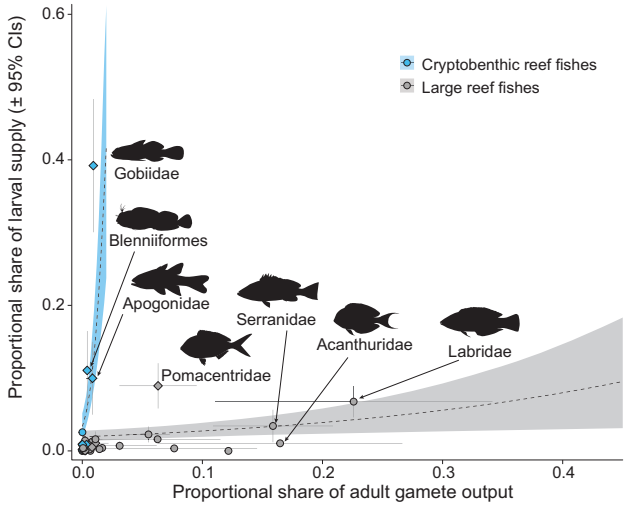


Fig. 1. Global dominance of cryptobenthic reef fishes in the near-reef ichthyoplankton. (**A**) Cryptobenthic larvae (blue) account for two-thirds (65.7%) of the larval reef fish pool <10 km from reefs, whereas large reef fish larvae (light gray) dominate >10 km from reefs. Crossbars represent predicted medians (±95% credible intervals) from a Bayesian beta-regression model; circles represent the raw data (*12*). (**B**) The high proportion of cryptobenthics (colored pie slices) in the near-reef ichthyoplankton is consistent across major biogeographic coral reef regions. Dots represent separate studies. (**C**) Average contribution of reef fish taxa to global near-reef ichthyoplankton. The three highest contributing taxa are cryptobenthic, represented by photographs of adult *Eviota infulata* (Gobiidae) (**D**), *Scartella cristata* (Blenniiformes–Blenniidae) (**E**), and *Cheilodipterus quinquelineatus* (Apogonidae) (**F**). Families contributing <0.1% were omitted for clarity.

Fig. 2. Differences in the relationship between larval supply and adult gamete output for cryptobenthic and large reef fishes. Dashed lines and ribbons represent predicted fits from a Bayesian beta-regression model [±95% credible intervals (CIs)]; circles (broadcast spawners) and diamonds (demersal brooders) represent raw data averaged across three sampling locations (±SE). Both axes represent proportional shares.



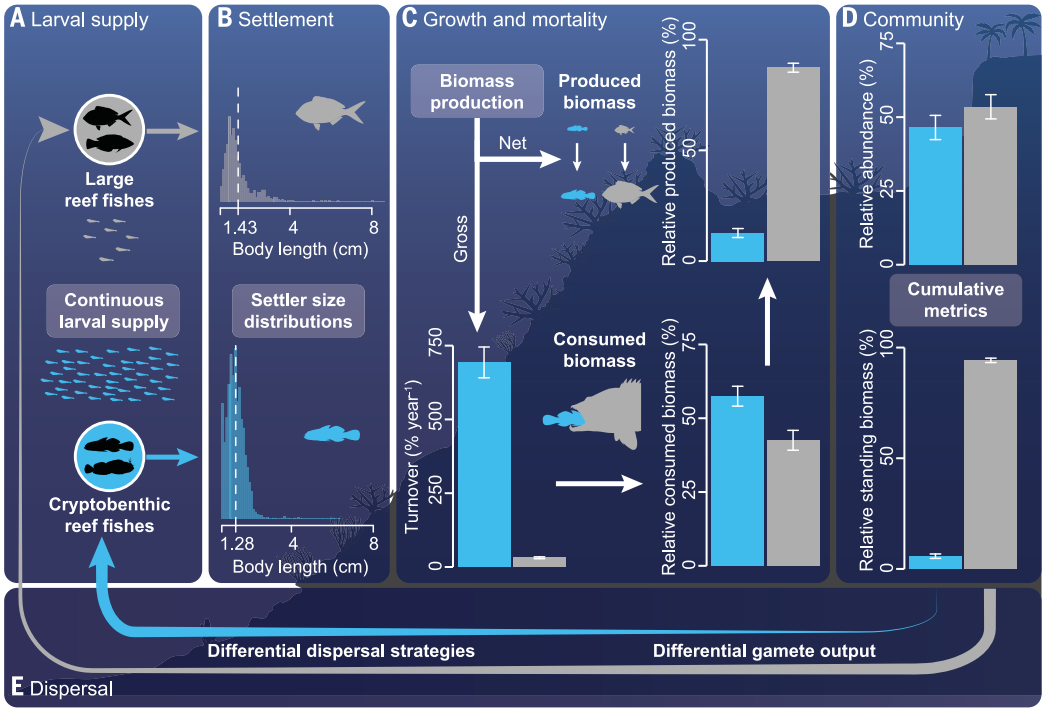


Fig. 3. Ecosystem-scale effects of the demographic dynamics of cryptobenthic reef fishes. (**A**) Cryptobenthics far outnumber large reef fishes in cohorts of larvae that recruit to reefs. (**B**) At settlement, large reef fish recruits are, on average, slightly larger than cryptobenthics. (**C**) Cryptobenthics contribute little (~13%) to net biomass production but produce almost 60% of consumed reef fish biomass because of exceptionally high turnover. (**D**) Standing stock biomass of large reef fishes far outweighs that of cryptobenthics, although adult abundances are approximately even. (**E**) Despite higher gamete output from large reef fishes, cryptobenthic larvae dominate the near-reef ichthyoplankton. This restarts the "crypto-pump" through rapid replenishment of consumed individuals. Uncertainty estimates in (C) and (D) are based on 100 iterations of the full model.

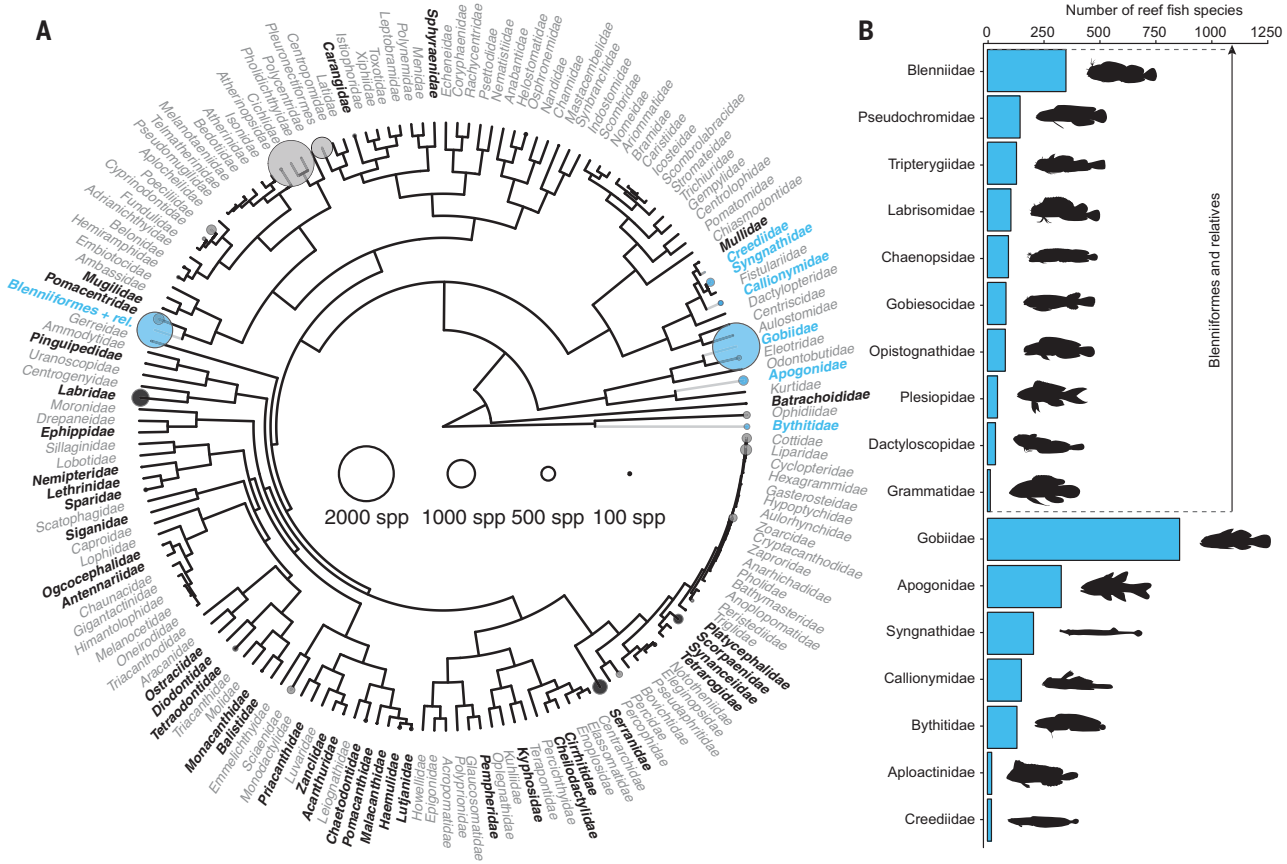


Fig. 4. Phylogenetic positioning and species richness in cryptobenthic and large reef fishes. (**A**) Cryptobenthic reef fishes (blue) have few independent origins but (**B**) account for almost half of all reef fish species (2799 species). Black, reef fish taxa; gray, non–reef fish taxa [as in Brandl *et al.* (*13*)]. Bubble sizes (A) and bars (B) represent species richness within all taxa (all fish species) and cryptobenthic families (reef-affiliated species only). Arrow in (B) indicates cumulative richness in the Blenniiformes.

cryptobenthic larvae have limited yolk sacs and ingest prey immediately after hatching, indicating dependence on resource-rich, near-reef environments (*8*, *29*). Finally, remaining close to natal reefs during development should result in fine-scale genetic structuring. With few exceptions (*30*), this is observed in cryptobenthic fishes (*10*, *27*, *31*). Collectively, this suggests that the "pelagic" larvae of most cryptobenthics remain close to their natal reefs (*32*), resulting in prodigious near-reef abundances and a unique role of cryptobenthics in coral reef ecosystem functioning.

Larval retention may lead to two evolutionary consequences: (i) rapid speciation through microallopatry arising from restricted gene flow among populations and frequent reproductive isolation (*10*, *33*, *34*) and (ii) a higher risk of extinction than commonly assumed for small marine fishes (*35*) because populations can easily become ephemeral and disappear after stochastic environmental changes (*36*). If these processes scale up to macroevolutionary levels, then cryptobenthics should be phylogenetically rare but species rich and this is indeed the case. Few reef fish families have successfully adopted cryptobenthic lifestyles and major diversification events are restricted to two lineages (Fig. 4): the larger Blenniiformes and the Gobiaria (gobies and apogonids). Nevertheless, these lineages are among the most rapidly diversifying clades of actinopterygian fishes (*37*) and cryptobenthics collectively account for almost half (44.5%) of total reef fish biodiversity (Fig. 4) (*13*).

In summary, through their extraordinary larval dynamics, rapid growth, and extreme mortality, the hyperdiverse consortium of abundant, tiny, and short-lived cryptobenthic species appears to be a critical functional group on coral reefs. The "dark productivity" provided by cryptobenthics underpins reef fish biomass production and supports the characteristic fast-paced dynamics of modern coral reefs.

REFERENCES AND NOTES

1. C. Crossland, B. Hatcher, S. Smith, *Coral Reefs* **10**, 55–64 (1991).
2. C. Richter, M. Wunsch, M. Rasheed, I. Kötter, M. I. Badran, *Nature* **413**, 726–730 (2001).
3. J. M. de Goeij *et al*., *Science* **342**, 108–110 (2013).
4. J. E. Allgeier, D. E. Burkepile, C. A. Layman, *Glob. Change Biol.* **23**, 2166–2178 (2017).
5. J. M. Leis, U. Siebeck, D. L. Dixson, *Integr. Comp. Biol.* **51**, 826–843 (2011).
6. J. E. Allgeier, K. E. Speare, D. E. Burkepile, *Ecosphere* **9**, e02216 (2018).
7. R. K. Cowen, C. B. Paris, A. Srinivasan, *Science* **311**, 522–527 (2006).
8. S. E. Swearer, J. E. Caselle, D. W. Lea, R. R. Warner, *Nature* **402**, 799–802 (1999).
9. O. C. Salles *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13245–13250 (2016).
10. M. S. Taylor, M. E. Hellberg, *Science* **299**, 107–109 (2003).
11. G. R. Almany *et al*., *Nat. Ecol. Evol.* **1**, 148 (2017).
12. Materials and methods are available in the supplementary materials.
13. S. J. Brandl, C. H. R. Goatley, D. R. Bellwood, L. Tornabene, *Biol. Rev.* **93**, 1846–1873 (2018).
14. K. Kasimatis, C. Riginos, *Coral Reefs* **35**, 387–397 (2016).
15. G. J. Edgar, R. D. Stuart-Smith, *Sci. Data* **1**, 140007 (2014).
16. C. D. Lefèvre, K. L. Nash, A. González-Cabello, D. R. Bellwood, *Coral Reefs* **35**, 399–409 (2016).
17. M. Caley *et al*., *Annu. Rev. Ecol. Syst.* **27**, 477–500 (1996).
18. M. Depczynski, C. J. Fulton, M. J. Marnane, D. R. Bellwood, *Oecologia* **153**, 111–120 (2007).
19. N. A. J. Graham, S. Jennings, M. A. MacNeil, D. Mouillot, S. K. Wilson, *Nature* **518**, 94–97 (2015).
20. M. Depczynski, D. R. Bellwood, *Ecology* **87**, 3119–3127 (2006).
21. M. A. Steele, G. E. Forrester, *Ecology* **83**, 1076–1091 (2002).
22. C. H. Goatley, A. González-Cabello, D. R. Bellwood, *Coral Reefs* **36**, 207–212 (2017).
23. R. W. Hiatt, D. W. Strasburg, *Ecol. Monogr.* **30**, 65–127 (1960).
24. B. Beukers-Stewart, G. Jones, *J. Exp. Mar. Biol. Ecol.* **299**, 155–184 (2004).
25. J. St John; J. St. John, *Mar. Ecol. Prog. Ser.* **180**, 233–246 (1999).
26. C. K. C. Wen, M. C. Bonin, H. B. Harrison, D. H. Williamson, G. P. Jones, *Coral Reefs* **35**, 451–455 (2016).
27. G. Gerlach, J. Atema, M. J. Kingsford, K. P. Black, V. Miller-Sims, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 858–863 (2007).
28. C. C. D'Aloia *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13940–13945 (2015).
29. J. E. Majoris, F. A. Francisco, J. Atema, P. M. Buston, *Aquaculture* **483**, 286–295 (2018).
30. G. M. Cooke, T. E. Schlub, W. B. Sherwin, T. J. Ord, *PLOS ONE* **11**, e0150991 (2016).
31. C. Riginos, B. C. Victor, *Proc. R. Soc. B.* **268**, 1931–1936 (2001).
32. M. W. Brogan, *Mar. Ecol. Prog. Ser.* **115**, 1–13 (1994).
33. R. Winterbottom, R. H. Hanner, M. Burridge, M. Zur, *ZooKeys* **381**, 79–111 (2014).
34. C. Riginos, Y. M. Buckley, S. P. Blomberg, E. A. Treml, *Am. Nat.* **184**, 52–64 (2014).
35. W. J. Ripple *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10678–10683 (2017).
36. D. R. Bellwood *et al*., *Oecologia* **170**, 567–573 (2012).
37. T. J. Near *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12738–12743 (2013).
38. S. J. Brandl, L. Tornabene, C. H. R. Goatley, J. M. Casey, R. A. Morais, I. M. Côté, C. Carole, Baldwin, V. Parravicini, N. M. D. Schiettekatte, D. R. Bellwood, Data and code for: Demographic dynamics of the smallest marine vertebrates fuel coral-reef ecosystem functioning, Version 1, Zenodo (2019); https://doi.org/10.5281/zenodo.2575005.

ACKNOWLEDGMENTS

We thank V. Huertas, S. Degregori, A. Mercière, and the staff of CRIOBE, Lizard Island, and Carrie Bow Caye for field support and S. Vignieri, K. Wilson, and R. Streit for editorial support and comments. **Funding:** Funding was provided by the BNP Paribas Foundation (V.P.), the Agence National de la Recherche (V.P., S.J.B., D.R.B.), the Smithsonian Institution's TMON (S.J.B.), the National Science and Engineering Research Council of Canada (I.M.C., S.J.B.), and the Australian Research Council (D.R.B.). This is contribution 35 from the Smithsonian's MarineGEO Network. **Author contributions:** S.J.B., C.H.R.G., L.T., and D.R.B. conceived the study; S.J.B., J.M.C., C.H.R.G., V.P., R.A.M., C.C.B., and N.M.D.S. collected data; S.J.B., R.A.M., N.M.D.S., and L.T. analyzed the data; S.J.B. wrote the first draft and all authors contributed to writing the paper. **Competing interests:** None declared. **Data availability:** Data and code are available on Zenodo (*38*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1189/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
Tables S1 to S6
References (*39*–*95*)

Raw Data and Code

6 September 2018; accepted 7 May 2019
Published online 23 May 2019
10.1126/science.aav3384

FISHERIES

The small world of global marine fisheries: The cross-boundary consequences of larval dispersal

Nandini Ramesh[1]*, James A. Rising[2], Kimberly L. Oremus[3]

Fish stocks are managed within national boundaries and by regional organizations, but the interdependence of stocks between these jurisdictions, especially as a result of larval dispersal, remains poorly explored. We examined the international connectivity of 747 commercially fished taxonomic groups by building a global network of fish larval dispersal. We found that the world's fisheries are highly interconnected, forming a small-world network, emphasizing the need for international cooperation. We quantify each country's dependence on its neighbors in terms of landed value, food security, and jobs. We estimate that more than $10 billion in annual catch from 2005 to 2014 is attributable to these international flows of larvae. The economic risks associated with these dependencies is greatest in the tropics.

Marine fisheries supply food and livelihoods to millions of people around the world (*1*). Though fisheries are typically managed at the scale of national exclusive economic zones (EEZs), many fish populations are connected beyond EEZ boundaries (*2*–*6*). Whereas pelagic species can be tracked across international borders as adults (*7*), nonpelagic populations connect primarily via the dispersal of fish eggs and larvae, forms that cannot yet swim by ocean currents (*2*, *8*). Larval connectivity patterns have been analyzed at both the regional (*6*, *9*–*12*) and global (*4*, *13*, *14*) levels and have been used to suggest changes for spatial management and conservation (*12*, *15*). However, the impact on fisheries of larval connectivity across EEZs is not well understood, even though more than 90% of the world's fish are caught within EEZs (*16*).

On the scale of a single species or region, this connectivity can be analyzed empirically through genetic testing (*9*, *10*). For analyses on larger scales, dispersal patterns can be estimated using biophysical models that combine oceanographic data with an understanding of the biology of the stocks (*4*, *14*, *17*). Such efforts can be challenging, because species vary widely in larval timing and duration and currents vary with the seasons; therefore, generalizations can be misleading. More realistic inputs can be achieved by using life history traits for each species, including time and place of spawning and larval duration. Sensitivity analyses can help to ensure that results are robust to changes in key assumptions (*14*), while empirical bounding can safeguard against predicting unrealistic dispersion outcomes (*6*).

Network analysis has previously been applied to marine systems to describe the connectivity of plankton communities (*18*), local fishing communities (*19*, *20*), and marine reserves (*14*). Networks of larval flows have been used to identify "hub" subpopulations for protection on a regional scale (*12*).

In this study, we combined oceanographic and life history data for 706 species and 434 genera of commercially harvested fish to estimate their connectivity across 249 EEZs and constructed a network representing the larval flows between

[1]Department of Earth and Planetary Science, University of California, Berkeley, Berkeley, CA 94720, USA. [2]Grantham Research Institute, London School of Economics, London, UK. [3]School of Marine Science and Policy, University of Delaware, Newark, DE 19716, USA.
*Corresponding author. Email: nandiniramesh@berkeley.edu

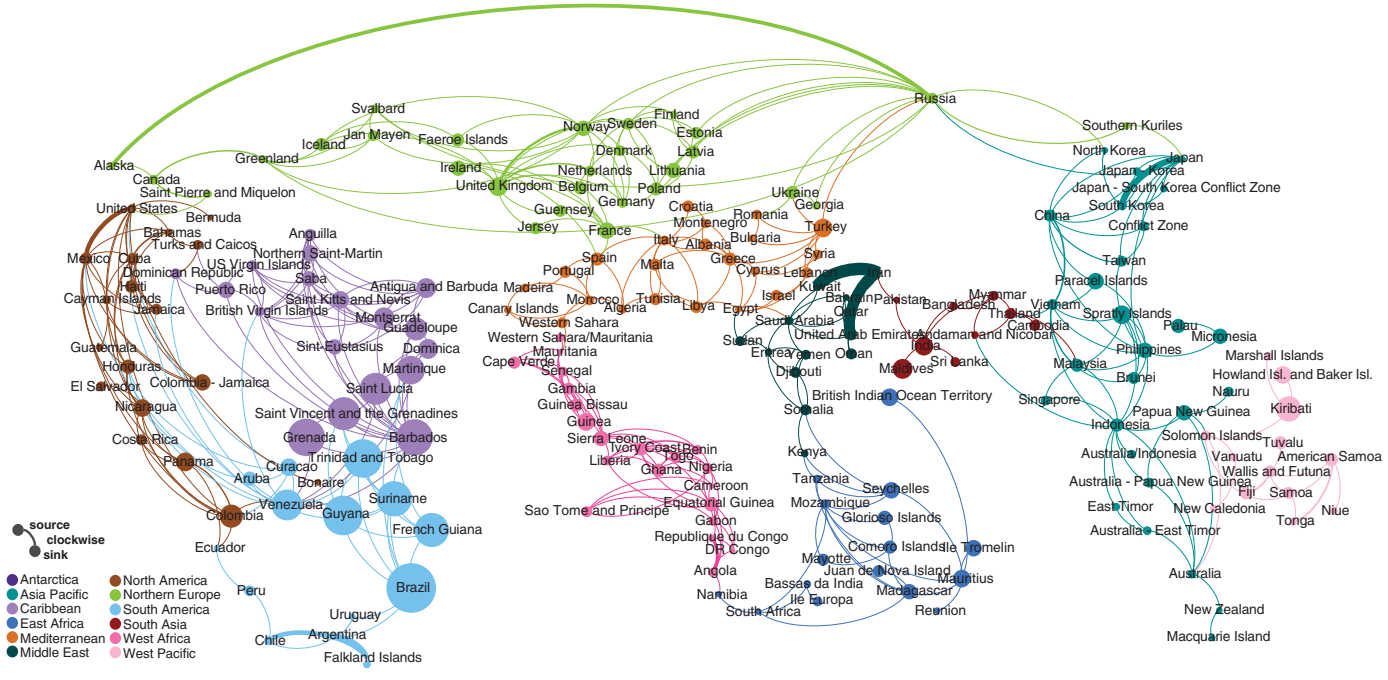


Fig. 1. The network of spawn-attributed catch flows between EEZs. Each EEZ is a node (circle) of the network and its color represents its network community. The connectors or edges in this network flow clockwise from source to sink, with their thicknesses representing the magnitude of the net flow of caught biomass between the EEZs. Only the edges in the upper tercile of edge weights are shown, for clarity (see SM 3.2 for the full network). The size of each node represents its out-degree, i.e., the number of other EEZs for which it acts as a source of fish larvae, including connections not shown in this image.

nations. Nations that depend heavily upon their neighbors for recruitment risk losing part of their catch if the fisheries in the source EEZs outside their jurisdiction are poorly managed. We quantified these risks in economic terms and identified regional "hotspots" of risk for catch, fishery employment, and food security.

We used a particle-tracking system (*21*) with time-varying ocean currents (*22*) and species-specific life histories (*23*) to simulate the dispersal of eggs and larvae through a dynamic ocean. We placed multiple simulated particles for each species based on the timing and location of that species' spawning and let them drift for their larval duration to obtain a probabilistic estimate of species-specific larval trajectories. We used a random-walk parameterization (*21*) that adds a small velocity at every time step to account for turbulent motion at small scales [see section 3.1.2 in the supplementary materials (SM 3.1.2)].

We empirically bounded our results by discarding particles that arrived in regions where the species is not present in observed catch data (*16*). For a given EEZ, catch is attributed based on the proportion of particles arriving there from each spawning country (see SM 1.1). This proportionality forms the core assumption of our model. We tested our main results with a series of analyses of sensitivity to this assumption. These included reducing spawn floating duration to account for uncertainties in spawning mortality (*2*, *24*), introducing return adult spawning migration (*25*) (see SM 3.6), and distinguishing different levels of recruitment limitation.

We estimated how much of each country's observed catch comes from its neighbors by constructing for each species a transition matrix that describes the probability of its offspring dispersing from one EEZ to another. This transfer of biomass between nations' EEZs is represented as a network in Fig. 1.

Each connector of the network represents net flows of fish from one country to another. Countries that depend on inflows of juvenile fish to maintain their local populations require international cooperation to ensure sustainable fisheries. Our analysis of these flows revealed that a large proportion of marine fisheries within EEZs form a single, global network (Fig. 1).

We found that the global network of marine fisheries is a scale-free, small-world network. The scale-free network property, common in natural systems (*26*), is characterized by an exponential distribution of the number of connections from each node (see SM 3.2). This exponential degree distribution results in a "hub-and-spoke" structure that is resilient to random disturbances because of the large number of less-connected countries from which disturbances do not easily propagate to other parts of the network. However, a disturbance to any of the highly connected hubs in a scale-free network can affect numerous surrounding nodes. In this context, habitat destruction, overfishing, or environmental change in a hub EEZ could have impacts that spread beyond its own boundaries. Conversely, targeted efforts to manage fisheries within these hub EEZs could benefit many nations.

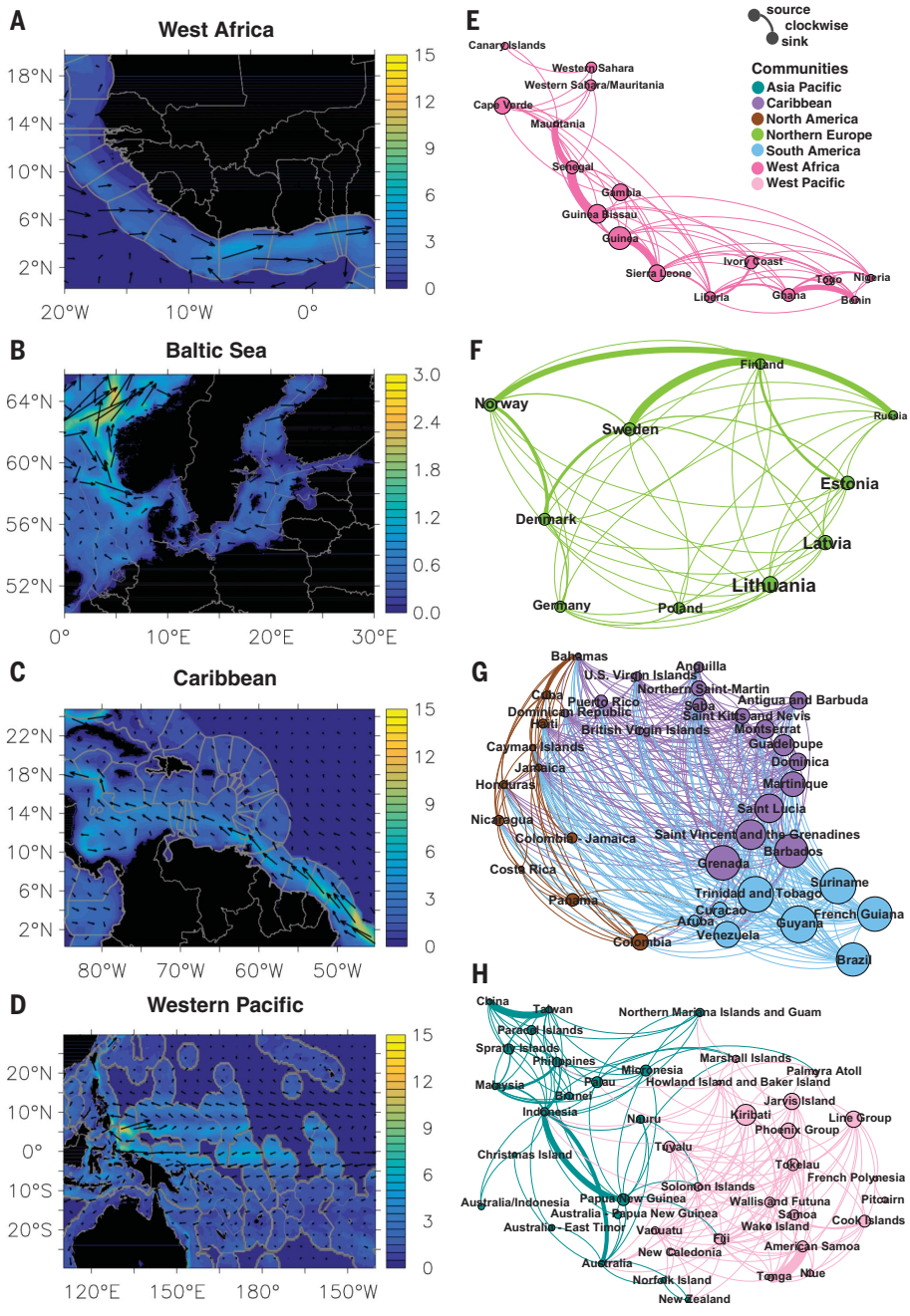


Fig. 2. Regional currents and community networks. (**A** to **D**) The speed (shown in colors, in centimeters per second) and direction (arrows) of ocean surface currents in four regions with interconnected fisheries (West Africa, Baltic Sea, the Caribbean, and Western Pacific) during the month of maximum spawning activity in each (August, May, June, and May, respectively). (**E** to **H**) The corresponding subset of the global network encompassed by the four regions. Colors, node sizing, and connector directions are as for Fig. 1. Nodes are arranged to approximately correspond to geographic locations of the EEZs.

To demonstrate the relationship between currents and the network of larval dispersal, we focused more closely on four regions (Fig. 2). The differences between the regional networks and average current speed arise from the details of current speeds during spawning, larval duration, and empirical observations of species presence or catch. The influence of the Guinea Current on the connectivity of West Africa's fisheries can be seen in the large number of EEZs that act as sources to their eastward neighbors, especially between Guinea-Bissau and Nigeria. While the strongest connections are typically

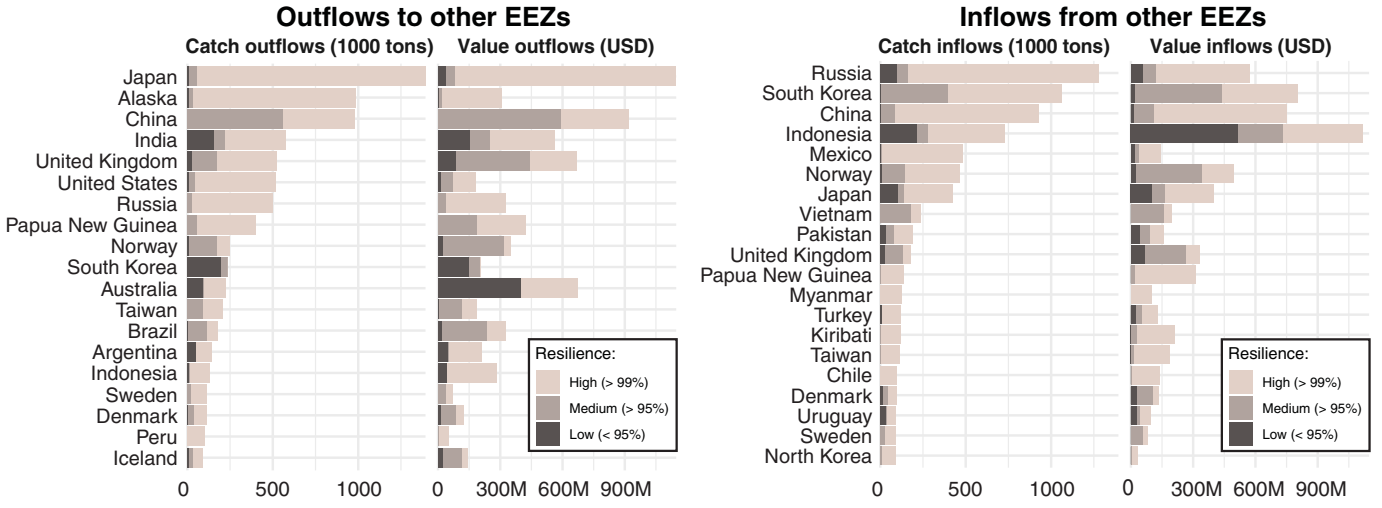


Fig. 3. Countries with highest outflowing and inflowing catches. (**Left**) Top 20 countries sorted by total outflowing catch (in thousands of tons) and value [in U.S. dollars (USD)] at risk. (**Right**) Top 20 countries sorted by total inflow of catch (thousands of tons) and value (USD) at risk. For both catch and landed values, data from 2005 to 2014 were used and attributed to larvae, by species. Resilience levels represent the estimated decline a population can endure without being considered vulnerable to local extinction.

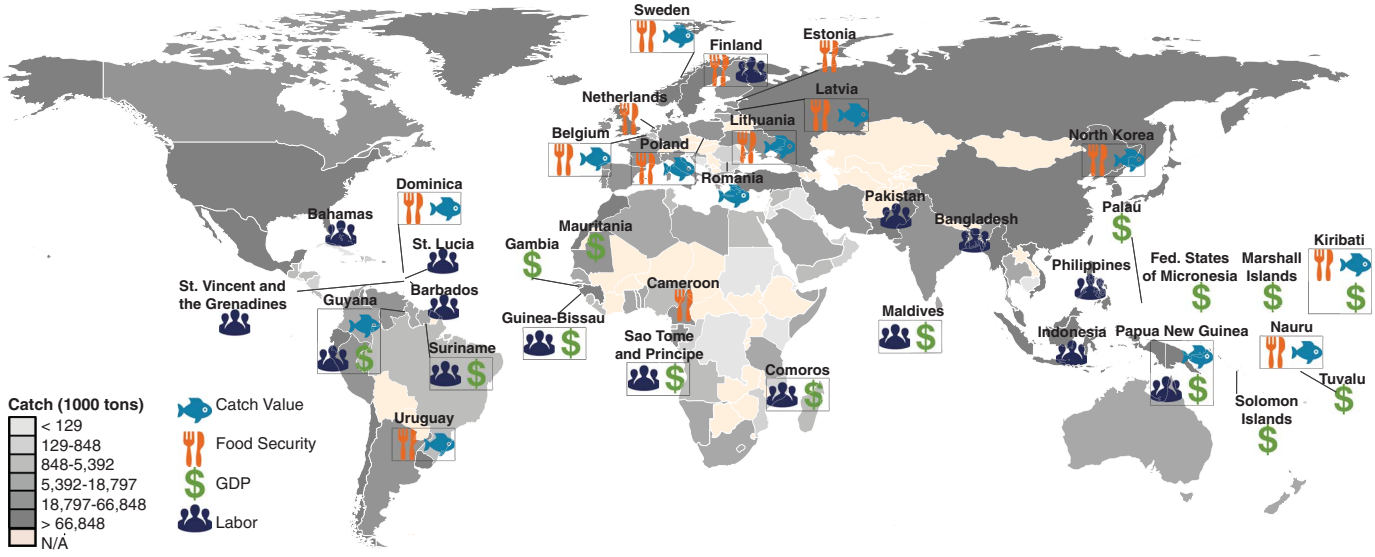


Fig. 4. Hotspot map showing fishing dependency on spawning grounds in neighboring waters, by country. Countries are shaded by catch (in thousands of metric tons) at risk, with darker shades representing higher catches. Icons depict EEZs that are the most dependent on their neighbors. The catch icon indicates that more than 30% of a country's catch value is dependent on neighboring spawning grounds, the GDP icon represents a risk to more than 0.8% of its GDP, the labor icon represents that more than 1.5% of its jobs are vulnerable, and the food security icon represents a food security dependence index >1.1%.

between adjacent EEZs, many connections also extend over longer distances. In contrast, the Baltic Sea has substantially weaker currents. There, the largest outward flows originate from Sweden and Norway, which have the region's longest coastlines. In the Caribbean, the North Brazil Current flows northwestward along the South American coast, and consequently many of the EEZs lying along this current act as sources for the Lesser Antilles. Within the Lesser Antilles, the density of small EEZs gives rise to a highly interconnected, complex network structure. The effect of the northward flow along this island chain can be inferred from the larger node sizes among the EEZs lying in its southern portion. In the Western Pacific, strong currents dominate in the equatorial ocean, with weaker currents at higher latitudes. The large areas encompassed by this region's EEZs mean that, unlike the other regions, most connections are between immediate neighbors.

The small-world property implies that it is possible to traverse the global network in a small number of steps, on average. Within this network, there exist smaller clusters or communities that are tightly connected. Most of these clusters internally exhibit the small-world property. In theory, this property of the global fisheries network suggests that disturbances to a large hub could propagate via cascading effects on the surrounding spokes.

A key question is whether disruptions to a given EEZ actually propagate in this manner. A stock's response to external shocks depends on both its population dynamics and mortality from fishing, which can be affected by management

(*27*). Some fish stocks are biologically capable of replenishing themselves when their numbers dwindle, provided that fishing pressure is relieved, reducing the likelihood that disturbances will propagate. However, "recruitment-limited" stocks are vulnerable to a decline in spawning population, making it more likely that disturbances will spread across the network even if the receiving fisheries are managed. We adopted FishBase's classification of stock resilience as a proxy for this type of density dependence. For high-resilience stocks, which are generally not recruitment limited, our measure of stock dependence overestimates the extent to which stocks will be reduced if recruitment inflows fail. For those stocks classified as having medium or low resilience, however, we found a strong correlation between our simulation's predictions and observed variances in stock levels (see SM 3.5). Even for countries whose fisheries mostly comprise non–density-dependent stocks, these larval inflows serve as a buffer against fishery collapse within their waters.

To contextualize our results, we estimated the economic significance of the network's international connections. First, we considered the amount and value of catch that flows in and out of each EEZ (Fig. 3). Japan, China, and Alaska are responsible for the greatest outflows, reflecting their productive waters. However, having fewer neighbors makes them smaller hubs (Fig. 1). Indonesia has the most landed value attributable to other countries, due to its high-value catch and many neighbors. The countries with the greatest catch inflows are generally those with the largest fisheries.

Next, we identified nations that are potentially most vulnerable, in socioeconomic terms, to the management of neighboring waters (see table S5). In Fig. 4, we highlight countries that depend the most on the spawning grounds of neighbors in terms of their total catch, gross domestic product (GDP), number of jobs in the fishery industry, and a fishery food security dependence index (*28*). The most vulnerable nations are concentrated in the hotspot regions of the Caribbean, West Africa, Northern Europe, and Oceania. The risks to national GDP and labor force are generally highest in the tropics. However, our measure of food security risk also identified a few European nations.

Our analysis showed that about $10 billion worth of annual marine catch may rely on transnational exchanges of fish offspring. These dependencies form a single global network, indicating that marine fisheries, even within national boundaries, constitute an interconnected, globally shared resource.

This network's scale-free and small-world properties imply that fish stocks from a small number of EEZs provide benefits to a large number of "downstream" countries. The most vulnerable nations are clustered in a few hotspot regions (Fig. 4). This pattern lends further support to the use of international frameworks, such as large marine ecosystems and marine protected area networks (*29*, *30*).

Further research is needed to understand how small-scale coastal processes, larval behavior, and fisheries management affect this connectivity. Beyond the spawning connections studied here, national fisheries are interdependent through the movement of adult fish, population shifts under climate change, and international fishing treaties. In particular, the role of adult fish migration in driving international connectivity remains an important question. While a more detailed analysis is required to accurately describe dispersal pathways of individual species, this study highlights the role of larval connectivity across international boundaries and the need for multilateral cooperation for sustainable management of these shared resources.

REFERENCES AND NOTES

1. FAO, "The state of world fisheries and aquaculture: Contributing to food security and nutrition for all" (Food and Agriculture Organization of the United Nations, 2016).
2. R K. Cowen, S. Sponaugle, *Annu. Rev. Mar. Sci.* **1**, 443–466 (2009).
3. A. Di Franco *et al.*, *Biol. Conserv.* **192**, 361–368 (2015).
4. E. Popova *et al.*, *Mar. Policy* **104**, 90–102 (2019).
5. B. P. Kinlan, S. D. Gaines, *Ecology* **84**, 2007–2020 (2003).
6. A. S. Kough, C. B. Paris, M. J. Butler 4th, *PLOS ONE* **8**, e64970 (2013).
7. B. A. Block *et al.*, *Nature* **434**, 1121–1127 (2005).
8. D. A. Siegel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8974–8979 (2008).
9. S. Planes, G. P. Jones, S. R. Thorrold, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5693–5697 (2009).
10. N. K. Truelove *et al.*, *Fish. Res.* **172**, 44–49 (2015).
11. M. Dubois *et al.*, *Glob. Ecol. Biogeogr.* **25**, 503–515 (2016).
12. J. R. Watson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, E907–E913 (2011).
13. S. Wood, C. B. Paris, A. Ridgwell, E. J. Hendy, *Glob. Ecol. Biogeogr.* **23**, 1–11 (2014).
14. M. Andrello *et al.*, *Nat. Commun.* **8**, 16039 (2017).
15. S. D. Gaines, C. White, M. H. Carr, S. R. Palumbi, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18286–18293 (2010).
16. D. Pauly, D. Zeller, Eds., *Sea Around Us: Concepts, Design and Data* (Univ. of British Columbia, 2015).
17. M. J. Fogarty, L. W. Botsford, *Oceanography* **20**, 112–123 (2007).
18. B. F. Jönsson, J. R. Watson, *Nat. Commun.* **7**, 11239 (2016).
19. E. C. Fuller, J. F. Samhouri, J. S. Stoll, S. A. Levin, J. R. Watson, *ICES J. Mar. Sci.* **74**, 2087–2096 (2017).
20. E. T. Addicott *et al.*, *Can. J. Fish. Aquat. Sci.* **76**, 56–68 (2019).
21. C. B. Paris, J. Helgers, E. van Sebille, A. Srinivasan, *Environ. Model. Softw.* **42**, 47–54 (2013).
22. J. A. Carton, B. S. Giese, *Mon. Weather Rev.* **136**, 2999–3017 (2008).
23. R. Froese, D. Pauly, FishBase, Version 11/2014 (2014); www.fishbase.se/search.php.
24. C. C. D'Aloia *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13940–13945 (2015).
25. A. Hastings, L. W. Botsford, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6067–6072 (2006).
26. D. J. Watts, S. H. Strogatz, *Nature* **393**, 440–442 (1998).
27. D. Pauly *et al.*, *Nature* **418**, 689–695 (2002).
28. M. Barange *et al.*, *Nat. Clim. Chang.* **4**, 211–216 (2014).
29. B. S. Halpern, S. E. Lester, K. L. McLeod, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18312–18317 (2010).
30. J. Lubchenco, *Stress, Sustainability, and Development of Large Marine Ecosystems during Climate Change: Policy and Implementation* (UNDP and GEF, 2013).
31. N. Ramesh, J. Rising, K. Oremus, The Small World of Global Marine Fisheries: The Cross-Boundary Consequences of Larval Dispersal, Version 1.0, Zenodo (2019); http://doi.org/10.5281/zenodo.2636745.

ACKNOWLEDGMENTS

The authors thank D. Dookie, M. Burgess, M. A. Cane, A. Chaintreau, A. Carlisle, J. Cohen, C. Costello, R. Defries, S. Gaines, S. Hsiang, C. Moffat, and C. Szuwalski for comments, suggestions, and references. This work was initiated and partly conducted at Columbia University in the City of New York, which provided computing resources. **Funding:** N.R. was partially supported by the National Aeronautics and Space Administration (NASA) Headquarters under the NASA Earth and Space Science Fellowship Program, grant NNX-14AK96H. **Author contributions:** N.R. performed the network analysis and Lagrangian modeling. J.A.R. performed the country-level risk analysis. N.R., J.A.R., and K.L.O. designed the study, collected data, and wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All newly organized data used in this study and the intermediate and final results data are publicly available at Zenodo (*31*). Analysis reproduction code is available at https://github.com/openmodels/small-world-fisheries.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6446/1192/suppl/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 to S8
References (*32*–*44*)

6 September 2018; accepted 23 May 2019
10.1126/science.aav3409

Wide-Mouth Bottle and Cap System

Designed to help scientists accommodate sample quantities from 250 mL to 20 L, the Media Bottle and Multiport Cap System offered by DWK Life Sciences is used for biological and chemical reactions, plant cell culture, photobiological studies, and sample collection and storage. The wide neck simplifies transfer of samples, including powders, pastes, granules, and liquids. Spoons, spatulas, tweezers, and large funnels are easily manipulated through the opening. Made from 33 expansion, low extractable borosilicate glass in conformance to U.S. Pharmacopeia Type I standards, the bottle is resistant to acids, corrosive chemicals, and thermal shock, withstanding temperatures as high as 500°C. When combined with the multiport cap, the bottle is suitable for dilutions, fermentation, and culturing. It is available in clear glass, an amber glass coating for light-sensitive materials, or with KimCote, a safety coating for added protection in the event of breakage.

DWK Life Sciences
For info: 800-225-1437
www.dwk.com/bioreactor

High-Yield Protein Production in CHO Cells

CHOgro High Yield Expression System from Mirus Bio offers a rapid means for developing protein-based biotherapeutics within a single-host cell type, with titers that exceed all other commercially available transient expression systems. Transient gene expression has been gaining popularity as a means for producing antibodies and other recombinant proteins in mammalian cells. As the demand for recombinant proteins in drug discovery continues to grow, and with the majority of therapeutic antibodies being produced in Chinese hamster ovary (CHO) cells, there is a need for a fast, reliable, and robust transient expression platform. CHOgro High Yield Expression System consists of CHOgro Expression Medium, L-glutamine and poloxamer 188 medium supplements, CHOgro Complex Formation Solution, TransIT-PRO Transfection Reagent, and CHOgro Titer Enhancer. This combination streamlines the protein expression process—high-antibody titers can be achieved in seven days.

Mirus Bio
For info: 888-530-0801
www.mirusbio.com

Automated Colony Counting System

AutoCOL from Synbiosis is an integrated system that allows microbiologists to load plates and count colonies on up to 100 plates in 30 min—ideal for improving throughput and accuracy in highly regulated laboratories. The plate feeder automatically identifies plates using a barcode reader and counts colonies on any agar plate. With AutoCOL, microbiologists can set pass/fail limits and the system will track and stack pass/fail plates in specific positions after reading them, ensuring accurate, walk-away counts for microbial applications, such as environmental monitoring, quality control labs, bioburden testing, and water analysis.

Synbiosis
For info: 800-686-4451
www.synbiosis.com

Automated NGS Library Prep

Free yourself from manual NGS library prep even on low-throughput experiments, with the epMotion 5073m NGS solution from Eppendorf. Until now, automated construction of next-generation sequencing (NGS) libraries has only been economical for preparations of at least 96 samples. Smaller sample numbers had to be processed manually using pipettes, which is timely and can incorporate pipetting errors. epMotion can process up to 24 samples simultaneously in a fully automated fashion. Smooth operation is guaranteed by ample space for pipette tips, three required single- and multichannel pipetting tools with a range of 5 µL–300 µL, and a special rack that accommodates all required solutions. The optimized epBlue software enables smart use of the single- and multichannel dispensing tools to optimize pipetting steps and tip use, even in the case of odd sample numbers. The system saves time and money, and reduces waste—all on a footprint of only six letter-sized pieces of paper.

Eppendorf
For info: 800-645-3050
www.eppendorf.com/ngs-compact

Microbioreactor System

The ambr 15 cell culture Generation 2 from Sartorius Stedim Biotech replicates laboratory-scale bioreactor performance at the microscale (10 mL–15 mL) and controls up to 48 single-use bioreactor cultures in parallel. The system offers increased flexibility and expanded capability for clone selection, media and feed optimization, and early process development work. It also provides a robust screening platform for development of cell- and gene-therapy processes, including HEK293 for viral vector production, T cells, induced pluripotent stem cells, and other immune-derived cell lines. The ambr clone selection software, included as a one-year license, provides simplified, streamlined multivariate data analysis for faster, more consistent cell-line screening and ranking. Adaptation of cell lines to different media for media screening studies can be performed in ambr 15 microbioreactors using new culture-passage steps and rapid vessel-drain functionality. New media-mixing steps automate the creation of media blends, eliminating the need to premix, saving time, and reducing errors.

Sartorius Stedim Biotech
For info: +49-(0)-551-308-0
www.sartorius.com

LC/MS and GC/MS Sample Preparation

Together with sister company J.G. Finneran Associates, Porvair Sciences offers an unmatched selection of innovative glass vials, seals, closures, and microplates to enhance liquid and gas chromatography (LC/GC) and mass spectrometry (MS) workflows. In addition, MS-certified kits for sample preparation and use in autosamplers are available to simplify and save time. The highly acclaimed P3 (Protein Precipitation Plate) sets the standard for efficiently and economically removing proteins from biological samples that can interfere with downstream LC/MS and GC/MS applications. For over 20 years, Porvair Sciences' Microlute 96-well SPE plates have been helping scientists increase the throughput of their MS analyses by speeding up sample preparation. The complete Microlute system provides a matched filter plate, vacuum manifold, and choice of collection plates with the option of a dedicated sample concentrator. In independent comparative testing, Porvair Sciences' deep-well plates have been repeatedly shown to contain near-zero contaminants, ensuring protection against leaching and interference with your LC/MS or GC/MS analyses.

Porvair Sciences
For info: 800-552-3696
www.microplates.com/resources/porvair-chromatography-catalogue

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.

VIRGINIA TECH™

Faculty Position in Cardiovascular Sciences

Virginia Tech (**https://vt.edu/**) is recruiting to fill tenure-track open rank faculty positions in cardiovascular sciences. The newly recruited faculty members will be part of a major initiative to expand the existing cardiovascular biomedical research team at the Fralin Biomedical Research Institute (FBRI) at Virginia Tech Carilion (**https://research.vtc.vt.edu/**) located on the Health Sciences and Technology campus in Roanoke, Virginia.

The research institute, led by Dr. Michael J. Friedlander and located in a 100,000 purpose designed facility is entering its tenth year with over 25 research teams, 350 employees and students with over $125M in extramural research funding. A new 140,000 SF research building representing a $90M investment from the state, Virginia Tech and Carilion Clinic and major additional investments from the university including a transformative $50M gift to the research institute, will open in spring of 2020. The new building will house state-of-the art labs and core facilities including those for *in vitro*, laboratory animal model, human subject and computational based research programs - high field (9.4T) MRI, micro PET, surgical research suites, metabolic chambers, scanning block face EM, full molecular biology, cellular imaging, histology, data analytics and research participant monitoring core facilities. Recruitment of outstanding faculty in **cardiovascular sciences** will build on existing research strengths in cardiomyocyte coupling, drug development, sudden cardiac death, arrhythmogenesis and vascular remodeling and regulation. Faculty will be embedded in a transdisciplinary research environment that will include additional new teams of investigators focused on metabolism and obesity, infectious disease/immunology, neuroscience and body-device interfaces and comparative oncology.

Applicants for all positions should have an earned Doctorate (PhD, MD, DVM or combined degrees), postdoctoral research experience in cardiovascular science, a demonstrated record of accomplishment of substantial research productivity through major publications and either active major extramural funding (preferably from NIH) or strong evidence of the likelihood of obtaining such funding in the near term. The new faculty will join the FRBRI's highly successful and nationally regarded Center for Heart and Reparative Medicine Research (CHRMR) led by Dr. Robert Gourdie (**https://research.vtc.vt.edu/heart_center/**). Collaborations will be encouraged within CHRMR, with the other groups at the FBRI and across Virginia Tech and with clinical colleagues at the adjacent Carilion Clinic, including the division of cardiology under the direction of Dr. David Sane.

Successful candidates will be expected to establish and maintain a high-quality, extramurally-funded research program and participate in teaching and mentoring of graduate students as well as mentoring of undergraduate and medical students. Highly competitive salary, space, and start-up packages will be provided.

Faculty members can draw from an outstanding and diverse pool of students to work with them in their research program. Specifically, Virginia Tech has over 6,000 graduate students with graduate programs in multiple biomedical disciplines including Translational Biology, Medicine, and Heath (**http://www.tbmh.vt.edu**), Molecular and Cellular Biology (**https://www.mcb.vt.edu**), Biomedical Engineering and Mechanics (**https://beam.vt.edu/graduate.html**) and Biomedical Sciences and Pathobiology (**https://www.vetmed.vt.edu/academics/bmvs/**). In addition, the Virginia Tech Carilion School of Medicine (**https://medicine.vtc.vt.edu/**) is one of the nation's most research intensive allopathic medical schools where every medical student carries out a four year hypothesis-driven research project under faculty mentorship.

The FBRI at VTC is located in Roanoke, Virginia and is a public-private partnership between the Commonwealth of Virginia's leading research-intensive university – Virginia Tech – and a major health care system – Carilion Clinic in Roanoke, VA. The research institute is located in the picturesque Roanoke Valley midway between Washington, D.C. and Charlotte, NC.

Applications will be reviewed continuously until the position is filled. It is expected that initial interviews will begin in mid-summer of 2019. To apply, please submit your application including curriculum vitae, statement of research accomplishments and future research plan, and statement of teaching/mentoring philosophy at **www.jobs.vt.edu**, posting **# TR0190068**. Also, have at least three references post their letters of support to the same site.

Virginia Tech recognizes the critical importance of diverse teams of scholars. It seeks to diversify its faculty along multiple dimensions and encourages applications from under-represented backgrounds. Virginia Tech is a public global land-grant university, committed to research, teaching and learning, and outreach to the Commonwealth of Virginia, the nation, and the world. Building on its motto of *Ut Prosim* (that I may serve), Virginia Tech is dedicated to InclusiveVT- **https://www.inclusive.vt.edu/** serving in the spirit of community, diversity, and excellence. We seek candidates who adopt and practice the Principles of Community, which are fundamental to our on-going efforts to increase access and inclusion and to create a community that nurtures learning and growth for all of its members.

Virginia Tech is an Equal Opportunity Employer.

For inquiries regarding non-discrimination policies, contact the Office of Equity and Access at 540-231-2010 or Virginia Tech, North End Center, Suite 2300 (0318), 300 Turner St. NW, Blacksburg, VA 24061.

DESDE 1902
INSTITUTO DE HIGIENE E MEDICINA TROPICAL
UNIVERSIDADE NOVA DE LISBOA

UNIVERSIDADE NOVA DE LISBOA

The Universidade NOVA de Lisboa invites applications for a full time position as Director of the Instituto de Higiene e Medicina Tropical (IHMT NOVA)

The position of Director of the Instituto de Higiene e Medicina Tropical (IHMT NOVA) of Universidade NOVA de Lisboa is open for applications.

The Director, whether of Portuguese or other nationality, must have a doctorate degree (PhD). Applications for Director must be from Full Professors or Research Coordinators from the Universidade Nova de Lisboa or from other teaching universities and research institutions, national or international. The Director must have scientific seniority in areas of interest to the mission of the IHMT NOVA (post-graduation education, scientific research, internationalization and cooperation for development) and relevant management experience. A good knowledge of spoken and written Portuguese and English languages is required.

The Director must have initiative and strategic vision, as well as capacity to implement the statutory mission of the IHMT NOVA, promoting national and international contacts, and representing the IHMT in the Universidade NOVA de Lisboa and in the national and international contexts.

The Director is also the main responsible for fundraising and promotion of the IHMT NOVA external image.

Applications will be treated confidentially and **must be sent before 12 pm on the 15th of July, 2019** to **eleicaodiretor2019@ihmt.unl.pt**

Other information may be found in the vacancy notice published at the IHMT site: https://www.ihmt.unl.pt/concurso-para-diretor-do-instituto-de-higiene-e-medicina-tropical-da-universidade-nova-de-lisboa/

Additional information can be requested to the Electoral Commission via the e-mail eleicaodiretor2019@ihmt.unl.pt

BURROUGHS WELLCOME FUND

President

Mission
The Burroughs Wellcome Fund is an independent private foundation dedicated to advancing the biomedical sciences by supporting research and other scientific and educational activities.

The Burroughs Wellcome provides approximately $40 million in annual funding across the biomedical research landscape and in STEM education in the Fund's home state of North Carolina. BWF believes that a diverse scientific workforce is essential to the process and advancement of research innovation, academic discovery, and public service.

The Opportunity
The Burroughs Wellcome Fund seeks an innovative and creative leader dedicated to advancing biomedical science, promoting career development for biomedical researchers, developing leaders in the research enterprise, and championing STEM education.

The President is responsible for developing, executing, and evaluating BWF competitive award programs and other program-related activities in targeted areas of biomedical research and STEM education. In particular, she or he will provide the visionary leadership for the Fund's grantmaking strategy. The President will also promote the mission of the Fund by engaging with various networks, audiences, and communication channels. Further information on the position can be found in the President's job description.

To assure full consideration, applicants must submit by **July 15, 2019,** a curriculum vitae, a statement of interest (maximum of two pages) that includes one's vision for directions for the Fund, and the name and contact information for three references. Interested candidates should submit these materials electronically to **Barbara Evans** at **bevans@bwfund.org**.

Questions regarding the details of this position should be directed to Russ Campbell, Secretary of the Board of Directors of the Burroughs Wellcome Fund, at rcampbell@bwfund.org.

BWF is an Equal Opportunity Employer and encourages women and candidates underrepresented groups to apply.

See www.bwfund.org for full details.

By **Pedro Resende**

Learning how to pivot

After 5 years as a postdoc and co–principal investigator in an academic research lab, I am about to start my own company. It's a big step, and I'm excited—and a little anxious—about how it will work out. I am very aware that things might not go the way I envision. But pivoting does not feel as intimidating as it could—because I've done it before, starting early in my career. In doing so, I've learned that career changes should not be feared. Instead, they are valuable opportunities for professional development and growth.

When I finished my undergraduate training 14 years ago, I wanted to continue in academic research. But none of my applications for Ph.D. programs was successful. I wondered whether spending time in an industry job might help me stand out. At the same time, I worried that going to industry would compromise my chances of ultimately pursuing an academic career. Colleagues and professors told me that a move to industry would be a path of no return, and that big career changes could be rough. It felt like a momentous decision that would set the tone for the rest of my career.

"Professors told me that a move to industry would be a path of no return."

Nonetheless, I decided to take the leap and accept a position at a biotech company. To my surprise, I loved working there. In fact, my experience was so positive that I changed plans. I did not want to do a Ph.D. anymore; I wanted a career in industry. I thought I had found my vocation, and that my path was clear.

But my position was just a 1-year contract, and when I started to look for my next industry job, I hit a bureaucratic obstacle. I was searching for jobs across Europe, and in many countries my undergraduate degree was not considered sufficient training for the positions I wanted. So much for that supposedly clear path.

It looked like it was time for another pivot—back to where I had started, applying to Ph.D. programs. Again, I wondered whether moving back and forth between industry and academia might put me at a disadvantage with potential future employers. But my previous transition had taught me I should be open-minded and embrace opportunities as they come, even if they are not my top choice at the time. You never know how things are going to turn out. So, I tried not to be too anxious about changing course again.

During one interview for a Ph.D. program, the selection committee focused on how uncommon it was for someone working in industry to want to come back to academia. I had played my cards right; I stood out! The program accepted me, and I enjoyed my Ph.D. experience. I loved the science, had a great mentor, and lived in a fantastic city. Even though I had pictured a different course, I didn't have any regrets.

When I finished my Ph.D., I still wanted to return to industry. I thought my previous industry experience coupled with my improved credentials would make the transition relatively easy, but that did not end up being the case. I wasn't getting any job offers. I was disappointed, but I took some solace in remembering how well pivoting had worked out for me in the past. I was now much more comfortable moving from plan A to plan B—or even plan C.

So, I started to apply for postdoctoral positions. I got a great offer to start an independent line of research in a group with an outstanding working environment, and my time there has been great. In addition to my research, I've developed my entrepreneurial and leadership skills as co-founder and president of my Ph.D. program's alumni association. Looking back, plan B feels like a plan A—just as it did so many times before in my back-and-forth career. And now I'm ready to embark on plan C: setting out as an entrepreneur.

I accepted a job in industry when my mind was in academia. I found a job in academia when all I wanted was a job in industry. Yet I feel happy with my career path, and I look forward to the pivots yet to come. ■

Pedro Resende is a research associate at i3S (the Institute for Research and Innovation in Health) in Porto, Portugal. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER