Are most human cancers due to bad luck? *pp. 12 & 78*

An exodus of scientists from eastern Ukraine *p. 14*

Optimizing chemistry at the microgram scale *p. 49*

Science

$10
2 JANUARY 2015
sciencemag.org

AAAS

An evolving threat

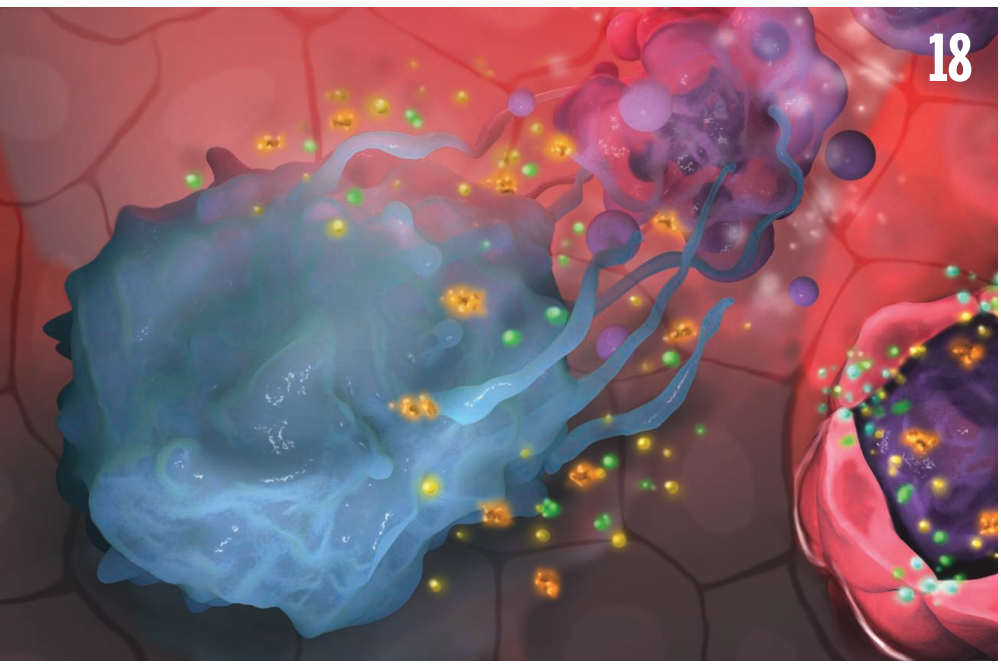
How gene flow sped the evolution of the malarial mosquito *pp. 23, 42, & 43*

CONTENTS

2 JANUARY 2015 • VOLUME 347 • ISSUE 6217

26 & 67

How fat cells kill *Staphylococcus aureus*



18

NEWS

IN BRIEF

IN DEPTH

FEATURES

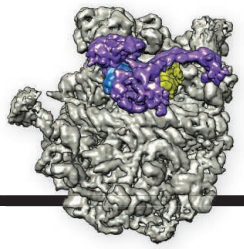
INSIGHTS

LETTERS

PERSPECTIVES

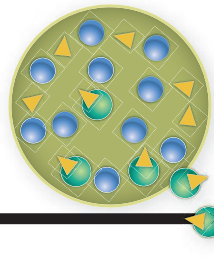
BOOKS *ET AL.*

Published by AAAS



75

Elongating proteins without mRNA



32 & 83

Signals from the lysosome regulate aging

RESEARCH

IN BRIEF

REVIEW

RESEARCH ARTICLES

MOSQUITO GENOMICS

REPORTS

DEPARTMENTS

ON THE COVER

Colored scanning electron micrograph of the mouthparts of a female *Anopheles gambiae* mosquito. Sheathed in a scaly lower lip ending in two hairy lobes (blue) are a pair of saw-toothed stylets (green) used to cut into the skin. Because only female *Anopheles* blood feed, only females transmit malaria—a disease that claims the lives of more than half a million children each year. See pages 27, 42, and 43. *Image: Science Photo Library*

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the 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.

Domestic individual membership and subscription (51 issues): $153 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $1282; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #1254 88122. 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:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. **Authorization to photocopy** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $30.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Published by AAAS

Science
AAAS
1200 New York Avenue, NW, Washington, DC 20005
Bateman House, 82-88 Hills Road, Cambridge, UK CB2 1LQ

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

INSTITUTIONAL SITE LICENSES 202-326-6755 **REPRINTS:** Author Inquiries 800-635-7181 **COMMERCIAL INQUIRIES** 803-359-4578 **PERMISSIONS** 202-326-6765, permissions@aaas.org **AAAS Member Services** 202-326-6417 or http://membercentral.aaas.org/discounts

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 of conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and books 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.

INFORMATION FOR AUTHORS See pages 680 and 681 of the 7 February 2014 issue or access www.sciencemag.org/about/authors

Published by AAAS

EDITORIAL

Data, eternal

Marcia McNutt
Editor-in-Chief
Science *Journals*

During 2014, *Science* worked with members of the research community, other publishers, and representatives of funding agencies on many initiatives to increase transparency and promote reproducibility in the published research literature. Those efforts will continue in 2015. Connected to that progress, and an essential element to its success, an additional focus will be on making data more open, easier to access, more discoverable, and more thoroughly documented. My own commitment to these goals is deeply held, for I learned early in my career that interpretations come and go, but data are forever.

During my qualifying exam to advance to Ph.D. candidacy, I drew a chalkboard cartoon of a then-new concept: that the weight of recently erupted oceanic volcanoes could elastically deform the surrounding seafloor, creating a deep depression and surrounding flexural arch. Afterward, H. W. Menard, the great marine geologist and a member of my exam committee, spread out a map of the Pacific. He pointed out places where there were older coral atolls (which marked former stands of sea level) that were either now uplifted or drowned in the vicinity of younger volcanoes. Using the distance from the young volcano to the atoll and the amount of uplift or depression, we were able to calibrate the long-term flexural strength of the Pacific seafloor under the weight of the volcanic loading. It was of no matter that Menard had published a paper years earlier using a subset of the uplifted atolls to argue for another hypothesis, which he now happily discarded in favor of the flexural warping one. It occurred to me that there was no database of "drowned and uplifted atolls" that one could access. Menard's prior publication provided only a biased sampling of all occurrences. Had he not been on my exam committee, that unique set of observations to constrain the flexural rigidity might never have presented itself.

"...interpretations come and go, but data are forever."

Data, particularly those collected with public funding, should be used so that they do the most good. When the greatest number of creative and insightful minds can find, access, and understand the essential features that led to the collection of a data set, the data reach their highest potential. Although the situation has improved some four decades after my student days in terms of the number of public data repositories, requirements for making data available, and metadata standards, there is still a long way to go. So what can *Science* do to help in this regard, given that it covers many disciplines but is not deeply embedded in any one field?

There are many publicly and privately funded data repositories worldwide, not all of which are being used to their full potential. In 2015, we want to work with authors and readers to identify which of those repositories *Science* should promote because they are well managed, have long-term support, and are responsive to community needs. For data that do not neatly fit into large-scale repositories, we will explore other available options. We also will evaluate different ways to tag data sets and integrate such tagging into our peer-review process. For example, one might associate a digital identifier for a data set with a figure in a paper. A reviewer could use such an identifier to find the particular data that are related to the figure. The hope is to work with repositories that allow bidirectional tagging so that it is easy for someone—a reviewer or reader—to identify the data used in a *Science* paper.

Along with improving the "discoverability" of data sets, *Science* hopes to inspire creative ways to visualize data sets to improve the communication of information and concepts and even facilitate the discoverability of new phenomena. What happens when you bring together those who collect large data sets with those who develop the tools to analyze and view them? Stay tuned!

– **Marcia McNutt**

10.1126/science.aaa5057

NEWS

> “It’s effectively a new machine, poised to set us on the path to new discoveries.”
>
> **CERN Director-General Rolf Heuer** on the Large Hadron Collider, which goes back online in early 2015 after 2 years of upgrades and maintenance.

IN BRIEF

2015: A LOOK AHEAD

What’s hot | What’s not

British chemist Humphry Davy once said that “nothing is so fatal to the progress of the human mind as to suppose that our views of science are ultimate ... that there are no new worlds to conquer.” In that spirit, *Science* takes a look at trends and ideas that preoccupied the scientific community last year—and makes some guesses at what new themes are likely to take hold in 2015. Our—subjective!—list, in no particular order, for your consideration:

A robot that went soft, at Harvard University.

Soft robots | Stiff robots

Sorry, R2-D2: Inspired by animals and armed with better materials and compact hardware, squishy robots take center stage.

Polio in Pakistan | Polio in Nigeria

As the disease disappears from its African stronghold, cases are soaring in Pakistan.

Senator Lamar Alexander (R–TN) Representative Lamar Smith (R–TX)

Reforming federal oversight of U.S. higher education may eclipse last year’s brawl over National Science Foundation peer review.

Europa or bust | Asteroid capture

Congress lights a fire under NASA to visit Jupiter’s moon.

European Political Strategy Centre E.U. science adviser

Anne Glover’s stint in Brussels ends; a new entity will give the European Commission scientific advice.

Next Generation Science Standards | Common Core

Science advocates hope to avoid missteps that have plagued math and reading standards.

iPS cell clinical trials | STAP cells

A simple recipe for stem cells was too good to be true, but reprogrammed adult cells move ahead.

ISRO, CNSA | NASA

NASA’s mission drifts while the ambitions of India’s and China’s space programs grow.

Chikungunya | MERS

Worries about the respiratory virus in the Arabian Peninsula ease, but a mosquito-borne agent is exploding in the Americas.

Paris climate talks Lima climate talks

The debate moves away from whether developing nations should cut carbon emissions ... to by how much.

Exoplanet atmospheres Exoplanet orbits

We know where extrasolar planets are—now the more penetrating questions begin.

“I’m not a scientist.” Direct attacks on science

U.S. politicians reframe their rhetorical assaults on climate change and evolution.

Reproducibility | Glamour journals

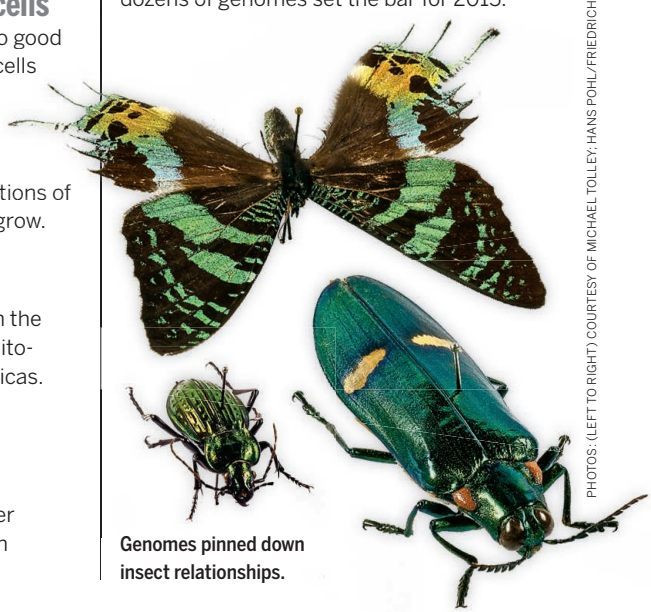
As retractions mount in high-profile journals such as *Science*, *Nature*, and *Cell*, the community pushes for reproducible experiments.

Ebola drug and vaccine trials U.S. Ebola panic

Efficacy trials in West Africa will likely determine whether an Ebola vaccine works—and can help end this epidemic.

Evolutionary trees with dozens of genomes | Trees with dozen of genes

More DNA means better family trees; recent bird and insect phylogenies built from dozens of genomes set the bar for 2015.



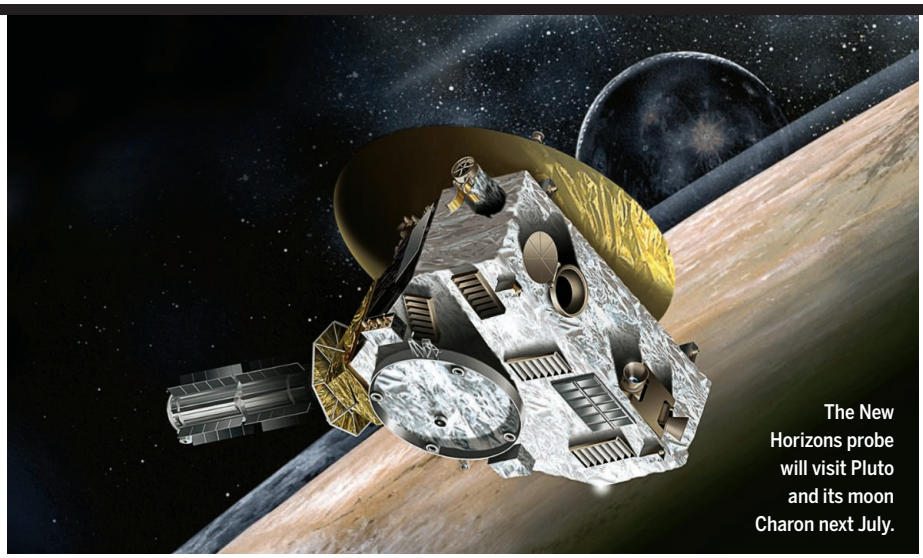
Genomes pinned down insect relationships.

PHOTOS: (LEFT TO RIGHT) COURTESY OF MICHAEL TOLLEY; HANS POHL/FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Published by AAAS

0.64 Predicted global mean temperature increase, in °C, for 2015 over the 1961 to 1990 average—which would make 2015 the warmest year on record, according to the U.K.-based Met Office.

50% Proportion of people with HIV who will be over 50 in 2015, a first. The shift presents physicians with new challenges.



The New Horizons probe will visit Pluto and its moon Charon next July.

Precision medicine
National Children's Study
After killing an ambitious health study, the National Institutes of Health (NIH) ponders a big investment in personalized treatments.

Drones doing science
Science making better drones
Hurdles remain (battery life, airspace regulations), but drones now collect data from poles to sea floor to clouds.

Carbon sinks | Carbon balances
Space measurements of fluorescence from photosynthesis offer a way to directly monitor carbon uptake and better calculate net atmospheric exchanges.

Standard three-neutrino model
Sterile neutrinos
The Planck spacecraft dealt a blow to the theorized fourth, "sterile" neutrino—by confirming predictions of the standard theory of cosmology.

Alaskan earthquakes
Induced earthquakes
Earthquakes due to wastewater injection may drop with the price of oil. But a major deployment of sensors will help monitor Alaska, the United States' most seismic state.

Dwarf planets | Comets
After one spacecraft landed on a comet, others head for Pluto and Ceres.

Twitter data grants
Facebook experiments
Facebook's "emotional contagion" study angered users, while Twitter gives a few select institutions access to its data.

Helpful microbes | Poop therapy
Fecal transplants can restore the gut's ecosystem—but future therapies cut out the middleman, delivering just the isolated beneficial bacteria.

South African fossils
Homo heidelbergensis
The supposed European ancestor of modern humans and Neandertals falls out of favor, while high-profile South African digs may yield new finds.

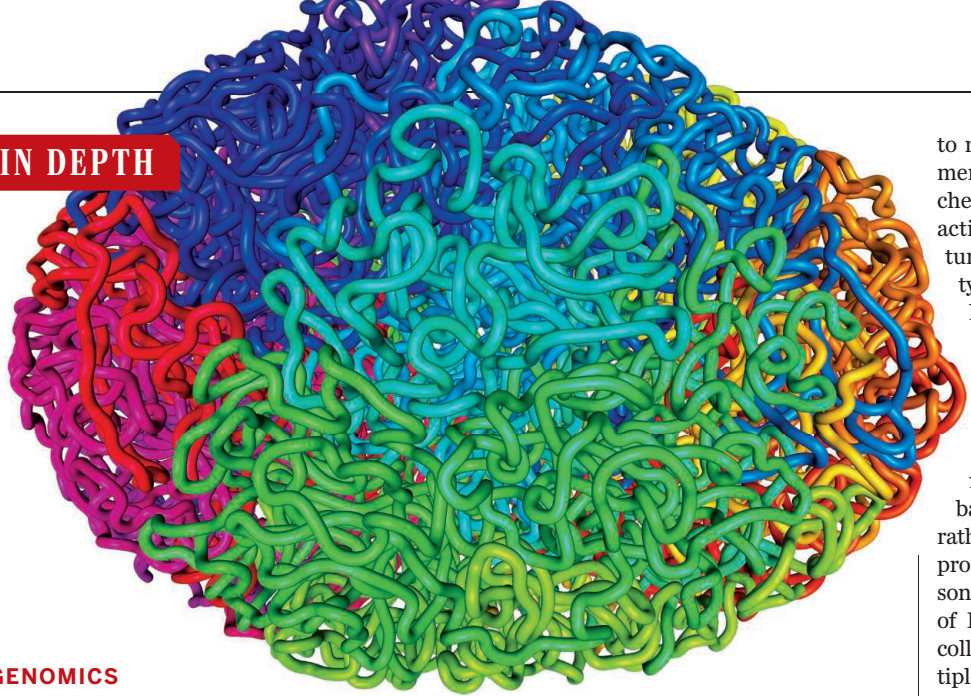
Representative Andy Harris (R–MD)
Senator Tom Harkin (D–IA)
NIH loses longtime champion, gets tea party member who wants more funding for young scientists.

Seal flu | Bird flu
Influenza's latest surprise is a massive die-off of European harbor seals from a subtype called H10N7.

A potential flu victim.

IMAGES: (TOP TO BOTTOM) JOHNS HOPKINS UNIVERSITY APPLIED PHYSICS LABORATORY/SOUTHWEST RESEARCH INSTITUTE; © BLICKWINKEL/ALAMY

Published by AAAS



GENOMICS

Inching toward the 3D genome

Maps of DNA's loops and folds advance—but may disagree

By **Elizabeth Pennisi**

The metaphors for DNA keep multiplying. It is a string of code, a spiral staircase, and, now, something very like origami. Just as folding a flat sheet of paper can transform it into a crane or lotus flower, researchers have come to realize that a complex pattern of loops and folds helps transform our genome into something meaningful. The twists and turns bring particular genes into close contact with distant stretches of DNA that regulate those genes' activity, spurring the gene expression that makes a bone, muscle, or brain cell—or fuels a cancer.

The elegance and potential of this idea has fascinated many biologists, but they've struggled to get good enough data to reliably understand these sinuous patterns. An 11 December report online in *Cell* revealed the most elaborate maps yet of how the 2 meters of DNA crammed into the cell nucleus fold up—the so-called nucleome. "[L]andmark work," said Francis Collins, the director of the National Institutes of Health (NIH), in a blog post.

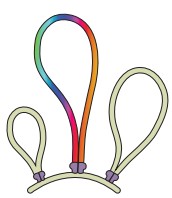
Four days later, however, a paper in *Genes & Development* reported that various DNA mapping techniques sometimes yield patterns that are radically different, raising questions about just what these maps really show. Now, through a newly funded NIH effort called the 4D Nucleome, researchers will set out to develop more reliable, accurate, and affordable ways to map and interpret the genome's elaborate folds. The program "is inviting people to come up with different and better ideas," says Job Dekker, a biologist at the University of Massachusetts Medical School in Worcester.

Over the past 2 decades, he; Erez Lieberman Aiden at Baylor College of Medicine in Houston, Texas; and others have used techniques such as chromosome conformation capture to look at genomes in 3D. The researchers chemically "freeze" any spots along the DNA chain where disparate stretches are in contact. Then they cut up all the DNA and glue the contacting DNA together. Based on the number of times two regions are in contact, the researchers estimate how far apart the regions are in 3D.

At first, researchers could look at one gene and a partner at a time, but ever more comprehensive methods soon appeared. A 2009 version, Hi-C, reveals how every piece of a genome's DNA interacts with every other piece. By figuring out how to process the DNA while it's still in a nucleus and then revamping the analysis techniques, Baylor's Suhas Rao and Harvard University graduate student Miriam Huntley, both in Aiden's lab, and colleagues got the resolution down from a million bases—way bigger than a gene—to 1000 bases—smaller than a gene.

The $3 million experiments reported in *Cell* generated more than 5 trillion sequenced bases and analyzed millions of human cells from eight cell lines, including cancer, and from one mouse cell line. The data revealed 10,000 loops, thousands to millions of bases long, and six compartments, regions where DNA with similar chemical modifications and levels of gene activity come together. Some of the structures are common to all the tested cell types, whereas others are unique to each kind of cell. "It opens up a new way of looking at biology," says Vishy Iyer, a molecular biologist at the University of Texas, Austin.

There's one wrinkle, as the *Genes & Development* paper showed: A different nucleome mapping technique based on direct observation of the DNA rather than on computational models can produce conflicting results. Iain Williamson and Wendy Bickmore of the University of Edinburgh in the United Kingdom and colleagues applied fluorescent labels to multiple pieces of DNA, using a different fluorescent probe for each one so that they could easily identify pieces that were close to each other. The researchers looked at a 1-million-base-long region of mouse chromosome 2, which contains a cluster of *Hox* genes that are key in development. For comparison, they analyzed the same DNA region using a computational technique similar to Hi-C. The two techniques agreed sometimes, for some parts of the *Hox* cluster, but in other cases, one indicated the DNA was stretched out whereas the other indicated it was tangled into a tight ball. "We don't know why and we don't know what method is right," says Ana Pombo, a cell biologist at the Berlin Institute for Medical Systems Biology. "We need to look carefully at what these methods are telling us."



A simulation (*top*) shows how a single DNA loop (*above*) twists and folds.

Aiden says his lab's new Hi-C results agree with the microscopy findings. Nonetheless, "we shouldn't fool ourselves into thinking that the Hi-C data is the be all and end all," he points out. "You want to have multiple lines of experimentation that can confirm one another or contradict one another."

NIH's 4D Nucleome program aims to provide that. This 5-year, $24-million-a-year effort, announced in July 2014, will improve the existing techniques and, possibly, come up with new ones. It's called 4D because the nucleome structure changes as cells age, differentiate, and divide, and researchers want to understand how and why. "The object is to make these techniques widely available," says Rafael Casellas, a molecular biologist at the National Institute of Arthritis and Musculoskeletal and Skin Diseases in Bethesda, Maryland. They are sorely needed, Dekker says. Just as origami paper comes to life only when folded, he says, "nothing in our genome makes sense except in 3D." ■

ILLUSTRATIONS: (TOP TO BOTTOM) ADRIAN SANBORN, MIRIAM HUNTLEY, AND EREZ LIEBERMAN AIDEN; M. ATAROD

SCIENTIFIC PUBLISHING

Geoscientists aim to magnify specialized Web searching

National Science Foundation project, GeoLink, could become search hub for the geoscience community

By **Jia You**

When oceanographer Peter Wiebe sat down recently to write a paper on his January Red Sea expedition, he wanted to examine all data sets on plankton in the region. Other researchers have been sampling the organisms for years, but he didn't know where to find those data sets. "These data centers are kind of black holes," says Wiebe, who works at the Woods Hole Oceanographic Institution in Massachusetts. "It's very hard to figure out what's in there and to get it out."

That could soon change. Wiebe is working with computer scientists to lay the groundwork for a smarter academic search engine that would help geoscientists find the exact data sets and publications they want in the blink of an eye, instead of spending hours scrolling through pages of irrelevant results. The group kicked off their project, called GeoLink, last month at the American Geophysical Union (AGU) meeting in San Francisco, California. The research effort is part of EarthCube, an initiative funded by the National Science Foundation (NSF) to upgrade cyberinfrastructure for the geosciences.

Over the next 2 years, Wiebe and colleagues will build computer programs that can extract information from AGU conference abstracts, NSF awards, and geoscience data repositories and then digitally connect these resources in ways that make them more accessible to scientists. A pilot project that concluded this year has already developed some of the underlying design. If GeoLink garners sufficient interest, the researchers could turn it into a one-stop search hub for the geoscience community, says computer scientist Tom Narock of Marymount University in Arlington, Virginia, another principal investigator on the project.

Projects like GeoLink are part of a growing effort to make literature reviews more efficient by leveraging the increasing ability of computers to process texts—a much needed service as millions of new papers come out every year. A similar initiative from the Allen Institute for Artificial Intelligence (AI2) in Seattle, Washington, is developing an intelligent academic search engine for computer science. Called Semantic Scholar, it is expected to be fully released by the end of 2015. Eventually, Semantic Scholar's coverage will include other subjects, says AI2 Chief Executive Officer Oren Etzioni.

> ***"We are imagining techniques that go way beyond just paper recommendation."***
>
> **Oren Etzioni**, Allen Institute for Artificial Intelligence

Existing academic search engines boast extensive coverage of scientific literature. (Google Scholar alone indexes about 160 million documents by some calculations.) Their reliance on keyword searches, however, often means users get more junk than treasure. Search engines also don't typically return raw data sets.

In contrast, GeoLink and Semantic Scholar attempt to build fine-grained, niche search engines catered to specific subject areas, by tapping into deeper semantic processing that helps computers establish scientifically meaningful connections between publications. If Wiebe types in "plankton in the Red Sea," for example, the search engine would understand it as a string of characters that show up on papers, but would also know the researchers who investigated the topic, the cruises they took, the instruments they used, and their data sets and papers. Google has applied similar techniques to improve its main search engine, but projects like GeoLink benefit from input from scientists with extensive knowledge in the subject area, who identify meaningful links that computer scientists then translate into code. Biomedical researchers have been forerunners in building fine-grained, subject-specific ontologies. PubMed, for example, uses the Medical Subject Headings ontology to streamline its query system.

The potential of these projects goes beyond helping scientists find the right papers quickly, says computer scientist C. Lee Giles of Pennsylvania State University, University Park. By extracting information on methods and results from papers and pooling such data together, search engines like Semantic Scholar could automate the process of literature review and comparison.

For example, Etzioni says, it would take a talented computer science graduate student weeks of extensive reading on dependent parsing (a task in natural language processing) to gain an overview of techniques used for it in the last 5 years, the data sets produced, and the accuracy rates. In contrast, Semantic Scholar could potentially compile the techniques and results within seconds. "We are imagining techniques that go way beyond just paper recommendation, to the point where we are really generating novel insights," Etzioni says.

Such instant overviews would especially benefit junior scientists and interdisciplinary scientists who enter a new field of study, says computer scientist Christina Lioma of the University of Copenhagen. It would also enable scientists to identify emerging trends in a field and adjust their directions accordingly, Giles says.

Realizing the technology's potential, however, partially depends on having accessible, text-minable literature for computers to read, an issue that has pitted scholars against some publishers. For now, the GeoLink project will mine only publicly available abstracts of studies. (Semantic Scholar will cover 4 million open-access computer science papers.) Computer scientists still have a lot of work to do to improve the accuracy of text processing, Giles says. Nonetheless, he believes that the semantic search approach "is the Web of the future." ■

A study of Red Sea plankton helped motivate smarter search engines.

PHOTO: CARLETON RAY/SCIENCE SOURCE

BIOMEDICINE

The bad luck of cancer

Analysis suggests most cases can't be prevented

By **Jennifer Couzin-Frankel**

Random mutations in healthy cells may explain two-thirds of cancers, like this one in the colon.

Why? That's the first word on many lips after a cancer diagnosis. "It's a perfectly reasonable question," says Bert Vogelstein, a cancer geneticist at Johns Hopkins University in Baltimore, Maryland, who has spent a lifetime trying to answer it. Thanks to his friendship with a recently minted Ph.D. in applied mathematics, the two now propose a framework arguing that most cancer cases are the result of biological bad luck.

In a paper published on page 78 this week in *Science*, Vogelstein and Cristian Tomasetti, who joined the biostatistics department at Hopkins in 2013, put forth a mathematical formula to explain the genesis of cancer. Here's how it works: Take the number of cells in an organ, identify what percentage of them are long-lived stem cells, and determine how many times the stem cells divide. With every division, there's a risk of a cancer-causing mutation in a daughter cell. Thus, Tomasetti and Vogelstein reasoned, the tissues that host the greatest number of stem cell divisions are those most vulnerable to cancer. When Tomasetti crunched the numbers and compared them with actual cancer statistics, he concluded that this theory explained two-thirds of all cancers.

"Using the mathematics of evolution, you can really develop an engineerlike understanding of the disease," says Martin Nowak, who studies mathematics and biology at Harvard University and has worked with Tomasetti and Vogelstein. "It's a baseline risk of being an animal that has cells that need to divide."

The idea emerged during one of the pair's weekly brainstorming sessions in Vogelstein's office. They returned to an age-old question: How much of cancer is driven by environmental factors, and how much by genetics? To solve that, Tomasetti reasoned, "I first need to understand how much is by chance and take that out of the picture."

By "chance" Tomasetti meant the roll of the dice that each cell division represents, leaving aside the influence of deleterious genes or environmental factors such as smoking or exposure to radiation. He was most interested in stem cells because they endure—meaning that a mutation in a stem cell is more likely to cause problems than a mutation in a cell that dies more quickly.

Tomasetti searched the literature to find the numbers he needed, such as the size of the stem cell "compartment" in each tissue. Plotting the total number of stem cell divisions over a lifetime against the lifetime risk of cancer in 31 different organs revealed a correlation. As the number of divisions rose, so did risk.

Colon cancer, for example, is far more common than cancer of the duodenum, the first stretch of the small intestine. This is true even in those who carry a mutated gene that puts their entire intestine at risk. Tomasetti found that there are about 10^{12} stem cell divisions in the colon over a lifetime, compared with 10^{10} in the duodenum. Mice, by contrast, have more stem cell divisions in their small intestine—and more cancers—than in their colon.

The line between mutations and cancer isn't necessarily direct. "It may not just be whether a mutation occurs," says Bruce Ponder, a longtime cancer researcher at the University of Cambridge in the United Kingdom. "There may be other factors in the tissue that determine whether the mutation is retained" and whether it triggers a malignancy.

That said, the theory remains "an extremely attractive idea," says Hans Clevers, a stem cell and cancer biologist at the Hubrecht Institute in Utrecht, the Netherlands. Still, he points out, the result "hinges entirely on how good the input data are."

Tomasetti was aware that some of the published data may not be correct. In 10,000 runs of his model, he skewed where various points on the graph were plotted. Always, "the result was still significant," he says, suggesting the big picture holds even if some of the data points do not. In mathematical jargon, the graph showed a correlation of 0.81. (A correlation of 1 means that by knowing the variable on the x-axis—in this case, the lifetime number of stem cell divisions—one can predict the y-axis value 100% of the time.) Squaring that 0.81 gives 0.65—an indicator of how much of the variation in cancer risk in a tissue is explained by variation in stem cell divisions (see graph).

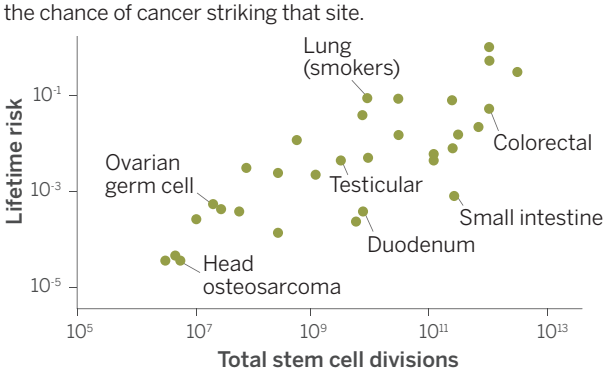
For Vogelstein, one major message is that cancer often cannot be prevented, and more resources should be funneled into catching it in its infancy. "These cancers are going to keep on coming," he says.

Douglas Lowy, a deputy director of the National Cancer Institute in Bethesda, Maryland, agrees, but also stresses that a great deal of "cancer is preventable" and efforts to avert the disease must continue.

Although the randomness of cancer might be frightening, those in the field see a positive side, too. The new framework stresses that "the average cancer patient … is just unlucky," Clevers says. "It helps cancer patients to know" that the disease is not their fault. ■

Charting cancer risk

As the number of stem cell divisions in a tissue rises, so does the chance of cancer striking that site.



CREDITS: (IMAGE) © SCIENCE SOURCE; (DATA) TOMASETTI *ET AL.*/*SCIENCE*

Targets of misconduct probe launch a legal counterattack

Heart researchers claim inquiry damaged their careers and derailed the sale of stem cell company

By **Kelly Servick**

Ongoing scientific misconduct investigations usually take place out of the public's view. An unusual lawsuit filed last month, however, sheds some light on a long-rumored inquiry. The complaint, filed in a federal district court by two prominent heart researchers, offers the first indication of just what is amiss in two papers they co-authored, which describe the heart's natural regenerative ability, and an effort to heal damaged hearts with stem cells.

The plaintiffs, Piero Anversa of Brigham and Women's Hospital (BWH) in Boston, an affiliate of Harvard Medical School, and Annarosa Leri, a Harvard associate professor in his lab, acknowledge that there are fictitious data points in a now-retracted 2012 paper that appeared in the journal *Circulation*. They also acknowledge that a much publicized 2011 paper in *The Lancet*, to which the journal had already attached an "expression of concern," contains altered figures. But they blame those problems on a third researcher. Besides raising questions about who bears final responsibility for possible misconduct, the lawsuit delves into another thorny issue: the obligation of research institutions to preserve the reputations of scientists implicated in an investigation.

The pair is suing Harvard and BWH over what they claim is a "procedurally and legally flawed" misconduct probe. The institutions, they argue, have wrongfully damaged their careers and cost them millions by derailing a deal to sell their stem cell company and by taking them out of the running for lucrative faculty positions.

> ***"I think it would be a very dangerous precedent to hold institutions culpable for doing due diligence in investigating allegations of problematic data."***
>
> **Ferric Fang**, University of Washington

Anversa and Leri claim they were unaware of any misconduct in their labs, which they blame on Jan Kajstura, the first author on the retracted paper and a former member of Anversa's lab. In the case of the *Circulation* paper, which reported a surprisingly high turnover rate for muscle cells in the adult heart, Leri and Anversa's lawsuit alleges that Kajstura apparently altered data from mass spectrometry experiments performed at Lawrence Livermore National Laboratory (LLNL) in California. (A researcher at LLNL contacted the lab after noticing that the *Circulation* paper contained 20 more data points than he had sent in a spreadsheet, the complaint explains.)

They further argue that Kajstura and an unnamed scientist under his supervision were responsible for the now-questioned figures in the *Lancet* paper, which reported results of a phase I clinical trial involving a stem cell treatment for heart failure. Anversa and Leri's complaint says they are willing to correct that paper, but that they must wait for approval from an institutional review board that oversaw the study. (The complaint also notes that 15 other papers were later added to the investigation.)

Kajstura could not be reached for comment. And it is not known whether he is actually implicated in the investigations under way at Harvard and BWH, because the institutions have not released any findings so far and have declined to comment on the inquiries. How much responsibility Anversa and Leri bear for any misconduct under their supervision is also an open question. "In the abstract, I think everyone agrees that a principal investigator has to take responsibility for whatever goes on in his or her lab," says Ferric Fang, a microbiologist at the University of Washington, Seattle, who has published several analyses of retractions, misconduct, and the scientific enterprise. But the community is often forgiving when misconduct slips past a principal investigator, he adds, as long as they are honest and forthcoming about the problems.

Anversa and Leri are also suing Gretchen Brodnicki, Harvard's dean for faculty and research integrity, who launched the initial inquiry into the lab, and Elizabeth Nabel, BWH's president. Brodnicki, they claim, exposed them to damaging criticism and speculation by calling for the two papers' retraction before the investigation concluded, without indicating that Kajstura was specifically responsible. They also argue that Nabel should have recused herself from the investigation because she has stock in and advises a firm that competes with therapies developed in their lab. They further allege that she inappropriately disclosed information about the investigation and personally maligned Anversa and Leri.

The claims of a confidentiality breach and conflict of interest, if they prove true, "are serious, and are not off-the-wall," says Paul Rothstein, a professor of torts, evidence, and civil litigation at Georgetown University Law Center in Washington, D.C. They could "give a court some serious issues to deal with and think about."

As for the damages over lost employment opportunities and the derailed business deal, Rothstein points out that courts require a high degree of proof that the defendants are directly responsible—and that the business deals would have definitely happened if not for the defendants' actions. That is often hard to demonstrate, Rothstein says, suggesting Anversa and Leri may face obstacles in making their case.

Fang worries that if the lawsuit is successful, it will deter misconduct probes. "Aside from the specifics of this particular case," he says, "I think it would be a very dangerous precedent to hold institutions culpable for doing due diligence in investigating allegations of problematic data." ■

PHOTO: MARIUS GEANTA/HIPOCRATE JOURNAL ROMANIA

Cardiac researcher Piero Anversa and a colleague are suing over an inquiry into their papers.

FEATURES

Exodus from the East

Thousands of scientists—along with entire universities—have fled war-torn eastern Ukraine. Others have staked their futures on the breakaway republics

By **Richard Stone**, *in Kyiv and Vinnytsya, Ukraine*

Refugees and a pro-Russian rebel at Donetsk National University in August; many professors and students later fled the campus.

Around 11 in the morning on 23 June, Volodymyr Semystyaga was at work in the State Archives of Luhansk, in conflict-riven eastern Ukraine, when his cellphone rang. The speaker claimed to have information and asked to meet him outside. There, two men approached. "One put a gun to my belly. The other put a gun to my back," says Semystyaga, a 65-year-old history professor at National University of Luhansk. "Nobody on the street recognized what was happening." They bundled him into a taxi and sped off.

His captors took Semystyaga to a building they called SMERSH, after the Soviet military counterintelligence unit created by Josef Stalin in World War II. In the basement, Semystyaga says, he was interrogated for weeks—injected with truth serum, burned with cigarettes, and beaten. "They had me put on a gas mask, then they blocked the air inlet and I felt like I was choking."

As his enemies had learned when they ransacked his home and office, Semystyaga was a key figure in the resistance to the region's Russian-backed separatists, who have carved out a breakaway republic. On his 55th day in detention, Semystyaga escaped with the help of a sympathetic guard and slipped out of separatist territory.

The conflict that nearly cost Semystyaga his life had flared in April, after the ouster 2 months earlier of Ukraine's pro-Russia President Viktor Yanukovych triggered Russia's annexation of the Crimean Peninsula. Centered in eastern Ukraine's Donets Basin, or Donbas, the turmoil has taken more than 4600 lives. It has also opened a bitter rift among academics and scientists. Some see a union with Russia as a chance to recover the prestige and resources that research and scholarship in the region enjoyed in Soviet days. Others, like Semystyaga, fear persecution and believe that westward-leaning Ukraine offers a brighter future.

A key concentration of Ukraine's scientific infrastructure—southeastern Ukraine hosts scores of universities and research facilities—is dissolving. In recent weeks, about 1500 scientists and professors and 100,000 students have fled rebel-held parts of the Luhansk and Donetsk regions in Donbas (see map, p. 16). In early autumn, the Ministry of Education and Science of Ukraine hurriedly began moving 11 universities out of rebel territory. Most relocated to other parts of Donbas that Ukraine still controls, in some cases leaving the original campuses in the hands of separatists. Donetsk National University (DonNU), one of Ukraine's leading universities, has set up shop 800 kilometers west in Vinnytsya.

Some of the displaced researchers have managed to spirit out valuable specimens and other research materials. But others have had to abandon their labs and life's work. A multitude of scholars remain in the war zone because they will not forsake elderly family members, students, or research projects—or because they are loyal to the separatist regimes, the so-called Donetsk People's Republic (DPR) and the Luhansk People's Republic.

The National Academy of Sciences of Ukraine (NASU), which manages a couple of dozen institutes, field stations, and other facilities in Donbas, has come under fire for its slow response to the crisis. "There was a period in which it would have been possible to get equipment out, and I urged them to do so," says physicist Maksym Strikha, Ukraine's deputy education and science minister. "But they were too conservative-minded and tried to avoid any decisions." When the insurrection in Donbas flared, "people assumed it was a fleeting problem," explains physicist Anton Naumovets, a NASU vice president who chairs a committee of the academy's leadership, or presidium, formed to tackle the problem.

Zoologist Igor Zagorodnyuk of the National University of Luhansk, who fled death threats from separatists in July, agrees. "Until the end of May, all of us believed the situation in Luhansk was absurd and temporary," he says. "But in reality, it was the beginning of the end."

"They had me put on a gas mask, then they blocked the air inlet and I felt like I was choking."

Volodymyr Semystyaga, National University of Luhansk

DONBAS HAS TILTED TOWARD RUSSIA for decades, ever since Soviet authorities built it up as a hub for industry and mining, especially for coal. Not surprisingly, much of the research that took root there was yoked to industry. Naumovets, whose uncle served in a tank battalion that liberated Donetsk from the Nazis during World War II, says that Donbas scientists have done pioneering work in nanotechnology to strengthen metals and create ceramic powders. But while applied science dominated the region, oases of basic research also formed, notably in mathematics and environmental sciences. By the late 1960s, Donetsk had become one of six designated scientific centers of Ukraine.

Many elderly residents there are nostalgic for Soviet times, when workers had stable salaries and a predictable rhythm of life. After Ukraine became independent, Donbas's subsidized economy lagged and its infrastructure decayed. Ukraine's yearning for closer ties with Europe also unsettled many people, scientists among them. "They were afraid that meant competing at a European level. They knew they would have to produce scientific articles and work hard. But allied to Russia, they could be equals," says a pro-Ukrainian geologist at the Ukrainian State R&D Institute of Mining Geology, Rock Mechanics and Mine Surveying in Donetsk who requested anonymity.

After Russia annexed Crimea, many in eastern and southern Ukraine hoped those regions would follow. Separatists in the Donetsk and Luhansk regions declared independence from Ukraine in April and held referendums on 11 May to lend the takeovers an aura of legitimacy. (The two rebel republics are recognized only by the shadow state of South Ossetia.) In short order, two major science institutes cast their lot with the breakaway republics. Alexander Kovalev, the director of the Institute of Applied Mathematics and Mechanics (IAMM) in Donetsk, declared his fealty to DPR. "He handed the institute to the terrorists on a silver platter," says one mathematician in Kyiv familiar with the situation. Kovalev's defection prompted 35 IAMM mathematicians to decamp to Kyiv, where they have formed a new division of the Institute of Mathematics. (Asked by e-mail whether he thinks IAMM is better off under DPR rule, Kovalev wrote back: "Come to Donetsk without weapons. See

PHOTOS: (OPPOSITE) PIERRE CROM; (RIGHT) RICHARD STONE

for yourself." He ignored other questions.) At the mining geology institute in Donetsk, Director Andrii Antsyferov also pledged his support, and the institute's array of costly equipment and instrumentation, to the breakaway republic.

Other directors and their staff fled the rebel-controlled regions, out of loyalty to Ukraine or fear of the separatists. At the Luhansk Nature Reserve, where ecologists study the steppe, forests, and wetlands of the Donets Basin, armed groups arrived in June and proceeded to camp near the Russian border. Director Vitaly Bondarev summoned the police. When they showed up in cars marked DPR, "it was clear it was all over," says Bondarev, who is a Ukrainian loyalist. The insurgents vacated the preserve, he says, but not before riddling it with land mines. He and his family left for Kyiv and he is now pursuing a doctoral degree in acarology—the study of mites and ticks—at the Institute of Zoology.

Roman Grynyuk had to evacuate an entire university. On Friday, the 20th of June, a group of armed men descended on the administration building of the DonNU campus in Donetsk and demanded to see Grynyuk, the rector. "I was afraid we would be shot," he says. The separatists evicted the administrators and commandeered the campus. "I tried to negotiate," Grynyuk says. "But there were many factions, and I wasn't sure who the leader was." He ordered most faculty and staff to take an extended vacation in July and August.

Even as fighting raged in August, Grynyuk held out hope that DonNU would be able to delay the fall term by only 1 month and resume classes on 1 October. But his hopes faded in early September, when DPR's education and science minister, Igor Kostenok, ordered Donetsk universities to cease teaching Ukrainian history, literature, language, and law. "It was apparent to me that the university wouldn't be able to continue to function," Grynyuk says.

The coup de grâce came on 17 September, when armed men escorted into Grynyuk's office a new acting rector: a former DonNU lecturer on Russian history named Sergey Baryshnikov. "Of course, my reaction was not positive," Grynyuk says. He got on the phone to the education ministry in Kyiv, and they began hashing out plans for relocating the university.

Casting around for affordable options, the cash-strapped ministry learned that

Fraying at the edges

Kyiv authorities have moved 11 universities out of rebel-held territory in Donetsk and Luhansk. Science assets in Crimea are beyond their grasp.

Ukraine mourns a lost science jewel

By **Richard Stone**

Months before Ukraine began losing facilities and scientists in breakaway regions of eastern Ukraine (see main text, p. 14), another of its key scientific centers slipped away: Crimea. The peninsula's mountains, Black Sea coast, and balmy climate had made it a magnet for science during and after the Soviet era. When Russia annexed it in March, "Ukraine lost some very precious assets," says Serhiy Komisarenko, director of the Palladin Institute of Biochemistry in Kyiv.

Ukraine had recently upgraded the 70-meter radio telescope in Yevpatoria.

Ukraine's astronomy community took the biggest losses, Komisarenko says.

- The Crimean Astrophysical Observatory in Nauchny has a 22-meter radio telescope that pools data with other radio telescopes around the world to form a virtual instrument thousands of kilometers wide. Other telescopes on the 550-meter ridge map sunspots and identify orbital debris.
- The Crimean Laser Observatory, a sister facility in Katsiveli, did pioneering work in the early 1960s to measure the distance between Earth and the moon. In recent years, it has studied geodynamics—probing irregularities in Earth's rotation, for example—by pinging satellites with laser pulses.
- The 70-meter radio telescope at the Center for Deep Space Communications in Yevpatoria has been used to communicate with Russian space missions and beam messages to other galaxies in hopes of making alien contact. Ukraine had recently spent millions of dollars upgrading the center, says Yaroslav Yatskiv, director of the Main Astronomical Observatory in Kyiv.

Ukraine's marine scientists are keenly feeling the loss of two facilities devoted to studying the unique biology and physical conditions in the Black Sea: the Marine Hydrophysical Institute (MHI) and the Institute of Biology of the Southern Seas, both in Sevastopol. Other losses include the Karadag Nature Reserve and the Crimean branch of Ukraine's Institute of Archaeology.

A handful of Ukrainian loyalists among the scientific personnel in Crimea have left the peninsula. Many other researchers there welcomed a reunion with Russia. Vladimir Fortov, the president of the Russian Academy of Sciences, said in July that the academy would form a department in Crimea and that institutes would be eligible to apply for Russian grants in 2015. Research will benefit, says Sergey Stanichny of MHI. "Corruption in Russian science is smaller than in Ukraine, and all procedures are more transparent," he claims.

In response, Ukraine's science ministry has proposed making Fortov persona non grata, says Maksym Strikha, a deputy science minister. "No official bilateral programs with Russia are possible now," he says. "I did my Ph.D. in Leningrad," Strikha adds. "Now I have no friends there." ■

IMAGES: (TOP TO BOTTOM) AVVECTOR/ISTOCKPHOTO.COM; S. KOROTKIY/WIKIMEDIA COMMONS

Crystal, a jewelry manufacturer in Vinnytsya, was struggling in the economic downturn, and DonNU could lease two derelict factory buildings for a pittance. Many DonNU faculty members had qualms about fleeing. DPR apparatchiks had warned of reprisals against anyone who tried. "They were told that there are very many cellars in Donetsk where they could end up and never be found," Grynyuk says. Nevertheless, by early October the hegira from DonNU had begun. On 3 November, Ukraine's science minister presided over the opening ceremony of DonNU-in-exile.

About 200 professors and lecturers (out of 700 before the split) are now nesting in spartan rooms almost devoid of furniture in the former jewelry factory. DonNU's budget barely covers salaries, and it now must fritter away modest student enrollment fees on temporary lodging for relocated faculty.

Baryshnikov has dismissed the exodus's impact, boasting that the rump DonNU he rules is a "powerful institution that cannot be transplanted like a ficus from one pot to another." The truth is quite different, says Liliya Hrynevych, a legislator in the Verkhovna Rada—Ukraine's parliament—and chair of its science and education committee. "It's clear the quality of education in the shadow universities in Donbas will be so dubious," she says, "that no one will want to employ their graduates."

Grynyuk has appealed for overseas help to outfit DonNU-in-exile, including funds to renovate a pair of ramshackle dormitories donated by a local college. Classes for the 7000-odd registered students (60% of last year's total), including about 500 residing in Vinnytsya, are online-only for the foreseeable future. "I hope in 3 or 4 years we'll be able to return to Donetsk and reclaim our campus," he says.

Other universities in occupied Donbas have shifted operations to cramped satellite campuses in regional cities under Ukrainian army control. "We thought moving them nearby would be less traumatic," says Strikha, the deputy education and science minister. For example, the National University of Luhansk is now headquartered in nearby Starobilsk, but only about 20 professors have taken up residence there, says the university's Zagorodnyuk. It, like DonNU, is now a virtual university. The education and science ministry will do its best to keep a scientific pilot light burning at the embattled universities by funding "weak" research proposals, Strikha says. "We need to give them a chance."

ABRUPTLY TRANSPLANTING a research career can be wrenching. DonNU-in-exile's chemistry dean, Alexander Shendrik, opens his laptop to show photos of the teaching and research labs he abandoned back in Donetsk, equipped with spectrometers, x-ray crystallography machines, and other pricey instruments. He slaps his hand on his desk. "The losses are catastrophic," he says. "For any chemist, his laboratory is his life."

Others have managed to smuggle out key specimens and data. In his cramped new office—just a gap between exhibits at Ukraine's National Museum of Natural History in Kyiv—Zagorodnyuk picks up a 200-gigabyte hard drive, wrapped in bubble sheet, containing all of his data and publications. "Guess how I got this out of Luhansk?" asks the zoologist. He passed through several checkpoints, where separatists rifled through his belongings, with the hard drive taped to his leg under his left sock.

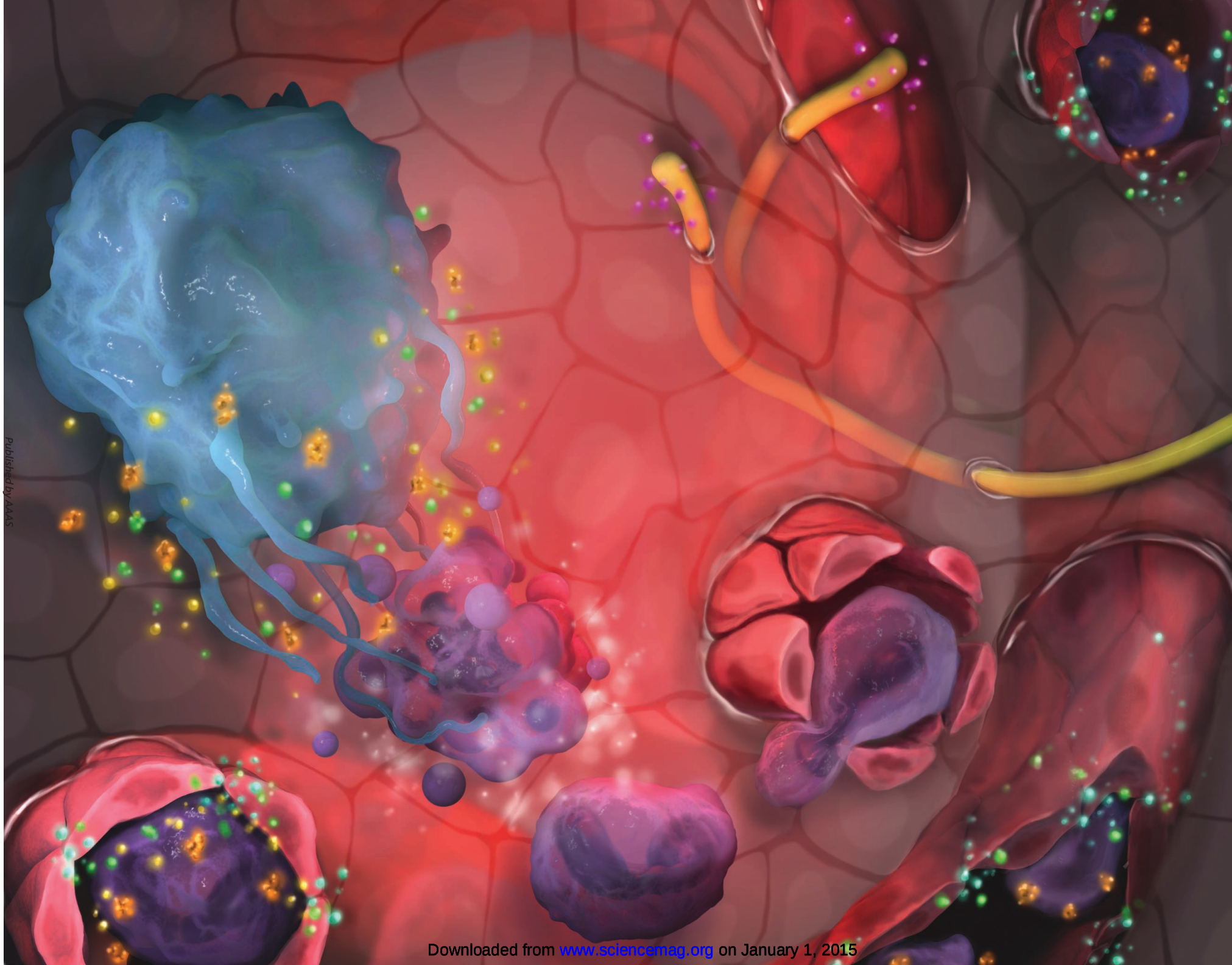
Other contraband has arrived just hours earlier. Zagorodnyuk picks up a brick of several matchboxes glued together and slides one open. Inside is the delicate chestnut-brown skull of a Strand's birch mouse. In another box nestle test tubes holding tiny skeletons. He opens one and plucks out a scrap of yellowing paper describing the specimen in ornate Cyrillic: It's a northern mole vole collected in 1912 in southern Ukraine. A few dozen more boxes of bones are stacked on a shelf. Three days earlier, one of Zagorodnyuk's graduate students who is riding things out in Luhansk had slipped into his mentor's apartment and retrieved the collections. The student paid a train conductor to deliver the specimens to Kyiv.

Maksym Netsvetov's research subjects—plants—are less portable. He's head of phytoecology at Donetsk Botanical Garden (DBG), one of Europe's largest botanical gardens. Established in 1964, DBG is known for pioneering work in "industrial botany"—ecological restoration for a region scarred by factories and mines—and for its world-class herbarium, where taxonomists have identified more than 200 species of medicinal plants. Netsvetov and others spirited out irreplaceable specimens to Kyiv. But he worries about how the remaining plants will fare in Donetsk. DBG staff members can't get near the arboretum, and winter cold—fatal to the tropical plants if the greenhouse is not heated—is closing in. "Terrorists set trip wires in the arboretum connected to bombs," Netsvetov says. "Before long, the tree collection will be lost," he says, putting his hand to his forehead and wincing. "It's hard not to get emotional."

Back in Donbas, the atmosphere is only growing more poisonous. In November, at Donetsk State Medical University, the DPR-appointed rector gathered remaining staff and challenged them to openly declare their loyalties. Of the 800-odd people in the room, 11 had the courage to pipe up that they are pro-Ukrainian. "Others were just scared," Netsvetov says. And at NASU's mining geology institute, DPR officials showed up recently with sacks of money, "as if they'd robbed a bank," says the pro-Ukrainian institute geologist. They doled out cash—but only to staff who support the separatists. ■

PHOTO: RICHARD STONE

Among the prize specimens spirited out of Luhansk is this northern mole vole collected in 1912 in southern Ukraine.



Inflammation's STOP SIGNALS

Inflammation doesn't just peter out. The body actively shuts it down, using signals that researchers hope to transform into therapies

By **Mitch Leslie**

Doctors never figured out what jabbed a hole in Charles Serhan's intestine during a trip to Asia in 1990. It might have been a kernel of undercooked rice or a splinter of shell in a seafood dish. Whatever the cause, by the time he'd flown back to Boston, where he's a biochemist and experimental pathologist at Harvard Medical School, he was suffering. "I could hardly stand up," Serhan recalls. He ascribed his misery to dehydration and bad jet lag because "I didn't have a fever."

Once Serhan's wife convinced him to go to the hospital, doctors discovered an abdominal abscess the size of a grapefruit. Intestinal bacteria had apparently spilled through the puncture into his abdominal cavity. And rather than helping, the body's natural frontline defensive reaction—the cascade of immune cells and molecules commonly called inflammation—had only made things worse. The glob of pus, white blood cells, and bacteria that formed at the puncture site had burst, and the inflammation had spread, causing a life-threatening case of peritonitis. Serhan was rushed straight from the CT scanner to the operating room.

Serhan had already been studying molecules involved in inflammation. But he says that thanks to this traumatic incident—he eventually needed three operations—"I learned firsthand how controlling inflammation is so important." It pushed him to delve deeper into how the body manages this powerful process.

ONLINE

A combined graphic is available at http://scim.ag/stopsignals.

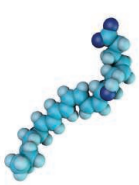
He and other scientists have found that, contrary to what they once thought, our bodies deploy an elaborate mechanism to shut down, or resolve, the inflammation that naturally arises in response to wounds and infections. Orchestrating this are proresolving molecules, a hodgepodge of more than 20 types of lipids, proteins, hormones, and other compounds (see graphic, below) that counteract the cells and molecules that stoke inflammation. "It makes sense that there would be an active program to resolve inflammation," says pulmonologist Bruce Levy of Harvard Medical School. "Otherwise, it would be hard to fine-tune the system," and inflammation would frequently run rampant.

The effects of proresolving molecules may extend beyond curtailing inflammation. In animal studies, they prompt repair of damaged tissues, and some evidence suggests they are essential for recovering from infectious diseases such as the flu. Researchers have found hints that these molecular stop signals falter in at least some of the common, chronic diseases that involve

Players in the endgame

An assortment of molecules shut down inflammation and promote tissue healing by targeting different cells.

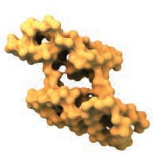
Lipoxins
Lipids whose jobs include stimulating macrophages and preventing neutrophils from slipping between endothelial cells to enter damaged tissue.



Resolvins
Family of lipids that block neutrophils' exit from the bloodstream and prod macrophages to eat cellular debris.

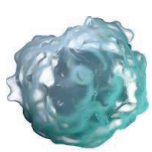
Maresins
Made by macrophages, lipids that spur tissue repair and act on nerves to ease pain.

Protectins
Lipids that curtail release of inflammation-promoting molecules and are protective in the nervous system.



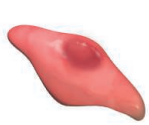
Annexin A1
A protein released by dying neutrophils, its functions include preventing other neutrophils from entering the injured site.

Hydrogen sulfide
Message-carrying gas that reduces pain and stimulates neutrophils to commit suicide.

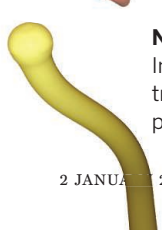


Macrophages
After clearing an infection, these immune cells consume proinflammatory cellular remains.

Neutrophils
First responders to wounds and infections, they release inflammatory cytokines.



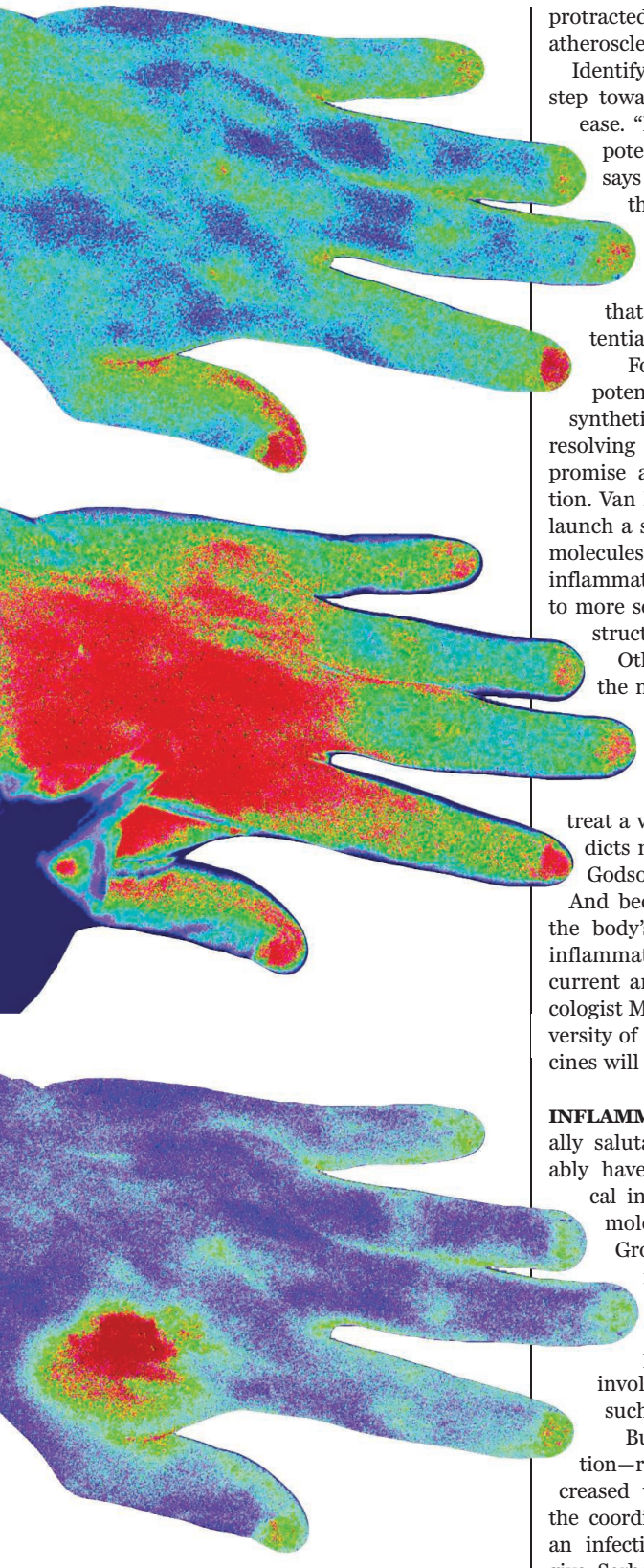
Endothelial cells
These cells form the walls of blood vessels and make H_2S.



Nerves
Inflammatory molecules trigger nerve cells, creating pain and itchiness.

ILLUSTRATIONS: V. ALTOUNIAN/*SCIENCE*

Published by AAAS



A thermograph shows the heat signature of inflammation as it rises (*middle*) and ebbs (*bottom*) after a hand is pricked with a needle.

protracted inflammation, including asthma, atherosclerosis, and Alzheimer's disease.

Identifying these stop signals is a first step toward exploiting them to treat disease. "People are starting to realize the potential of this class of compounds," says Thomas Van Dyke, a dentist at the Forsyth Institute in Cambridge, Massachusetts, who studies periodontitis, a condition in which inflammation destroys the bone that anchors the teeth. "I'd say the potential was enormous."

For a few of the molecules, that potential could soon become reality. A synthetic version of one inflammation-resolving compound has already shown promise against a type of eye inflammation. Van Dyke and colleagues are about to launch a safety study of other proresolving molecules in people who have gingivitis, an inflammation of the gums that can progress to more serious problems such as bone destruction and tooth loss.

Other researchers anticipate that the molecules could eventually be deployed against everything from cancer to diabetes. "We will be using drugs that promote the resolution of inflammation" to treat a variety of chronic conditions, predicts molecular cell biologist Catherine Godson of University College Dublin. And because such drugs would activate the body's natural pathways for curbing inflammation, they should be safer than current anti-inflammatories, says pharmacologist Mauro Perretti of Queen Mary University of London. "Resolution-based medicines will be better medicines."

INFLAMMATION IS INEVITABLE and usually salutary. Right now, all of us "probably have hundreds of sites of subclinical inflammation" in our bodies, says molecular pharmacologist Karsten Gronert of the University of California (UC), Berkeley. Every shaving cut, skinned knee, embedded sliver, mosquito bite, and pimple provokes an explosive reaction involving immune cells and molecules such as cytokines.

But the familiar signs of inflammation—redness, swelling, pain, and increased temperature—usually abate after the coordinated response has helped clear an infection or heal a wound. Scientists give Serhan credit for helping focus their attention on why. "We just didn't put much thought into how inflammation goes away," Godson says. The general view was that once the trigger—bacteria or another insult—was eliminated, inflammation petered out as proinflammatory molecules dissipated and immune cells died or dispersed. Most treatments for inflammation, such as nonsteroidal anti-inflammatory drugs (NSAIDs) like ibuprofen and naproxen, reflect this preconception. They dial down the severity of the reaction by preventing synthesis of inflammation-promoting molecules, but they don't directly stop the process.

The research that led Serhan to reconsider inflammation began during his postdoc with biochemist Bengt Samuelsson at the Karolinska Institute in Stockholm. Samuelsson had shared the 1982 Nobel Prize in physiology or medicine for working out how cells manufacture a suite of molecules that includes prostaglandins, lipid messengers that can trigger inflammation and other physiological responses. Looking for additional lipids that serve as signals, Serhan and colleagues in 1984 isolated from white blood cells a previously unknown molecule that they named lipoxin.

"We had no idea what it did," Serhan says. Other, more familiar lipid signals incited inflammation in lab studies, but experiments over the next several years, which Serhan continued after moving to Harvard in 1987, revealed that lipoxin instead stanches the reaction. In another surprise, he and his colleagues found in 1992 that aspirin triggers cells to produce lipoxin, suggesting that the molecule accounts for some of the drug's long-known anti-inflammatory effects. "I couldn't have predicted that if I'd tried," Serhan says.

As more data on what lipoxin does and how it interacts with inflammation-stimulating molecules rolled in, Serhan and his colleagues realized that it was part of a larger mechanism to shut off inflammation. Then they began searching for other chemicals with the same power. "You could say that we are molecule hunters," Serhan says.

In the late 1990s, his team began using a different snare for catching lipid signals. The researchers inject bubbles of air into the skin on the backs of mice and then inject inflammation-stimulating bacteria or molecules. The tiny pearls of pus that form inside the air pockets teem with lipids, which the researchers sort and identify based on their mass and charge using techniques such as liquid chromatography tandem mass spectrometry. Within a decade, Serhan's lab bagged three other groups of proresolving lipids: resolvins, protectins, and maresins. Other teams have expanded the list of inflammation-ending molecules to include proteins such as annexin A1, hormones such as cortisol, cytokines such as interleukin-10, and possibly gases such as hydrogen sulfide.

IMAGES: JESMOND DALLI AND C. N. SERHAN

Studies by Serhan, Perretti, and other researchers have revealed some of the mechanisms by which these molecules work. If a person cuts a finger, for instance, white blood cells called neutrophils slither through the walls of blood vessels and enter the site of the injury, where they spill inflammation-promoting compounds. Proresolving molecules prompt these neutrophils to commit suicide and prevent others from exiting the blood vessels. Some of the molecules direct the cleanup of the wreckage left behind when inflammation ends. They spur immune cells called macrophages to clear away debris such as the remnants of suicidal neutrophils, which can trigger dangerous chronic inflammation if they aren't removed.

They also help with another key task—repairing the tissue damage that inflammation typically produces. Van Dyke, Serhan, and colleagues provided a striking example in periodontitis. Almost half of Americans will eventually develop the condition, and many of them will require dentures or tooth implants if the inflammation erodes enough of the bone-teeth connection. "It is very difficult to regrow bone that is lost to disease," Van Dyke says. But the researchers revealed online in a November 2014 paper in the *Journal of Dental Research* that they could do just that. They operated on the gums of pigs to induce periodontitis. Then they squirted the pigs' inflamed gums with nanoparticles carrying a synthetic version of lipoxin and saw that some of the lost bone regenerated.

A brush with death motivated Charles Serhan to further explore the molecular signals the body uses to naturally shut down inflammation.

PHOTO: JEREMY WINKLER

These successes in animal studies raise a question: If proresolving molecules are so powerful, why do so many people suffer from diseases driven by chronic inflammation, such as atherosclerosis, diabetes, asthma, and arthritis? Some early research suggests that the levels of proresolving molecules are abnormally low in these individuals. In one study, Harvard's Levy and colleagues captured and analyzed the labored breaths of people who had asthma. As they detailed in 2013 in *The Journal of Allergy and Clinical Immunology*, the concentrations of lipoxin, relative to the levels of inflammation-stimulating lipids, were lowest in exhaled air from the patients who had the most severe breathing difficulties.

Yet some researchers are not convinced that these inflammatory stop signals are as potent as Serhan and others believe. The compounds occur at extremely low concentrations in the body, too low to convince lipid biologist Garret FitzGerald of the University of Pennsylvania that they have much of an impact. "I'm not saying they are not important in resolution [of inflammation], I'm saying that the evidence is incomplete," he says. That doesn't mean the compounds won't lead to treatments, he acknowledges. At higher concentrations, the molecules—or drugs derived from them—might trigger a beneficial effect.

RESEARCHERS ARE NOW TESTING that promise in the clinic. The limitations of current anti-inflammatories provide a powerful motivation. Some of them, such as corticosteroids and antibodies for treating rheumatoid arthritis, can suppress the overall immune system if used over the long term, leaving patients vulnerable to infections and even cancer. Proresolving molecules, in contrast, don't reduce the immune system's ability to fight infections and cancer.

That makes atherosclerosis a good target for the compounds, says cell biologist Ira Tabas of Columbia University. Patients with rheumatoid arthritis are willing to risk the side effects of powerful anti-inflammatories because the alternative is crippling joint inflammation, he notes. Patients with atherosclerosis, however, are often reluctant to take the drugs because the condition typically doesn't cause symptoms. In addition, current methods for opening up plugged arteries, such as angioplasty and implanted stents, don't cure the underlying inflammation, says Michael Conte, a vascular surgeon and vascular biologist at UC San Francisco. They often irritate the vessel lining further and spur the artery to clog again, a condition called restenosis. Two years ago, Conte and his colleagues revealed in *The FASEB Journal* that injecting one type of resolvin into the arteries of rabbits who had undergone angioplasties reduced the degree of narrowing by 29%.

Conte, Serhan, and colleagues are now working to create stents or other devices that could deliver proresolving molecules to the inflamed arterial lining. "We want to turn angioplasty more into a mosquito bite," Conte says. The team is about 2 years from testing one of these devices in people, he predicts.

Other potential uses of the molecules are further along. In 2013, for example, researchers in China reported that a cream containing lipoxin soothed eczema, an inflammatory skin disease, in babies. One synthetic form of resolvin has already demonstrated some effectiveness in a trial for dry eye, a type of irritation that's common in the elderly. Auven Therapeutics of the U.S. Virgin Islands, which bought the development rights to this form of resolvin from a firm Serhan had helped found, has started a second round of trials, which will evaluate whether the compound can reduce eye inflammation and pain after cataract surgery.

Van Dyke and colleagues plan to start their own clinical trial early in 2015, testing lipoxin and a lipoxin derivative that will be given as a mouthwash to middle-aged volunteers with gingivitis. This form of gum inflammation doesn't directly cause tooth loss but can lead to periodontitis. The molecules caused no side effects in animal studies, and Van Dyke expects them to be safe in humans. They stimulate receptors that cells produce only during inflammation, he says, so "if you have no inflammation, they have no effect."

Proresolving molecules are difficult to synthesize and to produce in sufficient quantities even for a clinical trial. But industry will quickly surmount those hurdles if the initial trials show promise, predicts UC Berkeley's Gronert. "It just takes one success of moving from bench to bedside, and then the field could explode." ■

LETTERS

Edited by **Jennifer Sills**

NextGen's course catalog

What was missing from your science education? In October, we asked young scientists to name and describe a course that would have better prepared them for their science careers. Below, we've created a course catalog featuring a sample of their responses, which range from serious to silly. To allow for as many voices as possible, in some cases we have printed excerpts of longer submissions (indicated by ellipses) and lightly copyedited original text for clarity. To read the complete versions, as well as many more, go to http://scim.ag/NG13_Results. Follow *Science*'s NextGen VOICES survey on Twitter with the hashtag #NextGenSci.

NEXTGEN VOICES

PSYCHOLOGY

PSY302: PREPARATION FOR SCIENCE

This course provides an introduction to the survival skills for a successful science career. The course covers the following topics: Never give up no matter how many times your manuscript has been rejected, stay optimistic when others are doubting your research, believe that science will love you back, prepare dinner and analyze data at the same time.... ***Li Dai***

School of Geographic and Environmental Sciences, Guizhou Normal University, Guiyang, Guizhou, 550001, China. E-mail: daili_helen@163.com

PSY501: FAKE IT TO MAKE IT

In this essential 1-credit seminar, you will learn to cultivate a hearty confidence while presenting ideas you barely understand, a stern resilience under repeated failure, and a keen cynicism about everything.

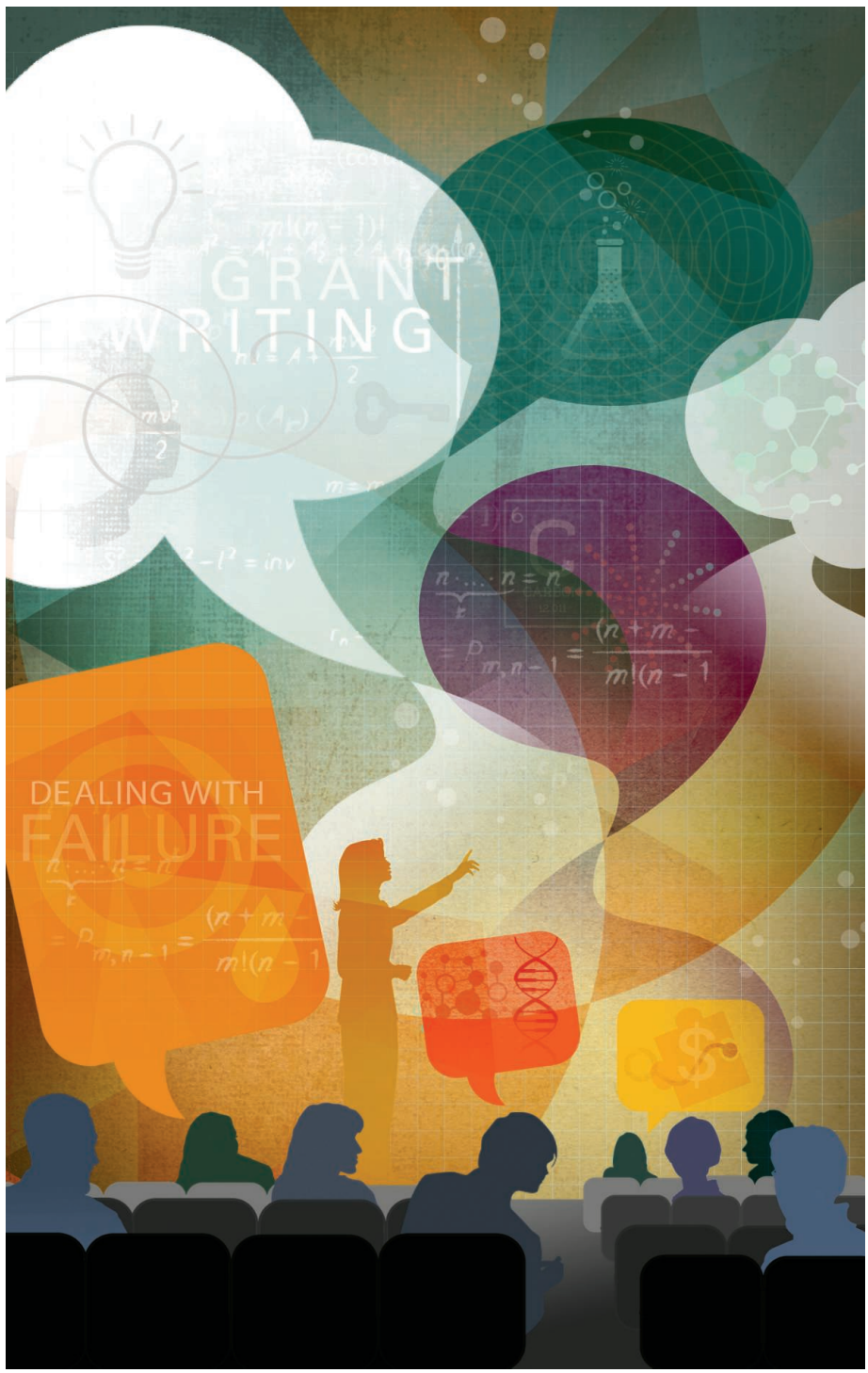


ILLUSTRATION: STEPHANIE DALTON COWAN; PHOTOS COURTESY OF THE AUTHORS

Published by AAAS

Minor skills introduced include the efficient brewing of strong coffee, task layering on Google calendar, and organizing Internet time for optimal productivity. Additionally, you will have access to Mindr to match you with a compatible graduate advisor. All students will be equipped with a BS radar upon completion of the course. Required reading: Ph.D. comics, #WhatShouldWeCallGradSchool.

Anna Beiler

Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85281, USA.
E-mail: anna.beiler@asu.edu

PSY516: DEALING WITH FAILURE

Think you have dealt with real failure as an undergrad? Think again. This advanced course analyzes the frustration, depression, and questioning-of-life-choices that accompany the months or years of repeated experimental failures you can expect to encounter in your scientific career. Through this course, you will develop helpful strategies for dealing with inevitable setbacks. Topics include setting realistic goals, building a supportive network of colleagues and mentors, learning from your mistakes, keeping your end goal in mind, and having fun along the way. The summative project is a personalized action plan for reevaluating demoralizing failures as opportunities that enable future growth. ***Yonatan Lipsitz***

Institute for Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON M5S 3E1, Canada.
E-mail: yonatan.lipsitz@mail.utoronto.ca

PSY549: YOU PROBABLY WON'T SAVE THE WORLD

This class is designed to set realistic expectations for you in your career, allowing you to skip the disenchantment phase and focus instead on your talents. Through presentations and peer-to-peer learning, you will learn that the world's problems are complex and require many, many dedicated scientists to tackle them, of which you will be only one. Learn how you can use your research or your teaching to improve, and even transform, the lives of people around you, and begin to build the skills to be satisfied with that. Prerequisites include Your Spouse Probably Won't Be a Supermodel. ***Toni Lyn Morelli***

Department of Environmental Conservation, University of Massachusetts, Amherst, MA 01035, USA. E-mail: morelli@umass.edu

COMMUNICATIONS

COM101: DON'T TALK NERDY TO ME: COMMUNICATING WITH THE PUBLIC

...In this class, public relations managers from the industry will share their insights into conveying the right message at the right time through the right channel, and we will coach communication skills by role-playing. Through effective communication, you can build the bridge connecting the ivory tower with society.

Kun-Hsing Yu

Biomedical Informatics Training Program, Stanford University, Stanford, CA 94305, USA.
E-mail: khyu@stanford.edu

COM201: HOW TO DESIGN CATCHY POSTERS AND WRITE ENGAGING ARTICLES

One of the fundamental aspects of a scientist's work is to present research results. This causes a bit of a problem when it comes to poster design (have you seen those walls of text?), PowerPoint presentations (ever fall asleep while reading your own slides?) or crazy-difficult articles that give you a headache. Let's stop that! In this course, we will drill writing, designing, and presenting so that no one ever falls asleep when confronted with your data.

Ilona Kotlewska

Laboratory of Neurophysiology, Nencki Institute for Experimental Biology, Polish Academy of Science, Warsaw, 02-093, Poland.
E-mail: i.kotlewska@nencki.gov.pl

STATISTICS AND LOGIC

STL13: BREAKING BIOLOGISTS' ARITHMOPHOBIA

13.1. NUMBERS: Learn that numbers are not adversaries to biologists and why numbers are an absolute necessity when studying biological systems....
13.2. ARITHMOPHOBIA: Discuss the fear of numbers, with particular emphasis on how biologists and clinicians are terrorized by anything remotely numerical.
13.3. STATISTICS: Learn that statistics is a tool for solving problems, not a problem itself, that $P < 0.05$ isn't a catchphrase or a style statement, that standard error is not an error in the data, that ANOVA and t tests are not interchangeable, and that the term significance has deeper meaning.
13.4. SOFTWARE: Explore the idea that using statistical software packages without knowing statistics is not an alternative to knowing statistics.

Gunjan Guha

Department of Biotechnology, School of Chemical and Bio Technology, SASTRA University, Thanjavur, Tamil, Nadu, 613401, India.
E-mail: gunjan.doc@gmail.com

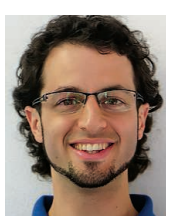
STL210: SCIENCE IN THE REAL WORLD: LEARNING TO LOVE IT MESSY

Scientific researchers often strive to find "clean" questions: Does treatment X cause effect Y? To do so, we need to carefully control our experiment and isolate our test subjects from all confounding factors. Unfortunately, in the real world, everything is a confounding factor and isolation is functionally impossible. This class will teach you how to cope when normal experimental practices cannot yield significant, interpretable answers. By using Big Data, qualitative research, natural experiments, and inductive logic, you will be prepared to address even the most confounding real-world problem. You will also learn how to communicate your results, without excessive over-qualification and undue focus on potential error. Despite the messiness of the world, science outside the laboratory can still lead to useful conclusions. Science in the Real World 210 will show you the way.

Colin W. Murphy

NextGen Climate America, San Francisco, CA 94104, USA. E-mail: cmurphy@nextgenamerica.org

Editor's Note: NextGen Climate America is not affiliated with Science*'s NextGen VOICES survey.*



STL227: HOW TO QUESTION

Science is all about questions: What are the fundamentals of natural phenomena? How can we improve our quality of life? Other courses cover the knowledge we've already acquired and focus on taking the next logical step. But how can we prepare ourselves to ask the questions nobody has asked before? In this course, you will

CREDITS: (YONATAN LIPSITZ) JENNIFER MA; (KUN-HSING YU) GUANG-MEI PHOTO; (ILONA KOTLEWSKA) ANNA MIRGOS; (COLIN W. MURPHY) DORIAN TOY; OTHER PHOTOS COURTESY OF THE AUTHORS

analyze the questioning process as you study a historic panorama of questions that generated a paradigm change in science. Activities will train you to use creativity and to look at problems and solutions from different perspectives. Better questions, better science.

Rolando Manuel Caraballo

Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, 1428, Argentina. E-mail: rcaraballo@qi.fcen.uba.ar

STL666: SCIENTIFIC SKEPTICAL THINKING

Learn to think like a skeptical scientist through an exploration of literature ranging from primary sources, newspaper and magazine articles, television, and talk radio, to blogs and memes on social media. Various logical fallacies and common misconceptions will be dissected so as to be easily recognizable. You will learn to debunk false claims in a concise and meaningful way. Finally, you will take a tour of common pseudoscience manifestations existing today....

Keah Schuenemann

Meteorology Program, Earth and Atmospheric Science Department, Metropolitan State University of Denver, Denver, CO 80217-3362, USA. E-mail: kschuene@msudenver.edu

COMPUTER SCIENCE

CPS198: DIGITAL LIBRARY ORGANIZATION

In today's digital world, scientists must maintain vast and ever-growing libraries of journal article PDFs. If there is no organizational system in place, these files often retain their original unhelpful names, such as "7320.full.pdf," and get hopelessly scattered throughout a computer system. In this course, you will learn to use citation management software to automatically rename and sort journal article PDFs based on useful metadata, such as author last names and year of publication. As an added bonus, you will learn to use the same software to insert citations and automatically generate bibliographies in any citation style.

Rosa Li

Department of Psychology and Neuroscience, Center for Cognitive Neuroscience, Duke University, Durham, NC 27708, USA. E-mail: rosa.li@duke.edu

CPS201: STONE-AGE TECHNOLOGY: COMPUTERS FROM THE 1990S AND 2000S

Scientists are often forced to deal with restrictive budgets and old computer equipment. In this course, learn about the ancient technology of our parents and grandparents in order to cope with the devices you may be required to use in the pursuit of knowledge. Topics include mice with rubber balls, dot matrix printers, cathode-ray tube monitors, floppy disks, and many more.

Aric Campling

McLean, VA 22102, USA. E-mail: ascampling@gmail.com

LABORATORY SCIENCE

LBS101: THE JOY OF SCIENCE

Do you think of science as complicated jargon and formulas to be memorized? What if a science course could be about curiosity and discovery? In this new hands-on course, you will use state-of-the-art science equipment to investigate real research questions. Your findings will contribute to current research projects and to the advancement of scientific knowledge. You will learn how to formulate research questions, identify suitable methods, analyze your data, and communicate your findings. The hands-on activities will align with online study material that provides the theoretical knowledge for the course. Experience the joy of science!

Beat A. Schwendimann

Computer-Human Interaction in Learning and Instruction, École Polytechnique Fédérale de Lausanne, 1015, Lausanne, Switzerland. E-mail: beat.schwendimann@gmail.com

LBS201: MACGYVER IT!

Can't figure out how to remove that pesky stir bar that just won't come out with the retriever? Rusty clamp won't hold your flask properly in the oil bath? Rotavap won't hold its vacuum? Temperature keeps creeping up on that water bath? Still see some solid on that filter paper from the Buchner that won't come off? Want to learn how to construct a glove box from cardboard and duct tape? Have no worries! This course will show you all the lab tricks to being a super-savvy chemist that the textbooks don't mention!

Masha G. Savelieff

Life Sciences Institute, University of Michigan, Ann Arbor, MI 48109, USA. E-mail: savelief@umich.edu

LBS501: HOW TO SET UP YOUR VERY OWN 500 SQUARE FEET OF LAB SPACE

Congratulations! You just got hired as an assistant professor. Your new lab has nothing but walls and old benches. Do you know which cylinder regulator you need for the incubators? Oh, yes, you need that "thingy." The huge Sorvall centrifuge will be left on the floor when delivered. Those tips aren't compatible with those pipettes: They will fall off right into your cell suspension. Do you know how and where to order liquid nitrogen? For everything you took for granted in your graduate lab, there's a specific diameter, catalog number, and safety regulation. Required course materials: measuring tape, wrench.

Irina Tiper

Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD 21201, USA. E-mail: itiper@umaryland.edu

BUSINESS AND FINANCE

BSF101: GRANT WRITING

The objective of this course is to provide you with the necessary skills to effectively market your scientific ideas to funding agencies and grant review panels. This foundation course will therefore be essential for survival as a scientist in an era of limited research funding. Topics will include developing professional connections, understanding reviewer psychology and bias, memorizing detailed grant instructions and guidelines, and maximizing page use and space in proposals. Course will be graded as pass/fail. However, only 10% of the students will pass during each term. Course can be repeated indefinitely.

Michael G. Kemp

Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599, USA. E-mail: michael_kemp@med.unc.edu

CREDITS: (MASHA G. SAVELIEFF) AIME MAGIERA. SENSORY RESEARCH: OTHER PHOTOS COURTESY OF THE AUTHORS

BSF301: RESEARCH FUNDING

...This course is designed to prepare you...to find, apply for, and secure research funding. Differences between fellowships, grants, and awards will also be covered. Guest speakers from local and international funding agencies will participate in the lectures. You will propose a funding application project as summative assessment.

Marie-Caroline Lefort

Bio-Protection Research Centre and Library, Teaching and Learning, Lincoln University, 7647, Christchurch, New Zealand.
E-mail: Marie-Caroline.Lefort@lincolnuni.ac.nz

BSF329: BUSINESS MANAGEMENT: FISCAL REALITIES OF SCIENCE

This course lays a foundation of basic business management for running a laboratory. Specifically, this course will include creating a business plan that aligns your budget with the pursuit of your hypotheses and the management of a fiscally responsible laboratory. Whereas most of your coursework focuses on core scientific principles, methodologies, and hypotheses driving scientific inquiry, this course will help you lay the fiscal framework to ensure your research stays within your fiscal boundaries. This course will also explore developing and working with intellectual property, common avenues of commercialization, and working with commercial partners. Mastering the fundamentals of fiscal responsibility will prepare you for a subsequent course in entrepreneurship. ***E. Loren Buhle Jr.***

Global Customer Success, Model N, Princeton, NJ 08540, USA. E-mail: lorenbuhle@gmail.com

BSF346: TRANSLATING RESEARCH INTO A COMMERCIAL PRODUCT

This course introduces the steps required to translate lab (i.e., in vitro) research into a viable commercial product. You will learn about intellectual property protections, development costs, attracting venture capital, and making inventions commercially viable. You will analyze real-world examples of products that have been commercialized from basic research and discuss how to attract grant funding from the government and private sources to develop products. The opportunities and challenges of interacting with companies as a graduate student will be emphasized. ***Ajay Kashi***

Rochester, NY 14620, USA.
E-mail: ajay.kashi@gmail.com

SUBMIT NOW:
THE SCIENTIST'S TOOLBOX

Add your voice to *Science*! Our new NextGen VOICES survey is now open:

Name and describe a currently nonexistent invention that would make you a more effective scientist. Your invention can be realistic, futuristic, or comical, and it can aid you in any aspect of your scientific process or career.

To submit, go to http://scim.ag/NG_14

Deadline for submissions is 13 February. A selection of the best responses will be published in the 3 April issue of Science. Submissions should be 100 words or less. Anonymous submissions will not be considered.

CAREERS

CAR501: SO YOU THINK YOU CAN BE A PI?

Prerequisite: academic excellence at the undergraduate level and an extensive publication profile at the Ph.D. level.
MODULE 1: Coping with no longer experimenting on a daily basis
MODULE 2: One style of supervision does not suit all—embracing the broad spectrum of personalities and ability levels of your research students
MODULE 3: Is everyone a competitor—what's happening to openness in solving important and interesting problems?
MODULE 4: Workplace politics—juggling the often disparate objectives of administrators, students, colleagues, and executives
MODULE 5: Hope—rediscovering the joy of scientific research ***Anthony O'Mullane***

School of Chemistry, Physics, and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD 4001, Australia.
E-mail: anthony.omullane@qut.edu.au

CAR502: TRANSLATING COMPETENCIES

So often graduate programs guide students toward an academic career, when in fact fewer than 23% of science Ph.D.s will take that road, according to recent NSF data. Most of those students will end up finding careers in other sectors such as government, industry, or NGOs. This course will enable graduate students to translate the academic achievements you spend years toiling to obtain (e.g., published results) into competencies that fit the jargon and framework (e.g., project management) of these other sectors in order to be both appreciated and understood by human resources and hiring offices.

Sarah M. Anderson

School of Biological Sciences, Washington State University, Pullman, WA 99164, USA.
E-mail: sarah.anderson2@email.wsu.edu

CAR503: PI OR BUST: LIFE BEYOND ACADEMIA

Can I be something other than a PI when I grow up? How do I decide what I want to do (even if I can be a PI)? How do I look my Ph.D. supervisor in the eye after moving over to the dark side (i.e., industry or god forbid, patent law)? This course will give you a broader perspective on job options outside of academia, with guest lectures from talented people who chose to go elsewhere. The pros and cons of each job and its accompanying lifestyle will be discussed. ***Noa Sher***

Bioinformatics Service Unit, University of Haifa, Haifa, 3498838, Israel. E-mail: nsher@univ.haifa.ac.il

INTERDISCIPLINARY CAPSTONE

INT600: SCIENCE WITHOUT BORDERS

Every research project has multiple facets, and expertise in each one of them is nearly impossible. This seminar exposes students to diverse ways of thinking and expertise through the inception and completion of interdisciplinary collaborative projects in small teams. Each stage of project development will be presented to the class. These presentations will offer the opportunity to share and discuss disciplinary dogmas that shape ethics, theories, methodology, and analysis. Tackling issues from multiple experts' points of view will allow for a more holistic approach in a world where everything is connected.

Félicia Olmeta-Schult

School of the Environment, Washington State University, Vancouver, WA 98686, USA.
E-mail: felicia.olmeta@email.wsu.edu

PERSPECTIVES

PHYSIOLOGY

Killer fat

Adipocytes in the skin release an antimicrobial factor to fight staphylococcus infection

By **John F. Alcorn**[1] *and* **Jay K. Kolls**[2]

The skin, the largest organ in the human body, plays a critical role as a barrier to pathogen entry into tissues. Its disruption can lead to invasive bacterial disease. When this does happen, many resident cells in the skin's dermal layers, including immune cells, limit bacterial colonization. The role of fat cells (adipocytes) in the skin's host defense function is only recently emerging. On page 67 in this issue, Zhang *et al.* (*1*) add to this view by showing that dermal adipocytes participate directly in innate immunity against *Staphylococcus aureus* (see the figure).

The skin is composed of an outermost stratified epidermis and an underlying dermis that is inundated with vascular tissue, fibroblasts, phagocytes, lymphocytes, and adipose tissue. Epidermal keratinocytes produce antimicrobial peptides that kill invading pathogens and, in concert with dermal phagocytes, promote pathogen clearance (*2*).

S. aureus is a commonly found commensal bacterium on human skin. Infection with methicillin-resistant *S. aureus* (MRSA) is responsible for more deaths in the United States than any other infectious pathogen (*3*). Susceptibility to *S. aureus* infection in the skin (and lung) (*2*, *4*) has been associated with decreased production of cytokines in these organs that regulate the production of antimicrobial peptides by epithelial and immune cells (*5*, *6*). *S. aureus* also triggers the production of interleukin-6 by adipocytes, a cytokine that stimulates the production of the bacteriostatic, iron-binding protein hepcidin (*7*). This suggests a role for these cells in host defense against this pathogen (*8*, *9*).

A breach in the epidermis can cause the underlying dermis to become infected with *S. aureus*, resulting in dangerous inflammation (cellulitis and fasciitis). To model this, Zhang *et al.* used subcutaneous injection of MRSA in mice to introduce infection directly into the underlying dermis. MRSA infection results in the recruitment of myeloid, lymphoid, and mast cells, all of which have been implicated in the bacterial clearance and successful host defense. However, the dermis is also characterized by connective tissue containing fibroblasts and adipocytes. The authors noted that MRSA infection caused a marked increase in dermal adipose tissue in part due to hypertrophy and proliferation of the adipocytes. Adipogenesis was partly due to expression of the transcription factor zinc finger protein 423 (ZFP423), whose expression controls another transcription factor called peroxisome proliferator-activated receptor gamma (PPAR-γ). Using mice with a mutation in ZFP423, or treating normal mice with a PPAR-γ inhibitor, Zhang *et al.* reveal the requirement for these transcription factors in the expansion of dermal adipose in response to MRSA infection. Blocking adipogenesis in mouse skin also impaired host defenses against MRSA infection.

Cathelicidin is an antimicrobial peptide with anti-staphylococcal activity. By showing that a murine adipocyte cell line and primary human adipocytes produce this peptide in response to *S. aureus* conditioned media or inactivated bacteria, Zhang *et al.* suggest that fat cells can directly sense the pathogen. As well, conditioned media from wild-type murine adipocytes, but not adipocytes from cathelicidin-deficient mice, controlled *S. aureus* growth in mice. Animals with deficient adipogenesis (mice lacking ZFP423 or mice treated with the PPAR-γ inhibitor) had impaired cathelicidin production upon *S. aureus* infection and were as susceptible to infection as cathelicidin-deficient mice. Moreover, PPAR-γ inhibition in cathelicidin-deficient mice did not exacerbate infection, suggesting that the major anti-staphylococcal protein controlled by adipogenesis is cathelicidin.

In addition to its well-known role in growth and metabolism, adipocytes play key roles in controlling soft tissue infection. From an evolutionary perspective, this makes sense, as this function would provide the host with an additional layer of defense against an abraded or trau-

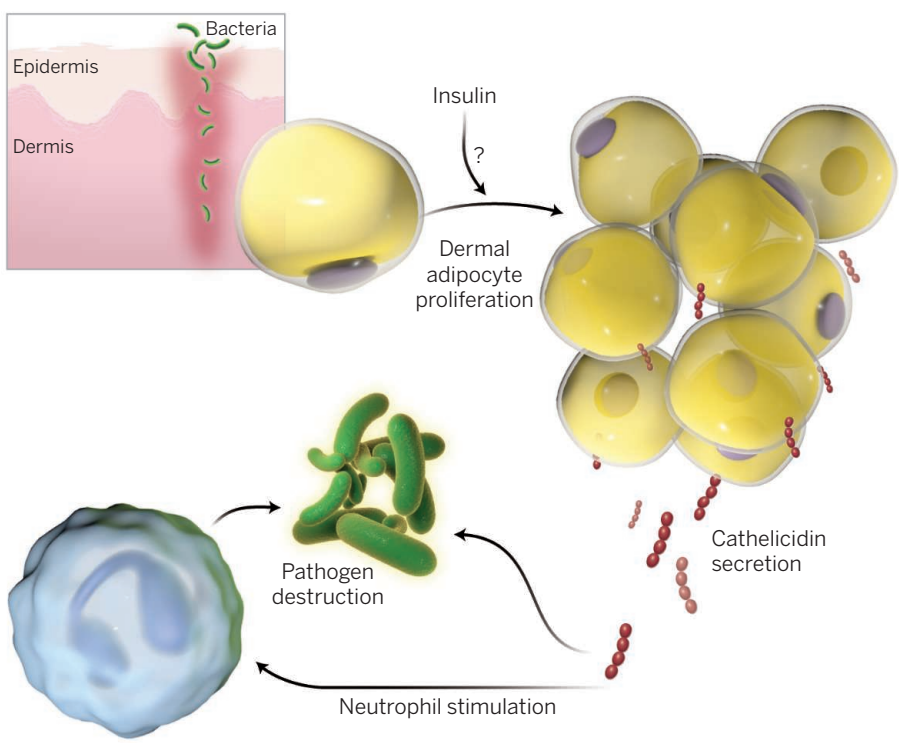


ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

Responding to the breach. Disruption of the epidermis can introduce pathogens into the dermis. This provokes the proliferation of adipocytes (involving transcription factors ZFP423 and PPAR-γ). Adipocytes secrete cathelicidin, whose anti-staphylococcal activity can control skin infections.

[1]*Department of Pediatrics, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA 15224, USA.* [2]*Richard King Mellon Foundation Institute for Pediatric Research, University of Pittsburgh, Pittsburgh, PA 15224, USA. E-mail: jay.kolls@chp.edu*

matic wound to the epidermis. However, there is likely a healthy amount of dermal fat and an unhealthy amount. Zhang *et al.* address this in part by studying a high-fat diet. Interestingly, induction of adipogenesis in mice through a high-fat diet also increased the production of cathelicidin by the proliferating adipocytes. However, mice harboring disabling mutations in the receptor for leptin—a hormone produced by fat cells that suppresses food intake—gain weight and develop type 2 diabetes, but are more susceptible to *S. aureus* infection (*10*). Likewise, in humans, obesity has been associated with an increased risk of skin and soft tissue infection (*11*). One possible explanation for this discrepancy is that insulin resistance or other aspects of metabolic syndrome perturb the infection-adipogenesis-cathelicidin pathway identified by Zhang *et al.* Thus, signaling by adipose-derived hormones that control energy expenditure (adipokines) could influence the expression of cathelicidin. This antimicrobial peptide also is posttranslationally cleaved to its active form, a process that that may also be influenced by obesity and metabolic syndrome.

The mechanism underlying the recognition of *S. aureus* by adipocytes remains unclear, although it likely involves toll-like receptor 2 (TLR2). Adipocytes express many members of the toll-like receptor family, including TLR2 (*9*, *12*), which recognizes lipopeptides produced by bacteria. This may be an operative pathway that controls cathelicidin production. Moreover, a TLR2-ZFP423-PPAR-γ-cathelicidin pathway might be augmented pharmacologically by PPAR-γ agonists, thereby increasing host resistance to infection in susceptible individuals such as those with diabetes and metabolic syndrome. ■

REFERENCES

1. L. j. Zhang *et al.*, *Science* **347**, 67 (2015).
2. M. H. Braff, M. Zaiou, J. Fierer, V. Nizet, R. L. Gallo, *Infect. Immun.* **73**, 6771 (2005).
3. L. S. Miller, J. S. Cho, *Nat. Rev. Immunol.* **11**, 505 (2011).
4. K. M. Robinson *et al.*, *J. Infect. Dis.* **209**, 865 (2014).
5. L. Maródi *et al.*, *J. Allergy Clin. Immunol.* **130**, 1019 (2012).
6. S. J. Aujla, J. F. Alcorn, *Biochim. Biophys. Acta* **1810**, 1066 (2011).
7. S. Coimbra, C. Catarino, A. Santos-Silva, *Obes. Rev.* **14**, 771 (2013).
8. E. Banke *et al.*, *Metabolism* **63**, 831 (2014).
9. K. M. Ajuwon, W. Banz, T. A. Winters, *J. Inflamm. (Lond.)* **6**, 8 (2009).
10. S. Park, J. Rich, F. Hanses, J. C. Lee, *Infect. Immun.* **77**, 1008 (2009).
11. P. Sreeramoju *et al.*, *Am. J. Surg.* **201**, 216 (2011).
12. J. Pietsch *et al.*, *Ann. N. Y. Acad. Sci.* **1072**, 407 (2006).

ACKNOWLEDGMENTS

J.F.A. acknowledges support from NIH grant R01HL107380; J.K.K. is supported by NIH grant R37HL079142.

10.1126/science.aaa4567

IMAGE: CDC/PHANIE/SCIENCE SOURCE

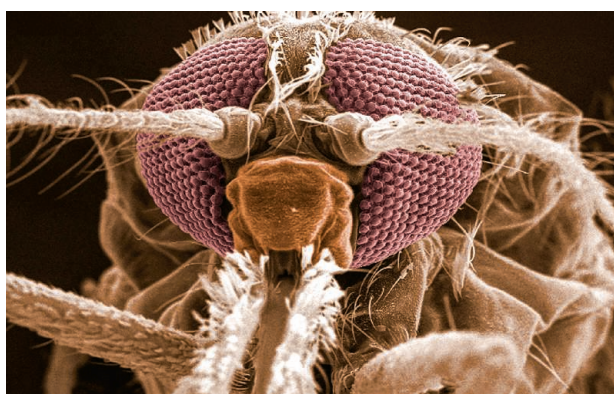
EVOLUTIONARY GENOMICS

Conundrum of jumbled mosquito genomes

Multiple *Anopheles* mosquito genome sequences reveal extreme levels of mixing

By **Andrew G. Clark**[1,2] *and* **Philipp W. Messer**[2]

Malaria is caused by injection of *Plasmodium* parasites into the human bloodstream via the bites of infected mosquitoes. This simple description overlooks a fantastic biological complexity: Some 60 anopheline mosquito species can serve as vectors for five distinct species of *Plasmodium* that produce varying levels of illness in many animal species. Comparative genomic studies may shed light on the mechanisms whereby *Anopheles gambiae* specifically target humans, why the mosquitoes can tolerate *P. falciparum* infection, and how the parasite has adapted to this lifestyle. In this issue, Neafsey *et al.* [(*1*), page 43] and Fontaine *et al.* [(*2*), page 42] analyze the genome sequences of 16 species of anopheline mosquitoes and reveal a complex pattern of evolution that defies the classic concept of a phylogenetic tree.



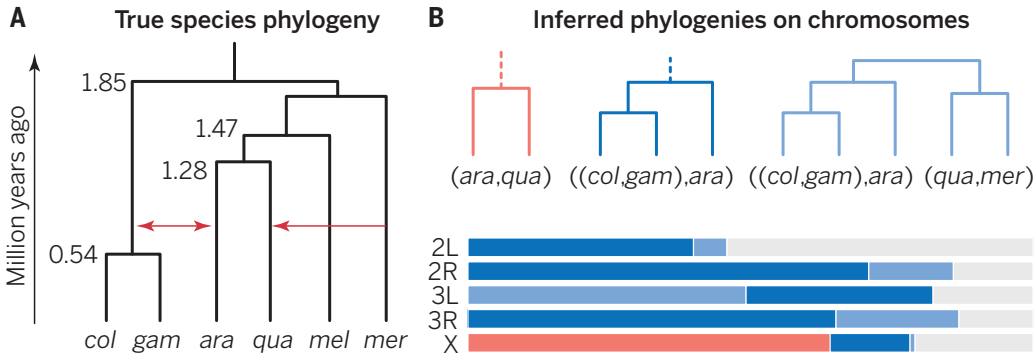
Anopheles gambiae.

Sequencing of multiple related species has revealed many attributes of the evolutionary pressures faced by those species (*3*–*6*). For example, multiple genome alignments can show which genes are most conserved and which evolve the fastest. In general, *Anopheles* genomes appear to evolve faster than do *Drosophila* genomes, perhaps because the former depend on hosts that may provide opportunities for coevolutionary arms races. This is especially evident in the families of closely related genes that formed from the duplication of a single original gene. Fontaine *et al.* show that the 16 *Anopheles* species gain and lose such gene family members at five times the rate of the 12 sequenced *Drosophila* species.

Anopheles genomics also sheds light on the genes involved in the specialization of *An. gambiae* on human hosts. Olfactory and gustatory receptors help the mosquitoes to identify and be attracted to hosts. Although these gene families are generally highly conserved across *Anopheles* genomes, *An. gambiae* shows a remarkable gain of 12 olfactory receptors, suggesting a possible role for these genes in guiding human host preference [as seen in *Aedes* mosquitoes (*7*)]. Many of the olfactory and gustatory receptors also display accelerated protein evolution, consistent with response to positive natural selection. Future studies should test the adaptive benefit of specific odorant chemicals and the specific associations between odorant chemicals and odorant receptors.

The genome sequences generated by Neafsey *et al.* provide the opportunity to investigate whether the observed evolutionary patterns in sequence divergences between the 16 mosquito species are consistent with a single phylogenetic tree. That the *Anopheles* phylogeny might be complex has been suspected since the first *An. gambiae* genome was sequenced (*8*) from a lab strain that included two distinct subtypes [today recognized as two separate species, *An. gambiae* and *An. coluzzi* (*9*)]. The observations that these two species readily hybridize and also have largely overlapping ranges suggest that there might be gene flow between them. Despite this, the *Anopheles* phylogeny has generally been described by a species tree, constructed from the informa-

[1]*Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA.* [2]*Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA. E-mail: ac347@cornell.edu*



Jumbled mosquito genomes. (**A**) The "true species tree" and major introgression events (red arrows) in the *An. gambiae* complex inferred from the X chromosomal sequences by Fontaine *et al.*: *An. coluzzii* (*col*), *An. gambiae* (*gam*), *An. arabiensis* (*ara*), *An. quadriannulatus* (*qua*), *An. melas* (*mel*), and *An. merus* (*mer*). (**B**) Locally inferred trees. The bars show the proportions of 50-kb windows on each chromosome that yield phylogenies consistent with one of the three topologies shown above. Large portions of the genome indicate that *ara* and *col+gam* are sister groups (dark blue); other regions of the genome also group together *qua* and *mer* (light blue), compatible with the two major introgression events. Only windows on the X chromosome predominantly recover the species topology, grouping together *ara* and *qua* (red). Altogether there are 85 tree topologies, and the gray areas correspond to other topologies distinct from the three depicted ones.

tion on whether crosses between species can produce fertile male progeny.

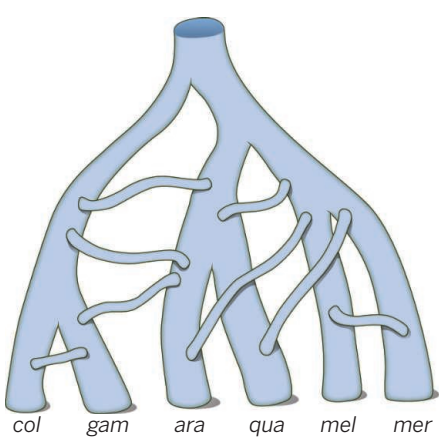
Fontaine *et al.* explored the consistency of an *Anopheles* species tree by partitioning the genome alignment of the six *gambiae* complex species into 50-kb windows and computationally inferring a separate phylogeny for each window. If the species split in an orderly fashion, then essentially every window ought to reflect the same tree. There are 85 possible topologies to the trees with six *gambiae* complex species and one outgroup, and in fact all 85 tree topologies were seen in at least one genomic window (see the first figure). Evidently the *gambiae* complex of *Anopheles* has not been respecting species boundaries.

There are two possible reasons that the overall pattern of divergence of genomic segments may differ from the overall species tree, which has one of the more rarely supported topologies genome-wide. First, there may be introgression when crosses between species produce fertile female hybrids that result in gene flow between species. Second, there may be incomplete lineage sorting, whereby an ancestral population splits into two daughter species and each becomes fixed for a different allele that was segregating at a previously polymorphic locus in the ancestral species. The local phylogeny at such a locus inferred in the two daughter species may then not necessarily be consistent with the branching pattern of the species tree.

Fontaine *et al.* come down firmly on the side of introgression (see the first figure). This process seems plausible, but the level of gene flow needed to essentially shuffle the autosomal (non–sex chromosome) variation is quite high. This much interspecific hybridization is surprising without more introgression on the X chromosome, especially in light of Neafsey's result that the X-to-autosome rate of transfer is unusually high in anopheline mosquitoes.

Furthermore, the timing of the *gambiae-arabiensis* introgression remains unclear. In particular, it is not clear whether introgression is still happening (in which case *arabiensis* must still be undergoing hybridization with both *gambiae* and *coluzzi*) or whether hybridization ceased some time ago. Fontaine *et al.* have done a marvelous job in highlighting the truly odd character of these genomes, and their explanation is consistent with the data, but it raises many additional questions that warrant deeper study.

The breakdown of tidy bifurcating trees with distinct species at the tips has been seen in many systems (*10*). When there is extensive exchange across species, the phylogeny is no longer treelike but rather has a web of crossing lineages in the form of a network (see the second figure). This so-called reticulate evolution is especially evident in bacteria, where genetic exchange can be so pervasive that the concept of species becomes quite slippery (*10*). Reticulate evolution has been seen in many other species groups, but the pattern in the *gambiae* complex of mosquitoes is so extreme that it, too, challenges any clear definition of species in this group. Fontaine *et al.* adhere to a classical view that there is a "true species tree," presumably the phylogeny that is shown by the genes that mediate male and female fertility. But given that the bulk of the genome has a network of relationships that is different from this true species tree, perhaps we should dispense with the tree and acknowledge that these genomes are best described by a network, and that they undergo rampant reticulate evolution.



More web than tree. The pattern of evolution seen in the *An. gambiae* species complex resembles a network more than a tree. This type of evolutionary network is referred to as reticulate evolution (*10*).

Beyond these two papers, additional tests of whether *Anopheles* mosquitoes have an accelerated rate of evolution with extensive introgression between species may come from contrasts of observed and expected patterns of polymorphism within species, requiring the sampling and sequencing of multiple individuals from within each species (*11*, *12*). Such a population genetic approach may be the simplest way to resolve the lingering puzzles about this system. In particular, the sizes and sequence diversity of introgressed segments could be used to model the past timing and extent of hybridization events. The ability to detect positive selection for genomic features that might confer human host adaptation would also be greatly improved with polymorphism data. Additional sequencing to characterize polymorphism in *An. gambiae* would answer some questions, but would undoubtedly also raise new ones. For now, the two papers succeed in dramatically advancing *Anopheles* genomics and providing baseline resources to answer many questions beyond those pursued here. ■

REFERENCES

1. D. E. Neafsey *et al.*, *Science* **347**, 1258522 (2015).
2. M. C. Fontaine *et al*, *Science* **347**, 1258524 (2015).
3. A. G. Clark *et al.*, *Nature* **450**, 203 (2007).
4. S. H. Martin *et al.*, *Genome Res.* **23**, 1817 (2013).
5. K. Lindblad-Toh *et al.*, *Nature* **478**, 476 (2011).
6. J. Rogers, R. A. Gibbs, *Nat. Rev. Genet.* **15**, 347 (2014).
7. C. S. McBride *et al.*, *Nature* **515**, 222 (2014).
8. R. A. Holt *et al.*, *Science* **298**, 129 (2002).
9. M. Coetzee *et al.*, *Zootaxa* **3619**, 246 (2013).
10. W. F. Doolittle, *Science* **284**, 2124 (1999).
11. J. E. Pool *et al.*, *PLOS Genet.* **8**, e1003080 (2012).
12. T. F. Mackay *et al.*, *Nature* **482**, 173 (2012).

10.1126/science.aaa3600

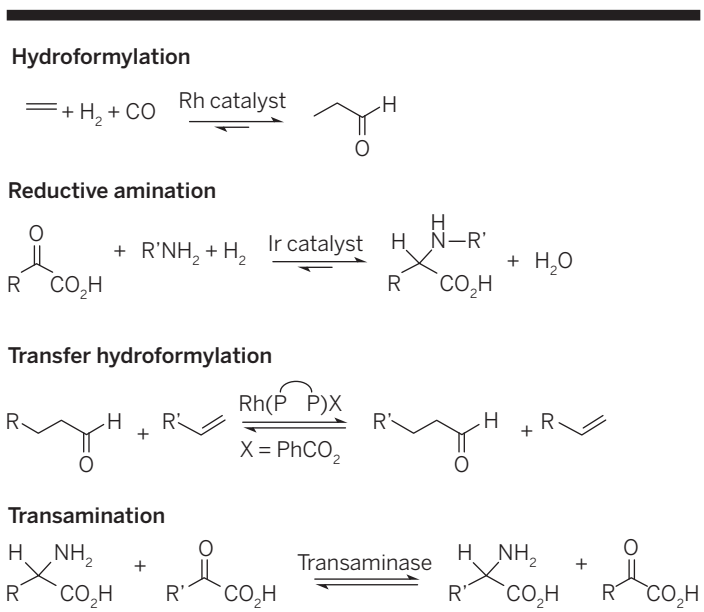
Construction and deconstruction of aldehydes by transfer hydroformylation

A soluble rhodium catalyst converts alkenes to aldehydes without the need for toxic or explosive gas-phase reactants

By **Clark R. Landis**

For more than 75 years, the chemical industry has relied on the hydroformylation reaction (the "oxo" process) to convert alkenes, CO, and H_2 into aldehydes (*1*) used in making plasticizers, alcohols, amines, and esters. Although "syngas," the 1:1 mixture of CO and H_2, is inexpensive and abundant, it is also toxic and highly flammable, so most commercial applications of hydroformylation are limited to commodity-scale processes in plants dedicated to handling superambient pressures of gases. More general use of the hydroformylation process would be facilitated by "gasless" group-transfer approaches to hydroformylation. On page 56 of this issue, Murphy *et al.* (*2*) report a rhodium catalyst that transfers the equivalent of H_2 and CO between an aldehyde and an alkene under mild conditions and without evolving gases.

Transfer hydroformylation shares essential characteristics with biological transferase-catalyzed transformations, such as transamination (*3*) (see the first figure). Transamination converts unsaturated ketoacids into amino acids by shuttling the amino group from a sacrificial amine. There is little thermodynamic driving force for transamination under standard conditions; control of the reaction direction rests in the selective addition or removal of the reaction components. Alternatively, amino acids can be synthesized by thermodynamically favorable reductive amination of a ketoacid with ammonia and H_2 (or its equivalent). Like transamination, the newly reported transfer hydroformylation process interchanges the functional group (aldehyde) between unsaturated molecules (alkenes) without liberating the small-molecule constituents (CO and H_2). Like transamination, transfer hydroformylation enables an abundant reactant (the donor aldehyde or alkene acceptor for transfer hydroformylation) to be combined with a more precious partner to create a product with high value. Similar considerations have spurred the development of a variety of transfer hydrogenation (*4*) and related hydrogen-scrambling reactions during the past few decades.



Related reactions. Transfer hydroformylation is related to hydroformylation and reductive amination as well as enzymatic transamination. Here, R and R′ are different organic groups, and Ph is phenyl.

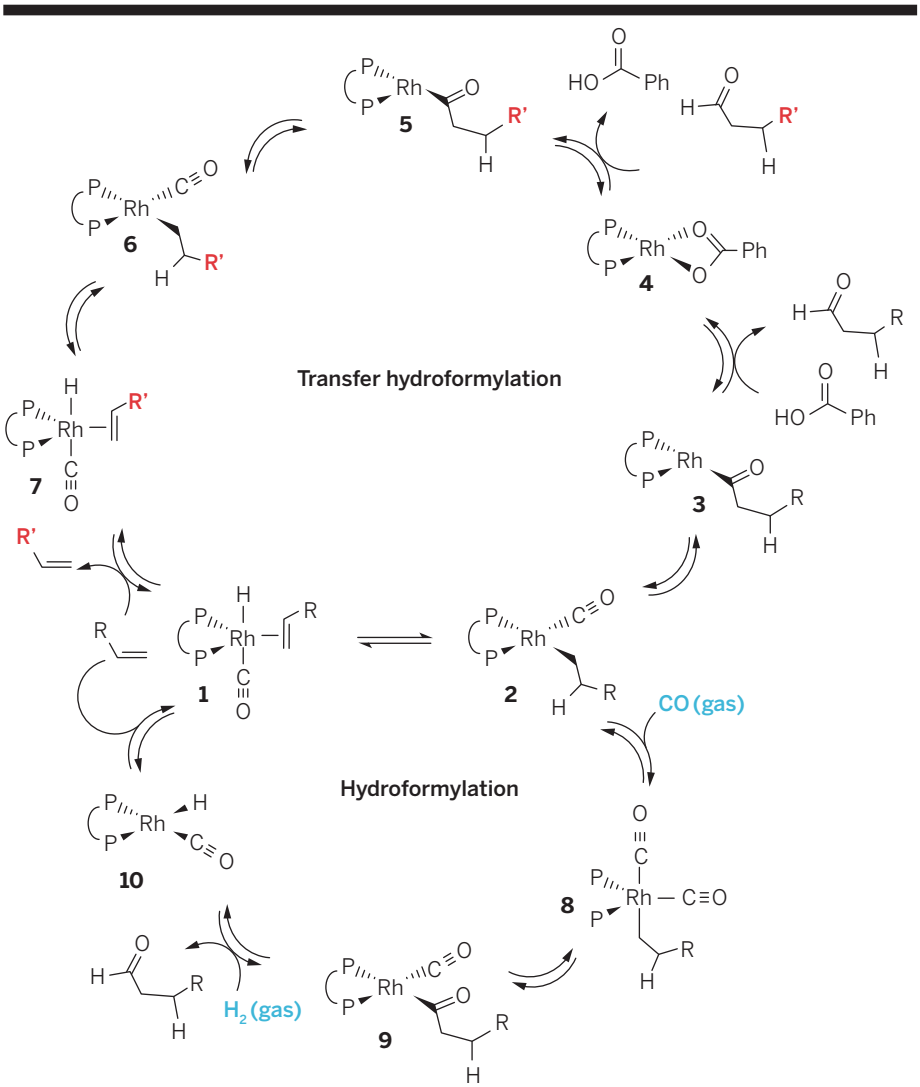
How can a hydroformylation catalyst be converted into a transfer hydroformylation catalyst? The transfer hydroformylation reported by Murphy *et al.* is a result of mechanistic insights from hydroformylation and hydroacylation [the insertion of an alkene into the formyl C-H bond of an aldehyde (*5*)] reactions as catalyzed by organotransition metal complexes. Proposed mechanisms for rhodium-catalyzed hydroformylation and transfer hydroformylation are shown in the second figure. For the two catalytic cycles, several intermediates have coordination geometries that are shared (structures **1** and **7**, or **2** and **8**) or are closely related (structures **3** and **9** differ by only one coordinated CO).

Critical to the development of the new transfer hydroformylation process is optimization of the transformation from **4** to **5**. This transformation comprises elementary steps of aldehyde C-H oxidative addition to a Rh(I) center—a critical step in the hydroacylation and decarbonylation reactions—and reductive elimination of an H-X bond (where X is a ligand bound to Rh) from the resulting acyl hydride. Empirical tuning revealed that when the X ligand is a benzoate, a thermodynamic and kinetic balance enables transfer hydroformylation to proceed without reverting the donor aldehyde all the way to alkene and syngas (*6*, *7*). As a result, transfer hydroformylation is achieved under mild conditions.

At this point, transfer hydroformylation is most effective for transformations with a clear thermodynamic driving force, such as donor aldehydes that yield stabilized alkenes and alkene acceptors that are strained. Thus, the demonstrated aldehyde donor–alkene acceptor pairs primarily comprise aldehydes that yield conjugated alkenes upon dehydroformylation and sacrificial alkenes such as norbornene derivatives. Because any catalytic intermediate with a Rh-CO bond is potentially susceptible to dissociation of CO and any intermediate with a Rh-H bond could lose H_2 either by a dinuclear elimination process or protonolysis by carboxylic acid, kinetics are also important. The acceptor alkene must trap hydride and carbonyl intermediates rapidly to prevent simple aldehyde decarbonylation and gas evolution.

Transfer hydroformylation has many important potential applications. One concerns the dehydroformylation of an abun-

Department of Chemistry, University of Wisconsin, Madison, WI 53706, USA. E-mail: landis@chem.wisc.edu



Gas versus liquid phase. Hydroformylation to generate aldehydes uses gas-phase reactants, where transfer hydroformylation developed by Murphy *et al.* proceeds in solution.

dant but complex aldehyde in the presence of an inexpensive donor alkene to yield a complex and valuable product alkene. For example, Murphy *et al.* demonstrate a three-step transformation of (+)-yohimbine to (+)-yohimbenone that features dehydroformylation with norbornene as the acceptor.

Alternatively, a simple alkene may be converted into a more valuable aldehyde by transfer from an abundant sacrificial aldehyde. Ultimately, this application is the more desirable because it would replace hydroformylation with a gasless equivalent. However, many questions must be addressed before transfer hydroformylation with inexpensive aldehydes provides a valuable alternative to hydroformylation under pressurized conditions. Can transfer hydroformylation proceed with bulky alkene acceptors? Can simple aldehydes such as propanal, or even formaldehyde or sugars, be used as aldehyde donors? Is useful control of regio- and enantioselectivity possible with different rhodium ligands, with acid promoters, or both? These new catalysts may provide the previously missing starting point for the development of general and completely gasless transfer hydroformylation processes. ■

REFERENCES AND NOTES

1. R. Franke, D. Selent, A. Börner, *Chem. Rev.* **112**, 5675 (2012).
2. S. K. Murphy, J.-W. Park, F. A. Cruz, V. M. Dong, *Science* **347**, 56 (2015).
3. A. E. Braunstein, M. G. Kritzman, *Enzymologia* **2**, 129 (1937).
4. S. Hashiguchi, A. Fujii, J. Takehara, T. Ikariya, R. Noyori, *J. Am. Chem. Soc.* **117**, 7562 (1995).
5. M. C. Willis, *Chem. Rev.* **110**, 725 (2010).
6. Dehydroformylation is thermodynamically uphill. Conversion of propanal to ethylene, CO, and H_2 has standard enthalpy and free energy of +31.2 and +13.9 kcal/mol, respectively. Data from the NIST Gas Phase Thermochemistry Database (*7*).
7. http://webbook.nist.gov/chemistry/guide/#thermo-gas

10.1126/science.aaa2329

MICROBIOLOGY

Persisters unmasked

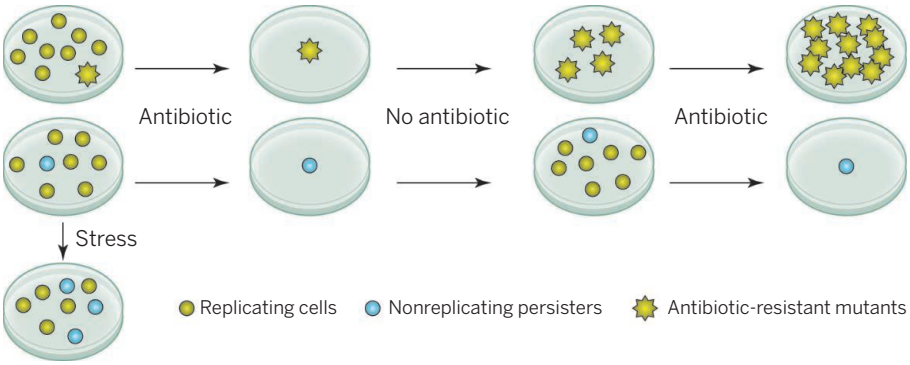
Intracellular toxins cause bacterial growth arrest and antibiotic tolerance

By **David W. Holden**

Penicillin may have saved more human lives than any other drug. Yet, almost as soon as it was introduced in the 1940s, researchers found that the antibiotic could not completely sterilize a culture of a *Staphylococcus aureus* strain sensitive to the drug (*1*). Shortly thereafter, Joseph Bigger showed that when the few cells that had survived an initial treatment were regrown in the absence of penicillin and then exposed again to the antibiotic, the proportion of survivors was similar to that found after the first treatment (see the first figure). Therefore, the survivors were not stable drug-resistant mutants, but transient drug-tolerant persisters (*2*). In the past decade, a resurgence of interest in persisters has revealed some of the molecular mechanisms that stimulate their formation. It has become clear that intracellular toxins present in virtually all bacteria control reversible bacterial growth arrest, explaining their antibiotic tolerance.

In genetically similar or identical populations of organisms that have the same phenotype, an infection or other stress that has the potential to kill one organism can in theory kill the entire population. The consequences of this frailty are exemplified by the vast monocultures of wheat devastated by fungal disease epidemics in the North American Great Plains in the early 20th century. Bacteria reproduce clonally and might therefore appear to be similarly vulnerable to attack from other bacteria or bacteriophages, sudden environmental changes, or exposure to antibiotics. These potentially lethal stresses are very likely to have imposed immense selective pressure for the evolution of phenotypic heterogeneity among bacteria. Since Bigger's experiment, numerous studies have found cell-to-cell variation involving different physiological processes in clonal populations of bacteria grown in the same environment (*3*, *4*).

Virtually all bacteria form antibiotic-tolerant persisters. Many aspects of their



Resistance and persisters. Bacterial populations can contain antibiotic-resistant mutants and/or nonreplicating persisters among replicating cells. Antibiotics kill replicating antibiotic-sensitive bacteria but not resistant mutants or persisters. Resistance is stably inherited among bacteria in the presence or absence of antibiotic, whereas persisters resume growth in the absence of antibiotic. Unstressed populations can continue to generate persisters at low frequency, while various stresses stimulate persister frequency.

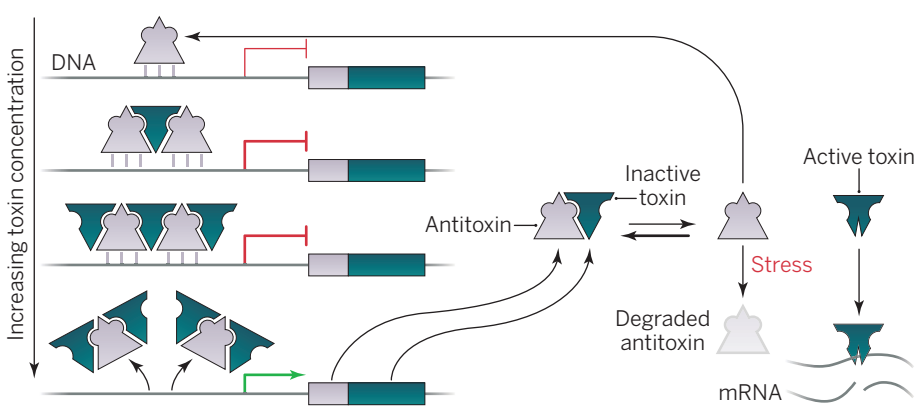
complex biology have now been revealed, especially with the laboratory workhorse, *Escherichia coli*. These advances have been facilitated by equipment and methods that enable single-cell analysis of bacteria.

Perhaps unsurprisingly, most persisters are slow- or nongrowing cells (*5*). Nongrowing cells tend to tolerate a wide variety of stresses. Their production in replicating populations of bacteria is unlikely to have evolved in response to selective pressure from antibiotics. But a lack of growth usually involves reduced metabolic activity, and metabolically inactive cells have greater tolerance to antibiotics than replicating cells. Persisters are present in low numbers in apparently nonstressed populations, before exposure to antibiotics (*5*). This can be viewed as a bacterial insurance against the arrival of a potentially lethal stress.

Entry into the persister state is often controlled by gene pairs encoding cognate toxin-antitoxins (TAs) (*6*, *7*). During bacterial replication, protein toxins are mainly bound to their protein or RNA antitoxins, preventing toxin activity. When liberated, the toxins variously inhibit DNA replication, cleave mRNA endoribonucleases, block protein translation, or interfere with the cell cytoskeleton (*8*) (see the second figure). The common physiological consequence of these activities is a transient arrest of cell growth.

Studies on the protein-based TA systems suggest that inherent instability of an antitoxin leads to random fluctuation in the amounts of free and bound toxin just below a critical threshold. In rare cases, the threshold is surpassed, generating sufficient free toxin to cause growth arrest in a very small part of the bacterial population (*9*). In addition, nutritional starvation and possibly other stresses stimulate persister formation by activating a pathway involving the intracellular signaling molecule ppGpp, polyphosphate, and the cellular protease Lon that degrades antitoxins (*10*).

A process called conditional cooperativity exerts exquisite control over the autoregulated expression of TA genes. For example, at low concentrations, the Doc toxin binds to its antitoxin (Phd) to form an efficient transcriptional repressor complex on the *phd-doc* operator. However, at higher concentrations, more Doc molecules bind to the antitoxin such that it can no longer bind DNA, thereby derepressing transcription (*11*). Thus, the toxin:antitoxin ratio fine-tunes transcription in response to changing levels of unbound toxin (see the second figure). Presumably, exit from the growth-arrested state is initiated by an increase in antitoxin levels and resultant inactivation of the toxin, but this has not yet been shown.



Dual functions of toxins and antitoxins. During normal cell growth, antitoxin and toxin are expressed from a DNA operon. Toxin activity is neutralized by binding antitoxin. Free antitoxin is a weak autotranscriptional repressor. Transcriptional repression is enhanced by increasing concentration of cognate toxin (which forms a complex with the antitoxin on the operator DNA). Above a toxin concentration threshold, toxin-antitoxin complexes dissociate from DNA, derepressing transcription. Thus, cooperative binding of toxin-antitoxin to DNA depends on toxin concentration. Antitoxin is degraded in response to stress, liberating activated toxin. This arrests cell growth—for example, by cleaving mRNA.

The near-ubiquitous occurrence of TA genes in bacteria and the numerous variations of these genes harbored by many bacteria show clearly that they are both ancient and of fundamental importance to bacteria. They are also likely to have important functions during infection, when bacterial pathogens encounter highly stressful environments in their hosts. Some strains of *Mycobacterium tuberculosis*, which is notorious for causing chronic and recurrent infections, contain up to 79 such genes, but detailed investigations into the functions of these and TA genes of other pathogens have only begun recently.

Salmonella Typhi and *Salmonella* Typhimurium also cause recurrent infections in humans. Whole-genome sequencing of *S.* Typhimurium isolates from patients has provided unambiguous evidence for reappearance of the original strain after repeated antibiotic treatment, directly implicating persisters (*12*). Use of fluorescence-based techniques with the mouse model of typhoid-like disease caused by *S.* Typhimurium has enabled direct observation and characterization of persisters during infection (*13*, *14*). *Salmonella* has adapted to survive and replicate in vacuoles after phagocytosis by macrophages. The low pH and poor nutritional status of the vacuole greatly enhance the frequency of TA-dependent nonreplicating persisters (*14*). Reporters of metabolic activity suggest a range of physiological states among nonreplicating intracellular *Salmonella*, from persisters primed for immediate resumption of growth in new macrophages, to cells that are metabolically inactive (dormant) and could require specific resuscitation factors (*14*).

Bigger's insights into persisters were remarkable. Using rudimentary microbio-

Section of Microbiology, MRC Centre for Molecular Bacteriology and Infection, Imperial College London, Armstrong Road, London SW7 2AZ, UK. E-mail: d.holden@imperial.ac.uk

logical techniques, he distinguished persisters from resistant mutants, showed that their levels could be enhanced by stress, and anticipated that they would be in a "dormant nondividing phase" and present among other bacterial pathogens (*2*). The recent exciting progress on mechanisms of TA function establishes these toxins as key inducers of the persister state. Future research should elucidate the many functions of TAs and how they work collectively during persistent bacterial infections. For example, it is unclear whether different stresses activate different TA subsets, and what the profiles of toxin activation are in individual bacterial cells. Some toxins have been shown to be sequence-specific ribonucleases, but whether this specificity has physiological implications is uncertain. It could be that bacteria perceive signals that trigger their exit from quiescence, but the mechanisms involved are unknown.

If persisters lead to recurrent infections requiring multiple courses of antibiotics, then they are likely to contribute appreciably to the current worldwide crisis of antibiotic resistance. Yet, surprisingly little is known about the relative usage of antibiotics for persistent infections and the degree to which persisters influence the emergence of resistance. In the long term, TAs and associated signaling molecules may provide targets for drugs that can either prevent persisters from being formed, or—perhaps more feasibly—coax them out of the nonreplicating state so that they resume susceptibility to antibiotics. This might finally enable complete eradication of an otherwise recurrent or persistent infection, so that, as Bigger put it, "the success of penicillin therapy will become more commensurate with its potentialities" (*2*). ■

REFERENCES AND NOTES

1. G. L. Hobby, K. Meyer, E. Chaffee, *Exp. Biol. Med.* **50**, 281 (1942).
2. J. W. Bigger, *Lancet* **244**, 497 (1944).
3. J. L. Spudich, D. E. Koshland Jr., *Nature* **262**, 467 (1976).
4. J. Casadesús, D. A. Low, *J. Biol. Chem.* **288**, 13929 (2013).
5. N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, S. Leibler, *Science* **305**, 1622 (2004).
6. H. S. Moyed, K. P. Bertrand, *J. Bacteriol.* **155**, 768 (1983).
7. E. Maisonneuve, L. J. Shakespeare, M. G. Jørgensen, K. Gerdes, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13206 (2011).
8. E. Maisonneuve, K. Gerdes, *Cell* **157**, 539 (2014).
9. E. Rotem *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12541 (2010).
10. E. Maisonneuve, M. Castro-Camargo, K. Gerdes, *Cell* **154**, 1140 (2013).
11. A. Garcia-Pino *et al.*, *Cell* **142**, 101 (2010).
12. C. K. Okoro *et al.*, *Clin. Infect. Dis.* **54**, 955 (2012).
13. B. Claudi *et al.*, *Cell* **158**, 722 (2014).
14. S. Helaine *et al.*, *Science* **343**, 204 (2014).

ACKNOWLEDGMENTS

I thank S. Helaine and T. Thurston for helpful comments.

10.1126/science.1262033

CELL BIOLOGY

Lysosomal lipid lengthens life span

A fatty acid moves from the lysosome to the nucleus, altering gene expression and extending longevity in the worm

By **Shuo Han** *and* **Anne Brunet**

Lysosomes were discovered more than 60 years ago as highly acidic cellular organelles containing many enzymes responsible for breaking down macromolecules (*1*). Since then, their roles have expanded. Lysosomes function in autophagy, the process that breaks down cellular components to allow cell survival and homeostasis in the face of starvation (*1*). These organelles also have emerged as a signaling hub for the enzyme mechanistic target of rapamycin (mTOR), a protein kinase involved in cellular and organismal growth responses to nutrient availability (*2*). We also now recognize links between aberrant lysosomal function and several diseases, including lysosomal storage diseases (e.g., Tay-Sachs disease) and neurodegenerative disorders (e.g., Parkinson's disease), and also with aging (*1*). On page 83 of this issue, Folick *et al.* (*3*) indicate how lysosomes play a role in the latter—by deploying a lipid molecule to the nucleus, whose impact on gene expression extends life span in an animal model (the nematode *Caenorhabditis elegans*). The study not only uncovers a lysosome-to-nucleus signaling pathway but also highlights the potential of lipids in mediating long-range physiological effects.

Lysosomes contain about 60 enzymes, including many well-conserved lipases involved in fatty acid breakdown. Defects in lysosomal acid lipase A (LIPA) lead to several human lysosomal storage diseases, including Wolman disease, a disorder characterized by metabolic defects and death in childhood (*4*). In *C. elegans*, the LIPA homolog LIPL-4 is highly expressed in specific conditions that are linked to life-span extension (*5*, *6*). However, the mechanism by which this enzyme modulates aging has remained elusive.

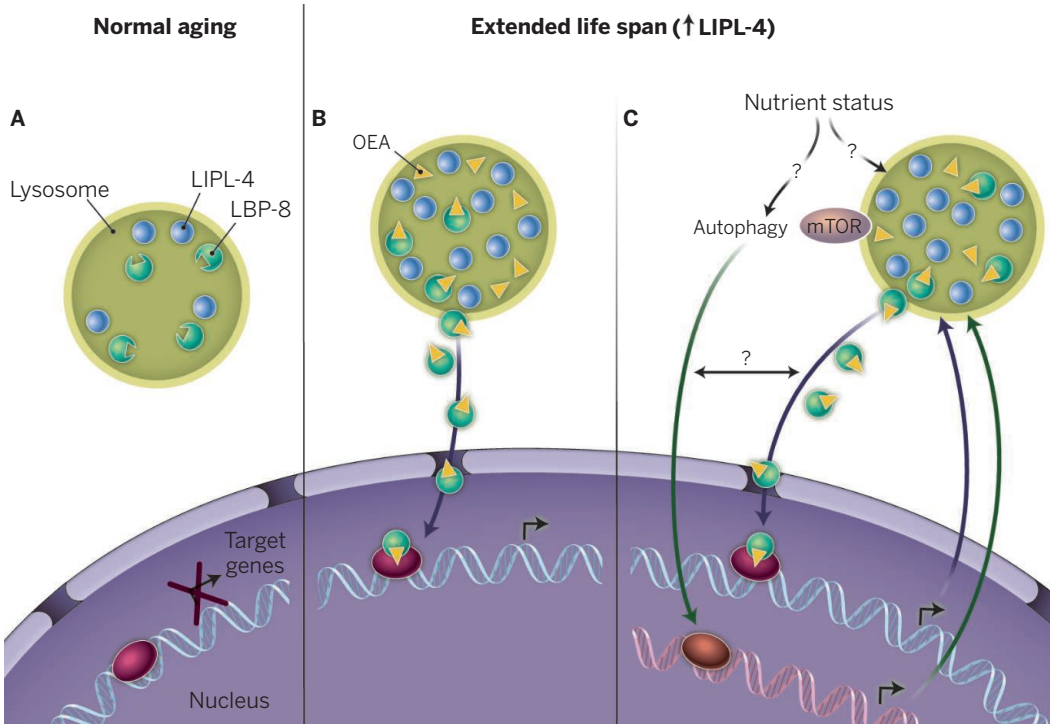
Using a combination of genetics, metabolomics, biochemistry, and immunocytochemistry, Folick *et al.* explored the molecular mechanisms by which lysosomal LIPL-4 activation regulates aging in *C. elegans*. They show that worms overexpressing LIPL-4 live substantially longer than normal worms and produce increased amounts of several bioactive lipids, notably the fatty acid oleoylethanolamide (OEA). OEA is likely generated by the breakdown of more complex lipids in the lysosome by LIPL-4. LIPL-4–overexpressing worms also exhibit an increased amount of a fatty acid binding protein called lipid-binding protein-8 [(LBP-8); the human homolog is fatty acid binding protein (FABP)]. Elegantly coupling fluorescence imaging with mutations that alter protein targeting to the lysosome, Folick *et al.* demonstrate that LIPL-4 must reside within the lysosome to extend life span. By contrast, LBP-8 translocates from the lysosome into the nucleus to ensure increased longevity. As LBP-8 can directly bind to OEA, these results suggest that LBP-8 is a lipid chaperone assisting OEA entry into the nucleus (see the figure).

What happens once OEA is shuttled into the nucleus? Folick *et al.* found that OEA physically binds to and activates conserved nuclear hormone receptors, thereby activating the transcription of target genes. Fatty acid ligands have been reported to control the transcriptional activity of subfamilies of nuclear receptors (*7*), and OEA can bind to the nuclear receptor peroxisome proliferator–activated receptor-α (PPARα) in mice (*8*). The authors report that two particular nuclear receptors—nuclear hormone receptor-49 (NHR-49) and NHR-80, the *C. elegans* homologs of mammalian PPARα and hepatic nuclear factor 4, respectively—are both required for LIPL-4–induced longevity, and that OEA can directly bind to NHR-80. This observation is consistent with previous reports that NHR-49 and NHR-80 play critical roles in life-span regulation in *C. elegans* (*9*, *10*).

> ***"...dietary modulation of fatty acids...has the potential to delay aging."***

What about dietary supplementation of OEA? Folick *et al.* found that feeding worms OEA during their adult life is sufficient to

Department of Genetics, Stanford University, Stanford, CA 94035, USA. E-mail: abrunet1@stanford.edu



A pathway to longevity. (**A**) During normal aging, the lipase LIPL-4 and lipid-binding protein LBP-8 reside in the lysosome. (**B**) In long-lived worms, there is an increase in LIPL-4 and the fatty acid OEA in the lysosome. OEA translocates into the nucleus by binding to the chaperone LBP-8. OEA binds to the nuclear receptor NHR-80 to affect the expression of genes that regulate longevity. (**C**) Nutrient status could also affect LIPL-4 activity. A positive feedback loop likely exists between nuclear transcription factors and the lysosome. The mTOR and autophagy pathways could modulate this lysosome-to-nucleus pathway for longevity.

activate these nuclear hormone receptors and promote longevity. OEA may therefore represent a key ligand that activates these nuclear receptors to modulate aging, although it remains to be determined whether OEA affects aging entirely by acting through these transcription factors or whether it has other effects on the organism. The ensemble of target genes downstream of these nuclear receptors that promote longevity also remains to be identified.

The findings of Folick *et al.* are exciting because they are the first to establish a lysosome-to-nucleus signal that functions in aging regulation and to show that dietary modulation of fatty acids such as OEA has the potential to delay aging. Because genes in this pathway are conserved, the findings also provide insights on the regulation of human nuclear receptors by lysosomal signaling. Given that OEA affects feeding behavior and body weight in mice by acting through PPARα (*8*), dietary OEA may also have an impact on aging in mammals. The availability of bioactive lipids such as OEA could depend on the internal nutritional state of the organism. Environmental interventions such as fasting or overfeeding could alter the availability and composition of the lipid pool, consequently changing the binding status of nuclear receptors, altering downstream transcription programs, and affecting aging.

Signaling between the lysosome and the nucleus is likely to be a two-way street. Indeed, recent reports in mammalian cells have established that lysosomal autophagy can be regulated by lipid-sensing transcription factors in the nucleus during the feeding and fasting cycles, namely the transcription factors farnesoid X receptor (FXR) (*11*), cyclic adenosine monophosphate responsive element binding protein (CREB) (*11*), and PPARα (*12*). Thus, while lysosomes release diffusible lipid messengers that affect transcription, lipid-sensing transcription factors could provide feedback regulation of the lysosome, maintaining metabolic homeostasis based on nutrient status.

Lysosomes are involved in controlling the activity of mTOR and the execution of autophagy (an intracellular mechanism that breaks down cellular components) in response to nutrient availability (*13*). LIPL-4 itself is important for inducing autophagy in *C. elegans* (*6*). Therefore, a key remaining question concerns the connection between this lysosome-to-nucleus signaling and the TOR-autophagy pathway. Could TOR and autophagy play a role in the longevity of LIPL-4–overexpressing animals? Conversely, does an increase in OEA extend life span because of inhibition of the TOR pathway? More generally, it will be important to determine whether this lysosome-to-nucleus signaling pathway is important for longevity conditions that have been shown to require LIPL-4, such as the lack of an intact germ line (*5*, *6*).

In a broader context, these findings raise the question of the site of action of this lysosome-to-nucleus signaling pathway and whether it is entirely cell-autonomous. Both LIPL-4 and LBP-8 are expressed in the gut of the worm, and presumably this is also the site of action of OEA. However, an interesting possibility could be that OEA, as a cell-permeable lipid, is secreted outside the gut to activate nuclear receptors in other tissues. In this way, an inter-tissue lipid signaling network may exist during the aging process and might be involved in systemic lifespan regulation (*14*).

In addition to OEA, other lipids or metabolites could act as diffusible signals between different organelles to orchestrate coordinated cellular responses. Unbiased metabolomic profiling is a promising discovery tool to decipher the mechanisms underlying many human metabolic diseases. This approach would also help to identify the elusive ligands for many nuclear receptors (*15*). Ultimately, modulations of bioactive lipids could be a therapeutic strategy for a wide range of human metabolic disorders and age-related diseases. ■

REFERENCES

1. C. Settembre, A. Fraldi, D. L. Medina, A. Ballabio, *Nat. Rev. Mol. Cell Biol.* **14**, 283 (2013).
2. L. Bar-Peled, D. M. Sabatini, *Trends Cell Biol.* **24**, 400 (2014).
3. A. Folick *et al.*, *Science* **347**, 83 (2015).
4. Ž. Reiner *et al.*, *Atherosclerosis* **235**, 21 (2014).
5. M. C. Wang, E. J. O'Rourke, G. Ruvkun, *Science* **322**, 957 (2008).
6. L. R. Lapierre, S. Gelino, A. Meléndez, M. Hansen, *Curr. Biol.* **21**, 1507 (2011).
7. A. Chawla, J. J. Repa, R. M. Evans, D. J. Mangelsdorf, *Science* **294**, 1866 (2001).
8. J. Fu *et al.*, *Nature* **425**, 90 (2003).
9. M. R. Van Gilst, H. Hadjivassiliou, K. R. Yamamoto, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13496 (2005).
10. J. Goudeau *et al.*, *PLOS Biol.* **9**, e1000599 (2011).
11. S. Seok *et al.*, *Nature* **516**, 108 (2014).
12. J. M. Lee *et al.*, *Nature* **516**, 112 (2014).
13. C. Settembre, A. Ballabio, *Trends Cell Biol.* **24**, 743 (2014).
14. A. Dillin, D. E. Gottschling, T. Nyström, *Curr. Opin. Cell Biol.* **26**, 107 (2014).
15. R. M. Evans, D. J. Mangelsdorf, *Cell* **157**, 255 (2014).

10.1126/science.aaa4565

ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

EDUCATION RESEARCH

Rebooting MOOC Research

Improve assessment, data sharing, and experimental design

By **Justin Reich**

The chief executive officer of edX, Anant Agarwal, declared that Massive Open Online Courses (MOOCs) should serve as "particle accelerator for learning" (*1*). MOOCs provide new sources of data and opportunities for large-scale experiments that can advance the science of learning. In the years since MOOCs first attracted widespread attention, new lines of research have begun, but findings from these efforts have had few implications for teaching and learning. Big data sets do not, by virtue of their size, inherently possess answers to interesting questions. For MOOC research to advance the science of learning, researchers, course developers, and other stakeholders must advance the field along three trajectories: from studies of engagement to research about learning, from investigations of individual courses to comparisons across contexts, and from a reliance on post hoc analyses to greater use of multidisciplinary, experimental design.

EDUCATION

CLICKING OR LEARNING? Few MOOC studies make robust claims about student learning, and fewer claim that particular instructional moves caused improved learning. We have terabytes of data about what students clicked and very little understanding of what changed in their heads.

Consider four recent studies conducted on Udacity, Khan Academy, Google Course Builder, and edX (*2*–*5*). Each study addressed a correlation between measures of student success (such as test scores or course completion) and measures of student activity. All four studies operationalized activity similarly, boiling down the vast data available to a simple, person-level summary variable: number of problems attempted (Udacity), minutes on site (Khan Academy), weekly activity completion (Google), and number of "clicks" per student in the event logs (edX). The complexity of student activity (*6*) captured by these platforms was lost. Using simple comparisons or regressions, all four concluded there is a positive correlation between student activity and success.

It does not require trillions of event logs to demonstrate that effort is correlated with achievement. As these are observational findings, the causal linkages between doing more and doing better are unclear. Beyond exhorting students to be more active, there are no practical implications for course design. The next generation of MOOC research needs to adopt a wider range of research designs with greater attention to causal factors promoting student learning.

WATCHING WITHOUT LEARNING. One reason that early MOOC studies have examined engagement or completion statistics is that most MOOCs do not have assessment structures that support robust inferences about learning. MOOC researchers would, ideally, have assessment data with three characteristics. First, assessments should take place at multiple time points. Pretesting is critical in MOOCs, because heterogeneous registrants include novices and domain experts (*7*). Second, assessments should capture multiple dimensions of learning, from procedural to conceptual. Students who earn high grades on quantitative exam questions often show no growth in their conceptual understanding or expert thinking (*8*). Finally, courses should include assessments that have been validated by prior research, so comparisons can be made to other settings. Some recent MOOC studies meet these criteria and offer insights about which learners benefit most from MOOCs and which course materials may best support learning (*9*). With greater attention to assessment in course design, researchers can make stronger claims about what students learn—not just what they do.

Distinguishing between engagement and learning is particularly crucial in voluntary online learning settings, because media that provoke confusion and disequilibrium can be productive for learners (*10*). Addressing misconceptions requires addressing the uncomfortable gap between our intuitions and scientific reality. Unfortunately, learners may prefer videos that present material more simply. For instance, students use more positive language to describe instructional videos that present straightforward descriptions of phenomena, even though students learn more from media that directly address misconceptions (*11*). Course developers optimizing for engagement statistics can create pleasurable media experiences that keep students watching without necessarily learning.

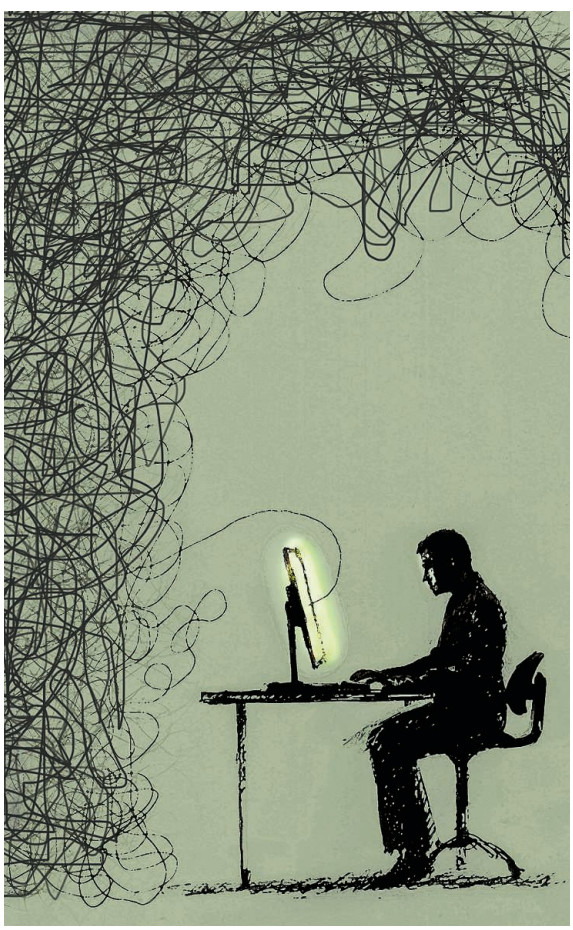


ILLUSTRATION: GARY WATERS/GETTY IMAGES

RETHINK DATA SHARING. Although MOOC researchers have data from thousands of students, few have data from many courses. Student privacy regulations, data protection concerns, and a tendency to hoard data conspire to curtail data sharing. As a result, researchers can examine variation between students but cannot make robust inferences about cross-course differences.

For example, Nesterko *et al.* found a modest positive correlation between frequent, intermediate due dates and MOOC completion rates (*12*). But the 10 courses they examined differed not only by their use of due dates but also by their enrollment size, subject matter, and other dimensions. Data from

HarvardX, Harvard University, Cambridge, MA 02476, USA. E-mail: justin_reich@harvard.edu

hundreds of courses will be necessary to conduct meaningful post hoc comparisons of instructional approaches.

Sharing learner data is no simple matter. Recent efforts to de-identify student data so as to meet privacy requirements demonstrate that the blurring and scrubbing required to protect student anonymity deform data to the point where they are no longer reliable for many forms of scientific inquiry (*13*). Enabling a shared science of MOOCs based on open data will require substantial policy changes and new technical innovations in social science data sharing. One policy approach would be to decouple privacy protections from efforts to maintain anonymity, which would allow researchers to share identifiable data in exchange for greater oversight of their data protection regimes. Technical solutions could include regimes based on differential privacy, where institutions would keep student data in a standardized format that allows researchers to query repositories, returning only aggregated results.

BEYOND A/B TESTS. In the absence of shared cross-course data, experimental designs will be central to investigating the efficacy of particular instructional approaches. From the earliest MOOC courses, researchers have implemented "A/B tests" and other experimental designs (*14*, *15*). These methods are poised to expand as MOOC platforms incorporate authoring tools for randomized assignment of course content.

The most common MOOC experimental interventions have been domain-independent "plug-in" experiments. In one study, students earned virtual "badges" for active participation in a discussion forum (*16*). Students randomly received different badge display conditions, some of which caused more forum activity than others. This experiment took place in a Machine Learning class, but it could have been conducted in American Literature or Biology. These domain-independent experiments, often inspired by psychology or behavioral economics, are widely under way in the field. HarvardX, for instance, has recently offered courses with embedded experiments that activate social supports and commitment devices and cause manipulations to increase perceptions of instructor rapport.

The signature advantage of plug-in experiments is that successful interventions to boost motivation, memorization, or other common facets of learning can be adapted to diverse settings. This universality is also a limitation: These studies cannot advance the science of disciplinary learning. They cannot identify how best to address a particular misconception or optimize a specific learning sequence. Boosting motivation in well-designed courses is good, but if a MOOC's overall pedagogical approach is misguided, then plug-in experiments can accelerate participation in ineffective practices. Discipline-based education research to understand domain-specific learning in MOOCs may be prerequisite to effectively leveraging domain-independent research.

There are fewer examples of domain-specific experiments that are "baked-in" to the architecture of MOOCs. Fisher randomly assigned students in his Copyright course to one of two curricula—one based on U.S. case law, the other on global copyright issues—to experimentally assess these approaches (*17*). He used final exam scores, student surveys, and teaching assistant feedback to evaluate the curricula and concluded that deep examination of a single copyright regime served students better than a survey of global approaches, providing actionable findings for open online legal education.

The next generation of MOOC research needs ... a wider range of research designs with greater attention to ... factors promoting student learning

Both domain-specific and domain-independent experiments will be important as MOOC research matures, but domain-specific endeavors may require more intentional nurturing. Plug-in experiments fit more easily in the siloed structures of academia, where psychologists and economists can generate interventions to be incorporated in courses developed by others. Domain-specific research requires multidisciplinary teams—content experts, assessment experts, and instructional designers—that are often called for in educational research (*18*) but remain elusive. More-complex MOOC research will require greater institutional support from universities and funding agencies to prosper.

RAISING THE BAR. In a new field, it is appropriate to focus on proof-of-concept demonstrations. For the first MOOC courses, getting basic course materials accessible to millions was an achievement. For the first MOOC researchers, getting data cleaned for any analysis was an achievement. In early efforts, following the path of least resistance to produce results is a wise strategy, but it runs the risk of creating path dependencies.

Using engagement data rather than waiting for learning data, using data from individual courses rather than waiting for shared data, and using simple plug-in experiments versus more complex design research are all sensible design decisions for a young field. Advancing the field, however, will require that researchers tackle obstacles elided by early studies.

These challenges cannot be addressed solely by individual researchers. Improving MOOC research will require collective action from universities, funding agencies, journal editors, conference organizers, and course developers. At many universities that produce MOOCs, there are more faculty eager to teach courses than there are resources to support course production. Universities should prioritize courses that will be designed from the outset to address fundamental questions about teaching and learning in a field. Journal editors and conference organizers should prioritize publication of work conducted jointly across institutions, examining learning outcomes rather than engagement outcomes, and favoring design research and experimental designs over post hoc analyses. Funding agencies should share these priorities, while supporting initiatives—such as new technologies and policies for data sharing—that have potential to transform open science in education and beyond. ■

REFERENCES

1. P. Stokes, *Inside Higher Ed* (2013); bit.ly/13deToN.
2. E. D. Collins, "SJSU plus augmented online learning environment: Pilot project report" (The Research & Planning Group for California Community Colleges, Sacramento, CA, 2013).
3. R. Murphy, L. Gallagher, A. Krumm, J. Mislevy, A. Hafter, "Research on the use of Khan Academy in schools" (SRI Education, Menlo Park, CA, 2014).
4. J. Wilkowski, A. Deutsch, D. M. Russell, in *Proceedings of the ACM Conference on Learning@Scale 2014*, Atlanta, GA, 4 and 5 March 2014 (ACM, New York, 2014), pp. 3–10.
5. J. Reich *et al.*, "HeroesX: The Ancient Greek Hero: Spring 2013 Course Report" (Working paper no. 3, Harvard–HarvardX, Cambridge, MA, 2014).
6. R. S. Siegler, K. Crowley, *Am. Psychol.* **46**, 606 (1991).
7. E. J. Emanuel, *Nature* **503**, 342 (2013).
8. A. Van Heuvelen, *Am. J. Phys.* **59**, 891 (1991).
9. K. Colvin *et al*, *IRRODL* **15**, no. 4 (2014).
10. S. D'Mello, B. Lehman, R. Pekrun, A. Graesser, *Learn. Instr.* **29**, 153 (2014).
11. D. A. Muller *et al.*, *Sci. Educ.* **92**, 278 (2008).
12. S. Nesterko *et al.*, in *Proceedings of the ACM Conference on Learning@Scale 2014*, Atlanta, GA, 4 and 5 March 2014 (ACM, New York, 2014), pp. 193–194.
13. J. P. Daries *et al.*, *Commun. ACM* **57**, 56 (2014).
14. R. F. Kizilcec, E. Schneider, G. Cohen, D. McFarland, *eLearning Pap.* **37**, 13 (2014).
15. D. Coetzee, A. Fox, M. A. Hearst, B. Hartmann, in *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing*, Baltimore, MD, 15 to 19 February 2014 (ACM, New York, 2014), pp. 1176–1187.
16. A. Anderson, D. Huttenlocher, J. Kleinberg, J. Leskovec, in *Proceedings of the 2014 International World Wide Web Conference*, Seoul, Korea, 7 to 11 April 2014 (ACM, New York, 2014), pp. 687–698.
17. W. W. Fisher, "HLS1X: CopyrightX" (Working paper no. 5, Harvard–HarvardX, Cambridge, MA, 2014).
18. G. Siemens, *J. Learn. Analyt.* **1**, 3 (2014).

10.1126/science.1261627

BOOKS *et al.*

TECHNOLOGY

Data meets design

Jennifer Golbeck enjoys a manifesto on the beauty of well-designed virtual interfaces

The Social Machine
Designs for Living Online
Judith Donath
MIT Press, 2014. 432 pp.

Judith Donath's new book, *The Social Machine*, sets out to inspire a change in the way we create online social networking platforms. In her own words, "This book … is a guide to understanding how existing systems influence behavior and a manifesto for designing radically new environments for social interaction."

The book's emphasis on style and presentation begins immediately, opening with an anecdote about a program Donath developed while in graduate school that allowed her to visualize the community of users who were logged in to a central computer in the MIT media lab. She considers and immediately rejects the idea of listing names in alphabetical order because it is "too reminiscent of the list of names on a committee program or memorial." Donath is clearly someone who carefully considers the importance of presentation on every level.

In many ways, this is a book about art before it is a book about technology or social systems. Data visualizations created as works of art appear alongside those created to support decision-making. The chapter on "Data Portraits"—depictions of people made from data by and about them—highlights the value Donath places on aesthetics. She calls these "portraits" rather than visualizations because, as she says, "a portrait is an artistic production."

That is not to say that *The Social Machine* fails in its role as a technology book; quite the contrary. It offers a sweeping literature review of design research related to social interaction. For anyone with interest in this field, either as a technology designer or just as someone who loves beautiful technology, this is destined to become the definitive text. It is eloquent, well organized, and thorough.

A long-standing tension in the space of data visualization exists between making attractive, engaging designs and making useful ones. I confess that in the past, I have used some of the images that Donath highlights in my own classes as examples of "pretty but not very helpful." Randal Munroe's well-known *Map of Online Communities* is one such example (*1*). These can be fun to look at, amusing, and witty. But if you're trying to analyze the data, it quickly becomes clear that they are much less effective than a simple, properly organized Excel spreadsheet. Readers interested in decision support, measurable impacts, and quantified results will not find their perspective here. Donath advocates strongly for beauty.

Still, there are many parts that will be of interest to anyone who is concerned with designing social technology. For example, Chapter 3, "Interfaces Make Meaning," is an excellent introduction to communication and the social meaning of design choices. Donath discusses fonts, color theory, and metaphor and provides lengthy descriptions of how social interpretation of these elements can affect what people see. Her discussion of what motivated physicists at CERN (European Organization for Nuclear Research) to choose the Comic Sans font for their presentation announcing the discovery of the Higgs boson is one of many interesting, thought-provoking examples. Not every designer is familiar with these lessons, but they all should be.

Aside from its focus on innovative, beautiful design, the book also seeks to explain the social impact that technologies have on the people who use them. Donath highlights theories and examples from psychology and the social sciences throughout the book to do this. This component of the book is interesting and often well reasoned. These ideas will open up new ways of thinking about social systems, especially for readers with a technology background.

One critique of the book is that, in some places, Donath's arguments feel a bit naïve or simplistic. For example, the opening chapter suggests that the problem of trolling—the act of posting deliberately inflammatory or off-topic messages on online forums with the intent to anger other users or disrupt normal discussion—might be solved by creating interfaces that "encourage people to take responsibility for their words and to engage with each other more cooperatively." That may be true in some cases, but as they say on the Internet: "trolls gonna troll" no matter how strongly they are encouraged to play nice.

Those who share the author's principles of valuing beautiful and thought-provoking designs for their own sake will fall in love with this book. Those who are motivated to design platforms to generate profits or support analysis and decision-making will find value here, too, although they are unlikely to be completely won over by all of Donath's opinions on radical, creative design. Then again, this is a common characteristic of manifestoes. I don't expect Karl Marx planned to win over everyone with his manifesto, either.

Anthropomorphic "data portraits" can help humanize the online experience, says author Judith Donath.

REFERENCES AND NOTES

1. R. Munroe, http://xkcd.com/256.

The reviewer is at University of Maryland, College Park, College Park, MD 20742, USA. E-mail: jgolbeck@umd.edu

10.1126/science.aaa0421

Designed for the harsh and unpredictable weather of the Arctic, this early inflatable dinghy doubled as a cloak and came with a sail that could also be used as an umbrella.

PHOTO CREDIT: PETER HALKETT

EXHIBITION

On top of the world

Ingenuity and Inuit knowledge led to an Arctic passage, finds Julia Fahrenkamp-Uppenbrink

The story of Arctic exploration is frequently portrayed as one of hardship and endurance, survival against all odds, and often tragic loss of life. This exhibition uses an exceptional set of maps, books, and artworks from the British Library's collection to provide a more nuanced view.

The exhibition begins by showing that well before the famous expeditions of the 19th and early 20th centuries, the search was on for riches in the Arctic and an easier route to China. In 1744, the British government offered £20,000 for the discovery of a Northwest Passage from the Atlantic to the Pacific Ocean. However, maps aiming to portray the passage as navigable were based more on fantasy than fact, and the long-held notion that open ocean water could not freeze proved hard to reconcile with reality.

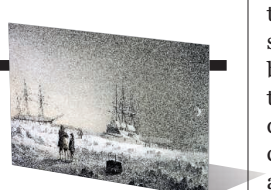
The most intriguing items in this part of the exhibition are a set of two wooden objects that look at first glance like abstract works of art but turn out to be three-dimensional maps of a coastline and a chain of islands. Given to explorers as mementos, the maps were created by the Inuit to be felt rather than seen. Rarely preserved, they are a fascinating reminder that maps can come in many forms, some of them ephemeral.

The reviewer is on staff at Science *magazine, AAAS Science International, Cambridge CB2 1LQ, UK. E-mail: jfahrenkamp@science-int.co.uk*

The central part of the exhibition focuses on the more familiar 19th-century expeditions. Following the Napoleonic Wars, the Royal Navy, with men and ships to spare, embarked on a concerted effort of Arctic exploration. The region also captured the popular imagination at this time and featured prominently in Mary Shelley's 1818 novel, *Frankenstein*. Exhibits here show the daily life of men overwintering. Far from the image of bravery and excitement, they can be seen playing cricket in the endless snow.

Human ingenuity was tested by the harsh and barren landscape and led to a number of marvelous inventions, including an inflatable dinghy that doubled as a cloak and umbrella. Other inventions included canned food, although, unfortunately for the explorers, the first can opener was not invented until decades later.

Danger was never far away, however. In 1845, Sir John Franklin, a Royal Navy officer and experienced explorer, set out to traverse the last unnavigated section of the Northwest Passage. The entire expedition was lost when the two ships became icebound in Victoria Strait. Although Inuit accounts were widely disbelieved and denounced—including, famously, by Charles Dickens—they turned out to be correct and ultimately aided in the recovery of one of Franklin's ships, the HMS Erebus, in September 2014. The Northwest Passage was finally navigated in 1906 by the Norwegian explorer Roald Amundsen, who relied on small teams and Inuit knowledge to complete the expedition.



Lines in the Ice
Seeking the Northwest Passage
Philip Hatfield and Tom Harper, Curators
British Library, London. 14 November 2014 to 29 March 2015. www.bl.uk/events/lines-in-the-ice-seeking-the-northwest-passage

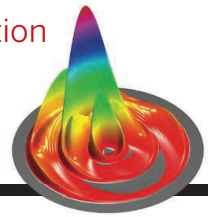
The exhibition places surprisingly little emphasis on climate change, providing just one set of maps that shows the melting sea ice cover in the Arctic. Yet a theme emerges from some of the most recent pieces, which illustrate Russia's use of seabed geology to claim Arctic territory and portray competing Canadian claims using sectors based on country borders. It becomes clear that as Arctic waters become increasingly navigable, the search for resources, this time largely fossil fuels, and for a northern shipping route between the Atlantic and Pacific Oceans continues to drive Arctic exploration, just as it did hundreds of years ago. Listening to recordings from the sound archive of whales and other marine animals that call the Arctic home, one cannot help but wonder how their already imperiled existence will be affected by this rush for Arctic riches.

10.1126/science.aaa3167

RESEARCH

Quick pauses in the action as Cl reacts with HD

Yang et al., *p. 60*



IN *SCIENCE* JOURNALS

Edited by **Melissa McCartney** *and* **Margaret Moerchen**

MUTAGENESIS

Men beware, when smoke gets in your Y's

The relationship between tobacco smoking and elevated cancer risk has been recognized for 60 years. Yet what smoking does to our genetic material is still not fully understood. New work suggests that men should be particularly concerned. In a study of over 6000 men, Dumanski *et al.* find that men who smoke are more than three times as likely as nonsmokers to show loss of the Y chromosome in their blood cells. Whether this is a causal factor in cancer development or simply a marker of more consequential damage on other chromosomes could not be deduced from the study. — PAK

Science, this issue p. 81

Y chromosomes are more susceptible to effects of smoking

REACTION DYNAMICS

A few very brief pauses in the action

Chemical reactions proceed by the cumulative effect of trillions upon trillions of collisions between atoms and molecules. Usually, a given collision bounces the participants right back out again, either in their original form or with the atoms shuffled around into distinct products. In certain cases, the reacting partners experience a brief lull, termed a resonance, before they rearrange. Yang *et al.* report the discovery of particularly short-lived resonances in certain reactive collisions of chlorine atoms with vibrationally excited hydrogen deuteride (HD). Their results suggest that similar, as yet overlooked, resonances may lurk in other reactions of vibrationally excited molecules. — JSY

Science, this issue p. 60

QUANTUM OPTICS

Engineering a shelter for quantum protection

In isolation, quantum states of matter can be stable entities. These states are often seen as useful when they can be made to interact in a controlled way. However, those interactions and the unavoidable interactions with their environment often correlate with decoherence and eventual loss of the quantum state. Kienzler *et al.* show that they can engineer the interactions between a quantum system (a trapped ion) and the environment to prepare stable quantum states. The generality of the technique implies applications for other interacting quantum systems. — ISO

Science, this issue p. 53

CANCER ETIOLOGY

Crunching the numbers to explain cancer

Why do some tissues give rise to cancer in humans a million times more frequently than others? Tomasetti and Vogelstein conclude that these differences can be explained by the number of stem cell divisions. By plotting the lifetime incidence of various cancers against the estimated number of normal stem cell divisions in the corresponding tissues over a lifetime, they found a strong correlation extending over five orders of magnitude. This suggests that random errors occurring during DNA replication in normal stem cells are a major contributing factor in cancer development. Remarkably, this "bad luck" component explains a far greater number of cancers than do hereditary and environmental factors. — PAK

Science, this issue p. 78

INNATE IMMUNITY

Skin infection triggers fat responses

Obesity is associated with chronic inflammation, but does fat tissue offer protection during infection? Zhang *et al.* noticed that the fat layers in the skin of mice thickened after inoculation with the pathogenic bacterium *Staphylococcus aureus* (see the Perspective by Alcorn and Kolls). Mutant mice incapable of forming new fat cells were more susceptible to infection. The differentiating fat cells secreted a small-molecule peptide called cathelicidin, specifically in response to the infection. By contrast, mature fat cells produce less cathelicidin, and are thus less protective. — CA

Science, this issue p. 67; see also p. 26

VIRUS STRUCTURE

Targeting EV-D68, a respiratory virus

A recent outbreak of respiratory illness in U.S. children was caused by entorovirus D68 (EV-D68). Enteroviruses also

include human pathogens such as human rhinovirus, which causes the common cold, and poliovirus. Most of these viruses are stabilized by a factor that binds in a hydrophobic pocket of the capsid protein VP1, and antiviral compounds can act by displacing this factor. Liu *et al*. report the crystal structure of EV-D68 and its complex with the antiviral compound peconaril. In EV-D68, the hydrophobic pocket contained a fatty acid that was displaced by peconaril. Peconaril efficiently inhibited EV-D68 infection of cells, making it a possible drug candidate against EV-D68. — VV

Science, this issue p. 71

ORGANIC CHEMISTRY

Breaking through the milligram floor

When chemists synthesize compounds, the threshold for success is at least a milligram of product. This has been true for decades—even though biochemical assays have long since descended into microgram territory—and results in part from the constraints of characterization methods. Buitrago Santanilla *et al*. present an automated dosing and characterization protocol for optimizing chemical reaction conditions on the microgram scale. This allowed them to screen numerous base and ligand combinations for catalytic C-N bond-forming reactions between complex pairs of compounds, in short supply, that resisted standard coupling conditions. — JSY

Science, this issue p. 49

IMMUNOLOGY

An immunological fountain of youth

mTOR signaling, a multipurpose pathway, controls all aspects of cell growth and motility and can also delay onset of aging-related diseases in many species. Mannick *et al.* now show that mTOR inhibition can benefit humans, too. They evaluated whether the mTOR inhibitor RAD001 could reverse the deterioration of immune function seen as people age. By assessing their elderly subjects' reaction to an influenza vaccination, the authors showed that RAD001 boosted their vaccine-induced immune defenses. — OMS

Sci. Transl. Med. **6**, 268ra179 (2014)

CANCER

Overcoming drug resistance in cancer

Cancer patients frequently develop drug resistance. Martz *et al.* devised a method of identifying pathways causing resistance in cancer cells and found that Notch signaling mediated resistance to drugs used in breast cancer and melanoma. Winter *et al*. also used this screening method for myeloproliferative neoplasms, which often have an activating mutation in the kinase JAK2 but are resistant to JAK inhibitors. They pinned the cause of this resistance to RAS, a signaling protein. Thus, screening entire signaling pathways instead of individual genes can identify new therapeutic targets that may be important in multiple types of drug-resistant cancers. — LKF

Sci. Signal. **7**, ra121 and ra122 (2014).

PHOTOS: (LEFT TO RIGHT) PHLOXII/ISTOCKPHOTO; © D. HURST/ALAMY

IN OTHER JOURNALS

Edited by **Kristen Mueller** *and* **Jesse Smith**

Avoiding nighttime eating may reduce the effects of an unhealthy diet

METABOLISM

You are not just what, but when you eat

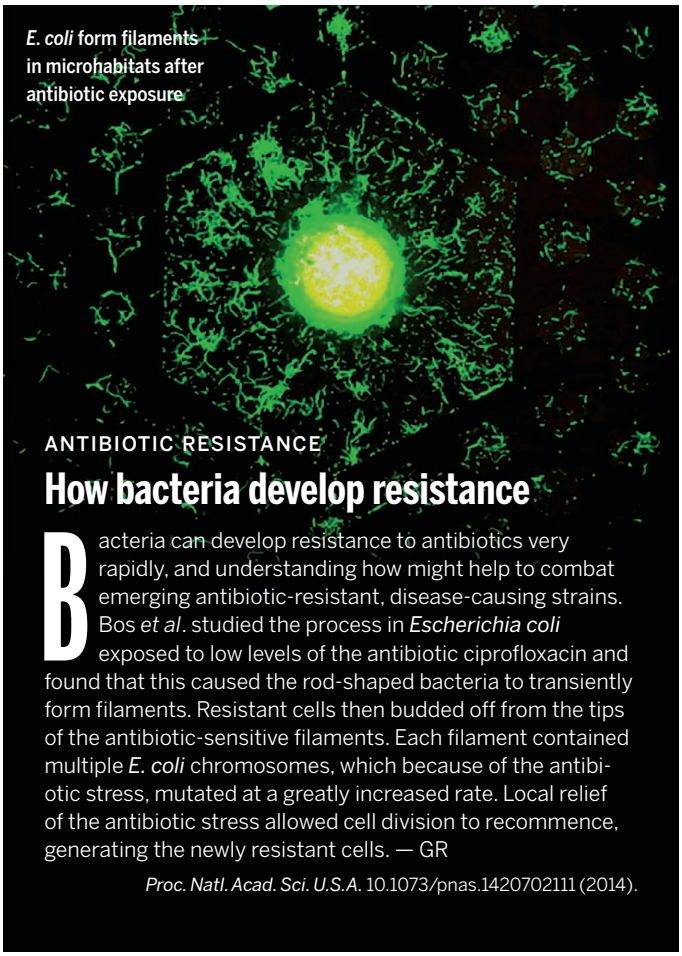
Limiting food intake to an 8-hour window that corresponds to a time of high activity protects mice from obesity and metabolic disease caused by a diet high in fat. Chaix *et al*. extended such studies to examine what would happen in a regimen more adaptable to peoples' lifestyles. Promisingly, they found protective effects from fasting periods as short as 12 hours. Even better, mice showed improved metabolic fitness even when they took the weekends off. This was most likely because the changes in gene expression caused by restricting food during the week continued even when mice had full access to food on the weekends. — LBR

Cell Metab. **20**, 991 (2014).

BIOGEOGRAPHY

Origins of the Southern Hemisphere flora

Which plants of the Southern Hemisphere represent descendants of the ancient flora of Gondwana, the southerly part of the supercontinent Pangaea? Evidence from molecular clocks suggests that many plant lineages descended from a more recent common ancestor. Wilf and Escapa challenge this view by comparing molecular origin dates with fossil dates for groups



E. coli form filaments in microhabitats after antibiotic exposure

ANTIBIOTIC RESISTANCE

How bacteria develop resistance

Bacteria can develop resistance to antibiotics very rapidly, and understanding how might help to combat emerging antibiotic-resistant, disease-causing strains. Bos *et al*. studied the process in *Escherichia coli* exposed to low levels of the antibiotic ciprofloxacin and found that this caused the rod-shaped bacteria to transiently form filaments. Resistant cells then budded off from the tips of the antibiotic-sensitive filaments. Each filament contained multiple *E. coli* chromosomes, which because of the antibiotic stress, mutated at a greatly increased rate. Local relief of the antibiotic stress allowed cell division to recommence, generating the newly resistant cells. — GR

Proc. Natl. Acad. Sci. U.S.A. 10.1073/pnas.1420702111 (2014).

of plants in Patagonia, such as cycads, that are incontrovertibly Gondwanan. They find that the molecular dates are artificially recent even for these ancient groups, and are a poor match for the reliable fossil dates. These findings suggest that biogeographers may have to temper reliance on molecular evidence for determining how flora dispersed. — AMS

New Phytol. 10.1111/nph.13114 (2014).

SURFACE CHEMISTRY

Assessing slow surface heating

Calorimetry can provide essential information about surface reactions, but the small amount of heat released from a surface makes experiments challenging. Pyroelectric detection can measure the heat produced when molecules are adsorbed by a thin metal surface, but calibration is simple only for rapid reactions (ones over in ~30 ms). For slower reactions, it is more difficult to assess how the experimental setup distorts the signal. Wolcott and Campbell report a fast Fourier transform method that can deconvolute the measured signal for slower reactions. They use it to reproduce successfully how heat is generated when methyl iodide is adsorbed by a platinum surface, as determined with a more complex modeling method. — PDS

Surf. Sci. 10.1016/j.susc.2014.11.005 (2014).

HEART DISEASE

Triglycerides, bedside to bench

Blood tests for heart disease risk measure cholesterol and fat molecules called triglycerides. Interest in triglycerides recently intensified with the discovery that people who carry mutations that disrupt the function of a glycoprotein called APOC3 (apolipoprotein C-III) have lower plasma triglyceride levels and a reduced risk of heart disease. Scientists thought APOC3 inhibited lipoprotein lipase, an enzyme catalyzing triglyceride breakdown. Gaudet *et al.* now reveal a more complicated mechanism. They found that three patients who had extremely high triglyceride levels because of a genetic deficiency in lipoprotein lipase nonetheless benefited from a drug that inhibits APOC3 synthesis. Clearly, APOC3 still has secrets to reveal. — PAK

N. Engl. J. Med. **371**, 2200 (2014).

ULTRAFAST OPTICS

Constraining the speed of tunneling

In contrast to expectation and experience in the classical world, a particle in the quantum world hitting a barrier can appear on the other side, having tunneled through the otherwise impassible obstruction. Debate about how quickly tunneling occurs has existed almost since the discovery of quantum mechanics. Now, Landsman *et al.* use ultrafast optical techniques to measure the amount of time it takes to ionize a helium atom. By following an electron on attosecond time scales as it passes through the energy barrier and leaves the bonds of its parent atom, the authors rule out certain theoretical possibilities, thereby providing a clearer picture of the tunneling process. — ISO

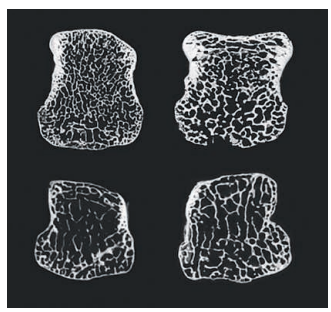
Optica **1**, 343 (2014); 10.1364/optica.1.000343 (2014).

PLANETARY ACCRETION

Tacking the mantle for planet formation

Earth's distance from the Sun and its mass help constrain planetary accretion models. Rubie *et al.* use the initial composition of the mantle as another constraint for accretion simulations, by considering the chemical effects of metallic core formation. The simulations that produce a realistic Earth require that impacting planetessimals span a range of compositions, metal-silicate element equilibration occurs at progressively greater depths as the planet grows, and only a small amount of interaction takes place between the metal impactor core and the proto-mantle. Adding terrestrial mantle composition to accretion simulations provides insights into the most important factors for planetary formation. — BG

Icarus 10.1016/j.icarus.2014.10.015 (2014).



CT scans of hand bones from (clockwise from top left): chimp, *Australopithecus*, Neandertal, modern human.

PALEOANTHROPOLOGY

Human skeleton became lighter over time

Chimp bones are packed with microscopic structures known as spongy bone; modern human bones aren't, increasing risk of fractures and osteoporosis. Two studies propose an explanation for this change: Chirchir *et al.* found that skeletons from modern chimpanzees, *Australopithecus africanus*, Neandertals, and early *Homo sapiens* all had higher densities of spongy bone than modern humans, suggesting that our sedentary lifestyle is to blame. Ryan and Shaw also found lower spongy bone density in the hip joints of ancient farmers compared with hips from nonhuman primates and ancient hunter-gatherers, supporting the idea that a lack of rigorous exercise, not evolutionary pressure, is responsible for our weak bones. — LW

Proc. Natl. Acad. Sci., 10.1073/pnas.1411696112(2014), 10.1073/pnas.1418646112(2014).

ALSO IN *SCIENCE* JOURNALS

Edited by **Melissa McCartney** *and* **Margaret Moerchen**

2D MATERIALS

Layered materials power the cause

Methods for storing and converting energy, including fuel cells, solar cells, and water splitting, often benefit from having materials with a large surface area. When combined with a high surface reactivity, high conductivity, or useful optical properties, two-dimensional layered materials become of notable interest for a range of applications. Bonaccorso *et al*. review the progress that has been made using graphene and other layered or two-dimensional materials at laboratory scales and the challenges in producing these materials in industrially relevant quantities. –MSL

Science, this issue p. 41

STRUCTURAL BIOLOGY

Energy conversion in complex 1

ATP, the energy source of the cell, is synthesized by a protein residing in the mitochondrial inner membrane. The synthesis is driven by a proton gradient generated by redox reactions that transfer electrons between a series of enzymes in the membrane. The largest complex in this electron transfer chain is the 1-MD complex 1. It couples electron transfer from NADH to ubiquinone to the translocation of four protons. Zickermann *et al*. report the crystal structure of a complex comprising the 14 central subunits and the largest accessory subunit of mitochondrial complex 1 from a yeast-genetic model at 3.6 Å resolution. The structure identifies four potential proton translocation pathways and gives insight into how energy from the redox reactions is transmitted to drive proton pumping. – VV

Science, this issue p. 44

ORGANIC CHEMISTRY

Shifting hydroformylation into reverse

The hydroformylation reaction is applied on large scale in the chemical industry to make aldehydes by adding hydrogen and carbon monoxide to olefins. The reverse process could also prove useful in modifying complex molecules for pharmaceutical research, but methods directed toward that end often strip off the CO without the hydrogen. Murphy *et al*. now show that a rhodium catalyst can achieve selective dehydroformylation of a diverse range of compounds under mild conditions (see the Perspective by Landis). The protocol relies on effective transfer of the CO and H_2 equivalents to a sacrificial strained olefin added to the mix. –JSY

Science, this issue p. 56; see also p. 29

PROTEIN SYNTHESIS

Tagging truncated proteins with CAT tails

During the translation of a messenger RNA (mRNA) into protein, ribosomes can sometimes stall. Truncated proteins thus formed can be toxic to the cell and must be destroyed. Shen *et al*. show that the proteins Ltn1p and Rqc2p, subunits of the ribosome quality control complex, bind to the stalled and partially disassembled ribosome. Ltn1p, a ubiquitin ligase, binds near the nascent polypeptide exit tunnel on the ribosome, well placed to tag the truncated protein for destruction. The Rqc2p protein interacts with the transfer RNA binding sites on the partial ribosome and recruits alanine- and threonine-bearing tRNAs. Rqc2p then catalyzes the addition of these amino acids onto the unfinished protein, in the absence of both the fully assembled ribosome and mRNA. These so-called CAT tails may promote the heat shock response, which helps buffer against malformed proteins. – GR

Science, this issue p. 75

LATERAL GENE TRANSFER

Killing, sex, and gene swaps in bacteria

The bacterial type VI secretion system (T6SS) is used by bacteria to inject toxins into neighboring cells to eliminate competition. This molecular machine is thus considered to be a mechanism by which bacteria can exert social control in complex microbial communities. Borgeaud *et al*. have discovered that in *Vibrio cholerae,* T6SS genes are co-regulated with genes involved in DNA uptake. Hence, T6SS-dependent killing of other bacteria is directed at neighboring cells, which release their DNA to be taken up by the killer, which can then integrate valuable genes and rapidly evolve, leading to antibiotic resistance or virulence. –CA

Science, this issue p. 63

AGING

Lysosomes signal the nucleus to control aging

Folick *et al.* propose a mechanism by which a lysosomal enzyme influences nuclear events that control longevity in the worm (see the Perspective by Shuo and Brunet). Increased expression of the lysosomal acid lipase LIPL-4 increased longevity, and this effect depended on the presence of the lysosomal lipid-binding protein LBP-8. LBP-8 acts as a chaperone that helps carry lipds to the nucleus. The authors identified the fatty acid oleoylethanolamide (OEA) as a potential signaling molecule whose transport to the nucleus could activate nuclear hormone receptors and transcription factors NHR-49 and NHR-80. The transcriptional targets of NHR-49 and NHR-80 in turn regulate longevity. – LBR

Science, this issue p. 83

MICROBIOLOGY

How bacterial populations get ready for stress

Populations of bacterial cells must be able to react to stressful situations, such as exposure to antibiotics, in order to survive. Holden explains that persisters, a very small subpopulation of bacterial cells that stop dividing, are key to this stress response. These cells are not genetically antibiotic-resistant but can nevertheless survive exposure to antibiotics before beginning to grow again. Single-cell analyses have provided insight into the molecular pathways through which bacterial cells enter the persister state. The results may help to develop successful treatments for recurrent infections. – JFU

Science, this issue p. 30

MOSQUITO GENOMICS

Mosquito adaptability across genomes

Virtually everyone has first-hand experience with mosquitoes. Few recognize the subtle biological distinctions among these blood-sucking flies that render some bites mere nuisances and others the initiation of a potentially life-threatening infection. By sequencing the genomes of several mosquitoes in depth, Neafsey *et al*. and Fontaine *et al*. reveal clues that explain the mystery of why only some species of one genus of mosquitoes are capable of transmitting human malaria (see the Perspective by Clark and Messer). – CA

Science, this issue p. 42 and p. 43; see also p. 27

REVIEW SUMMARY

2D MATERIALS

Graphene, related two-dimensional crystals, and hybrid systems for energy conversion and storage

Francesco Bonaccorso,* Luigi Colombo, Guihua Yu, Meryl Stoller, Valentina Tozzini, Andrea C. Ferrari, Rodney S. Ruoff, Vittorio Pellegrini

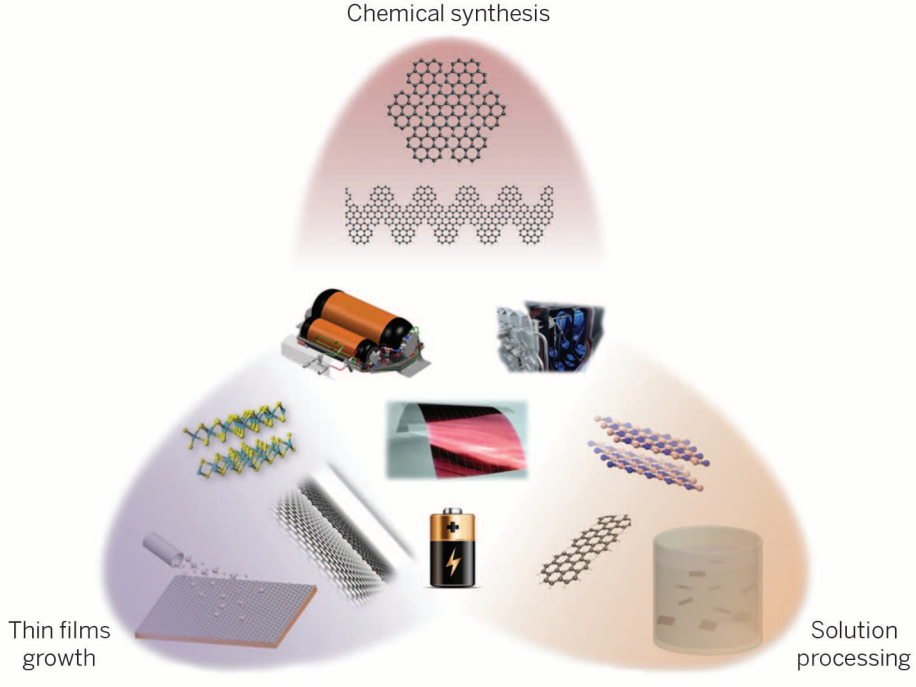
BACKGROUND: The integration of graphene in photovoltaic modules, fuel cells, batteries, supercapacitors, and devices for hydrogen generation offers opportunities to tackle challenges driven by the increasing global energy demand. Graphene's two-dimensional (2D) nature leads to a theoretical surface-to-mass ratio of ~2600 m^2/g, which combined with its high electrical conductivity and flexibility, gives it the potential to store electric charge, ions, or hydrogen. Other 2D crystals, such as transition metal chalcogenides (TMDs) and transition metal oxides, are also promising and are now gaining increasing attention for energy applications. The advantage of using such 2D crystals is linked to the possibility of creating and designing layered artificial structures with "on-demand" properties by means of spin-on processes, or layer-by-layer assembly. This approach exploits the availability of materials with metallic, semiconducting, and insulating properties.

ADVANCES: The success of graphene and related materials (GRMs) for energy applications crucially depends on the development and optimization of production methods. High-volume liquid-phase exfoliation is being developed for a wide variety of layered materials. This technique is being optimized to control the flake size and to increase the edge-to-surface ratio, which is crucial for optimizing electrode performance in fuel cells and batteries. Micro- or nanocrystal or flake edge control can also be achieved through chemical synthesis. This is an ideal route for functionalization, in order to improve storage capacity. Large-area growth via chemical vapor deposition (CVD) has been demonstrated, producing material with high structural and electronic quality for the preparation of transparent conducting electrodes for displays and touchscreens, and is being evaluated for photovoltaic applications. CVD growth of other multicomponent layered materials is less mature and needs further development. Although many transfer techniques have been developed successfully, further improvement of high-volume manufacturing and transfer processes for multilayered heterostructures is needed. In this context, layer-by-layer assembly may enable the realization of devices with on-demand properties for targeted applications, such as photovoltaic devices in which photon absorption in TMDs is combined with charge transport in graphene.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1246501

OUTLOOK: Substantial progress has been made on the preparation of GRMs at the laboratory level. However, cost-effective production of GRMs on an industrial scale is needed to create the future energy value chain. Applications that could benefit the most from GRMs include flexible electronics, batteries with efficient anodes and cathodes, supercapacitors with high energy density, and solar cells. The realization of GRMs with specific transport and insulating properties on demand is an important goal. Additional energy applications of GRMs comprise water splitting and hydrogen production. As an example, the edges of MoS_2 single layers can oxidize fuels—such as hydrogen, methanol, and ethanol—in fuel cells, and GRM membranes can be used in fuel cells to improve proton exchange. Functionalized graphene can be exploited for water splitting and hydrogen production. Flexible and wearable devices and membranes incorporating GRMs can also generate electricity from motion, as well as from water and gas flows. ■



GRMs for energy applications. The ability to produce GRMs with desired specific properties paves the way to their integration in a variety of energy devices. Solution processing and chemical vapor deposition are the ideal means to produce thin films that can be used as electrodes in energy devices (such as solar panels, batteries, fuel cells, or in hydrogen storage). Chemical synthesis is an attractive route to produce "active" elements in solar cells or thermoelectric devices.

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: francesco.bonaccorso@iit.it
Cite this article as F. Bonaccorso *et al.*, *Science* **347**, 1246501 (2015). DOI: 10.1126/science.1246501

REVIEW

2D MATERIALS

Graphene, related two-dimensional crystals, and hybrid systems for energy conversion and storage

Francesco Bonaccorso,[1,2]* Luigi Colombo,[3] Guihua Yu,[4] Meryl Stoller,[5] Valentina Tozzini,[6] Andrea C. Ferrari,[2] Rodney S. Ruoff,[7] Vittorio Pellegrini[1,6]

Graphene and related two-dimensional crystals and hybrid systems showcase several key properties that can address emerging energy needs, in particular for the ever growing market of portable and wearable energy conversion and storage devices. Graphene's flexibility, large surface area, and chemical stability, combined with its excellent electrical and thermal conductivity, make it promising as a catalyst in fuel and dye-sensitized solar cells. Chemically functionalized graphene can also improve storage and diffusion of ionic species and electric charge in batteries and supercapacitors. Two-dimensional crystals provide optoelectronic and photocatalytic properties complementing those of graphene, enabling the realization of ultrathin-film photovoltaic devices or systems for hydrogen production. Here, we review the use of graphene and related materials for energy conversion and storage, outlining the roadmap for future applications.

The development of reliable and environmentally friendly approaches for energy conversion and storage is one of the key challenges that our society is facing. Wearable energy conversion and storage devices require flexible, lightweight, conductive materials with a large surface-to-mass ratio [specific surface area (SSA)(m^2/g)] to allow storing and releasing of "particles" (such as lithium ions, hydrogen atoms or molecules, or electric charges).

A sheet of graphene has a theoretical SSA = 2630 m^2/g (*1*). This is much larger than that reported to date for carbon black [typically smaller than 900 m^2/g (*2*)] or for carbon nanotubes (CNTs), from ≈100 to 1000 m^2/g (*1*), and is similar to activated carbon (carbon processed with oxygen to make it porous) (*3*). The large SSA of graphene —when combined with its high electrical conductivity (*4*), high mechanical strength (*5*), ease of functionalization (*6*), and potential for mass production (*7*)—makes it an ideal platform for energy applications, such as a transparent conductive electrode for solar cells or as flexible high-capacity electrode in lithium-ion batteries and supercapacitors. Moreover, the combination of chemical functionalization and curvature control opens new opportunities for hydrogen storage (*8*, *9*).

Other two-dimensional (2D) crystals, such as the transition metal dichalcogenides (TMDs) (for example, WS_2, MoS_2, and WSe_2), display insulating, semiconducting (with band gaps in the visible region of the spectrum), and metallic behavior and can enable novel device architectures also in combination with graphene (*10*). As for the case of graphene, these materials can be integrated on flexible surfaces and can be mass-produced. Another class of 2D crystals is the MXenes (*11*, *12*), derived by exfoliating the so-called MAX phases: layered, hexagonal carbides and nitrides that can accommodate various ions and molecules between their layers by intercalation (*11*, *12*). MXene sheets are promising for energy applications, such as lithium-ion batteries (*11*), supercapacitors (*12*), and hydrogen storage (*13*).

Some 2D crystals are also promising for fuel cells and in water-splitting applications because of the large photocatalytic properties of their edges (*14*). The creation of hybrids with graphene and other nanomaterials, such as CNTs, can find applications in energy storage devices, such as supercapacitors (*15*), but also in photovoltaics. For simplicity, we will refer to graphene, other 2D crystals, and hybrid systems as graphene and related materials (GRMs) (*16*).

The challenge is to develop GRMs with properties tailored to create new devices that can be assembled for large-scale energy conversion (photovoltaics, thermoelectric, or fuel cells), and storage (supercapacitors, batteries, or hydrogen storage). This will require the production of high-quality material in high volumes by means of liquid-phase exfoliation (LPE)—for example, via ultrasonication (*7*, *17*) or shear mixing (*18*). Large-area GRMs grown by means of chemical vapor deposition can also play a role because they can have better morphological and optical/electric properties than those of LPE materials. Chemical synthesis (*19*) is also a possible route to tailor the shape of graphene flakes with atomic precision, but the scale-up remains challenging (*7*). A review of GRMs production is provided in (*7*).

Energy conversion in solar cells, thermoelectric devices, and fuel cells

In a photovoltaic (PV) device or solar cell, the incoming radiation creates electron-hole pairs in the active material. These are then separated and transported to electrodes [for example, Fig. 1A refers to a dye-sensitized solar cell (DSSC)] (*20*). Because graphene does not have a band gap and absorbs 2.3% of the incoming radiation almost independent of wavelength (*21*), it can capture a much broader spectrum than can semiconductors used today [for comparison, a silicon layer with the same thickness as graphene would absorb ~0.03% of the incident radiation at a wavelength of 500 nm (*22*)]. Graphene can perform different functions in inorganic and organic solar cells, such as transparent conductive electrodes (*23*, *24*) and counter-electrodes (*25*–*27*). Other layered materials (LMs) with a band gap in the visible region of the electromagnetic spectrum [(such as MoS_2 (Fig. 1B)] and chemically functionalized graphene can be used as photosensitizers, which transform absorbed photons into electrons (*20*). Graphene quantum dots (*28*) or graphene nanoribbons (GNRs) also enable a higher optical absorption close to their band gap (*28*).

In a thermoelectric device, a potential difference between electron and hole-doped crystals is created by a temperature gradient, as shown in Fig. 1C. GNRs or graphene with engineered defects can potentially improve the conversion efficiency (the ratio between the energy provided to the external load and the thermal energy absorbed) with respect to conventional thermoelectric materials based on PbTe or Bi_2Te_3 and their alloys (*29*), in addition to decreasing the environmental impact and cost.

Last, fuel cell devices in which electrical energy is generated by the conversion of chemical energy via redox reactions at the anode and cathode (Fig. 1D) (*30*, *31*), can also take advantage of GRMs as catalysts, so as to replace more expensive noble metals, such as platinum (*30*), with the added value of enabling more flexible and lighter devices.

Solar cells

The key figures of merit of solar cells are (*32*) the internal photocurrent efficiency, or the fraction of absorbed photons converted into electrical current; the external quantum efficiency, or the fraction of incident photons converted into electrical current; and the energy conversion efficiency $\eta = P_{max}/P_{inc}$, where P_{inc} is the incident power and $P_{max} = V_{OC} \times I_{SC} \times FF$, where V_{OC} is the maximum open-circuit voltage, I_{SC} is the maximum short-circuit current, and FF is the fill factor, defined as $(V_{max} \times I_{max})/(V_{OC} \times I_{SC})$, with V_{max} and I_{max} the maximum voltage and current, respectively (*32*).

[1]Istituto Italiano di Tecnologia, Graphene Labs, Via Morego 30, I-16163 Genova, Italy. [2]Cambridge Graphene Centre, University of Cambridge, Cambridge CB3 0FA, UK. [3]Texas Instruments, Dallas, TX 75243, USA. [4]Materials Science and Engineering Program and Department of Mechanical Engineering, University of Texas at Austin, Austin, TX 78712–0292, USA. [5]nCarbon, Austin, TX 78727, USA. [6]National Enterprise for nanoScience and nanoTechnology, Istituto Nanoscienze–CNR and Scuola Normale Superiore, I-56126 Pisa, Italy. [7]Center for Multidimensional Carbon Materials, Institute for Basic Science, Department of Chemistry, Ulsan National Institute of Science & Technology, UNIST-gil 50, Eonyang-eup, Ulju-gun, Ulsan 689-798, Republic of Korea.
*Corresponding author. E-mail: francesco.bonaccorso@iit.it

Silicon is the most widely used absorber to date (20) and currently dominates the PV device market. State-of-the-art silicon-PV devices based on p-n junctions, often referred to as first-generation solar cells (20), have an efficiency of up to~25% (32). The development of second-generation PVs, based on thin-film technologies, has been driven by the need to increase efficiency (32). However, to date the efficiency of second-generation PVs is below that of silicon (32). Third-generation PVs rely on the exploitation of emerging organic PV cells (33), DSSCs (Fig. 1A) (34), and quantum dots solar cells (QDSCs) (35), which may be less expensive, more versatile, and perhaps more environmentally friendly (34). However, they have lower efficiency [~12 and ~13% for organic PV cells (36) and DSSCs (27), respectively], low stability, and lower strength as compared with those of first- and second-generation PV cells. An important recent development is the meso-super-structured solar cell (37), based on an organic halide perovskite LMs (such as $CH_3NH_3PbX_3$, where X is chlorine, bromine iodine, or their combination) as photosensitizer (37, 38), and an organic hole-transport material (38). An efficiency of 15.6% in a meso-super-structured (perovskite) solar cell for an un-optimized device was reported in (38), whereas an efficiency of 20.1% has been recently developed at KRICT (Korean Research Institute of Chemical Technology) (39). However, these LMs may not satisfy sustainability requirements because of their lead content.

Driven by the need for new "environmentally friendly" materials, that can further improve efficiency and/or reduce cost of photovoltaic devices, GRMs are being developed as transparent conductors (TCs) (23, 24, 40), photosensitizers (10, 28), channels for charge transport (41, 42), and catalysts (25, 43). The use of GRMs for TCs to replace indium tin oxide (ITO), and catalysts to replace platinum, can improve the performance/cost ratio. For example, the use of graphene nanoplatelets as electro-catalysts for the polypyridine complexes of Co(III)/(II) in DSSCs allowed the achievement of the new record of efficiency of 13% (27) (Fig. 1A). The replacement of platinum, which is routinely used as an electro-catalyst in DSSCs, with GRMs may result in almost four orders of magnitude cost reduction. (The costs are based on Sigma-Aldrich values available at www.sigmaaldrich.com.)

Transparent conductive window

Transparent conducting films can act both as windows to the photosensitizer and as an ohmic contact (32). The key requirements for transparent conductive windows in PV systems are low sheet resistance [R_s<10 ohms per square (44)] and high transmittance (Tr) [>90% (44)]. R_s has units of ohms, as resistance does, but it is historically quoted in "ohms per square," which is defined as $R = R_s \times L/W$, where L/W is defined as the number of squares of side W that can be superimposed on the resistor without overlapping (21). The search for previously unidentified and less expensive conductive materials with good chemical stability, high Tr, and high electric conductivity is crucial for cost reduction. Today, the conductive support [such as ITO, or fluorine-doped tin oxide (FTO)] is one of the most expensive component of a DSSC (45). Beyond cost, the need for flexibility limits the use of current TC substrates. ITO and FTO are usually deposited at temperatures higher than the thermal stability of the polymeric substrates; additionally, their brittleness makes it difficult to use them when flexibility is a requirement (21).

Although the combined R_s and Tr targets have not been achieved yet, graphene-based TC windows were implemented in a variety of solar cell systems: inorganic (46), organic (23), DSSCs (40), and hybrid organic/inorganic (24). Given the continued progress in both quality [such as growth of graphene single crystals >1 cm (47)] and scalability [such as development of roll-to-roll (48) production lines], graphene-based TCs are an appealing alternative to ITO and FTO (21). R_s ~ 30 ohms per square and Tr ~ 90% were achieved via doping (49). TCs based on graphene doped with bis(trifluoromethanesulfonyl)-amide [$((CF_3SO_2)_2NH)$] resulted in graphene/n-silicon Schottky junction solar cells with a η = 8.6% (50), whereas TCs based on graphene doped with nitric acid used in an organic/silicon cell yielded η ~ 10.34% (24). Hybrid structures, such as graphene/metal grids (51), have also been considered. Graphene/metal grids were reported with R_s = 20 ohms per square and Tr = 90% (51). Transparent conductive windows based on other LMs (such as Bi_2Se_3) have been fabricated on mica with R_s = 330 ohms per square and Tr = 70% (52), which is still well below state-of-the-art graphene-based TCs (49, 51). More work, however, is needed for a conclusive assessment on their applicability as TCs.

Photosensitizers

The key requirements of a photosensitizer depend on the type of solar cell. In general, an efficient sensitizer should have the ability to absorb light over a wide energy range (32–34), high carrier mobility (32), and thermal and photochemical stability (33, 34). There are other more specific requirements for the various types of solar cells regarding, for example, the charge separation between donor/acceptor materials in organic PVs (33) and the efficiency of electronic injection from the photosensitizer into the TiO_2 in DSSCs (34).

Transition metal coordination compounds such as ruthenium complexes (53) and synthetic organic dyes (54) are used as sensitizers in DSSCs. However, the preparation routes for

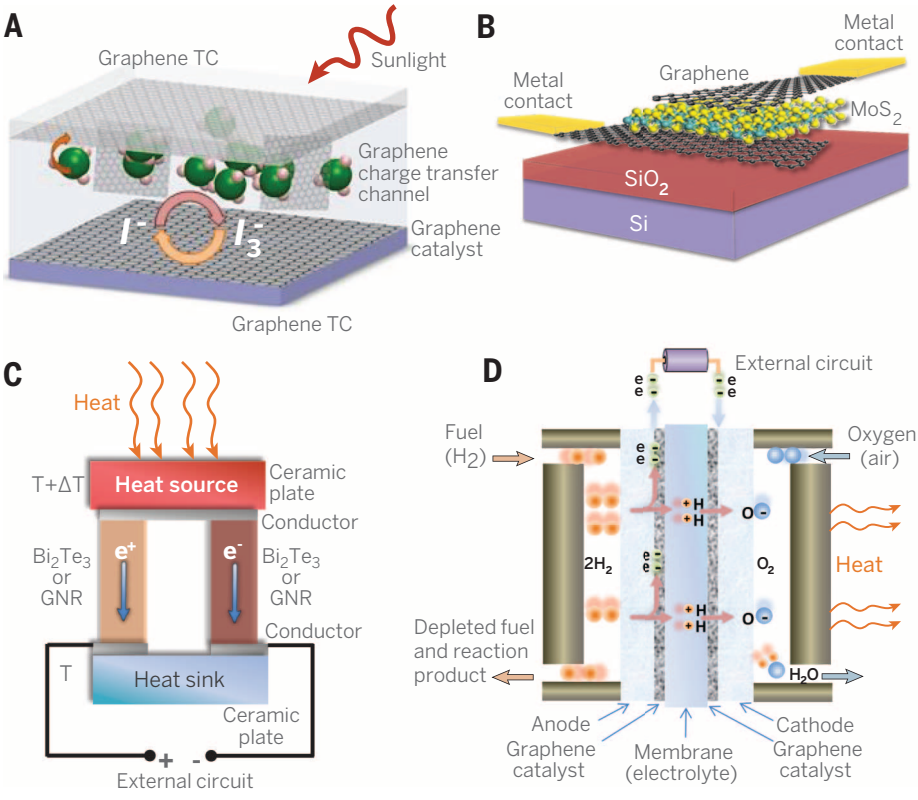


Fig. 1. Energy conversion devices. (**A**) Schematic of a dye-sensitized solar cell (DSSC) with graphene used in several components, as detailed in the text. (**B**) Heterostructure (graphene/MoS_2/graphene) photovoltaic device. (**C**) Schematic illustration of power generation in a thermoelectric device based on LM Bi_2Te_3 or GNRs. A temperature gradient, ΔT, causes charge carriers in the material to diffuse, resulting in current flow through the external circuit. (**D**) Proton exchange membrane fuel cells (31). Fuel (for example, H_2) channeled from one side of the cell is split by the catalyst (a GRM electrode) into H^+ ions and e^-. Electrons generate a current in the external circuit and then combine with H^+ and the oxidant (O_2) at the cathode, forming water and heat.

metal complexes are multi-step, involving long and expensive chromatographic purification, whereas organic dyes have a narrow spectral absorption and low electric conductivity. GRMs, on the other hand, have superior optoelectronic properties (*21*), which can be tuned by means of chemical synthesis (*19*, *28*) or post-production functionalization (*55*).

Chemically functionalized graphene (with organic molecules, conjugated polymers, rare-earth components, and inorganic semiconductors) (*55*), chemically synthesized GNRs (*56*), and quantum dots (*28*) have been used as photosensitizers (*28*, *55*). However, although graphene quantum dots have molar extinction coefficient [absorbance of light per unit path length (in centimeters) and per unit of concentration (moles per liter)] of ~1×10^5 M^{-1} cm^{-1} (*28*), which is about one order of magnitude larger than inorganic dyes (such as ruthenium complexes) (*28*), the energy conversion efficiency is still too low as a consequence of low current density (200 μA/cm^2) (*28*). This is due to low chemical affinity between the graphene quantum dots and the TiO_2 surface, resulting in poor charge injection (*28*). Calculations based on equivalent electric circuits for organic PVs (*56*) indicate that η ~ 12%, which is comparable with the state-of-the-art organic PVs (*36*), could be achieved with GNR photosensitizers (*56*).

TMDs are also potential photosensitizers (Fig. 1B) because of their large optical absorption [up to 10% or more of the incident light in a thickness of less than 1 nm (*57*), when in resonance], band gaps in the visible region, and chemical stability. For example, graphene/WS_2 vertical hybrid structures were studied for PV applications, with WS_2 acting as a photosensitizer (*10*). The van Hove singularities in the electronic density of states of WS_2 allowed large photon absorption and electron-hole creation with an external quantum efficiency (the ratio of the number of charge carriers collected by the solar cell to the number of photons) of ~33%. An alternative route is to combine metal nanoparticles with graphene, a method that can increase its light-harvesting capacity by more than one order of magnitude (*58*), making this hybrid structure a candidate photosensitizer.

Channel for charge transport

Charge-collection and transport are other important issues in PV devices. The transport of photo-generated electrons across the TiO_2 nanoparticle network (*34*) in DSSCs competes with charge recombination (*34*), a major bottleneck for increasing efficiency. In order to suppress charge recombination and increase photo-generated carriers, 1D materials such as CNTs can be used, but the cost and 1D nature (limiting the point contact between TiO_2 nanoparticles and CNTs) call for a better alternative. Graphene, with its high electron mobility (*4*), could be integrated with TiO_2 films to enhance the electron transfer properties of the photoanode.

Reduced graphene oxide [RGO (*7*), or rG-O (*59*)], was incorporated into nanostructured TiO_2 (*41*) and ZnO (*42*) photo-anodes in DSSCs in order to enhance the charge transport rate by preventing charge carrier recombination. RGO allows the use of thicker photo-anodes (*42*) [higher dye loading and consequently higher light harvesting (*34*)], thus improving efficiency with respect to conventional DSSCs (*41*). An energy conversion efficiency η = 5.86% was reported in (*42*) for a DSSC composed of a 9-μm-thick ZnO photoanode with 1.2 weight % RGO loading, which is higher than DSSCs with "conventional" photoanodes of the same thickness (*41*).

The electron collection layer is also important in perovskite solar cells (*37*), in which high-temperature sintered n-type TiO_2 electron-selective contacts are used (*38*), but this increases the cost and hinders the use of plastic substrates (*38*). Replacing the sintered TiO_2 should make perovskite solar cells a more versatile technology for inorganic PVs. Few-layers graphene (FLG) flakes, prepared via solution processing and incorporated in TiO_2 nanoparticles, were used as electron collection layer in perovskite-based solar cells (*60*), achieving η ~ 15.6% owing to the superior charge-collection of the FLG-TiO_2 composite, with respect to bare TiO_2 (η = 10%) (*60*). This η matches perovskite solar cells (*38*) and is the highest among graphene-based solar cells reported to date (Fig. 2).

Charge transport and collection also have a fundamental role in organic PVs (OPV). For example, in a poly-3-hexyl thiophene (P3HT)/phenyl-C61-butyric acid methyl ester (PCBM) solar cell, both the donor (P3HT) and acceptor (PCBM) materials are in direct electrical contact with the cathode (back electrode) and anode (ITO) electrodes, leading to carrier recombination (*61*). To reduce such a negative effect, electron blocking and hole transport layers are usually deposited on top of ITO (*61*). Currently, the most popular hole transport layers are wide–band gap p-type materials, such as NiO, MoO_3, V_2O_5, and poly(3,4-ethylenedioxythiophene) poly(styrenesulfonate) (PEDOT:PSS) (*33*, *61*). However, inorganic hole transport layers are deposited by vacuum techniques, incompatible with the roll-to-roll processes used in OPV, whereas PEDOT:PSS is usually deposited from highly acidic (pH = 1) aqueous dispersions (*61*). These corrode the ITO and can also introduce water into the active layer (processed in organic solvents), thus degrading device performance (*61*).

GRMs have been used as hole transport layers in OPV (*61*–*63*). OPV devices based on GO as the hole transport layer have shown η = 3.5%, which is comparable with devices fabricated with PEDOT:PSS (η = 3.6%) (*61*), whereas OPV exploiting RGO as hole transport layers were reported with η = 3.98%, which is superior to PEDOT:PSS (η = 3.6%) (*62*). Graphene quantum dots can also be efficient hole transport layers for OPVs, with η = 6.82% (*63*), showing longer lifetime and more reproducible performance with respect to PEDOT:PSS–based cells (*63*).

Counter-electrode

The role of the counter-electrode in a DSSC is twofold: It (i) back-transfers electrons arriving from the external circuit to the redox system and (ii) catalyzes the reduction of the charge mediator (*34*). The key requirements for counter-electrodes are high exchange current density (the rate of electron transfer between electrolyte and electrode), low charge-transfer resistance (the electrode-electrolyte interface resistance), and high SSA (*59*). Currently, DSSC counter-electrodes are made of platinum layers, which are expensive, deposited onto conductive electrodes, ITO or FTO (*34*). Moreover, platinum tends to degrade over

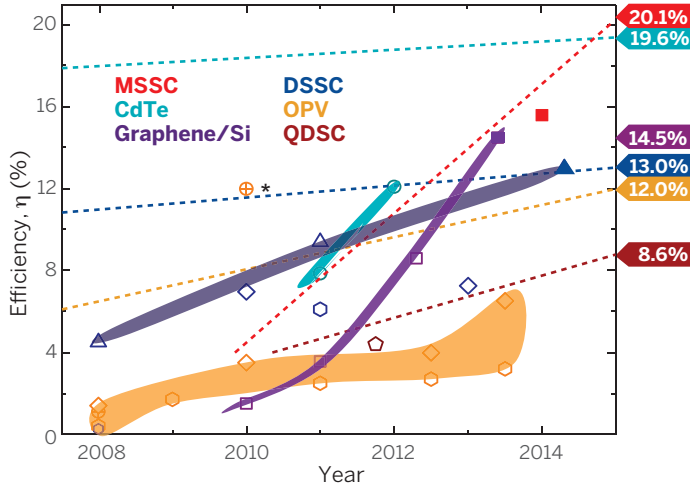


Fig. 2. Solar energy conversion efficiency of GRM-based photovoltaic devices. Colors define different classes of GRM-based photovoltaic devices: purple, graphene/silicon (*46*, *50*, *68*); orange, organic photovoltaics (OPVs) (*23*, *24*, *56*); dark blue, DSSCs (*25*, *41*–*43*); light blue, CdTe (*67*); dark red, QDSCs (*35*); light red, meso-super-structured solar cells (MSSCs) (*60*). Symbols are linked to different functions of the GRMs for each photovoltaic device: hexagons (*23*, *24*), transparent conductor; triangles (*25*–*27*), counter-electrode; rhombuses (*41*, *42*), charge transport. Orange closed areas cluster different GRM functions in OPVs. The data on the right axis refer to state-of-the-art PV efficiency (*32*), with the dashed lines representing the performance timeline of devices based on non-GRM materials [data taken from (*32*)]. The asterisk close to the crossed circle refers to a theoretical work for graphene nanoribbons as a photosensitizer in OPVs (*56*).

time when in contact with the tri-Iodide/Iodide (I_3^-/I^-) liquid electrolyte, reducing the efficiency. Thus, the replacement of such elements with lower cost and/or more reliable materials (leading to high-efficiency devices) is needed. Graphene can satisfy all the counter-electrode requirements because of its high SSA (*59*), which is essential to help the I_3^- reduction, high electric conductivity (*4*), low charge-transfer resistance (*25*), and lower cost than platinum.

Graphene oxide (*43*) and hybrid structures of RGO-CNTs (*64*) have been used as counter electrodes in DSSCs, with results close to the state of the art with platinum (Fig. 2). Graphene nanoplatelets (sheets of functionalized graphene with an overall thickness ranging from ~2 to ~15 nm) are now emerging as the best performing counter-electrodes in DSSCs, with (*27*) reporting the highest η to date of 13%.

The need to develop a platinum-free counter-electrode has seen a rising interest also in inorganic LMs such as transition metal oxide (TMO) and metal carbides, nitrides, and sulfides (*65*). Thin flakes of MoS_2 and WS_2 counter-electrodes were used in (*65*), with the I_3^-/I^- redox couple, achieving η = 7.59 and 7.73%, respectively, which is close to that of platinum counter-electrodes. In particular, platinum was outperformed as a counter-electrode by MoS_2 (4.97%) and WS_2 (5.24%) in DSSCs using an organic disulfide/thiolate (T_2/T^-) redox couple (*65*). Hybrid systems, such as graphene-MoS_2, were also used as counter-electrodes in DSSCs, achieving η = 5.81% (*66*). Thus, although to date η is lower than the best reported for platinum (12.3%) and graphene nanoplatelets (13%) (*27*), with further optimization 2D crystals (*65*) and hybrids (*66*) could play a key role as counter-electrodes in DSSCs.

Outlook

GRMs exploited as counter-electrodes in DSSCs (*65*, *66*) or in CdTe (*67*) solar cells show encouraging results. The efficiency of PV devices based on GRMs is progressing at a pace superior to those based on conventional materials (*32*). The highest η = 13% to date for DSSCs was recently achieved by using graphene nanoplatelets as a counter-electrode. Graphene/silicon hybrid solar cells, although first reported in 2010 (*46*), already have η = 14.5% (*68*), whereas graphene-based perovskite solar cells have η = 15.6% (*60*) for low-temperature (<150°C) processing, matching that reported for high-temperature (>500°C) cells (*38*), thus with an advantage in processing and cost reduction. In Fig. 2, we compare η of GRM and conventional non-GRM–based PV devices. The results to date could enable integration in existing devices with higher η and the development of new-concept devices, such as graphene/silicon solar cells.

Thermoelectric devices

About half of the energy generated worldwide is lost as waste heat (*69*, *70*). Thermoelectrics, solid-state devices (Fig. 1C) that generate electricity from a temperature gradient, are ideal to recover waste thermal energy (*69*, *70*). Thermoelectric devices can also convert heat produced by concentrated or unconcentrated sunlight, into electricity (*69*, *70*). This is important because infrared radiation with photon energies below the band gap of the photosensitizers is not absorbed in conventional PV cells and generates only waste heat (*69*, *70*).

In a typical thermoelectric device, a junction is formed between two different n- and p-doped conducting materials (Fig. 1C). A heat source at the junction causes carriers to flow away from it, resulting in a "thermo-electric" generator (exploiting the Seebeck effect, resulting in a voltage induced by a temperature gradient). In a thermoelectric device, many of these junctions are connected electrically in series and thermally in parallel. They can also work inversely, using electricity to generate or remove heat. When a current is passed in the appropriate direction through a junction, both types of charge carriers move from the junction and transport heat away, thus cooling the junction (Peltier effect). Thermoelectric devices are appealing, but their low efficiencies limit their widespread use.

The effectiveness of a thermoelectric device is assessed in two ways: by its Carnot efficiency (the fraction of absorbed heat that is converted into work) and by a material-dependent figure of merit, known as zT; $zT = TS^2\sigma/k$ (*69*), where S is the Seebeck coefficient, T is the temperature, σ is the electric conductivity, κ is the thermal conductivity, and $z = S^2\sigma/k$ (*69*). Thus, thermoelectric materials require high S and σ and low κ (*69*). In order to optimize zT, phonons must experience a high scattering rate, thus lowering thermal conductivity [like in a glass (*69*, *70*)], whereas electrons must experience very little scattering, maintaining high electric conductivity (as in a crystal) (*70*).

The majority of explored materials in thermoelectric devices have zT ~ 1 (*69*). LMs such as Bi_2Te_3, PbTe, and their alloys (*29*) and, in particular, the $(Bi_{1-x}Sb_x)_2(Se_{1-y}Te_y)_3$ alloy family have been in commercial use for several decades because of their room-temperature zT ~ 1 and Carnot conversion efficiencies ~5 to 6% (*69*). State-of-the-art thermoelectric materials design relies on engineering of the scattering mechanisms for phonons (*70*) and charge carriers (*70*). Currently, superlattices of Bi_2Te_3/Sb_2Te_3 (*71*) and quantum dots fabricated by means of atomic layer deposition designed to disrupt the phonon mean free path, while still allowing good electron mobilities, have the highest zT ~ 2.4 to 2.9 at 300 to 400 K (*71*) and 3.5 at 575 K, respectively (*72*).

Graphene has both high electric (*4*) and thermal (*73*) conductivity, a combination not ideal for thermoelectric devices. However, it is possible to tailor the thermal transport properties of graphene by nano-structuring techniques, such as defect (*74*) and isotope (*75*) engineering or edge roughness (*74*), or by introducing periodic nano-holes (*76*). The combination of geometrical structuring, GNRs with predefined geometries (*19*), and isotopic enrichment with ^{13}C (*75*) can reduce thermal conductivity by up to two orders of magnitude with respect to pristine graphene (*77*). It has been estimated that zT up to 3.25 can be achieved by exploiting GNRs that have a chevron-like geometry (*77*). However, scaling up of GNRs via chemical synthesis (*19*) still poses a challenge. Nevertheless, the modulation of geometric factors determining electric and thermal conductivity might be achieved via LPE. This technique also allows the blending of LMs with CNTs in order to increase the electrical conductivity while not reducing the Seebeck coefficient.

Advances in nanostructuring (*74*–*76*) to create hybrid structures on demand, with high electrical conductivity and low thermal conductivity, could accelerate the development of high-performance GRMs for thermoelectric devices.

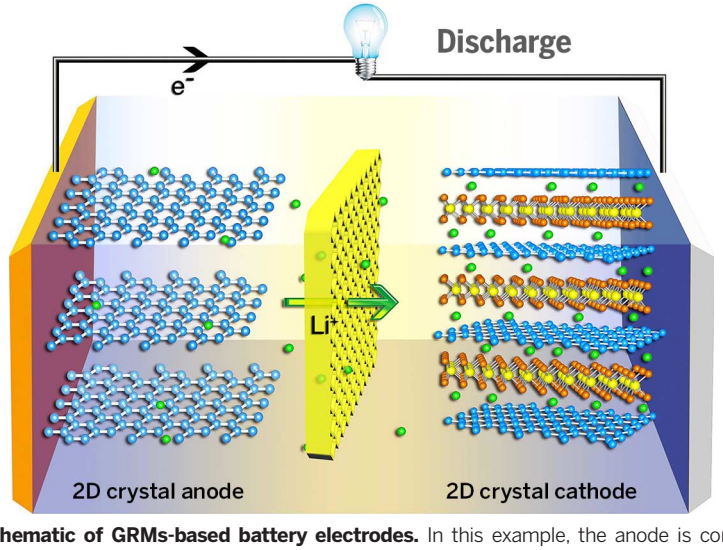


Fig. 3. Schematic of GRMs-based battery electrodes. In this example, the anode is composed of graphene flakes, but other 2D crystals can also be used, alone or in hybrid structures, as detailed in the text. The cathode is a hybrid graphene-lithium compound (such as $LiCoO_2$ or $LiFePO_4$), designed to enhance electron transport kinetics compared with graphene-free lithium compounds.

Fuel cells

Fuel cells convert chemical energy from a fuel into electricity via a reaction with oxygen or other oxidizing agents (Fig. 1D) (*31*). Their development goes hand-by-hand with hydrogen production and storage. In the next section, we will outline the use of GRMs for production and storage of hydrogen. Here, we focus on the possible use of GRMs in conversion of hydrogen into electrical energy in fuel cells.

The integration of fuel cells in electronics faces several challenges: (i) electrodes suitable for flexible electronics; (ii) replacement of expensive noble metals such as platinum, ruthenium, gold, and their alloys as electro-catalysts; and (iii) the need to avoid metal electrode poisoning (*31*). To address these challenges, a new class of materials with low cost, high efficiency (both for fuel oxidation at the anode and oxygen reduction reaction at the cathode), and durability have to be developed before fuel cells can be considered as a serious technology for energy conversion in electronic devices. GRMs are an ideal material platform to address these challenges and facilitate the transport of electrons produced during both fuel oxidation and oxygen reduction reaction (*30*). Moreover, GRMs have been demonstrated to be attractive candidates as proton membranes owing to their high proton conductivity (*78*). This, coupled with impermeability to water, H_2, and methanol, might solve the problem of fuel cross-over and electrode poisoning (*78*). Several GRMs are being investigated both as alternative to metal catalysts (*30*) or to be used in conjunction with platinum in hybrid structures (*79*). As demonstrated in (*79*), graphene-supported platinum and Pt–Ru nanoparticles have higher methanol and ethanol oxidations in comparison with those of the widely-used Vulcan XC-72R carbon black catalyst. As demonstrated in (*80*), RGO modifies the properties of platinum electro-catalysts supported on it. Platinum/graphene hybrid electrocatalysts were shown to have a high activity for methanol oxidation compared with commercial platinum/carbon black (*80*).

GRMs have also shown catalytic properties for oxygen reduction reactions at the cathode. It was reported in (*30*) that edge-halogenated graphene nanoplatelets, produced via high-volume scalable ball-milling, outperformed commercial platinum catalysts. LMs such as perovskites also have good catalytic activity at the cathode surface in solid-oxide fuel cells, thanks to their mixed ionic-electronic conductivity (*81*).

The possibility to tune the GRM lateral size and thickness (*7*), thus increasing their edge/bulk atoms ratio, can improve the catalytic activity because a higher number of active catalytic sites are present at the edges (*14*) for fuel oxidation at the anode and oxygen reduction at the cathode. This will be a step forward in making GRMs a key set of materials for the development of cheaper and more efficient fuel cells.

Energy storage

Current energy storage devices are based on the capture and release of lithium ions, electric charges, or hydrogen atoms or molecules. For example, lithium-ion batteries (*82*), now ubiquitous in portable electronics, consist of an intercalated lithium compound cathode, a graphitic anode, and an electrolyte. Crucial to the performance of these rechargeable batteries is the gravimetric capacity to store lithium ions (the charge stored per gram of battery weight). Compared with graphite, graphene and other related materials have a larger theoretical gravimetric capacity (*83*). Additionally, the use of graphene enables flexible and/or stretchable battery devices (*84*). Similar advantages also apply to electrochemical double-layer capacitors (EDLCs), which are currently formed via electrode/electrolyte systems based on two symmetric activated carbon electrodes impregnated with electrolytes (*85*). Other relevant properties of GRMs are the electrochemical and thermal stability within the device's operational temperature range (–50 to 100° C).

Batteries

Most of the commercial rechargeable lithium-ion batteries are based on a $LiCoO_2$ cathode and a graphite anode. This combination has a theoretical energy density of 387 Wh kg^{-1} (*86*) and a measured energy density of 120 to ~150 Wh kg^{-1} (*86*), which is higher than that of other batteries, such as lead acid [~30 Wh/kg (*87*)] and nickel metal hydrides (45 to 68 Wh/kg) (*87*). Potential oxide host structures, not yet commercialized, include ordered olivine $Li_{1-x}MPO_4$ (M = Fe, Mn, or Ni) (*88*), layered $Li_{1-x}MO_2$ (M = Ni, Mn, or Co) (*89*), and spinel $LiMn_2O_4$ (*90*).

GRMs can improve the gravimetric capacity and energy density compared with current technology owing to GRMs' high electrical conductivity (*4*), high SSA (*59*), large number of active sites for Li^+ storage, and short Li^+ diffusion distances (*91*). Indeed, GRMs are appealing both as cathodes (*92*) and anodes (Fig. 3) (*83*). Graphene, in particular, has a theoretical specific capacity [total ampere-hours (Ah) available when the battery is discharged at a certain discharge current, per unit weight] of 744 mAh g^{-1} assuming lithium adsorbed on both sides of graphene to form Li_2C_6 (*83*).

The discharge current is often expressed as a C-rate in order to normalize against battery-specific capacity. The C-rate is a measure of the rate at which a battery is discharged relative to its maximum capacity, hence the name (*82*). For example, at 1C the battery will discharge in 1 hour. A specific capacity of 540 mAh g^{-1} for RGO-based electrodes was reported in (*93*), and up to 730 and 784 mAh g^{-1} in RGO-CNT and RGO-C_{60} hybrid systems, respectively. Edges and defects could act as reversible lithium storage sites, thus contributing to the specific capacity (*83*). The importance of edges for lithium uptake has been demonstrated in (*91*), in which an anode containing <100-nm LPE flakes, deposited by inks, achieved a specific capacity of ~1500 mAh g^{-1} at a discharge current of 100 mAg^{-1}. The anode has also shown a specific capacity of 165 mAh g^{-1} at 1C when assembled in a full-battery configuration (*91*). Graphene-based hybrid electrodes (*94*), in which graphene is used as a substrate for electrochemically active nanoparticles (such as $Li_{1-x}MPO_4$ or $LiMn_{1-x}Fe_xPO_4$) have been exploited to increase electron transport, specific capacity, C-rate, and cyclability (the number of charge/discharge cycles before the battery-specific capacity falls below 60% of the nominal value). Graphene was also used as a substrate for the growth of anode/cathode nanomaterials [for example, olivine-type phosphates (*94*)] to achieve higher-rate-performance electrodes with respect to nonconducting materials (*94*). For example, $LiMn_{1-x}Fe_xPO_4$ nanorods grown on RGO flakes have shown only a 1.9% degradation for 100 cycles of the nominal capacity ~100 mAh g^{-1} at 50 C. This improved electrochemical performance, with respect to graphite or RGO, is attributed to Li^+ rapid diffusion along the radial direction of the nanocrystals, in addition to facilitated electron transport between RGO and nanocrystals. A similar approach to create GRMs-based hybrid electrodes was applied to different materials, such as other olivine-type phosphates (*95*), and spinels (*96*).

Another pathway to increase the charge/discharge capacity and C-rate of lithium-ion batteries is to confine the electrochemically active particles (such as sulfur, Co_3O_4, Fe_3O_4, or Li_3VO_4) within the graphene flakes (*97*). For example, in hybrid electrodes, graphene flakes enwrapping Co_3O_4 nanoparticles can suppress nanoparticle aggregation and accommodate their volume expansion/contraction upon lithiation/de-lithiation, in addition to ensuring high electrical conductivity (*90*). Thus, the specific capacity and cycling performance of hybrid RGO/Fe_3O_4 (*98*) or RGO/Li_3VO_4 (*99*), for example, improves as compared with electrodes made of nanoparticles alone (*98*, *99*).

The third approach for electrode optimization targets flexible and/or stretchable battery devices (*84*), which are able to accommodate large strain while retaining their function. GO flakes have been exploited to fabricate a flexible, layer-by-layer assembled conducting scaffold with tolerance to structural deformation (*100*). Such 3D flexible scaffolds loaded with silicon nanoparticles have shown a specific capacity of 1100 mAh g^{-1} at a discharge current of 8 A g^{-1}, degrading ~0.34% per cycle for 150 cycles (*100*).

TMDs, TMOs, and TMHs (transition metal hydroxides) are also promising for batteries (*101*). Some of the TMDs are accessible for lithium intercalation and exhibit fast ionic conductivity (*101*). Examples include TiS_2 as an electrochemically active material (*102*) and exfoliated MoS_2 flakes (*103*). A ~750mAh g^{-1} specific capacity was reported in (*103*) when using turbostratically restacked MoS_2 single layers as battery electrodes. The restacking enlarges the *c* axis parameter—the space between layers—thus increasing the accessible SSA (*103*). ZrS_2 colloidal nanodisks with diameters of ~20 nm delivered a specific capacity~600 mAh g^{-1} (*104*). Hybrid WS_2/RGO composites were used as electrodes, achieving ~450 mAh g^{-1} at 0.1 A g^{-1} and ~240 mAh g^{-1} at~4 A g^{-1}, respectively (*105*).

MXenes, such as Ti_2AlC, have shown lithiation/delithiation peaks at 1.6 and 2 V versus Li^+/Li (*11*). At 1C, the specific capacity was 110 mAh g^{-1} after 80 cycles. Compared with materials currently

used in lithium and sodium ion battery anodes, MXenes show promise in increasing overall battery performance (*11*).

Supercapacitors

Supercapacitors store energy using either ion adsorption (EDLC) (*85*) or redox reactions (*106*), in which most of the charge is transferred at or near the surface of the electrode material [pseudocapacitors (*106*)]. Supercapacitors are ideal for applications in which high power density [at least 10 kW Kg^{-1} (*85*, *106*), one order of magnitude larger than lithium-ion batteries] is needed, such as in the case of energy recapture and delivery in hybrid vehicles, mass transit, load cranes, load leveling, and backup power for electric utilities and factories (*106*).

EDLC supercapacitors

Almost all commercial EDLCs are based on two symmetric electrodes impregnated with electrolytes comprising tetraethylammonium tetrafluoroborate salts in organic solvents [acetonitrile (AN) and propylene carbonate (PC)] (*85*). Current commercial packaged EDLC supercapacitors, with organic electrolytes operating at 2.7 V, reach energy densities ~5 to 8 Wh/kg and 7 to 10 Wh/liter (*85*). Another type of supercapacitor based on lithium-ion hybrid cells [in which a graphite lithium-ion anode is coupled with an AC supercapacitor cathode (*107*)] is also offered commercially with energy densities of ~10 to 14 Wh/kg and 18 to 25 Wh/liter (*107*).

In an EDLC, energy is stored by forming an electrical double layer of electrolyte ions on the surface of conductive electrodes (Fig. 4A) (*108*). EDLCs are not limited by the electrochemical charge transfer kinetics of batteries and thus can operate at charge/discharge rates of the order of seconds, and with lifetimes of >1 million cycles (*106*). The EDLC energy density is determined by the square of the operating voltage (V_0) and the specific capacitance [capacitance per unit mass (farad per gram) or volume (farad per cubic centimeter)] of the electrode/electrolyte system: Wh/Kg = ξ(F/g) × $(V_0)^2$, where ξ is a constant (*85*). The specific capacitance in turn is related to the electrode's SSA accessible by the electrolyte, its interfacial double-layer capacitance (farad per square centimeter), and the electrode material density (*109*). The need to maintain electrochemical stability limits the operating voltage with organic electrolytes to ~2.7 V (*106*) because higher voltages result in electrolyte breakdown.

As in the case of batteries, electrode materials for EDLCs must be produced in tons and processed into electrodes 100 to 200 μm thick (*109*, *110*) to be commercially viable. Because the weight of the active electrode material when used as a thin coating is negligible compared with the support material, energy and power densities measured at the active material level do not translate to current commercial EDLC performances when scaled to full-size devices (*109*, *110*). Specific capacitances as high as ~190 F/g (*111*) in aqueous electrolytes and ~120 F/g in organic electrolytes were obtained with RGO produced by different routes (*112*), but all with SSAs <700 m^2/g, which is far short of the theoretical 2630 m^2/g (*59*). One approach to increase the SSA accessible to the electrolyte is the use of graphene-based platelets with spacer materials such as CNTs (*15*), mesoporous carbon spheres (*113*), water and ionic liquids (*114*), and resin that is subsequently chemically activated to create a porous structure (*115*). The reported SSAs range from 421 (*15*) to 1810 m^2/g (*115*). The activation of microwave expanded graphite oxide with KOH forms a porous material comprised of highly curved single-layer sheets of *n*-membered rings of carbon, with *n* varying between 5 and 8, and with measured SSA of 3100 m^2/g (*3*). Hydrothermal carbonization of either biomass or polymer mixed with dispersed GO, followed by chemical activation, yielded a 3D structure with nanoscale pores and a SSA of 3523m^2/g (*116*). However, although these values are higher than that of graphene, the measured SSA should be considered as an apparent or equivalent area because the Brunauer-Emmett-Teller (BET) method used for the determination of SSA is not applicable to microporous solids (*3*).

The interfacial capacitance of high-SSA GRMs comprises both the quantum capacitance (the response of the charge to the conduction and valence band movement, proportional to the electronic density of states) and the capacitance of the solid electrolyte interface (*107*). High SSA alone is, however, not sufficient to further increase performance of EDLC electrodes (*107*). High SSA and interfacial capacitance do not necessarily translate into high specific capacitance. A low packing density (<0.5 g/cm^3), for example, leads to empty space in the electrode that will be flooded by the electrolyte, increasing the cost and weight of the device, without adding capacity (*109*). Larger densities (~1.58 g/cm^3) were achieved by evaporation drying of graphene hydrogel, yielding specific capacitances of ~167 F/cm^3 in organic electrolyte (*117*). Capillary compression of RGO gave electrode densities of ~1.25 g/cm^3 and a specific capacitance of ~206 F/cm^3 in ionic liquids (*118*). Another method to increase energy storage capacity is to increase the operating voltage. To this end, graphene-based electrodes with ionic liquid electrolytes operating at voltages up to 3.5V and in a wide temperature range (−50 to 200°C) are currently being investigated (*119*, *120*).

Intercalation of cations (such as Na^+, K^+, Mg_2^+, NH_4^+, and Al_3^+) from aqueous salt solutions between Ti_3C_2 MXene was reported (*12*). A specific capacitance in excess of 300 F/cm^3, higher than in porous carbons, was reported in (*12*).

Hybrid and Pseudocapacitors

A different type of supercapacitor contains at least one electrode material with redox reactions that occur close to the electrode surface (pseudocapacitor) or a secondary battery electrode. Lithium-ion hybrid supercapacitors combine the rapid charge/discharge and long cycle life of an EDLC electrode with the higher energy storage capacity of a lithium-ion battery anode (*121*). However, the higher energy density currently comes with the trade-off of slower charge/discharge rates, lower efficiency, and reduced cycle life. Activated microwave expanded graphite oxide EDLC electrodes with lithium-ion battery electrodes comprising graphite (*122*), $Li_4T_5O_{12}$ (*122*), and Fe_3O_4

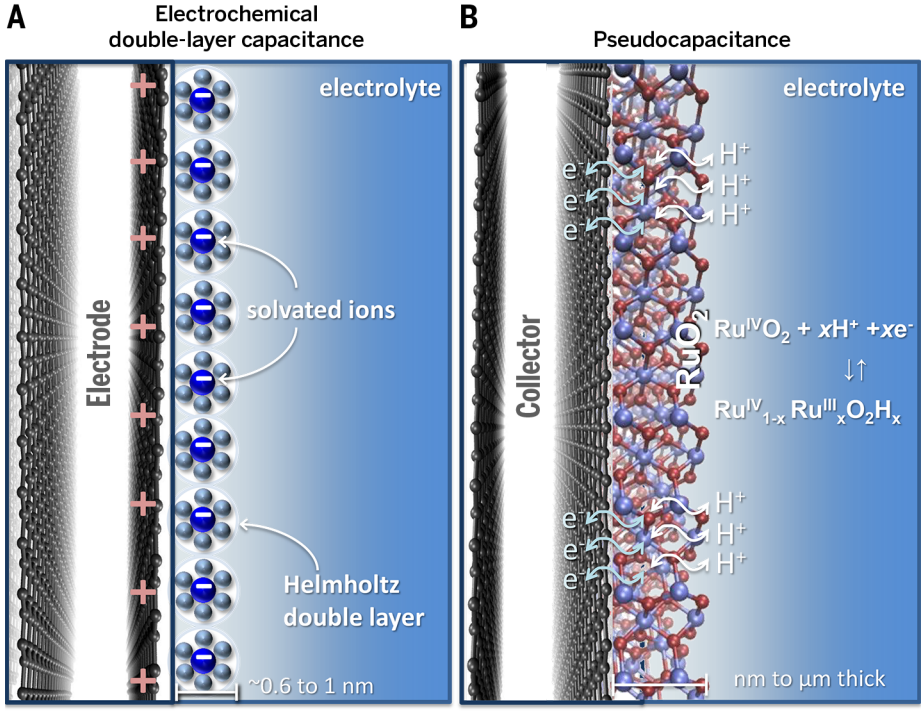


Fig. 4. Schematic of charge storage in supercapacitors. (**A**) Ion adsorption at the electrode surface (EDLC). (**B**) Charge transfer near the surface of the electrode (pseudo-capacitance).

(123) were studied. Also, electrodes containing metal oxides such as RuO_2, MnO_2, MoO_3 and conducting polymers were used to increase the specific capacitance via redox reactions (Fig. 4B) (108). In these systems, graphene was used as conductive support for composites with MnO_2 (124) and with conductive polymers such as polyaniline (125).

Hydrogen production and storage

The chemical energy density of hydrogen is 142 MJ/kg, which is more than three times that of gasoline. The by-product of its combustion is water. Thus, hydrogen is one of the most interesting "green" fuels. GRMs can help address the two main issues related to the use of hydrogen as fuel: (i) production and (ii) storage/transportation.

The key mechanism for the production of hydrogen gas is the hydrogen evolution reaction (HER; $2H^+ + e^- \rightarrow H_2$). The edges of 2D crystals, such as MoS_2 and WS_2 (126), and hybrid systems (such as MoS_2/RGO) (127) are active catalytic sites, making them promising electrodes for the HER (126, 127). However, the HER mechanism varies from material to material (126); thus, its understanding is fundamental for the optimized use of GRMs (126, 128). Resistive losses are one of the key problems for nonmetallic electrodes (31). Thus, combinations of 2D crystals with CNTs (128) and graphene (126) are being explored to further enhance HER by improving the electron transport efficiency.

Storage is also a challenge in hydrogen technology. The approach based on liquefying and pressurizing hydrogen presents safety issues. Solid-state storage is thus being investigated as an alternative.

Carbon-based structures are particularly attractive for hydrogen storage because carbon is a light element, and graphene in particular has potentially the most favorable gravimetric density among the carbon-based materials [the weight percentage of stored hydrogen (Fig. 5)] (129). The storage of molecular hydrogen in graphene relies on the van der Waals forces (binding energy of ~0.01 to 0.06 eV/molecule (129), leading to theoretical estimates of gravimetric density of ~3.3% (129), which is increased up to 8% in multi-layers spaced by pillar structures or CNTs (130) at cryogenic temperatures and/or high pressure. The estimated range of gravimetric density at ambient conditions is 1 to 4% (130). Experimental data are lower, in the range of 2 to 6% for low temperature and/or high pressures and ~1% at ambient conditions (Fig. 5, orange band) (131). Semiconductor LMs, such as InSe and GaSe (132), are reported to have gravimetric densities of up to 3 to 4%, obtained by a combination of electrochemical and thermal treatments (Fig. 5, blue and purple bands). In this case, hydrogen intercalates between the layers (132).

Decorating graphene with alkaline (133) or transition metals (133) can increase the hydrogen adsorption energy, leading to a theoretical gravimetric density up to 10%. Stronger binding is obtained by chemisorption (Fig. 5, gray band) leading to the formation graphane (8, 134). The use of chemisorption as a storage mechanism requires overcoming the high H_2 chemi(de)sorption barriers [~1.5 eV/atom (8)] to achieve loading/release kinetics at room temperature. Possible catalytic strategies for hydrogen adsorption/desorption involve the functionalization of graphene with metals such as palladium (135), known to catalyze the dissociation of hydrogen molecules into ions onto the graphene surface, or the combined effect of nitrogen-substitutional doping and an electric field normal to the sheet, which is predicted to produce dissociation-adsorption of H_2 (136).

The peculiar structural and mechanical properties of graphene enable alternative strategies for adsorption/desorption. It was theoretically (137) and experimentally shown (134) that the hydrogen affinity is enhanced on the convex areas and reduced on the concave areas of rippled graphene. The possibility of controlling load and release by modifying the local curvature was predicted with density functional theory (Fig. 5, structures at top) (8). To this end, mechanisms to control the curvature of graphene should be identified: the use of transverse acoustic phonons generated by a piezoelectric substrate (9) or inducing piezoelectricity within graphene by means of specific doping or decoration (138) were suggested.

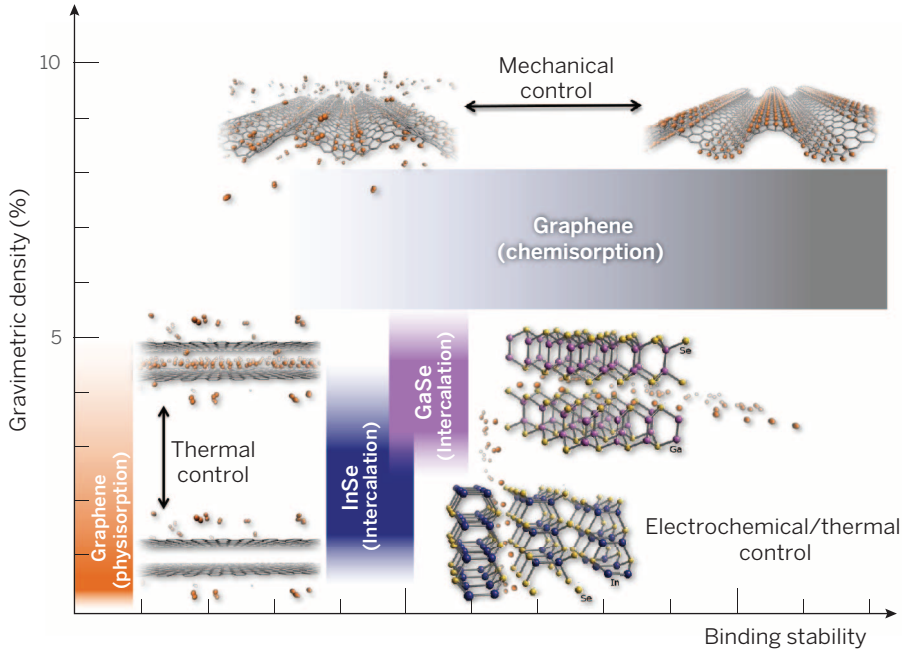


Fig. 5. Hydrogen storage in GRMs. The bands indicate the gravimetric density and binding stability (the log of the desorption energy) of different GRMs and for different adsorption mechanisms. In the case of graphene, two different storage procedures are illustrated, based on physisorption or chemisorption, with uptake/release controlled thermally and mechanically, respectively.

Perspective

Graphene, related 2D crystals, and hybrid systems might play a major role in future energy conversion and storage technologies. The ability to produce these GRMs, and control their properties, might enable a range of device characteristics, with optimized energy/power densities, lifetime, safety, and potentially reducing cost while minimizing environmental impact. To be commercially viable, GRMs must substantially surpass the performance of existing materials at comparable manufacturing costs. For example, GRMs have been reported with specific capacitances of ~300 F/cm^3, which is much higher than chemically activated state-of-the-art carbons. The ability to create stacked hetero-structures of metallic, semiconducting, and insulating 2D crystals might enable an even broader spectrum of device structures, perhaps with tunable properties. This might enable bulk thermoelectric materials with on-demand band structures and transport properties, or photosensitizers with broadband photon absorption. Owing to the GRMs intrinsic flexibility, we also envision applications such as wearable energy devices and energy harvesting from water or gas flows. Additionally, because GRMs can perform different functions, they may enable the realization of affordable energy systems with integrated conversion, storage, and sensing modules. In the future, it might be possible to target flexible photovoltaic cells with efficiencies of 12% and cost of ~0.5€/Wpeak (peak power output), fuel cells with 10 kW per gram of platinum, and energy storage devices with an energy density of at least 250 Wh/kg and cyclability up to 5000 cycles for batteries and a power density of 100kW/kg for supercapacitors. For hydrogen storage, the challenge is to achieve a gravimetric storage of 5.5%.

REFERENCES AND NOTES

1. A. Peigney, Ch. Laurent, E. Flahaut, R. R. Bacsa, A. Rousset, Specific surface area of carbon nanotubes and bundles of carbon nanotubes. *Carbon* **39**, 507–514 (2001). doi: 10.1016/S0008-6223(00)00155-X

2. J. B. Donnet, R. C. Bansal, M.-J. Wang, *Carbon Black Science and Technology* (Marcel Dekker, New York, 1993).
3. Y. Zhu *et al*., Carbon-based supercapacitors produced by activation of graphene. *Science* **332**, 1537–1541 (2011). doi: 10.1126/science.1200770; pmid: 21566159
4. A. K. Geim, K. S. Novoselov, The rise of graphene. *Nat. Mater.* **6**, 183–191 (2007). doi: 10.1038/nmat1849; pmid: 17330084
5. C. Lee, X. Wei, J. W. Kysar, J. Hone, Measurement of the elastic properties and intrinsic strength of monolayer graphene. *Science* **321**, 385–388 (2008). doi: 10.1126/science.1157996; pmid: 18635798
6. S. Park, R. S. Ruoff, Chemical methods for the production of graphenes. *Nat. Nanotechnol.* **4**, 217–224 (2009). doi: 10.1038/nnano.2009.58; pmid: 19350030
7. F. Bonaccorso *et al*., Production and processing of graphene and 2d crystals. *Mater. Today* **15**, 564–589 (2012). doi: 10.1016/S1369-7021(13)70014-2
8. D. C. Elias *et al*., Control of graphene's properties by reversible hydrogenation: Evidence for graphane. *Science* **323**, 610–613 (2009). doi: 10.1126/science.1167130; pmid: 19179524
9. V. Tozzini, V. Pellegrini, Reversible hydrogen storage by controlled buckling of graphene layers. *J. Phys. Chem. C* **115**, 25523–25528 (2011). doi: 10.1021/jp208262r
10. L. Britnell *et al*., Strong light-matter interactions in heterostructures of atomically thin films. *Science* **340**, 1311–1314 (2013). doi: 10.1126/science.1235547; pmid: 23641062
11. M. Naguib *et al*., MXene: A promising transition metal carbide anode for lithium-ion batteries. *Electrochem. Commun.* **16**, 61–64 (2012). doi: 10.1016/j.elecom.2012.01.002
12. M. R. Lukatskaya *et al*., Cation intercalation and high volumetric capacitance of two-dimensional titanium carbide. *Science* **341**, 1502–1505 (2013). doi: 10.1126/science.1241488; pmid: 24072919
13. Q. Hu *et al*., MXene: A new family of promising hydrogen storage medium. *J. Phys. Chem. A* **117**, 14253–14260 (2013). doi: 10.1021/jp409585v; pmid: 24261885
14. Q. Xiang, J. Yu, M. Jaroniec, Synergetic effect of MoS_2 and graphene as cocatalysts for enhanced photocatalytic H_2 production activity of TiO_2 nanoparticles. *J. Am. Chem. Soc.* **134**, 6575–6578 (2012). doi: 10.1021/ja302846n; pmid: 22458309
15. Q. Cheng *et al*., Graphene and carbon nanotube composite electrodes for supercapacitors with ultra-high energy density. *Phys. Chem. Chem. Phys.* **13**, 17615–17624 (2011). doi: 10.1039/c1cp21910c; pmid: 21887427
16. A. C. Ferrari *et al*., Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems. *Nanoscale* 10.1039/C4NR01600A (2014). doi: 10.1039/C4NR01600A
17. Y. Hernandez *et al*., High-yield production of graphene by liquid-phase exfoliation of graphite. *Nat. Nanotechnol.* **3**, 563–568 (2008). doi: 10.1038/nnano.2008.215; pmid: 18772919
18. K. R. Paton *et al*., Scalable production of large quantities of defect-free few-layer graphene by shear exfoliation in liquids. *Nat. Mater.* **13**, 624–630 (2014). doi: 10.1038/nmat3944; pmid: 24747780
19. J. Cai *et al*., Atomically precise bottom-up fabrication of graphene nanoribbons. *Nature* **466**, 470–473 (2010). doi: 10.1038/nature09211; pmid: 20651687
20. D. Chapin, C. Fuller, G. Pearson, A new silicon p-n junction photocell for converting solar radiation into electrical power. *J. Appl. Phys.* **25**, 676 (1954). doi: 10.1063/1.1721711
21. F. Bonaccorso, Z. Sun, T. Hasan, A. C. Ferrari, Graphene photonics and optoelectronics. *Nat. Photonics* **4**, 611–622 (2010). doi: 10.1038/nphoton.2010.186
22. M. A. Green, M. J. Keevers, Optical properties of intrinsic silicon at 300 K. *Prog. Photovolt. Res. Appl.* **3**, 189–192 (1995). doi: 10.1002/pip.4670030303
23. L. Gomez De Arco *et al*., Continuous, highly flexible, and transparent graphene films by chemical vapor deposition for organic photovoltaics. *ACS Nano* **4**, 2865–2873 (2010). doi: 10.1021/nn901587x; pmid: 20394355
24. Y. Wu *et al*., Graphene transparent conductive electrodes for highly efficient silicon nanostructures-based hybrid heterojunction solar cells. *J. Phys. Chem. C* **117**, 11968–11976 (2013). doi: 10.1021/jp402529c
25. W. Hong, Y. Xu, G. Lu, C. Li, G. Shi, Transparent graphene/PEDOT–PSS composite films as counter electrodes of dye-sensitized solar cells. *Electrochem. Commun.* **10**, 1555–1558 (2008). doi: 10.1016/j.elecom.2008.08.007
26. L. Kavan, J.-H. Yum, M. Grätzel, Graphene nanoplatelets outperforming platinum as the electrocatalyst in co-bipyridine-mediated dye-sensitized solar cells. *Nano Lett.* **11**, 5501–5506 (2011). doi: 10.1021/nl203329c; pmid: 22103554
27. S. Mathew *et al*., Dye-sensitized solar cells with 13% efficiency achieved through the molecular engineering of porphyrin sensitizers. *Nat. Chem.* **6**, 242–247 (2014). doi: 10.1038/nchem.1861; pmid: 24557140
28. X. Yan, X. Cui, B. Li, L. S. Li, Large, solution-processable graphene quantum dots as light absorbers for photovoltaics. *Nano Lett.* **10**, 1869–1873 (2010). doi: 10.1021/nl101060h; pmid: 20377198
29. B. Poudel *et al*., High-thermoelectric performance of nanostructured bismuth antimony telluride bulk alloys. *Science* **320**, 634–638 (2008). doi: 10.1126/science.1156446; pmid: 18356488
30. I.-Y. Jeon *et al*., Facile, scalable synthesis of edge-halogenated graphene nanoplatelets as efficient metal-free eletrocatalysts for oxygen reduction reaction. *Sci. Rep.* **3**, 1810 (2013). doi: 10.1038/srep01810; pmid: 23736800
31. H. A. Liebhafsky, E. J. Cairns, *Fuel Cells and Fuel Batteries: A Guide to Their Research and Development* (Wiley, New York, 1969).
32. M. A. Green, K. Emery, Y. Hishikawa, W. Warta, E. D. Dunlop, Solar cell efficiency tables (version 42). *Prog. Photovolt. Res. Appl.* **21**, 827 (2013). doi: 10.1002/pip.2352
33. H. Hoppe, N. S. Sariciftci, Organic solar cells: An overview. *J. Mater. Res.* **19**, 1925 (2004). doi: 10.1557/JMR.2004.0252
34. B. O'Regan, M. Grätzel, A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO_2 films. *Nature* **353**, 737–740 (1991). doi: 10.1038/353737a0
35. J. G. Radich, R. Dwyer, P. V. Kamat, Cu2S reduced graphene oxide composite for high-efficiency quantum dot solar cells. Overcoming the redox limitations of S_2^{-}/S_n^{2-} at the counter electrode. *J. Phys. Chem. Lett* **2**, 2453–2460 (2011). doi: 10.1021/jz201064k
36. www.heliatek.com/newscenter/latest_news/neuer-weltrekord-fur-organische-solarzellen-heliatek-behauptet-sich-mit-12-zelleffizienz-als-technologiefuhrer/?lang=en.
37. A. Kojima, K. Teshima, Y. Shirai, T. Miyasaka, Organometal halide perovskites as visible-light sensitizers for photovoltaic cells. *J. Am. Chem. Soc.* **131**, 6050–6051 (2009). doi: 10.1021/ja809598r; pmid: 19366264
38. M. Liu, M. B. Johnston, H. J. Snaith, Efficient planar heterojunction perovskite solar cells by vapour deposition. *Nature* **501**, 395–398 (2013). doi: 10.1038/nature12509; pmid: 24025775
39. www.nrel.gov/ncpv/images/efficiency_chart.jpg.
40. X. Wang, L. Zhi, K. Müllen, Transparent, conductive graphene electrodes for dye-sensitized solar cells. *Nano Lett.* **8**, 323–327 (2008). doi: 10.1021/nl072838r; pmid: 18069877
41. N. Yang, J. Zhai, D. Wang, Y. Chen, L. Jiang, Two-dimensional graphene bridges enhanced photoinduced charge transport in dye-sensitized solar cells. *ACS Nano* **4**, 887–894 (2010). doi: 10.1021/nn901660v; pmid: 20088539
42. F. Xu *et al*., Graphene scaffolds enhanced photogenerated electron transport in ZnO photoanodes for high-efficiency dye-sensitized solar cells. *J. Phys. Chem. C* **117**, 8619–8627 (2013). doi: 10.1021/jp312379b
43. J. D. Roy-Mayhew, D. J. Bozym, C. Punckt, I. A. Aksay, Functionalized graphene as a catalytic counter electrode in dye-sensitized solar cells. *ACS Nano* **4**, 6203–6211 (2010). doi: 10.1021/nn1016428; pmid: 20939517
44. M. W. Rowell, M. D. McGehee, Transparent electrode requirements for thin film solar cell modules. *Energy Environ. Sci* **4**, 131 (2011). doi: 10.1039/c0ee00373e
45. J. Kalowekamo, E. Baker, Estimating the manufacturing cost of purely organic solar cells. *Sol. Energy* **83**, 1224–1231 (2009). doi: 10.1016/j.solener.2009.02.003
46. X. Li *et al*., Graphene-on-silicon Schottky junction solar cells. *Adv. Mater.* **22**, 2743–2748 (2010). doi: 10.1002/adma.200904383; pmid: 20379996
47. Y. Hao *et al*., The role of surface oxygen in the growth of large single-crystal graphene on copper. *Science* **342**, 720–723 (2013). doi: 10.1126/science.1243879; pmid: 24158906
48. T. Kobayashi *et al*., Production of a 100-m-long high-quality graphene transparent conductive film by roll-to-roll chemical vapor deposition and transfer process. *Appl. Phys. Lett.* **102**, 023112 (2013). doi: 10.1063/1.4776707
49. S. Bae *et al*., Roll-to-roll production of 30-inch graphene films for transparent electrodes. *Nat. Nanotechnol.* **5**, 574–578 (2010). doi: 10.1038/nnano.2010.132; pmid: 20562870
50. X. Miao *et al*., High efficiency graphene solar cells by chemical doping. *Nano Lett.* **12**, 2745–2750 (2012). doi: 10.1021/nl204414u; pmid: 22554195
51. Y. Zhu, Z. Sun, Z. Yan, Z. Jin, J. M. Tour, Rational design of hybrid graphene films for high-performance transparent electrodes. *ACS Nano* **5**, 6472–6479 (2011). doi: 10.1021/nn201696g; pmid: 21774533
52. H. Peng *et al*., Topological insulator nanostructures for near-infrared transparent flexible electrodes. *Nat. Chem.* **4**, 281–286 (2012). doi: 10.1038/nchem.1277; pmid: 22437712
53. M. K. Nazeeruddin *et al*., Conversion of light to electricity by cis-X2bis(2,2'-bipyridyl-4,4'-dicarboxylate)ruthenium(II) charge-transfer sensitizers (X = Cl^-, Br^-, I^-, CN^-, and SCN^-) on nanocrystalline titanium dioxide electrodes. *J. Am. Chem. Soc.* **115**, 6382–6390 (1993). doi: 10.1021/ja00067a063
54. J. H. Yum *et al*., Efficient far red sensitization of nanocrystalline TiO_2 films by an unsymmetrical squaraine dye. *J. Am. Chem. Soc.* **129**, 10320–10321 (2007). doi: 10.1021/ja0731470; pmid: 17672464
55. H. X. Wang, Q. Wang, K. G. Zhou, H. L. Zhang, Graphene in light: Design, synthesis and applications of photo-active graphene and graphene-like materials. *Small* **9**, 1266–1283 (2013). doi: 10.1002/smll.201203040; pmid: 23554268
56. V. Yong, J. M. Tour, Theoretical efficiency of nanostructured graphene-based photovoltaics. *Small* **6**, 313–318 (2010). doi: 10.1002/smll.200901364; pmid: 19943249
57. M. Bernardi, M. Palummo, J. C. Grossman, Extraordinary sunlight absorption and one nanometer thick photovoltaics using two-dimensional monolayer materials. *Nano Lett.* **13**, 3664–3670 (2013). doi: 10.1021/nl401544y; pmid: 23750910
58. T. J. Echtermeyer *et al*., Strong plasmonic enhancement of photovoltage in graphene. *Nat. Commun.* **2**, 458 (2011). doi: 10.1038/ncomms1464; pmid: 21878912
59. M. D. Stoller, S. Park, Y. Zhu, J. An, R. S. Ruoff, Graphene-based ultracapacitors. *Nano Lett.* **8**, 3498–3502 (2008). doi: 10.1021/nl802558y; pmid: 18788793
60. J. T.-W. Wang *et al*., Low-temperature processed electron collection layers of graphene/TiO_2 nanocomposites in thin film perovskite solar cells. *Nano Lett.* **14**, 724–730 (2014). doi: 10.1021/nl403997a; pmid: 24341922
61. S.-S. Li, K.-H. Tu, C.-C. Lin, C.-W. Chen, M. Chhowalla, Solution-processable graphene oxide as an efficient hole transport layer in polymer solar cells. *ACS Nano* **4**, 3169–3174 (2010). doi: 10.1021/nn100551j; pmid: 20481512
62. Y.-J. Jeon, J.-M. Yun, D.-Y. Kim, S.-I. Na, S.-S. Kim, High-performance polymer solar cells with moderately reduced graphene oxide as an efficient hole transporting layer. *Sol. Energy Mater. Sol. Cells* **105**, 96–102 (2012). doi: 10.1016/j.solmat.2012.05.024
63. M. Li *et al*., Graphene quantum dots as the hole transport layer material for high-performance organic solar cells. *Phys. Chem. Chem. Phys.* **15**, 18973–18978 (2013). doi: 10.1039/c3cp53283f; pmid: 24097209
64. H. Choi, H. Kim, S. Hwang, W. Choi, M. Jeon, Dye-sensitized solar cells using graphene-based carbon nano composite as counter electrode. *Sol. Energy Mater. Sol. Cells* **95**, 323–325 (2010). doi: 10.1016/j.solmat.2010.04.044
65. M. Wu *et al*., Economical and effective sulfide catalysts for dye-sensitized solar cells as counter electrodes. *Phys. Chem. Chem. Phys.* **13**, 19298–19301 (2011). doi: 10.1039/c1cp22819f; pmid: 21984309
66. J.-Y. Lin, C.-Y. Chan, S.-W. Chou, Electrophoretic deposition of transparent MoS2-graphene nanosheet composite films as counter electrodes in dye-sensitized solar cells. *Chem. Commun. (Camb.)* **49**, 1440–1442 (2013). doi: 10.1039/c2cc38658e; pmid: 23321629
67. J. Liang, H. Bi, D. Wan, F. Huang, Novel Cu nanowires/graphene as the back contact for CdTe solar cells. *Adv. Funct. Mater.* **22**, 1267–1271 (2012). doi: 10.1002/adfm.201102809
68. E. Shi *et al*., Colloidal antireflection coating improves graphene-silicon solar cells. *Nano Lett.* **13**, 1776–1781 (2013). pmid: 23517083
69. F. J. DiSalvo, Thermoelectric cooling and power generation. *Science* **285**, 703–706 (1999). doi: 10.1126/science.285.5428.703; pmid: 10426986
70. G. A. Slack, *CRC Handbook of Thermoelectrics*, D. M. Rowe, Ed. (CRC Press, Boca Raton, FL, 1995).
71. R. Venkatasubramanian, E. Siivola, T. Colpitts, B. O'Quinn, Thin-film thermoelectric devices with high room-temperature figures of merit. *Nature* **413**, 597–602 (2001). doi: 10.1038/35098012; pmid: 11595940
72. C. B. Vining, An inconvenient truth about thermoelectrics. *Nat. Mater.* **8**, 83–85 (2009). doi: 10.1038/nmat2361; pmid: 19165205

73. A. A. Balandin *et al*., Superior thermal conductivity of single-layer graphene. *Nano Lett.* **8**, 902–907 (2008). doi: 10.1021/nl0731872; pmid: 18284217
74. J. Haskins *et al*., Control of thermal and electronic transport in defect-engineered graphene nanoribbons. *ACS Nano* **5**, 3779–3787 (2011). doi: 10.1021/nn200114p; pmid: 21452884
75. S. Chen *et al*., Thermal conductivity of isotopically modified graphene. *Nat. Mater.* **11**, 203–207 (2012). doi: 10.1038/nmat3207; pmid: 22231598
76. T. Gunst, T. Markussen, A.-P. Jauho, M. Brandbyge, Thermoelectric properties of finite graphene antidot lattices. *Phys. Rev. B* **84**, 155449 (2011). doi: 10.1103/PhysRevB.84.155449
77. H. Sevinçli, C. Sevik, T. Cain, G. Cuniberti, A bottom-up route to enhance thermoelectric figures of merit in graphene nanoribbons. *Sci. Rep.* **3**, 1228 (2013). pmid: 23390578
78. S. Hu *et al*., Proton transport through one-atom-thick crystals. *Nature* **516**, 227–230 (2014). doi: 10.1038/nature14015
79. L. Dong, R. R. S. Gari, Z. Li, M. M. Craig, S. Hou, Graphene-supported platinum and platinum–ruthenium nanoparticles with high electrocatalytic activity for methanol and ethanol oxidation. *Carbon* **48**, 781–787 (2010). doi: 10.1016/j.carbon.2009.10.027
80. E. Yoo *et al*., Enhanced electrocatalytic activity of Pt subnanoclusters on graphene nanosheet surface. *Nano Lett.* **9**, 2255–2259 (2009). doi: 10.1021/nl900397t; pmid: 19405511
81. A. Tarancon, M. Burriel, J. Santiso, S. J. Skinnerc, J. A. Kilner, Advances in layered oxide cathodes for intermediate temperature solid oxide fuel cells. *J. Mater. Chem.* **20**, 3799 (2010). doi: 10.1039/b922430k
82. D. Linden, T.B. Reddy, *Handbook of Batteries* (McGraw Hill, New York, ed. 4, 2010)
83. G. Wang, X. Shen, J. Yao, J. Park, Graphene nanosheets for enhanced lithium storage in lithium ion batteries. *Carbon* **47**, 2049–2053 (2009). doi: 10.1016/j.carbon.2009.03.053
84. G. Zhou, F. Li, H.-M. Cheng, Progress in flexible lithium batteries and future prospects. *Energy Environ. Sci* **7**, 1307 (2014). doi: 10.1039/c3ee43182g
85. A. Burke, M. Miller, H. Zhao, "Ultracapacitors in Hybrid Vehicle Applications: Testing of New High Power Devices and Prospects for Increased Energy Density." Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-12-06 (2012).
86. P. G. Bruce, S. A. Freunberger, L. J. Hardwick, J.-M. Tarascon, Li-O2 and Li-S batteries with high energy storage. *Nat. Mater.* **11**, 19–29 (2012). doi: 10.1038/nmat3191; pmid: 22169914
87. K. C. Divya, J. Ostergaard, Battery energy storage technology for power systems—An overview. *Electr. Power Syst. Res.* **79**, 511–520 (2009). doi: 10.1016/j.epsr.2008.09.017
88. A. K. Padhi, K. S. Nanjundaswamy, J. B. Goodenough, Phospho-olivines as positive-electrode materials for rechargeable lithium batteries. *J. Electrochem. Soc.* **144**, 1188 (1997). doi: 10.1149/1.1837571
89. T. Ohzuku, Y. Makimura, Layered lithium insertion material of $LiCo_{1/3}Ni_{1/3}Mn_{1/3}O_2$ for lithium-ion batteries. *Chem. Lett.* **30**, 642–643 (2001). doi: 10.1246/cl.2001.642
90. M. M. Thackeray, W. David, P. G. Bruce, J. B. Goodenough, Lithium insertion into manganese spinels. *Mater. Res. Bull.* **18**, 461–472 (1983). doi: 10.1016/0025-5408(83)90138-1
91. J. Hassoun *et al*., An advanced lithium-ion battery based on a graphene anode and a lithium iron phosphate cathode. *Nano Lett.* **14**, 4901–4906 (2014). doi: 10.1021/nl502429m; pmid: 25026051
92. B. Lung-Hao Hu, F. Y. Wu, C. T. Lin, A. N. Khlobystov, L. J. Li, Graphene-modified $LiFePO_4$ cathode for lithium ion battery beyond theoretical capacity. *Nat. Commun.* **4**, 1687 (2013). doi: 10.1038/ncomms2705; pmid: 23575691
93. E. Yoo *et al*., Large reversible Li storage of graphene nanosheet families for use in rechargeable lithium ion batteries. *Nano Lett.* **8**, 2277–2282 (2008). doi: 10.1021/nl800957b; pmid: 18651781
94. H. Wang *et al*., $LiMn_{1-x}Fe_xPO4$ nanorods grown on graphene sheets for ultrahigh-rate-performance lithium ion batteries. *Angew. Chem. Int. Ed.* **123**, 7502–7506 (2011). doi: 10.1002/ange.201103163
95. H. Kim *et al*., Graphene-based hybrid electrode material for high-power lithium-ion batteries. *J. Electrochem. Soc.* **158**, A930 (2011). doi: 10.1149/1.3599632
96. S.-M. Bak *et al*., Spinel $LiMn_2O_4$/reduced graphene oxide hybrid for high rate lithium ion batteries. *J. Mater. Chem.* **21**, 17309 (2011). doi: 10.1039/c1jm13741g
97. S. Yang, X. Feng, S. Ivanovici, K. Müllen, Fabrication of graphene-encapsulated oxide nanoparticles: Towards high-performance anode materials for lithium storage. *Angew. Chem. Int. Ed. Engl.* **49**, 8408–8411 (2010). doi: 10.1002/anie.201003485; pmid: 20836109
98. G. Zhou *et al*., Graphene-wrapped Fe_3O_4 anode material with improved reversible capacity and cyclic stability for lithium ion batteries. *Chem. Mater.* **22**, 5306–5313 (2010). doi: 10.1021/cm101532x
99. Y. Shi *et al*., Hollow structured Li_3VO_4 wrapped with graphene nanosheets in situ prepared by a one-pot template-free method as an anode for lithium-ion batteries. *Nano Lett.* **13**, 4715–4720 (2013). doi: 10.1021/nl402237u; pmid: 24024651
100. X. Zhao, C. M. Hayner, M. C. Kung, H. H. Kung, In-plane vacancy-enabled high-power Si–graphene composite electrode for lithium-ion batteries. *Adv. Energy Mater.* **1**, 1079–1084 (2011). doi: 10.1002/aenm.201100426
101. B. Ammundsen, J. Paulsen, Novel lithium-ion cathode materials based on layered manganese oxides. *Adv. Mater.* **13**, 943–956 (2001). doi: 10.1002/1521-4095(200107)13:12/13<943::AID-ADMA943>3.0.CO;2-J
102. M. S. Whittingham, Electrical energy storage and intercalation chemistry. *Science* **192**, 1126–1127 (1976). doi: 10.1126/science.192.4244.1126; pmid: 17748676
103. G. Du *et al*., Superior stability and high capacity of restacked molybdenum disulfide as anode material for lithium ion batteries. *Chem. Commun. (Camb.)* **46**, 1106–1108 (2010). doi: 10.1039/b920277c; pmid: 20126728
104. J. T. Jang *et al*., Ultrathin zirconium disulfide nanodiscs. *J. Am. Chem. Soc.* **133**, 7636–7639 (2011). doi: 10.1021/ja200400n; pmid: 21539379
105. K. Shiva, H. S. S. R. Matte, H. B. Reajendra, A. J. Bhattacharyya, C. N. R. Rao, Employing synergistic interactions between few-layer WS_2 and reduced graphene oxide to improve lithium storage, cyclability and rate capability of Li-ion batteries. *Nano Energy* **2**, 787–793 (2013). doi: 10.1016/j.nanoen.2013.02.001
106. B. E. Conway, *Electrochemical Supercapacitors: Scientific Fundamentals and Technological Applications* (Plenum Publishers, New York, 1999).
107. M. Stoller *et al*., Interfacial capacitance of single layer graphene. *Energy Environ. Sci.* **4**, 4685 (2011). doi: 10.1039/c1ee02322e
108. J. W. Longa *et al*., Asymmetric electrochemical capacitors—Stretching the limits of aqueous electrolytes. *MRS Bull.* **36**, 513–522 (2011). doi: 10.1557/mrs.2011.137
109. M. D. Stoller, R. S. Ruoff, Best practice methods for determining an electrode material's performance for ultracapacitors. *Energy Environ. Sci.* **3**, 1294 (2010). doi: 10.1039/c0ee00074d
110. Y. Gogotsi, P. Simon, Materials science. True performance metrics in electrochemical energy storage. *Science* **334**, 917–918 (2011). doi: 10.1126/science.1213003; pmid: 22096182
111. Y. W. Zhu *et al*., Microwave assisted exfoliation and reduction of graphite oxide for ultracapacitors. *Carbon* **48**, 2118–2122 (2010). doi: 10.1016/j.carbon.2010.02.001
112. S. R. Vivekchand, S. R. Chandra, K. S. Subrahmanyam, A. Govindaraj, C. N. R. Rao, Graphene-based electrochemical supercapacitors. *J. Chem. Sci.* **120**, 9–13 (2008). doi: 10.1007/s12039-008-0002-7
113. Z. Lei, N. Christova, X. S. Zhao, Intercalation of mesoporous carbon spheres between reduced graphene oxide sheets for preparing high-rate supercapacitor electrodes. *Energy Environ. Sci.* **4**, 1866 (2011). doi: 10.1039/c1ee01094h
114. M. A. Pope, S. Korkut, C. Punckt, I. A. Aksay, Supercapacitor electrodes produced through evaporative consolidation of graphene oxide-water-ionic liquid gels. *J. Electrochem. Soc.* **160**, A1653–A1660 (2013). doi: 10.1149/2.017310jes
115. Y. Li, Z. Li, P. K. Shen, Simultaneous formation of ultrahigh surface area and three-dimensional hierarchical porous graphene-like networks for fast and highly stable supercapacitors. *Adv. Mater.* **25**, 2474–2480 (2013). doi: 10.1002/adma.201205332; pmid: 23495046
116. L. Zhang *et al*., Porous 3D graphene-based bulk materials with exceptional high surface area and excellent conductivity for supercapacitors. *Sci. Rep.* **3**, 1408 (2013). pmid: 23474952
117. Y. Tao *et al*., Towards ultrahigh volumetric capacitance: Graphene derived highly dense but porous carbons for supercapacitors. *Sci. Rep.* **3**, 2975 (2013). doi: 10.1038/srep02975; pmid: 24131954
118. X. Yang, C. Cheng, Y. Wang, L. Qiu, D. Li, Liquid-mediated dense integration of graphene materials for compact capacitive energy storage. *Science* **341**, 534–537 (2013). doi: 10.1126/science.1239089; pmid: 23908233
119. W.-Y. Tsai *et al*., Outstanding performance of activated graphene based supercapacitors in ionic liquid electrolyte from −50 to 80 °C. *Nano Energy* **2**, 403–411 (2013). doi: 10.1016/j.nanoen.2012.11.006
120. R. S. Borges *et al*., Supercapacitor operating at 200 degrees celsius. *Sci. Rep.* **3**, 2572 (2013). doi: 10.1038/srep02572; pmid: 23999206
121. F. Beguin, V. Khomenko, E. Raymundo-Pinero, High-energy density graphite/AC capacitor in organic electrolyte. *J. Power Sources* **177**, 643–651 (2008). doi: 10.1016/j.jpowsour.2007.11.101
122. M. D. Stoller *et al*., Activated graphene as a cathode material for Li-ion hybrid supercapacitors. *Phys. Chem. Chem. Phys.* **14**, 3388–3391 (2012). doi: 10.1039/c2cp00017b; pmid: 22298158
123. F. Zhang *et al*., A high-performance supercapacitor-battery hybrid energy storage device based on graphene-enhanced electrode materials with ultrahigh energy density. *Energy Environ. Sci.* **6**, 1623 (2013). doi: 10.1039/c3ee40509e
124. Z. S. Wu *et al*., High-energy MnO_2 nanowire/graphene and graphene asymmetric electrochemical capacitors. *ACS Nano* **4**, 5835–5842 (2010). doi: 10.1021/nn101754k; pmid: 20857919
125. Q. Wu, Y. Xu, Z. Yao, A. Liu, G. Shi, Supercapacitors based on flexible graphene/polyaniline nanofiber composite films. *ACS Nano* **4**, 1963–1970 (2010). doi: 10.1021/nn1000035; pmid: 20355733
126. J. Yang, H. S. Shin, Recent advances in layered transition metal dichalcogenides for hydrogen evolution reaction. *J. Mater. Chem. A* **2**, 5979 (2014). doi: 10.1039/c3ta14151a
127. X. Zheng *et al*., Space-confined growth of MoS_2 nanosheets within graphite: The layered hybrid of MoS_2 and graphene as an active catalyst for hydrogen evolution reaction. *Chem. Mater.* **26**, 2344–2353 (2014). doi: 10.1021/cm500347r
128. D. Voiry *et al*., Conducting MoS_2 nanosheets as catalysts for hydrogen evolution reaction. *Nano Lett.* **13**, 6222–6227 (2013). doi: 10.1021/nl403661s; pmid: 24251828
129. S. Patchkovskii *et al*., Graphene nanostructures as tunable storage media for molecular hydrogen. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10439–10444 (2005). doi: 10.1073/pnas.0501030102; pmid: 16020537
130. G. K. Dimitrakakis, E. Tylianakis, G. E. Froudakis, Pillared graphene: A new 3-D network nanostructure for enhanced hydrogen storage. *Nano Lett.* **8**, 3166–3170 (2008). doi: 10.1021/nl801417w; pmid: 18800853
131. G. Srinvas *et al*., Synthesis of graphene-like nanosheets and their hydrogen adsorption capacity. *Carbon* **48**, 630–635 (2010). doi: 10.1016/j.carbon.2009.10.003
132. Y. I. Zhirko, Z. D. Kovalyuk, M. M. Pyrlja, V. B. Boledzyuk, in *Hydrogen Materials Science and Chemistry of Carbon Nanomaterials series: NATO Security through Science Series A: Chemistry and Biology* (Springer, Netherlands, 2007).
133. G. Kim, S.-H. Jhi, N. Park, M. L. Cohen, S. G. Louie, Optimization of metal dispersion in doped graphitic materials for hydrogen storage. *Phys. Rev. B* **78**, 085408 (2008). doi: 10.1103/PhysRevB.78.085408
134. S. Goler *et al*., The influence of graphene curvature on hydrogen adsorption: A study for future hydrogen storage devices. *J. Phys. Chem. C* **117**, 11506–11513 (2013). doi: 10.1021/jp4017536
135. V. B. Parambhath, R. Nagar, K. Sethupathi, S. Ramaprabhu, Investigation of spillover mechanism in palladium decorated hydrogen exfoliated functionalized graphene. *J. Phys. Chem. C* **115**, 15679–15685 (2011). doi: 10.1021/jp202797q
136. Z. M. Ao, F. M. Peeters, Electric field activated hydrogen dissociative adsorption to nitrogen-doped graphene. *J. Phys. Chem. C* **114**, 14503–14509 (2010). doi: 10.1021/jp103835k
137. D. W. Boukhvalov, M. I. Katsnelson, Enhancement of chemical activity in corrugated graphene. *J. Phys. Chem. C* **113**, 14176–14178 (2009). doi: 10.1021/jp905702e
138. M. T. Ong, E. J. Reed, Engineered piezoelectricity in graphene. *ACS Nano* **6**, 1387–1394 (2012). doi: 10.1021/nn204198g; pmid: 22196055

ACKNOWLEDGMENTS

We thank M. Prato for useful discussions. We acknowledge funding from European Union (EU) Graphene Flagship (contract 604391); Newton International Fellowship; faculty start-up support from the University of Texas–Austin; the Welch Foundation Award F-1861; European Research Council grant Hetero2D; a Royal Society Wolfson Research Merit Award, EU projects GENIUS, CARERAMM, and MEM4WIN; Engineering and Physical Sciences Research Council grants EP/K01711X/1, EP/K017144/1, and EP/L016087/1; and Lancaster University's Distinguished Visitor Program in Physical Sciences.

10.1126/science.1246501

RESEARCH ARTICLE SUMMARY

MOSQUITO GENOMICS

Extensive introgression in a malaria vector species complex revealed by phylogenomics

Michael C. Fontaine,* James B. Pease,* Aaron Steele, Robert M. Waterhouse, Daniel E. Neafsey, Igor V. Sharakhov, Xiaofang Jiang, Andrew B. Hall, Flaminia Catteruccia, Evdoxia Kakani, Sara N. Mitchell, Yi-Chieh Wu, Hilary A. Smith, R. Rebecca Love, Mara K. Lawniczak, Michel A. Slotman, Scott J. Emrich, Matthew W. Hahn,† Nora J. Besansky†

INTRODUCTION: The notion that species boundaries can be porous to introgression is increasingly accepted. Yet the broader role of introgression in evolution remains contentious and poorly documented, partly because of the challenges involved in accurately identifying introgression in the very groups where it is most likely to occur. Recently diverged species often have incomplete reproductive barriers and may hybridize where they overlap. However, because of retention and stochastic sorting of ancestral polymorphisms, inference of the correct species branching order is notoriously challenging for recent speciation events, especially those closely spaced in time. Without knowledge of species relationships, it is impossible to identify instances of introgression.

RATIONALE: Since the discovery that the single mosquito taxon described in 1902 as *Anopheles gambiae* was actually a complex of several closely related and morphologically indistinguishable sibling species, the correct species branching order has remained controversial and unresolved. This Afrotropical complex contains the world's most important vectors of human malaria, owing to their close association with humans, as well as minor vectors and species that do not bite humans. On the basis of ecology and behavior, one might predict phylogenetic clustering of the three highly anthropophilic vector species. However, previous phylogenetic analyses of the complex based on a limited number of markers strongly disagree about relationships between the major vectors, potentially because of historical introgression between them. To investigate the history of the species complex, we used whole-genome reference assemblies, as well as dozens of resequenced individuals from the field.

RESULTS: We observed a large amount of phylogenetic discordance between trees generated from the autosomes and X chromosome. The autosomes, which make up the majority of the genome, overwhelmingly supported the grouping of the three major vectors of malaria, *An. gambiae*, *An. coluzzii*, and *An. arabiensis*. In stark contrast, the X chromosome strongly supported the grouping of *An. arabiensis* with a species that plays no role in malaria transmission, *An. quadriannulatus*. Although the whole-genome consensus phylogeny unequivocally agrees with the autosomal topology, we found that the topology most often located on the X chromosome follows the historical species branching order, with pervasive introgression on the autosomes producing relationships that group the three highly anthropophilic species together. With knowledge of the correct species branching order, we are further able to uncover introgression between another species pair, as well as a complex history of balancing selection, introgression, and local adaptation of a large autosomal inversion that confers aridity tolerance.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1258524

CONCLUSION: We identify the correct species branching order of the *An. gambiae* species complex, resolving a contentious phylogeny. Notably, lineages leading to the principal vectors of human malaria were among the first in the complex to radiate and are not most closely related to each other. Pervasive autosomal introgression between these human malaria vectors, including nonsister vector species, suggests that traits enhancing vectorial capacity can be acquired not only through de novo mutation but also through a more rapid process of interspecific genetic exchange.

Time-lapse photographs of an adult anopheline mosquito emerging from its pupal case.

PHOTO CREDIT: JAMES GATHANY, CENTERS FOR DISEASE CONTROL AND PREVENTION

RELATED ITEMS IN *SCIENCE*

D. E. Neafsey *et al.*, *Science* **347**, 1258522 (2015)
DOI: 10.1126/science.1258522

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. E-mail: mwh@indiana.edu (M.W.H.); nbesansk@nd.edu (N.J.B.)
Cite this article as M. C. Fontaine *et al.*, *Science* **347**, 1258524 (2015). DOI: 10.1126/science.1258524

RESEARCH ARTICLE

MOSQUITO GENOMICS

Extensive introgression in a malaria vector species complex revealed by phylogenomics

Michael C. Fontaine,[1,2]*† James B. Pease,[3]* Aaron Steele,[4] Robert M. Waterhouse,[5,6,7,8] Daniel E. Neafsey,[6] Igor V. Sharakhov,[9,10] Xiaofang Jiang,[10] Andrew B. Hall,[10] Flaminia Catteruccia,[11,12] Evdoxia Kakani,[11,12] Sara N. Mitchell,[11] Yi-Chieh Wu,[5] Hilary A. Smith,[1,2] R. Rebecca Love,[1,2] Mara K. Lawniczak,[13]‡ Michel A. Slotman,[14] Scott J. Emrich,[2,4] Matthew W. Hahn,[3,15]§ Nora J. Besansky[1,2]§

Introgressive hybridization is now recognized as a widespread phenomenon, but its role in evolution remains contested. Here, we use newly available reference genome assemblies to investigate phylogenetic relationships and introgression in a medically important group of Afrotropical mosquito sibling species. We have identified the correct species branching order to resolve a contentious phylogeny and show that lineages leading to the principal vectors of human malaria were among the first to split. Pervasive autosomal introgression between these malaria vectors means that only a small fraction of the genome, mainly on the X chromosome, has not crossed species boundaries. Our results suggest that traits enhancing vectorial capacity may be gained through interspecific gene flow, including between nonsister species.

The notion that species boundaries can be porous to introgression is increasingly accepted. Charismatic cases, such as gene flow between Neanderthals and anatomically modern humans (*1*) or between *Heliconius* butterflies (*2*, *3*), show that introgression can transfer beneficial alleles between closely related species. Yet the broader role of introgression in evolution remains contentious and poorly documented, partly because of the challenges involved in accurately identifying introgression in the very groups where it is most likely to occur. Recently diverged species often have incomplete reproductive barriers, hence, may hybridize in sympatry. However, another feature of rapid radiations is that ancestral polymorphism predating lineage splitting may be sorted stochastically among descendant lineages in a process known as incomplete lineage sorting (ILS). Alleles shared through ILS can be difficult to distinguish from those shared through secondary contact and introgression. Newly developed methods can differentiate these two processes (*1*, *4*) but only if the correct species branching order is known. Because both introgression and ILS cause discordance between gene trees and the species tree, inference of the correct species phylogeny (i.e., the historical branching order of the taxa) is notoriously challenging for recent radiations (*5*–*7*).

Since the discovery that the single mosquito taxon described in 1902 as *Anopheles gambiae* (*8*) was actually a complex of several closely related and morphologically indistinguishable sibling species (known as the *An. gambiae* complex) (*9*), the correct species branching order has remained controversial and unresolved. This Afrotropical complex (*10*–*13*) contains three widely distributed and extensively sympatric species that rank among the world's most important vectors of human malaria, owing to their association with humans (*An. gambiae sensu stricto*, its closest relative and sister species, *Anopheles coluzzii*, and *Anopheles arabiensis*) (Fig. 1A). *Anopheles merus* and *Anopheles melas*, salt-tolerant species that breed in brackish coastal waters of eastern and western Africa, respectively, are minor vectors. *Anopheles quadriannulatus* plays no role in malaria transmission despite vector competence for *Plasmodium falciparum*, as it tends to bite animals other than humans. On the basis of ecology and behavior, one might predict phylogenetic clustering of the highly anthropophilic vector species. Yet such clustering has not been supported by chromosomal inversion phylogenies (*10*, *14*). The apparent phylogenetic affinity between *An. arabiensis* and *An. gambiae* supported by molecular markers and shared chromosomal inversion polymorphisms was instead attributed to introgression (*15*, *16*). Introgression is plausible between any geographically overlapping pair of species in the complex, as reproductive isolation is incomplete: Adult female F_1 hybrids—although uncommon in nature—are fertile and vigorous (only F_1 hybrid males are sterile) (*10*). Nonetheless, comprehensive evidence for introgression between *An. arabiensis* and the other major vector lineage (*An. gambiae* + *An. coluzzii*) has been lacking, until now, because of insufficient genomic resources.

Important malaria vectors make up a small fraction of the genus *Anopheles* but almost invariably are embedded in species complexes similar to that of *An. gambiae sensu lato* (*17*). Thus, resolving the phylogeny of the *An. gambiae* complex has the potential to yield insights into the origin and evolution of traits that are associated with highly successful malaria vectors across the genus as a whole. Here, we have used newly available whole-genome reference assemblies (*18*) to infer the species phylogeny. We found that the two most important malaria vector lineages in the complex are not the most closely related, and we have uncovered pervasive introgression between them. The extent of introgression is such that the phylogenetic tree inferred from whole-genome alignments supports the incorrect species branching order with high confidence.

The X chromosome reflects the species branching order

As a first step in phylogenomic analysis, alignments were generated for existing (*19*, *20*) and newly available genome assemblies, representing six species of the *An. gambiae* complex and two Pyretophorus out-group species (*An. christyi* and *An. epiroticus*) (supplementary text S1 and fig. S1) (*21*). Among the in-group species, the alignment included ~60% of nongap and nonmasked base pairs of the *An. gambiae* PEST reference assembly. Across all assemblies, the proportion of aligned base pairs was lower (~53%) but, nevertheless, spanned more than 93 megabase pairs (Mb) (~40% of the euchromatic genome).

As the *An. gambiae* complex reference genomes were assembled from laboratory colony–derived sequencing template [except for *An. melas* (*18*)], we also performed whole-genome shotgun sequencing of individual field-collected specimens sampled from at least one population of each of the six in-group species (supplementary text S2). Sequencing reads were aligned to the

[1]Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA. [2]Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN 46556, USA. [3]Department of Biology, Indiana University, Bloomington, IN 47405, USA. [4]Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN 46556, USA. [5]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139, USA. [6]The Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142, USA. [7]Department of Genetic Medicine and Development, University of Geneva Medical School, rue Michel-Servet 1, 1211 Geneva, Switzerland. [8]Swiss Institute of Bioinformatics, rue Michel-Servet 1, 1211 Geneva, Switzerland. [9]Department of Entomology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. [10]The Interdisciplinary PhD Program in Genetics, Bioinformatics, and Computational Biology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. [11]Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA 02115, USA. [12]Dipartimento di Medicina Sperimentale e Scienze Biochimiche, Università degli Studi di Perugia, Perugia, Italy. [13]Department of Life Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, UK. [14]Department of Entomology, Texas A&M University, College Station, TX 77843, USA [15]School of Informatics and Computing, Indiana University, Bloomington, IN 47405, USA.
*These authors contributed equally to this work. †Present address: Centre for Ecological and Evolutionary Studies (MarECon group), University of Groningen, Nijenborgh 7, 9747 AG Groningen, Netherlands. ‡Present address: Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. §Corresponding author. E-mail: mwh@indiana.edu (M.W.H.); nbesansk@nd.edu (N.J.B.)

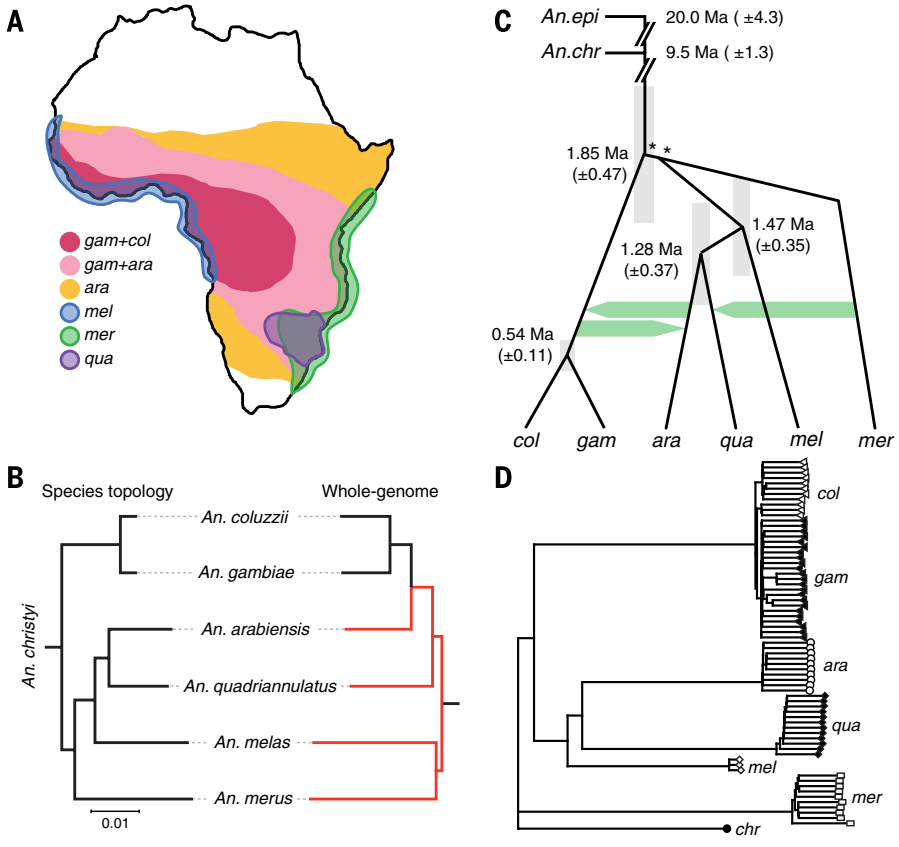


Fig. 1. Distribution and phylogenetic relationships of sequenced members of the *An. gambiae* complex. (**A**) Schematic geographic distribution of *An. gambiae* (gam, formerly *An. gambiae* S form), *An. coluzzii* [col, formerly *An. gambiae* M form (*13*)], *An. arabiensis* (ara), *An. quadriannulatus* (qua), *An. merus* (mer), and *An. melas* (mel). (**B**) Species topology as estimated from a ML phylogeny of the X chromosome (see text) compared with the ML phylogeny estimated from the whole-genome sequence alignment. The scale-bar denotes nucleotide divergence as calculated by RAxML. Red branches of the latter tree highlight topological differences with the species tree. (**C**) Schematic of the species topology rooted by *An. christyi* (*An. chr*) and *An. epiroticus* (*An. epi*), showing the major introgression events (green arrows) and the approximate divergence and introgression times [Ma ± 1 standard deviation (SD)], if one assumes a substitution rate of 1.1×10^{-9} per site, per generation, and 10 generations per year. The gray bar surrounding each node shows ±2 SD. Nodes marked by asterisks are not fully resolved owing to high levels of ILS. (**D**) Neighbor-joining tree displaying the Euclidian distance between individuals from population samples of each species, calculated using sequence data from the X chromosome. Each species is represented by a distinct symbol shape or shade.

species-appropriate reference assembly to avoid considerable interspecific mapping bias (supplementary text S2, figs. S3 to S4, and table S3), and single-nucleotide variant positions were converted to a common coordinate system relative to the reference genome alignment (supplementary texts S2 and S3 and fig. S2). The results presented below are based on the field-collected samples, but analyses were performed in parallel on the reference genome assemblies, and our findings were consistent in all cases.

To infer the correct species branching order in the face of anticipated ILS and introgression, maximum-likelihood (ML) phylogenies were constructed from 50-kilobase (kb) nonoverlapping windows across the alignments (referred to here as "gene trees" regardless of their protein-coding content), considering six in-group species rooted alternatively with *An. christyi* or *An. epiroticus* (n = 4063 windows) (supplementary text S3). As the choice of out-group did not materially alter our results, we present our findings based on the more closely related species, *An. christyi*. When the 85 tree topologies observed at least once across the genome were sorted by chromosome arm and relative frequency, the most commonly observed trees were strongly discordant between the X chromosome and the autosomes (Fig. 2, table S9, and fig. S16). Although weak disagreement among tree topologies concerning the branching order of basal nodes was a consequence of poor resolution due to ILS (Fig. 1C and fig. S16B), the striking discordance between the X chromosome and the autosomes was not a trivial consequence of lack of phylogenetic signal; conflicting topologies had strong bootstrap support (fig. S16B). Because the major disagreement between the X chromosome and the autosomes concerned the relative phylogenetic positions of *An. arabiensis* and *An. quadriannulatus*, for simplicity, we grouped the most frequently observed topologies into three sets (Fig. 2 and fig. S16), (i) A+CG: those that supported *An. arabiensis* clustering with *An. coluzzii* + *An. gambiae*; (ii) A+CG, R+Q: those that supported both *An. arabiensis* (*An. coluzzii* + *An. gambiae*) and *An. merus* + *An. quadriannulatus*; and (iii) A+Q: those that supported the clustering of *An. arabiensis* and *An. quadriannulatus*. On the X chromosome, all three of the most frequently observed topologies (inferred from 64% of windows) strongly supported the relationship [*melas* (*arabiensis*, *quadriannulatus*)], which indicated a sister-taxon relationship between *An. arabiensis* and *An. quadriannulatus* (i.e., A+Q) (orange shades in Fig. 2). This relationship was shared among all field-collected samples (Fig. 1D). Notably, the X chromosomal windows supporting these topologies are concentrated distal to the centromere (Fig. 3D and supplementary text S3), in an ~15-Mb region corresponding to the Xag inversion whose orientation is ancestral to the *An. gambiae* complex and shared by *An. gambiae*, *An. coluzzii*, and *An. merus* (see supplementary text S5).

In stark contrast to the X chromosome, the overwhelming majority of window-based topologies across the autosomes supported *An. arabiensis* as sister to *An. gambiae* + *An. coluzzii* (green and purple shades in Fig. 2). On chromosomes 3 and 2R, a subset of these topologies also supported a sister-taxon relationship between *An. quadriannulatus* and *An. merus*, in further disagreement with the X chromosome (purple shades, Fig. 2).

Autosomal introgression between *An. arabiensis* and the ancestor of *An. gambiae* + *An. coluzzii* has long been postulated (*10*, *22*) and could explain the strong discordance between the dominant tree topologies of the X and autosomes. However, before this study, the correct species branching order was unresolved, which precluded definitive interpretation of these conflicting signals. To infer the correct historical branching order, we applied a strategy based on sequence divergence (supplementary text S3 and fig. S16). Because introgression will reduce sequence divergence between the species exchanging genes, we expect that the correct species branching order revealed by gene trees constructed from non-introgressed sequences will show deeper divergences than those constructed from introgressed sequences. If the hypothesis of autosomal introgression is correct, this implies that the topologies supported by the X chromosome should show significantly higher divergence times between *An. arabiensis* and either *An. gambiae* or *An. coluzzii* than topologies supported by the autosomes.

To test this hypothesis, we used *An. arabiensis*, *An. gambiae*, and *An. melas*, as these three species show strongly discordant trees on the X and autosomes. For this trio, there are three possible phylogenetic relationships and two divergence times for each, T_1 and T_2 (Fig. 3A). Using the metrics T_1 and T_2, we conducted two different tests. Initially,

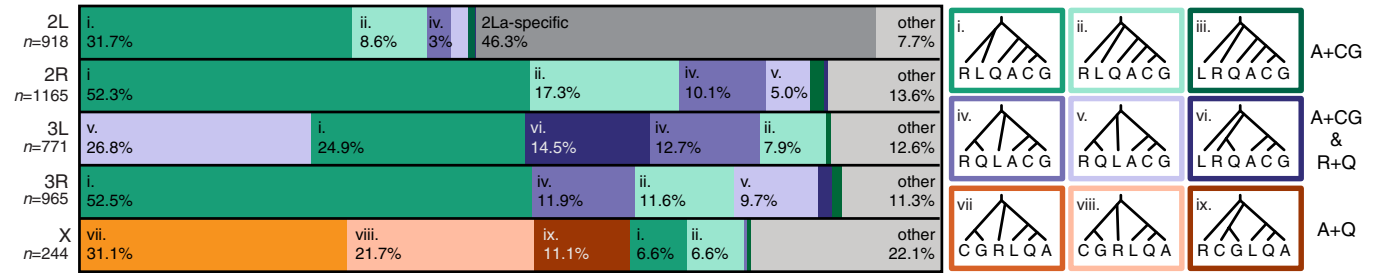
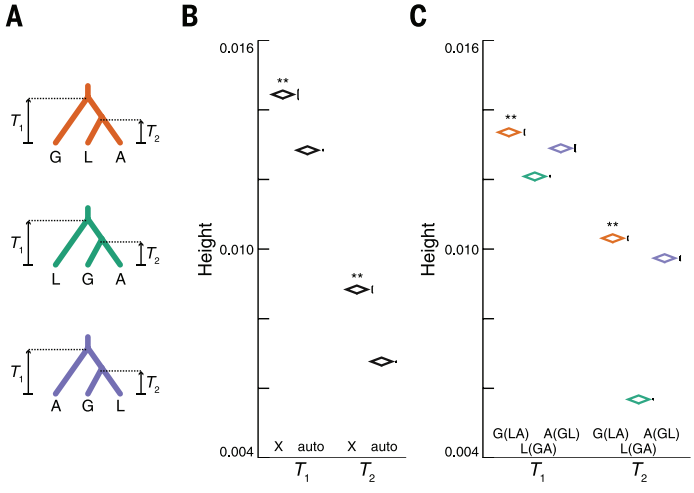


Fig. 2. Phylogenies inferred from regions on the autosomes differ strongly from phylogenies inferred from regions on the X chromosome. ML-rooted phylogenies were inferred from *n* = 4063 50-kb genomic windows for *An. arabiensis* (A), *An. coluzzii* (C), *An. gambiae* (G), *An. melas* (L), *An. merus* (R), and *An. quadriannulatus* (Q), with *An. christyi* as out-group. The nine most commonly observed topologies across the genome (trees i to ix) are indicated on each of the five chromosome arms (if found) by correspondingly colored blocks whose length represents the proportion of all 50-kb windows on that arm that support the topology. Topologies specific to 2La are represented by the dark gray block on 2L; all other topologies observed on each chromosome arm were pooled, and their combined frequencies are indicated by the light gray blocks. The X chromosome most often indicates that A and Q are sister taxa (trees vii to ix), whereas the autosomes indicate that A and C+G are sister groups (trees i to vi). Large portions of the autosomes (particularly 3L and 3R) indicate that R and Q are sister taxa (trees iv to vi). Additional diversity in inferred trees is the result of rearrangement of three groups (R, C+G, and L+A+Q) because of ILS. The most common phylogeny on the X chromosome (vii) represents the most likely species branching order. The 2L arm has a markedly different distribution of phylogenies because of the unique history of the 2La inversion region (see Fig. 5), which creates unusual phylogenetic topologies found nowhere else in the genome. The autosome-like trees on the X chromosome (i and ii) are entirely found in the pericentromeric region (15 to 24 Mb) (see Fig. 3D), where introgression between *An. arabiensis* and *An. gambiae* + *An. coluzzi* has previously been implicated.



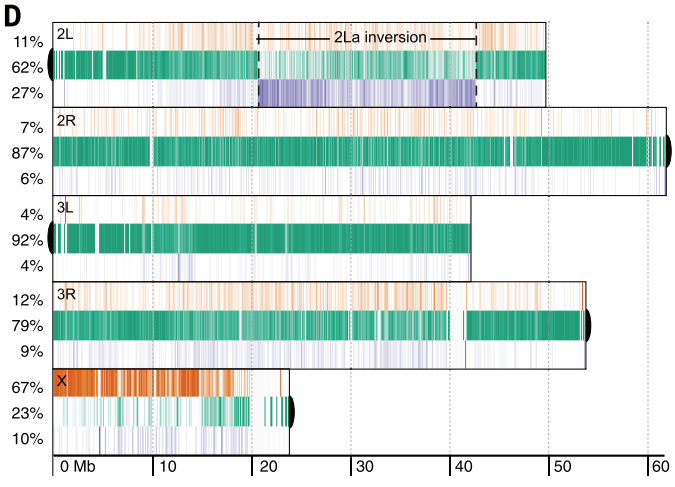


Fig. 3. Tree height reveals the true species branching order in the face of introgression. (**A**) Color-coded trees show the three possible rooted phylogenetic relationships for *An. arabiensis* (A), *An. gambiae* s.s. (G), and *An. melas* (L), with out-group *An. christyi*. For any group of three taxa, there are two species divergence times (T_1 and T_2). When introgression has occurred, a strong decrease is expected in these divergence times. (**B**) Trees on the X chromosome have significantly higher mean values of T_1 and T_2 than the autosomes, which indicates that introgression is more likely to have occurred on the autosomes than the X chromosome (diamond for mean, whiskers ± SEM offset to the right for visual clarity; $^{**}P < 1.0 \times 10^{-40}$). (**C**) Even among only autosomal loci, regions with the X majority relationship G(LA) (orange) have significantly higher T_1 and T_2 ($^{**}P < 1.0 \times 10^{-40}$), which indicates that the autosomal majority topology (green) is the result of widespread introgression between *An. arabiensis* and *An. gambiae*. The X majority relationship G(LA) therefore represents the true species branching relationships. (**D**) The gene tree distributions or "chromoplots" for all five chromosomal arms show that the autosomal and X chromosome phylogenies strongly disagree. Three-taxon phylogenies were inferred from 10-kb chromosomal windows across the genome, and colored vertical bars represent the relative abundance of the three alternative topologies (A) in a 200-kb window. The proportions of the three 10-kb gene trees for each chromosome arm (left) indicate that the autosomes all strongly show a closer relationship between *An. gambiae* and *An. arabiensis* (green); the X indicates a closer relationship between *An. arabiensis* and *An. melas* (orange). Black semicircles indicate the location of centromeres.

we compared mean sequence divergence between X chromosome–based versus autosome-based topologies. As predicted, the estimated mean values of T_1 and T_2 for trees inferred from 10-kb windows on the X chromosome were significantly higher than their counterparts on the autosomes (both $P < 1.0 \times 10^{-40}$) (Fig. 3B). However, the possibility exists that this result was confounded or entirely driven by other factors that differ between the X and autosomes, including a faster rate of evolution on the X (*23*). Indeed, we found evidence supporting faster evolution of genes on the X chromosome (supplementary text S3 and fig. S23). To avoid this problem, we conducted a second test on the autosomes alone, focusing on trees inferred from 10-kb windows and the mean divergence levels among those supporting the three possible phylogenetic relationships illustrated in Fig. 3A. The result, congruent with the first test but providing unequivocal evidence for the correct species branching order, was that the set of autosomal trees supporting the majority X chromosome topology [*gambiae* (*melas*, *arabiensis*)] again had the highest mean values of T_1 and T_2 (both $P < 1.0 \times 10^{-40}$) (Fig. 3C). This indicates that the species branching order inferred from the X chromosome is the correct one (Fig. 1B), despite extensive amounts of *An. gambiae–An. arabiensis* autosomal introgression (Fig. 3D and fig. S16).

The total-evidence approach to phylogenetic reconstruction (*24*) is premised on the notion that the best estimate of species relationships arises

from character congruence when all available data are considered simultaneously. Alternative approaches seek taxonomic congruence between different data sets or data partitions, analyzed separately (25). Either way, standard phylogenomic practice entails some form of "majority rule," based on the assumption that as the amount of data increases so will the probability of converging on the correct species branching order [apart from exceptional cases (26)]. Because of both ILS and introgression, the historical species branching order for the *An. gambiae* complex is represented by only 1.9% of 50-kb windows across the entire genome (Fig. 2). As a result, when we inferred a ML tree for the *An. gambiae* complex on the basis of whole-genome alignments, we recovered the wrong species branching order supported by 100% of the bootstrap replicates at each node (Fig. 1B, supplementary text S3, and figs. S17A and S18A). The extent of autosomal introgression in the *An. gambiae* complex has the paradoxical effect that, as an increasing amount of the genome is sampled, support for the incorrect species branching order is maximized.

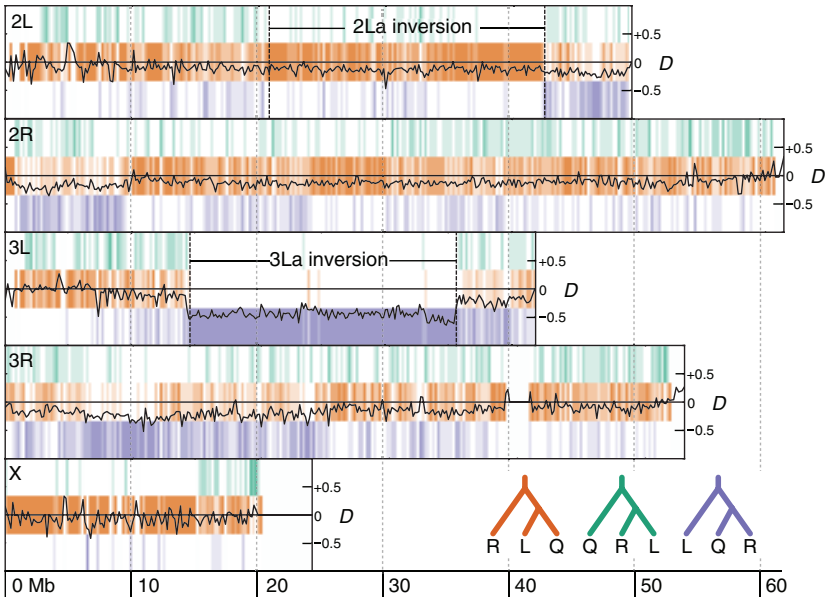


Fig. 4. Introgression between *An. merus* and *An. quadriannulatus*. Chromoplots for all five chromosomal arms show a highly spatially heterogeneous distribution of phylogenies inferred from 50-kb genomic regions, particularly on 3L. The three possible rooted phylogenetic relationships for *An. quadriannulatus* (Q), *An. melas* (L), and *An. merus* (R), with out-group *An. christyi* are shown, colored according to the key in the lower right. The region on 3L corresponding to the 3La inversion shows strong evidence of R-Q introgression and a strong negative deviation of the *D* statistic. The region on 2L corresponding to the 2La inversion is highly enriched for the R(LQ) relationship, as expected, given that L and Q both have the 2L+^{a} orientation, whereas R has 2La (see Fig. 5). Across all the autosomes, the *D* statistic generally trends toward negative values, which indicates weak or ancient R-Q introgression may have been occurring across the autosomes (see supplementary text S3).

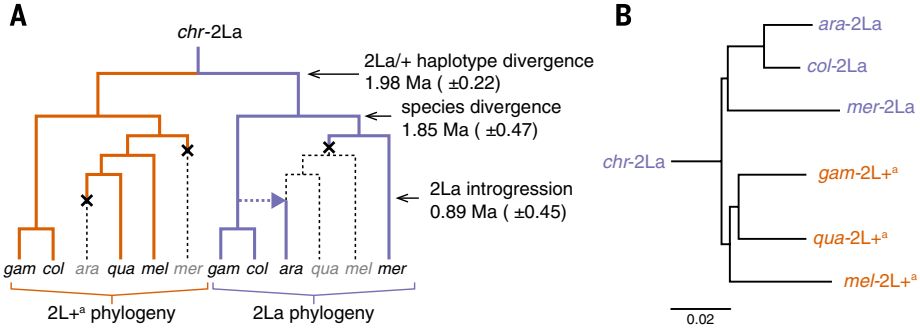


Fig. 5. Ancient trans-specific polymorphism of an inversion predates radiation of the *An. gambiae* complex. (**A**) All species in the complex are fixed for either the 2La (*An. arabiensis*, *ara*; *An. merus*, *mer*) or 2L+^{a} (*An. melas*, *mel*) orientation of the 2La rearrangement except *An. gambiae* (gam) and *An. coluzzii* (col), which remain polymorphic for both orientations. The unique phylogenies observed in the 2La region (Fig. 2) are the result of differential loss (×) of the 2La and 2L+^{a} orientations and the introgression of 2La from the ancestral population of *An. gambiae* + *coluzzii* into *An. arabiensis* (dotted arrow). The overall divergence between the 2La and 2L+^{a} orientations inferred from sequences inside the inversion breakpoints is higher, on average, than the predicted divergence time of the species complex. (**B**) ML phylogeny inferred from sequences of the two different orientations of the 2La region (2La and 2L+^{a}) shows that the divergence between opposite orientations is greater than the divergence between the same orientation present in different species (all nodes, 100% bootstrap support). Particularly notable is the separation of the sister taxa *An. gambiae*-2L+^{a} and *An. coluzzii*-2La, as these are known sister taxa. The scale bar denotes nucleotide divergence as calculated by RAxML.

Autosomal permeability of species boundaries

Early cytotaxonomic evidence (10, 12, 27), as well as more recent ribosomal DNA–based evidence (28, 29), supports rare occurrences (<0.1%) of natural female F_1 hybrids between *An. arabiensis*–*An. gambiae* + *An. coluzzii*, *An. arabiensis*–*An. quadriannulatus*, and *An. melas*–*An. gambiae* + *An. coluzzii*, although there are no reports of hybrids involving *An. merus*. The evolutionary importance of these rare hybrids as bridges to interspecific gene flow has remained controversial. Inference of the correct species branching order for the *An. gambiae* complex (Fig. 1B) allowed a systematic analysis of introgression across the genomes of six members of this complex using the *D* (1, 4) and D_{FOIL} (30) statistics (supplementary text S4).

As expected, such tests revealed pervasive introgression across all autosomes between *An. arabiensis* and the ancestor of *An. gambiae* + *An. coluzzii* (figs. S24 and S25). Although introgression was detected in both directions, the majority involved genetic transfer from *An. arabiensis* into the ancestor of *An. gambiae* + *An. coluzzii*. This recent and massive episode of introgression impedes our ability to detect older introgression events between these species. Unexpectedly, we also found evidence of extensive autosomal introgression between another species pair, *An. merus* and *An. quadriannulatus* (Fig. 4, supplementary text S3, and figs. S24 and S25). One of the most striking of the introgressed regions was a contiguous block of genes coincident with the ~22-Mb 3La chromosomal inversion (31). The corresponding sequence originally present in ancestral populations of *An. quadriannulatus* has been entirely replaced by its counterpart from *An. merus*, a conclusion supported by the clustering of *An. merus* with *An. quadriannulatus* in gene trees constructed from sequences in the 3La inversion (figs. S19, C and D, and S21B). Extant populations of both of these species, and indeed all recognized species in the *An. gambiae* complex, are fixed for the standard (3L+^{a}) orientation except *An. melas* and its putative sister species *An. bwambae*, both fixed for the 3La orientation (31). Considering that the exact ~22-Mb 3La region was replaced between species whose contemporary populations are collinear for 3L+^{a}, it is conceivable that ancestral *An. quadriannulatus* populations originally carried 3La before the 3L+^{a} introgression. The expected reduced recombination between

inverted and standard arrangements in inversion heterozygotes may explain why the introgressed region is coincident with the entire 22-Mb 3La region.

The introgression profile was consistent between samples from natural populations and the reference genome assemblies, involving the same species pairs and the same chromosomal locations (supplementary text S4 and fig. S24). Moreover, based on population samples from multiple geographic locations in Africa, patterns of introgression between *An. arabiensis* and *An. gambiae* + *An. coluzzii*, and between *An. merus* and *An. quadriannulatus*, are similar across their geographic range (fig. S25 and table S10). These findings refute the possibility that introgression was an unnatural artifact of colonization and laboratory maintenance of multiple species. The lack of geographic variation in patterns of introgression also suggests that autosomal introgression occurred sufficiently long ago to have spread across subpopulations. By contrast, mitochondrial DNA (mtDNA) revealed patterns consistent with ongoing gene flow between *An. arabiensis* and *An. gambiae* or *An. coluzzii* (supplementary text 3.3 and fig. S22).

Transspecific inversion polymorphism

The unusually high sequence divergence between alternative orientations of a chromosomal inversion polymorphic within *An. gambiae* and *An. coluzzii* (2La and 2L+^{a}) has been noted previously (*32*, *33*) but has not been adequately explained. Other species in the complex are fixed for either 2La (*An. arabiensis* and *An. merus*) or 2L+^{a} (*An. quadriannulatus* and *An. melas*) (*31*) (Fig. 5A). In *An. gambiae* and *An. coluzzii*, the 2La arrangement has been shown to confer superior resistance to desiccating environments (*34*, *35*) relative to 2L+^{a}, and its frequency correlates with environmental gradients of aridity (*10*, *32*). It has been argued that 2La introgressed from *An. arabiensis* into the presumed 2L+^{a} ancestor of *An. gambiae* + *An. coluzzii* (*36*), a crucial step facilitating the range expansion of a presumed forest-adapted species into drier savannas. Sequence divergence-based estimates of the age of 2La and 2L+^{a} arrangements relative to the age of the species complex (Fig. 1C) suggest a radically different scenario in which 2La/2L+^{a} is an ancient inversion polymorphism that predates the initial diversification of the entire complex but is maintained as polymorphic only in *An. gambiae* and *An. coluzzii* (Fig. 5). Consistent with this scenario, the topology of the gene tree built from sequences inside the inversion boundaries indicates that species are grouped by their 2La or 2L+^{a} karyotype (Fig. 5B). Furthermore, contrary to the longstanding assumption (*36*), our data suggest that 2La introgressed from ancestral *An. gambiae* into *An. arabiensis*, not the other way around—eventually replacing the *An. arabiensis* 2L+^{a} arrangement (Fig. 5A). Our inference about the direction of 2La introgression, as well as the underlying phylogenetic hypothesis for the *An. gambiae* complex, are consistent with the genome-enabled chromosomal inversion phylogeny of the *An. gambiae* complex reconstructed from ancestral and derived gene orders at the breakpoints of 10 fixed chromosomal inversions (supplementary text S5 and fig. S27).

Functional insights from differential genetic exchange

We found that introgression mainly involved the autosomes. Our data suggest that the X chromosome is largely resistant to introgression, consistent with studies in this and other systems indicating that the X (or Z) disproportionately harbors factors responsible for reproductive isolation (*3*, *6*, *37*–*41*). The nature, number, and chromosomal organization of these X-linked factors are unsolved puzzles for future research, but our data offer one tantalizing clue. *An. gambiae* males deliver to females large amounts of 20-hydroxyecdysone (20E) (*42*), a steroid hormone that increases female fertility (*43*) and fecundity (*44*) and regulates mating behavior and success (*45*). The cytochrome p450 gene *CYP315A1* (AGAP000284) that synthesizes the 20E precursor ecdysone is located near the distal end of the X chromosome. Furthermore, this region is associated with male hybrid sterility between *An. gambiae* and *An. arabiensis*, with the *An. gambiae* X chromosome causing inviability in an *An. arabiensis* genetic background (*40*). Combined with our data showing that male *An. arabiensis* produce significantly less 20E than *An. gambiae* (supplementary text S6 and fig. S28), these observations prompt the hypothesis that divergence in 20E function between the two species may have a role in speciation through possible effects on the reproductive fitness of hybrid males.

Pervasive autosomal introgression between *An. arabiensis* and the *An. gambiae*–clade ancestor is consistent with the paucity of sterility factors across much of the autosomes, although several autosomal quantitative trait loci have been mapped in both species (*40*). Accordingly, we explored the small subset of autosomal genes (n = 485) that showed no indication of introgression (supplementary text S6), as these are candidates contributing to reproductive isolation. We found a remarkable overrepresentation of genes encoding cyclic-nucleotide phosphodiesterases, enzymes that regulate the levels of the messengers cyclic adenosine monophosphate and cyclic guanosine monophosphate (*46*), which in turn control (among other processes) ecdysone synthesis (*47*, *48*) (table S16).

Implications for the evolution of vectorial capacity

Initial radiation of the *An. gambiae* complex was both recent and rapid. Counter to the traditional view (*49*), it is now clear that the ancestor of the principal malaria vectors *An. gambiae* and *An. coluzzii* separated from other species in the group approximately 2 million years ago (Ma) and that *An. gambiae* and *An. coluzzii* are distantly related to the other primary vector in the group, *An. arabiensis*. Extant populations of *An. gambiae* and *An. coluzzii* are highly anthropophilic vectors, dependent upon humans for blood meals, adult shelter, and larval breeding sites, yet anthropogenic influences are unlikely to have triggered their cladogenesis an estimated 0.5 Ma. Instead, anthropophilic traits are likely to have developed in conjunction with the expansion of Neolithic human populations that occurred more recently. Despite a history of extensive introgression with *An. arabiensis*, *An. gambiae* and *An. coluzzii* are behaviorally, physiologically, ecologically, and epidemiologically distinct from *An. arabiensis*; the same is true for *An. merus* and *An. quadriannulatus*. Notably, experimental introgressions of certain autosomal inversions result in stable heterotic polymorphisms, whereas other introgressed autosomal and X chromosome inversions are rapidly eliminated (*22*), consistent with a role for natural selection in the fate of introgressed regions. Given evidence that the 2La inversion polymorphism is maintained by selection in *An. gambiae* and *An. coluzzii* (*32*, *50*), it seems likely that its introgression into *An. arabiensis* was adaptive, and bidirectional introgressions across the genome between these species probably contributed to their wide ecological flexibility and their vectorial capacity. Broad overlap exists between geographic ranges of these species, and the potential for ongoing hybridization and introgression remains, including the opportunity for introgression of insecticide resistance alleles (*51*). Our study establishes a foundation for further study of adaptive introgression in this species complex and its role in shaping vectorial capacity in this and other malaria vector species complexes.

REFERENCES AND NOTES

1. R. E. Green *et al*., A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010). doi: 10.1126/science.1188021; pmid: 20448178
2. K. K. Dasmahapatra *et al*., Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012). doi: 10.1038/nature11041; pmid: 22722851
3. S. H. Martin *et al*., Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* **23**, 1817–1828 (2013). doi: 10.1101/gr.159426.113; pmid: 24045163
4. 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
5. D. Garrigan *et al*., Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. *Genome Res.* **22**, 1499–1511 (2012). doi: 10.1101/gr.130922.111; pmid: 22534282
6. M. A. White, C. Ané, C. N. Dewey, B. R. Larget, B. A. Payseur, Fine-scale phylogenetic discordance across the house mouse genome. *PLOS Genet.* **5**, e1000729 (2009). doi: 10.1371/journal.pgen.1000729; pmid: 19936022
7. N. Patterson, D. J. Richter, S. Gnerre, E. S. Lander, D. Reich, Genetic evidence for complex speciation of humans and chimpanzees. *Nature* **441**, 1103–1108 (2006). doi: 10.1038/nature04789; pmid: 16710306
8. G. M. Giles, *A handbook of the gnats or mosquitoes giving the anatomy and life history of the Culicidae together with descriptions of all species noticed up to the present date*. (John Bale, Sons & Danielsson, Ltd., London, 1902).
9. G. Davidson, *Anopheles gambiae* complex. *Nature* **196**, 907 (1962). doi: 10.1038/196907a0
10. M. Coluzzi, A. Sabatini, V. Petrarca, M. A. Di Deco, Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans. R. Soc. Trop. Med. Hyg.* **73**, 483–497 (1979). doi: 10.1016/0035-9203(79)90036-1; pmid: 394408
11. B. J. White, F. H. Collins, N. J. Besansky, Evolution of *Anopheles gambiae* in relation to humans and malaria. *Annu. Rev. Ecol. Evol. Syst.* **42**, 111–132 (2011). doi: 10.1146/annurev-ecolsys-102710-145028

12. G. B. White, *Anopheles gambiae* complex and disease transmission in Africa. *Trans. R. Soc. Trop. Med. Hyg.* **68**, 278–301 (1974). doi: 10.1016/0035-9203(74)90035-2; pmid: 4420769
13. M. Coetzee *et al*., *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. *Zootaxa* **3619**, 246–274 (2013). doi: 10.11646/zootaxa.3619.3.2
14. N. J. Besansky *et al*., Molecular phylogeny of the *Anopheles gambiae* complex suggests genetic introgression between principal malaria vectors. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6885–6888 (1994). doi: 10.1073/pnas.91.15.6885; pmid: 8041714
15. B. A. García, A. Caccone, K. D. Mathiopoulos, J. R. Powell, Inversion monophyly in African anopheline malaria vectors. *Genetics* **143**, 1313–1320 (1996). pmid: 8807303
16. N. J. Besansky *et al*., Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: Evidence from multilocus DNA sequence variation. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10818–10823 (2003). doi: 10.1073/pnas.1434337100; pmid: 12947038
17. R. E. Harbach, in *Anopheles Mosquitoes: New Insights into Malaria Vectors*, S. Manguin, Ed. (InTech, DOI: 10.5772/54695, 2013), chap. 1, pp. 3–55. doi: 10.5772/54695
18. D. E. Neafsey *et al*., *Science* **346**, 10.1126/science.1258522 (2014).
19. R. A. Holt *et al*., The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**, 129–149 (2002). doi: 10.1126/science.1076181; pmid: 12364791
20. M. K. Lawniczak *et al*., Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. *Science* **330**, 512–514 (2010). doi: 10.1126/science.1195755; pmid: 20966253
21. Supplementary text, figures, and tables are available on *Science* Online.
22. A. della Torre, L. Merzagora, J. R. Powell, M. Coluzzi, Selective introgression of paracentric inversions between two sibling species of the *Anopheles gambiae* complex. *Genetics* **146**, 239–244 (1997). pmid: 9136013
23. B. Charlesworth, J. A. Coyne, N. H. Barton, The relative rates of evolution of sex chromosomes and autosomes. *Am. Nat.* **130**, 113–146 (1987). doi: 10.1086/284701
24. D. J. Eernisse, A. G. Kluge, Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Mol. Biol. Evol.* **10**, 1170–1195 (1993). pmid: 8277850
25. S. V. Edwards, Is a new and general theory of molecular systematics emerging? *Evolution* **63**, 1–19 (2009). doi: 10.1111/j.1558-5646.2008.00549.x; pmid: 19146594
26. J. H. Degnan, N. A. Rosenberg, Discordance of species trees with their most likely gene trees. *PLOS Genet.* **2**, e68 (2006). doi: 10.1371/journal.pgen.0020068; pmid: 16733550
27. J. H. Bryan, Observations on the member species of the *Anopheles gambiae* complex in The Gambia, West Africa. *Trans. R. Soc. Trop. Med. Hyg.* **73**, 463–466 (1979). doi: 10.1016/0035-9203(79)90179-2; pmid: 555076
28. E. A. Temu, R. H. Hunt, M. Coetzee, J. N. Minjas, C. J. Shiff, Detection of hybrids in natural populations of the *Anopheles gambiae* complex by the rDNA-based, PCR method. *Ann. Trop. Med. Parasitol.* **91**, 963–965 (1997). doi: 10.1080/00034989760383; pmid: 9579219
29. C. Costantini *et al*., Living at the edge: Biogeographic patterns of habitat segregation conform to speciation by niche expansion in *Anopheles gambiae*. *BMC Ecol.* **9**, 16 (2009). doi: 10.1186/1472-6785-9-16; pmid: 19460144
30. J. B. Pease, M. W. Hahn, Detection and polarization of introgression in a five-taxon phylogeny. *BioRxiv* 10.1101/004689 (2014). doi: 10.1101/004689
31. M. Coluzzi, A. Sabatini, A. della Torre, M. A. Di Deco, V. Petrarca, A polytene chromosome analysis of the *Anopheles gambiae* species complex. *Science* **298**, 1415–1418 (2002). doi: 10.1126/science.1077769; pmid: 12364623
32. C. Cheng *et al*., Ecological genomics of *Anopheles gambiae* along a latitudinal cline: A population-resequencing approach. *Genetics* **190**, 1417–1432 (2012). doi: 10.1534/genetics.111.137794; pmid: 22209907
33. B. J. White *et al*., The population genomics of trans-specific inversion polymorphisms in *Anopheles gambiae*. *Genetics* **183**, 275–288 (2009). doi: 10.1534/genetics.109.105817; pmid: 19581444
34. C. Fouet, E. Gray, N. J. Besansky, C. Costantini, Adaptation to aridity in the malaria mosquito *Anopheles gambiae*: Chromosomal inversion polymorphism and body size influence resistance to desiccation. *PLOS ONE* **7**, e34841 (2012). doi: 10.1371/journal.pone.0034841; pmid: 22514674
35. E. M. Gray, K. A. Rocca, C. Costantini, N. J. Besansky, Inversion 2La is associated with enhanced desiccation resistance in *Anopheles gambiae*. *Malar. J.* **8**, 215 (2009). doi: 10.1186/1475-2875-8-215; pmid: 19772577
36. J. R. Powell, V. Petrarca, A. della Torre, A. Caccone, M. Coluzzi, Population structure, speciation, and introgression in the *Anopheles gambiae* complex. *Parassitologia* **41**, 101–113 (1999). pmid: 10697841
37. S. Sankararaman *et al*., The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357 (2014). doi: 10.1038/nature12961; pmid: 24476815
38. M. Meyer *et al*., A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012). doi: 10.1126/science.1224344; pmid: 22936568
39. J. A. Coyne, H. A. Orr, *Speciation* (Sinauer Associates, Sunderland, MA, 2004).
40. M. Slotman, A. Della Torre, J. R. Powell, The genetics of inviability and male sterility in hybrids between *Anopheles gambiae* and *An. arabiensis*. *Genetics* **167**, 275–287 (2004). doi: 10.1534/genetics.167.1.275; pmid: 15166154
41. S. M. O'Loughlin *et al*., Genomic analyses of three malaria vectors reveals extensive shared polymorphism but contrasting population histories. *Mol. Biol. Evol.* **31**, 889–902 (2014). doi: 10.1093/molbev/msu040; pmid: 24408911
42. E. Pondeville, A. Maria, J. C. Jacques, C. Bourgouin, C. Dauphin-Villemant, *Anopheles gambiae* males produce and transfer the vitellogenic steroid hormone 20-hydroxyecdysone to females during mating. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19631–19636 (2008). doi: 10.1073/pnas.0809264105; pmid: 19060216
43. W. R. Shaw *et al*., Mating activates the heme peroxidase HPX15 in the sperm storage organ to ensure fertility in *Anopheles gambiae*. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 5854–5859 (2014). doi: 10.1073/pnas.1401715111; pmid: 24711401
44. F. Baldini *et al*., The interaction between a sexually transferred steroid hormone and a female protein regulates oogenesis in the malaria mosquito *Anopheles gambiae*. *PLOS Biol.* **11**, e1001695 (2013). doi: 10.1371/journal.pbio.1001695; pmid: 24204210
45. P. Gabrieli *et al*., Sexual transfer of the steroid hormone 20E induces the postmating switch in *Anopheles gambiae*. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 16353–16358 (2014). doi: 10.1073/pnas.1410488111; pmid: 25368171
46. A. T. Bender, J. A. Beavo, Cyclic nucleotide phosphodiesterases: Molecular regulation to clinical use. *Pharmacol. Rev.* **58**, 488–520 (2006). doi: 10.1124/pr.58.3.5; pmid: 16968949
47. B. A. Cooke, Signal transduction involving cyclic AMP-dependent and cyclic AMP-independent mechanisms in the control of steroidogenesis. *Mol. Cell. Endocrinol.* **151**, 25–35 (1999). doi: 10.1016/S0303-7207(98)00255-X; pmid: 10411317
48. G. Manière, E. Vanhems, F. Gautron, J. P. Delbecque, Inhibition of ovarian steroidogenesis by cyclic GMP in a fly. *J. Endocrinol.* **177**, 35–44 (2003). doi: 10.1677/joe.0.1770035; pmid: 12697035
49. F. J. Ayala, M. Coluzzi, Chromosome speciation: Humans, *Drosophila*, and mosquitoes. *Proc. Natl. Acad. Sci. U.S.A.* **102** (suppl. 1), 6535–6542 (2005). doi: 10.1073/pnas.0501847102; pmid: 15851677
50. B. J. White *et al*., Localization of candidate regions maintaining a common polymorphic inversion (2La) in *Anopheles gambiae*. *PLOS Genet.* **3**, e217 (2007). doi: 10.1371/journal.pgen.0030217; pmid: 18069896
51. H. D. Mawejje *et al*., Insecticide resistance monitoring of field-collected *Anopheles gambiae s.l.* populations from Jinja, eastern Uganda, identifies high levels of pyrethroid resistance. *Med. Vet. Entomol.* **27**, 276–283 (2013). doi: 10.1111/j.1365-2915.2012.01055.x; pmid: 23046446

ACKNOWLEDGMENTS

We thank M. Kern and B. J. White for early assistance with this project. Genome assemblies and sequence reads generated by the Broad and the BGI have been submitted to National Center for Biotechnology Information, NIH, under two umbrella BioProject IDs: PRJNA6751 and PRJNA254046. The mitochondrial genome alignments, VCF files of field individual samples, the Multiple Alignment Format *Threaded-Blockset Aligner* (MAF TBA) genome alignment of reference assemblies, and the MAF genome alignment from high depth field samples are available on Dryad: doi:10.5061/dryad.f4114. Sequencing was funded at the Broad Institute by a grant from the National Human Genome Research Institute. Individual laboratories were funded by the NIH [R01AI076584 (N.J.B., M.W.H.), R21AI101459 (N.J.B., H.A.S.), R21AI094289 and R21AI099528 (I.V.S.), R01AI104956 (F.C.), R01AI085079 (M.A.S.), HHSN272200900039C (S.J.E.), T32GM007757 (J.B.P.)], Foundation for the National Institutes of Health through the Vector-based Control of Transmission Discovery Research program of the Grand Challenges in Global Health Initiative (N.J.B., M.C.F.), Marie Curie International Outgoing Fellowship (R.M.W.), Department of Education Graduate Assistance in Areas of National Need fellowship (A.S.), Richard and Peggy Notebaert Premier Fellowship and Jack Kent Cooke Foundation Graduate Scholarship (R.R.L.), Fralin Life Science Institute of Virginia Tech (I.V.S.), U.K. Medical Research Council Career Development Award (M.K.L.), European Research Council FP7 (Starting Grant 260897 to F.C.). Author contributions: conceived project: N.J.B.; coordinated project: N.J.B., M.C.F.; supervised project at the Broad Institute: D.E.N.; genome alignments: R.M.W.; genomic data processing, read mapping, development of pipelines for variant calling and scripts for coordinate conversion, mtDNA genome reconstruction: A.S., M.C.F., S.J.E.; provided access to *An. gambiae* genomic sequences from Mali: R.R.L.; window-based phylogeny reconstruction, development and implementation of chromoplots, development and implementation of divergence-based method to infer species branching order; estimation of sequence-based divergence times: J.B.P., M.W.H.; introgression testing: J.B.P., M.C.F., M.W.H.; polymorphism and divergence analyses, phylogenomic analyses of nuclear and mtDNA genomes: M.C.F.; orthology predictions and gene-based phylogenetic reconstruction: R.M.W., Y.-C.W.; identification of inversion breakpoints, chromosomal inversion phylogeny, estimation of divergence time based on rates of rearrangement: I.V.S., A.B.H., X.J.; genomic library construction, H.A.S., M.C.F.; functional enrichment analyses: H.A.S., M.C.F., J.B.P.; ecdysteroid quantification: F.C., E.K., S.N.M.; molecular evolution: D.E.N., M.K.L., M.A.S.; wrote paper: N.J.B., M.C.F., M.W.H. with input from other authors.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/1258524/suppl/DC1
Supplementary Text
Figs. S1 to S28
Tables S1 to S16
References (*52*–*110*)

9 July 2014; accepted 14 November 2014
Published online 27 November 2014;
10.1126/science.1258524

RESEARCH ARTICLE SUMMARY

MOSQUITO GENOMICS

Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes

Daniel E. Neafsey,*† Robert M. Waterhouse,* *et al.*

INTRODUCTION: Control of mosquito vectors has historically proven to be an effective means of eliminating malaria. Human malaria is transmitted only by mosquitoes in the genus *Anopheles*, but not all species within the genus, or even all members of each vector species, are efficient malaria vectors. Variation in vectorial capacity for human malaria among *Anopheles* mosquito species is determined by many factors, including behavior, immunity, and life history.

RATIONALE: This variation in vectorial capacity suggests an underlying genetic/genomic plasticity that results in variation of key traits determining vectorial capacity within the genus. Sequencing the genome of *Anopheles gambiae*, the most important malaria vector in sub-Saharan Africa, has offered numerous insights into how that species became highly specialized to live among and feed upon humans and how susceptibility to mosquito control strategies is determined. Until very recently, similar genomic resources have not existed for other anophelines, limiting comparisons to individual genes or sets of genomic markers with no genome-wide data to investigate attributes associated with vectorial capacity across the genus.

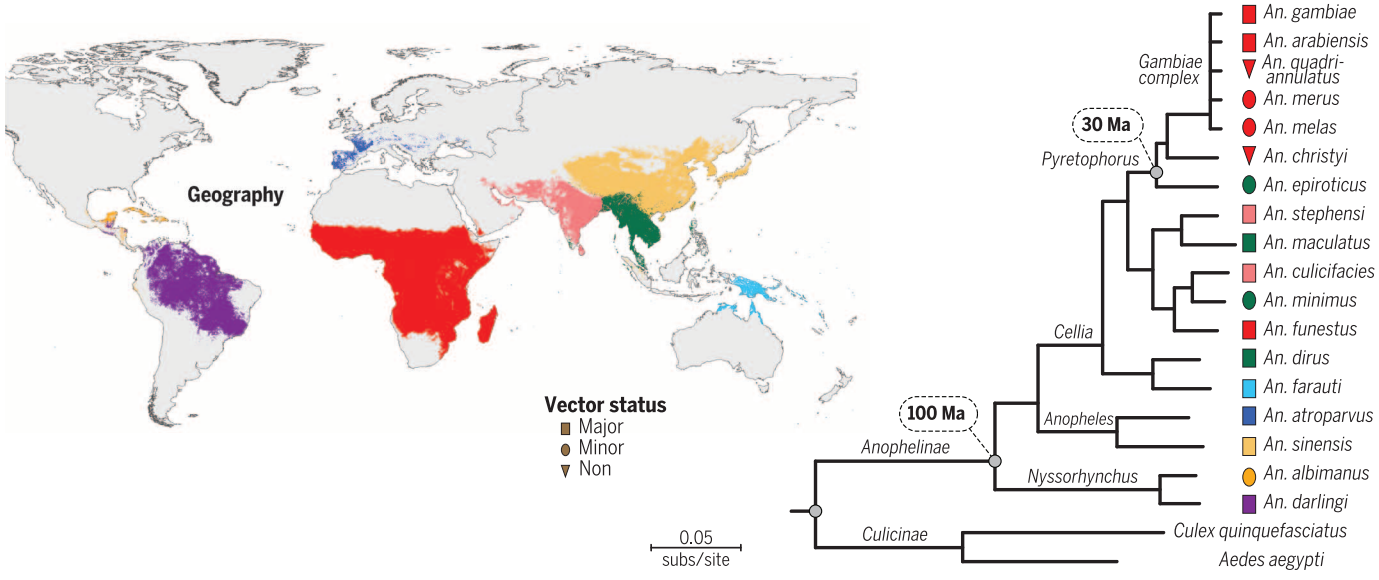
RESULTS: We sequenced and assembled the genomes and transcriptomes of 16 anophelines from Africa, Asia, Europe, and Latin America, spanning ~100 million years of evolution and chosen to represent a range of evolutionary distances from *An. gambiae*, a variety of geographic locations and ecological conditions, and varying degrees of vectorial capacity. Genome assembly quality reflected DNA template quality and homozygosity. Despite variation in contiguity, the assemblies were remarkably complete and searches for arthropod-wide single-copy orthologs generally revealed few missing genes. Genome annotation supported with RNA sequencing transcriptomes yielded between 10,738 and 16,149 protein-coding genes for each species. Relative to *Drosophila,* the closest dipteran genus for which equivalent genomic resources exist, *Anopheles* exhibits a dynamic genomic evolutionary profile. Comparative analyses show a fivefold faster rate of gene gain and loss, elevated gene shuffling on the X chromosome, and more intron losses in *Anopheles*. Some determinants of vectorial capacity, such as chemosensory genes, do not show elevated turnover but instead diversify through protein-sequence changes. We also document evidence of variation in important reproductive phenotypes, genes controlling immunity to *Plasmodium* malaria parasites and other microbes, genes encoding cuticular and salivary proteins, and genes conferring metabolic insecticide resistance. This dynamism of anopheline genes and genomes may contribute to their flexible capacity to take advantage of new ecological niches, including adapting to humans as primary hosts.

CONCLUSIONS: Anopheline mosquitoes exhibit a molecular evolutionary profile very distinct from *Drosophila*, and their genomes harbor strong evidence of functional variation in traits that determine vectorial capacity. These 16 new reference genome assemblies provide a foundation for hypothesis generation and testing to further our understanding of the diverse biological traits that determine vectorial capacity. ■

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1258522

The complete list of authors and affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. E-mail: neafsey@broadinstitute.org (D.E.N.); nbesansk@nd.edu (N.J.B.)
Cite this article as D. E. Neafsey *et al.*, *Science* **347**, 1258522 (2015). DOI: 10.1126/science.1258522



Geography, vector status, and molecular phylogeny of the 16 newly sequenced anopheline mosquitoes and selected other dipterans. The maximum likelihood molecular phylogeny of all sequenced anophelines and two mosquito outgroups was constructed from the aligned protein sequences of 1085 single-copy orthologs. Shapes between branch termini and species names indicate vector status and are colored according to geographic ranges depicted on the map. Ma, million years ago.

RESEARCH ARTICLE

MOSQUITO GENOMICS

Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes

Daniel E. Neafsey,[1*†] Robert M. Waterhouse,[2,3,4,5*] Mohammad R. Abai,[6] Sergey S. Aganezov,[7] Max A. Alekseyev,[7] James E. Allen,[8] James Amon,[9] Bruno Arcà,[10] Peter Arensburger,[11] Gleb Artemov,[12] Lauren A. Assour,[13] Hamidreza Basseri,[6] Aaron Berlin,[1] Bruce W. Birren,[1] Stephanie A. Blandin,[14,15] Andrew I. Brockman,[16] Thomas R. Burkot,[17] Austin Burt,[18] Clara S. Chan,[2,3] Cedric Chauve,[19] Joanna C. Chiu,[20] Mikkel Christensen,[8] Carlo Costantini,[21] Victoria L. M. Davidson,[22] Elena Deligianni,[23] Tania Dottorini,[16] Vicky Dritsou,[24] Stacey B. Gabriel,[25] Wamdaogo M. Guelbeogo,[26] Andrew B. Hall,[27] Mira V. Han,[28] Thaung Hlaing,[29] Daniel S. T. Hughes,[8,30] Adam M. Jenkins,[31] Xiaofang Jiang,[32,27] Irwin Jungreis,[2,3] Evdoxia G. Kakani,[33,34] Maryam Kamali,[35] Petri Kemppainen,[36] Ryan C. Kennedy,[37] Ioannis K. Kirmitzoglou,[16,38] Lizette L. Koekemoer,[39] Njoroge Laban,[40] Nicholas Langridge,[8] Mara K. N. Lawniczak,[16] Manolis Lirakis,[41] Neil F. Lobo,[42] Ernesto Lowy,[8] Robert M. MacCallum,[16] Chunhong Mao,[43] Gareth Maslen,[8] Charles Mbogo,[44] Jenny McCarthy,[11] Kristin Michel,[22] Sara N. Mitchell,[33] Wendy Moore,[45] Katherine A. Murphy,[20] Anastasia N. Naumenko,[35] Tony Nolan,[16] Eva M. Novoa,[2,3] Samantha O'Loughlin,[18] Chioma Oringanje,[45] Mohammad A. Oshaghi,[6] Nazzy Pakpour,[46] Philippos A. Papathanos,[16,24] Ashley N. Peery,[35] Michael Povelones,[47] Anil Prakash,[48] David P. Price,[49,50] Ashok Rajaraman,[19] Lisa J. Reimer,[51] David C. Rinker,[52] Antonis Rokas,[52,53] Tanya L. Russell,[17] N'Fale Sagnon,[26] Maria V. Sharakhova,[35] Terrance Shea,[1] Felipe A. Simão,[4,5] Frederic Simard,[21] Michel A. Slotman,[54] Pradya Somboon,[55] Vladimir Stegniy,[12] Claudio J. Struchiner,[56,57] Gregg W. C. Thomas,[58] Marta Tojo,[59] Pantelis Topalis,[23] José M. C. Tubio,[60] Maria F. Unger,[42] John Vontas,[41] Catherine Walton,[36] Craig S. Wilding,[61] Judith H. Willis,[62] Yi-Chieh Wu,[2,3,63] Guiyun Yan,[64] Evgeny M. Zdobnov,[4,5] Xiaofan Zhou,[53] Flaminia Catteruccia,[33,34] George K. Christophides,[16] Frank H. Collins,[42] Robert S. Cornman,[62] Andrea Crisanti,[16,24] Martin J. Donnelly,[51,65] Scott J. Emrich,[13] Michael C. Fontaine,[42,66] William Gelbart,[67] Matthew W. Hahn,[68,58] Immo A. Hansen,[49,50] Paul I. Howell,[69] Fotis C. Kafatos,[16] Manolis Kellis,[2,3] Daniel Lawson,[8] Christos Louis,[41,23,24] Shirley Luckhart,[46] Marc A. T. Muskavitch,[31,70] José M. Ribeiro,[71] Michael A. Riehle,[45] Igor V. Sharakhov,[35,27] Zhijian Tu,[27,32] Laurence J. Zwiebel,[72] Nora J. Besansky[42†]

Variation in vectorial capacity for human malaria among *Anopheles* mosquito species is determined by many factors, including behavior, immunity, and life history. To investigate the genomic basis of vectorial capacity and explore new avenues for vector control, we sequenced the genomes of 16 anopheline mosquito species from diverse locations spanning ~100 million years of evolution. Comparative analyses show faster rates of gene gain and loss, elevated gene shuffling on the X chromosome, and more intron losses, relative to *Drosophila*. Some determinants of vectorial capacity, such as chemosensory genes, do not show elevated turnover but instead diversify through protein-sequence changes. This dynamism of anopheline genes and genomes may contribute to their flexible capacity to take advantage of new ecological niches, including adapting to humans as primary hosts.

Malaria is a complex disease, mediated by obligate eukaryotic parasites with a life cycle requiring adaption to both vertebrate hosts and mosquito vectors. These relationships create a rich coevolutionary triangle. Just as *Plasmodium* parasites have adapted to their diverse hosts and vectors, infection by *Plasmodium* parasites has reciprocally induced adaptive evolutionary responses in humans and other vertebrates (*1*) and has also influenced mosquito evolution (*2*). Human malaria is transmitted only by mosquitoes in the genus *Anopheles*, but not all species within the genus, or even all members of each vector species, are efficient malaria vectors. This suggests an underlying genetic/genomic plasticity that results in variation of key traits determining vectorial capacity within the genus.

In all, five species of *Plasmodium* have adapted to infect humans and are transmitted by ~60 of the 450 known species of anopheline mosquitoes (*3*). Sequencing the genome of *Anopheles gambiae*, the most important malaria vector in sub-Saharan Africa, has offered numerous insights into how that species became highly specialized to live among and feed upon humans and how susceptibility to mosquito control strategies is determined (*4*). Until very recently (*5*–*7*), similar genomic resources have not existed for other anophelines, limiting comparisons to individual genes or sets of genomic markers with no genome-wide data to investigate attributes associated with vectorial capacity across the genus.

Thus, we sequenced and assembled the genomes and transcriptomes of 16 anophelines from Africa, Asia, Europe, and Latin America. We chose these 16 species to represent a range of evolutionary distances from *An. gambiae*, a variety of geographic locations and ecological conditions, and varying degrees of vectorial capacity (*8*) (Fig. 1, A and B). For example, *An. quadriannulatus*, although extremely closely related to *An. gambiae*, feeds preferentially on bovines rather than humans, limiting its potential to transmit human malaria. *An. merus, An. melas*, *An. farauti*, and *An. albimanus* females can lay eggs in salty or brackish water, instead of the freshwater sites required by other species. With a focus on species most closely related to *An. gambiae* (*9*), the sampled anophelines span the three main subgenera that shared a common ancestor ~100 million years ago (Ma) (*10*).

Materials and methods summary

Genomic DNA and whole-body RNA were obtained from laboratory colonies and wild-caught specimens (tables S1 and S2), with samples for nine species procured from newly established isofemale colonies to reduce heterozygosity. Illumina sequencing libraries spanning a range of insert sizes were constructed, with ~100-fold paired-end 101–base pair (bp) coverage generated for small (180 bp) and medium (1.5 kb) insert libraries and lower coverage for large (38 kb) insert libraries (table S3). DNA template for the small and medium input libraries was sourced from single female mosquitoes from each species to further reduce heterozygosity. High-molecular-weight DNA template for each large insert library was derived from pooled DNA obtained from several hundred mosquitoes. ALLPATHS-LG (*11*) genome assemblies were produced using the "haploidify" option to reduce haplotype assemblies caused by high heterozygosity. Assembly quality reflected DNA template quality and homozygosity, with a mean scaffold N50 of 3.6 Mb, ranging to 18.1 Mb for *An. albimanus* (table S4). Despite variation in contiguity, the assemblies were remarkably complete and searches for arthropod-wide single-copy orthologs generally revealed few missing genes (fig. S1) (*12*).

Genome annotation with MAKER (*13*) supported with RNA sequencing (RNAseq) of transcriptomes (produced from pooled male and female larvae, pupae, and adults) (table S5) and comprehensive noncoding RNA gene prediction (fig. S2) yielded relatively complete gene sets

(fig. S3), with between 10,738 and 16,149 protein-coding genes identified for each species. Gene count was generally commensurate with assembly contiguity (table S6). Some of this variation in total gene counts may be attributed to the challenges of gene annotations with variable levels of assembly contiguity and supporting RNAseq data. To estimate the prevalence of erroneous gene model fusions and/or fragmentations, we compared the new gene annotations to *An. gambiae* gene models and found an average of 3.3% and 9.7% potentially fused and fragmented gene models, respectively. For analyses described below that may be sensitive to variation in gene model accuracy or gene set completeness, we have conducted sensitivity analyses to rule out confounding results from these factors (*12*).

Rapidly evolving genes and genomes

Orthology delineation identified lineage-restricted and species-specific genes, as well as ancient genes found across insect taxa, of which universal single-copy orthologs were employed to estimate the molecular species phylogeny (Fig. 1, B and C, and fig. S4). Analysis of codon frequencies in these orthologs revealed that anophelines, unlike drosophilids, exhibit relatively uniform codon usage preferences (fig. S5).

Polytene chromosomes have provided a glimpse into anopheline chromosome evolution (*14*). Our genome-sequence–based view confirmed the cytological observations and offers many new insights. At the base-pair level, ~90% of the non-gapped and nonmasked *An. gambiae* genome (i.e., excluding transposable elements, as detailed in table S7) is alignable to the most closely related species, whereas only ~13% aligns to the most distant (Fig. 1D, fig. S6, and table S8), with reduced alignability in centromeres and on the X chromosome (Fig. 1D). At chromosomal levels, mapping data anchored 35 to 76% of the *An. stephensi*, *An. funestus*, *An. atroparvus*, and *An. albimanus* genome assemblies to chromosomal arms (tables S9 to S12). Analysis of genes in anchored regions showed that synteny at the whole-arm level is highly conserved, despite several whole-arm translocations (Fig. 2A and table S13). In contrast, small-scale rearrangements disrupt gene colinearity within arms over time, leading to extensive shuffling of gene order over a time scale of 29 million years or more (*10*, *15*) (Fig. 2B and fig. S7). As in *Drosophila*, rearrangement rates are higher on the X chromosome than on autosomes (Fig. 2C and tables S14 to S16). However, the difference is significantly more pronounced in *Anopheles*, where X chromosome rearrangements are more frequent by a factor of 2.7 than autosomal rearrangements; in *Drosophila*, the corresponding ratio is only 1.2 (*t* test, $t_{10} = 7.3$; $P < 1 \times 10^{-5}$) (fig. S8). The X chromosome is also notable for a significant degree of observed gene movement to other chromosomes relative to *Drosophila* (one sample proportion test, $P < 2.2 \times 10^{-16}$) (Fig. 2D and tables S17 and S18), as was previously noted for *Anopheles* relative to *Aedes* (*16*), further underscoring its distinctive evolutionary profile in *Anopheles* compared with other dipteran genera.

Such dynamic gene shuffling and movement may be facilitated by the multiple families of DNA transposons and long terminal repeat (LTR) and non-LTR retroelements found in all genomes (table S7), as well as a weaker dosage compensation phenotype in *Anopheles* compared with *Drosophila* (*17*). Despite such shuffling, comparing genomic locations of orthologs can be successfully employed to reconstruct ancestral chromosomal arrangements (fig. S9) and to confidently improve assembly contiguity (tables S19 to S21).

Copy-number variation in homologous gene families also reveals striking evolutionary dynamism. Analysis of 11,636 gene families with CAFE 3 (*18*) indicates a rate of gene gain/loss higher by a factor of at least 5 than that observed for 12 *Drosophila* genomes (*19*). Overall, these *Anopheles* genomes exhibit a rate of gain or loss per gene per million years of 3.12×10^{-3} compared with 5.90×10^{-4} for *Drosophila*, suggesting substantially higher gene turnover within anophelines relative to fruit flies. This fivefold greater gain/loss rate in anophelines holds true under models that account for uncertainty in

[1]Genome Sequencing and Analysis Program, Broad Institute, 415 Main Street, Cambridge, MA 02142, USA. [2]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139, USA. [3]The Broad Institute of Massachusetts Institute of Technology and Harvard, 415 Main Street, Cambridge, MA 02142, USA. [4]Department of Genetic Medicine and Development, University of Geneva Medical School, Rue Michel-Servet 1, 1211 Geneva, Switzerland. [5]Swiss Institute of Bioinformatics, Rue Michel-Servet 1, 1211 Geneva, Switzerland. [6]Department of Medical Entomology and Vector Control, School of Public Health and Institute of Health Researches, Tehran University of Medical Sciences, Tehran, Iran. [7]George Washington University, Department of Mathematics and Computational Biology Institute, 45085 University Drive, Ashburn, VA 20147, USA. [8]European Molecular Biology Laboratory, European Bioinformatics Institute, EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK. [9]National Vector Borne Disease Control Programme, Ministry of Health, Tafea Province, Vanuatu. [10]Department of Public Health and Infectious Diseases, Division of Parasitology, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185 Rome, Italy. [11]Department of Biological Sciences, California State Polytechnic–Pomona, 3801 West Temple Avenue, Pomona, CA 91768, USA. [12]Tomsk State University, 36 Lenina Avenue, Tomsk, Russia. [13]Department of Computer Science and Engineering, Eck Institute for Global Health, 211B Cushing Hall, University of Notre Dame, Notre Dame, IN 46556, USA. [14]Inserm, U963, F-67084 Strasbourg, France. [15]CNRS, UPR9022, IBMC, F-67084 Strasbourg, France. [16]Department of Life Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, UK. [17]Faculty of Medicine, Health and Molecular Science, Australian Institute of Tropical Health Medicine, James Cook University, Cairns 4870, Australia. [18]Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK. [19]Department of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada. [20]Department of Entomology and Nematology, One Shields Avenue, University of California–Davis, Davis, CA 95616, USA. [21]Institut de Recherche pour le Développement, Unités Mixtes de Recherche Maladies Infectieuses et Vecteurs Écologie, Génétique, Évolution et Contrôle, 911, Avenue Agropolis, BP 64501 Montpellier, France. [22]Division of Biology, Kansas State University, 271 Chalmers Hall, Manhattan, KS 66506, USA. [23]Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology, Hellas, Nikolaou Plastira 100 GR-70013, Heraklion, Crete, Greece. [24]Centre of Functional Genomics, University of Perugia, Perugia, Italy. [25]Genomics Platform, Broad Institute, 415 Main Street, Cambridge, MA 02142, USA. [26]Centre National de Recherche et de Formation sur le Paludisme, Ouagadougou 01 BP 2208, Burkina Faso. [27]Program of Genetics, Bioinformatics, and Computational Biology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. [28]School of Life Sciences, University of Nevada, Las Vegas, NV 89154, USA. [29]Department of Medical Research, No. 5 Ziwaka Road, Dagon Township, Yangon 11191, Myanmar. [30]Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA. [31]Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, USA. [32]Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. [33]Harvard School of Public Health, Department of Immunology and Infectious Diseases, Boston, MA 02115, USA. [34]Dipartimento di Medicina Sperimentale e Scienze Biochimiche, Università degli Studi di Perugia, Perugia, Italy. [35]Department of Entomology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. [36]Computational Evolutionary Biology Group, Faculty of Life Sciences, University of Manchester, Oxford Road, Manchester M13 9PT, UK. [37]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA 94143, USA. [38]Bioinformatics Research Laboratory, Department of Biological Sciences, New Campus, University of Cyprus, CY 1678 Nicosia, Cyprus. [39]Wits Research Institute for Malaria, Faculty of Health Sciences, and Vector Control Reference Unit, National Institute for Communicable Diseases of the National Health Laboratory Service, Sandringham 2131, Johannesburg, South Africa. [40]National Museums of Kenya, P.O. Box 40658-00100, Nairobi, Kenya. [41]Department of Biology, University of Crete, 700 13 Heraklion, Greece. [42]Eck Institute for Global Health and Department of Biological Sciences, University of Notre Dame, 317 Galvin Life Sciences Building, Notre Dame, IN 46556, USA. [43]Virginia Bioinformatics Institute, 1015 Life Science Circle, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. [44]Kenya Medical Research Institute-Wellcome Trust Research Programme, Centre for Geographic Medicine Research - Coast, P.O. Box 230-80108, Kilifi, Kenya. [45]Department of Entomology, 1140 East South Campus Drive, Forbes 410, University of Arizona, Tucson, AZ 85721, USA. [46]Department of Medical Microbiology and Immunology, School of Medicine, University of California Davis, One Shields Avenue, Davis, CA 95616, USA. [47]Department of Pathobiology, University of Pennsylvania School of Veterinary Medicine, 3800 Spruce Street, Philadelphia, PA 19104, USA. [48]Regional Medical Research Centre NE, Indian Council of Medical Research, P.O. Box 105, Dibrugarh-786 001, Assam, India. [49]Department of Biology, New Mexico State University, Las Cruces, NM 88003, USA. [50]Molecular Biology Program, New Mexico State University, Las Cruces, NM 88003, USA. [51]Department of Vector Biology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK. [52]Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN 37235, USA. [53]Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA. [54]Department of Entomology, Texas A&M University, College Station, TX 77807, USA. [55]Department of Parasitology, Faculty of Medicine, Chiang Mai University, Chiang Mai 50200, Thailand. [56]Fundação Oswaldo Cruz, Avenida Brasil 4365, RJ Brazil. [57]Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil. [58]School of Informatics and Computing, Indiana University, Bloomington, IN 47405, USA. [59]Department of Physiology, School of Medicine, Center for Research in Molecular Medicine and Chronic Diseases, Instituto de Investigaciones Sanitarias, University of Santiago de Compostela, Santiago de Compostela, A Coruña, Spain. [60]Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. [61]School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool L3 3AF, UK. [62]Department of Cellular Biology, University of Georgia, Athens, GA 30602, USA. [63]Department of Computer Science, Harvey Mudd College, Claremont, CA 91711, USA. [64]Program in Public Health, College of Health Sciences, University of California, Irvine, Hewitt Hall, Irvine, CA 92697, USA. [65]Malaria Programme, Wellcome Trust Sanger Institute, Cambridge CB10 1SJ, UK. [66]Centre of Evolutionary and Ecological Studies (Marine Evolution and Conservation group), University of Groningen, Nijenborgh 7, NL-9747 AG Groningen, Netherlands. [67]Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA. [68]Department of Biology, Indiana University, Bloomington, IN 47405, USA. [69]Centers for Disease Control and Prevention, 1600 Clifton Road NE MSG49, Atlanta, GA 30329, USA. [70]Biogen Idec, 14 Cambridge Center, Cambridge, MA 02142, USA. [71]Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, 12735 Twinbrook Parkway, Rockville, MD 20852, USA. [72]Departments of Biological Sciences and Pharmacology, Institutes for Chemical Biology, Genetics and Global Health, Vanderbilt University and Medical Center, Nashville, TN 37235, USA.

*These authors contributed equally to this work. **†Corresponding author. E-mail: neafsey@broadinstitute.org (D.E.N.); nbesansk@nd.edu (N.J.B.)**

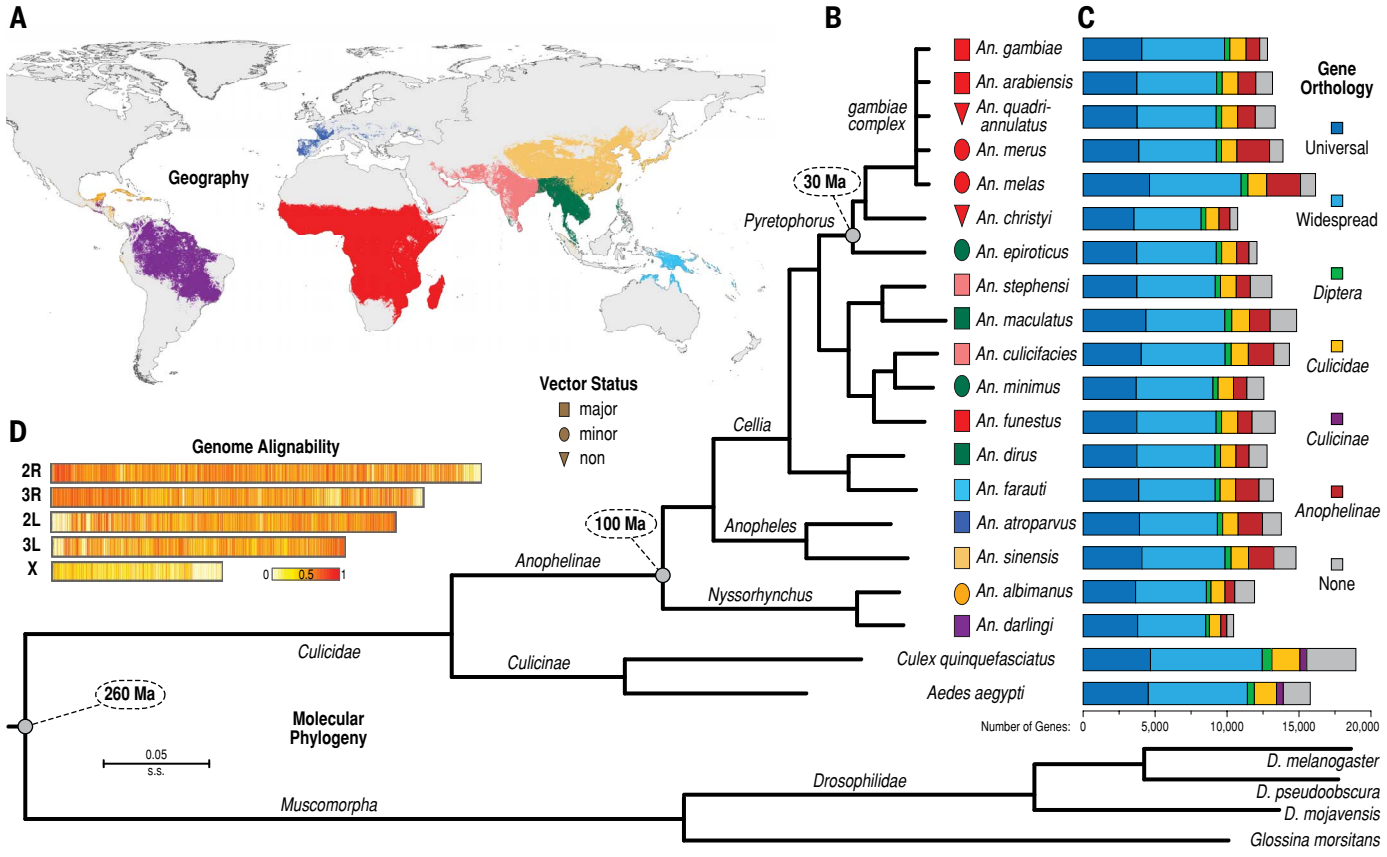


Fig. 1. Geography, vector status, molecular phylogeny, gene orthology, and genome alignability of the 16 newly sequenced anopheline mosquitoes and selected other dipterans. (**A**) Global geographic distributions of the 16 sampled anophelines and the previously sequenced *An. gambiae* and *An. darlingi*. Ranges are colored for each species or group of species as shown in (B), e.g., light blue for *An. farauti*. (**B**) The maximum likelihood molecular phylogeny of all sequenced anophelines and selected dipteran outgroups. Shapes between branch termini and species names indicate vector status (rectangles, major vectors; ellipses, minor vectors, triangles, nonvectors) and are colored according to geographic ranges shown in (A). (**C**) Bar plots show total gene counts for each species partitioned according to their orthology profiles, from ancient genes found across insects to lineage-restricted and species-specific genes. (**D**) Heat map illustrating the density (in 2-kb sliding windows) of whole-genome alignments along the lengths of *An. gambiae* chromosomal arms: from white where *An. gambiae* aligns to no other species to red where *An. gambiae* aligns to all the other anophelines.

gene family sizes at the tips of the species tree due to annotation or assembly errors and is not sensitive to inclusion or exclusion of taxa affecting the root age of the tree nor to the exclusion of taxa with the poorest assemblies and gene sets (fig. S10 and tables S22 and S23). Examples include expansions of cuticular proteins in *An. arabiensis* and neurotransmitter-gated ion channels in *An. albimanus* (table S24).

The evolutionary dynamism of *Anopheles* genes extends to their architecture. Comparisons of single-copy orthologs at deeper phylogenetic depths showed losses of introns at the root of the true fly order Diptera and revealed continued losses as the group diversified into the lineages leading to fruit flies and mosquitoes. However, anopheline orthologs have sustained greater intron loss than drosophilids, leading to a relative paucity of introns in the genes of extant anophelines (fig. S11 and table S25). Comparative analysis also revealed that gene fusion and fission played a substantial role in the evolution of mosquito genes, with apparent rearrangements affecting an average of 10.1% of all genes in the genomes of the 10 species with the most contiguous assemblies (fig. S12). Furthermore, gene boundaries can be flexible; whole genome alignments identified 325 candidates for stop-codon readthrough (fig. S13 and table S26).

Because molecular evolution of protein-coding sequences is a well-known source of phenotypic change, we compared evolutionary rates among different functional categories of anopheline orthologs. We quantified evolutionary divergence in terms of protein sequence identity of aligned orthologs and the d_N/d_S statistic (ratio of nonsynonymous to synonymous substitutions) computed using PAML (*12*, *20*). Among curated sets of genes linked to vectorial capacity or species-specific traits against a background of functional categories defined by Gene Ontology or InterPro annotations, odorant and gustatory receptors show high evolutionary rates and male accessory gland proteins exhibit exceptionally high d_N/d_S ratios (Fig. 3, figs. S14 and S15, and tables S27 to S29). Rapid divergence in functional categories related to malaria transmission and/or mosquito control strategies led us to examine the genomic basis of several facets of anopheline biology in closer detail.

Insights into mosquito biology and vectorial capacity

Mosquito reproductive biology evolves rapidly and presents a compelling target for vector control. This is exemplified by the *An. gambiae* male accessory gland protein (Acp) cluster on chromosome 3R (*21*, *22*), where conservation is mostly lost outside the *An. gambiae* species complex (fig. S16). In *Drosophila*, male-biased genes such as Acps tend to evolve faster than loci without male-biased expression (*23*–*25*). We looked for a similar pattern in anophelines after assessing each gene for sex-biased expression using microarray and RNAseq data sets for *An. gambiae* (*12*). In contrast to *Drosophila,* female-biased genes show dramatically faster rates of evolution across the genus than male-biased genes (Wilcoxon rank sum test, $P = 5 \times 10^{-4}$) (fig. S17).

Differences in reproductive genes among anophelines may provide insight into the origin and function of sex-related traits. During

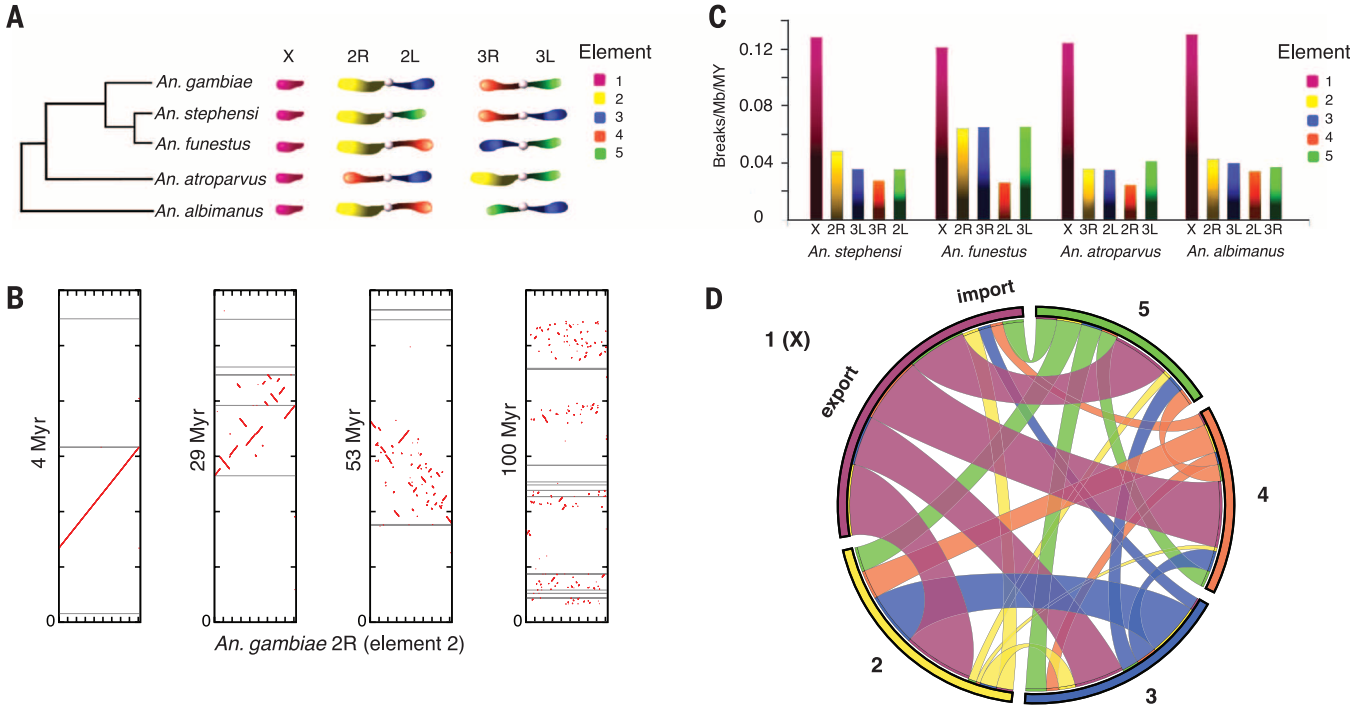


Fig. 2. Patterns of anopheline chromosomal evolution. (**A**) Anopheline genomes have conserved gene membership on chromosome arms ("elements"; colored and labeled 1 to 5). Unlike *Drosophila*, chromosome elements reshuffle between chromosomes via translocations as intact elements and do not show fissions or fusions. The tree depicts the supported molecular topology for the species studied. (**B**) Conserved synteny blocks decay rapidly within chromosomal arms as the phylogenetic distance increases between species. Moving left to right, the dot-plot panels show gene-level synteny between chromosome 2R of *An. gambiae* (*x* axis) and inferred ancestral sequences (*y* axes; inferred using PATHGROUPS) at increasing evolutionary time scales (million years ago) estimated by an ultrametric phylogeny. Gray horizontal lines represent scaffold breaks. Discontinuity of the red lines/dots indicates rearrangement. (**C**) Anopheline X chromosomes exhibit higher rates of rearrangement ($P < 1 \times 10^{-5}$), measured as breaks per Mb per million years, compared with autosomes, despite a paucity of polymorphic inversions on the X. (**D**) The anopheline X chromosome also displays a higher rate of gene movement to other chromosomal arms than any of the autosomes. Chromosomal elements are labeled around the perimeter; internal bands are colored according to the chromosomal element source and match element colors in (A) and (C). Bands are sized to indicate the relative ratio of genes imported versus exported for each chromosomal element and the relative allocation of exported genes to other elements.

copulation, *An. gambiae* males transfer a gelatinous mating plug, a complex of seminal proteins, lipids, and hormones that are essential for successful sperm storage by females and for reproductive success (*26*–*28*). Coagulation of the plug is mediated by a seminal transglutaminase (*TG3*), which is found in anophelines but is absent in other mosquito genera that do not form a mating plug (*26*). We examined *TG3* and its two paralogs (*TG1* and *TG2*) in the sequenced anophelines and investigated the rate of evolution of each gene (Fig. 4A). Silent sites were saturated at the whole-genus level, making d_S difficult to estimate reliably, but *TG1* (the gene presumed to be ancestral owing to broadest taxonomic representation) exhibited the lowest rate of amino acid change (d_N = 0.20), *TG2* exhibited an intermediate rate (d_N = 0.93), and the anopheline-specific *TG3* has evolved even more rapidly (d_N = 1.50), perhaps because of male/male or male/female evolutionary conflict. Note that plug formation appears to be a derived trait within anophelines, because it is not exhibited by *An. albimanus* and intermediate, poorly coagulated plugs were observed in taxa descending from early-branching lineages within the genus (table S30). Functional studies of mating plugs will be necessary to understand what drove the origin and rapid evolution of *TG3*.

Proteins that constitute the mosquito cuticular exoskeleton play important roles in diverse aspects of anopheline biology, including development, ecology, and insecticide resistance, and constitute approximately 2% of all protein-coding genes (*29*). Comparisons among dipterans have revealed numerous amplifications of cuticular protein (CP) genes undergoing concerted evolution at physically clustered loci (*30*–*33*). We investigated the extent and time scale of gene cluster homogenization within anophelines by generating phylogenies of orthologous gene clusters (fig. S18 and table S31). Throughout the genus, these gene clusters often group phylogenetically by species rather than by position within tandem arrays, particularly in a subset of clusters. These include the 3RB and 3RC clusters of CP genes (*30*), the CPLCG group A and CPLCW clusters found elsewhere on 3R (*32*), and six tandemly arrayed genes on 3L designated *CPFL2* through *CPFL7* (*34*). CPLCW genes occur in a head-to-head arrangement with CPLCG group A genes and exhibit highly conserved intergenic sequences (fig. S19). Furthermore, transcript localization studies using in situ hybridization revealed identical spatial expression patterns for CLPCW and CPLCG group A gene pairs suggestive of coregulation (fig. S19). For these five gene clusters, complete grouping by organismal lineage was observed for most deep nodes as well as for many individual species outside the shallow *An. gambiae* species complex (Fig. 4B), consistent with a relatively rapid (less than 20 million years) homogenization of sequences via concerted evolution. The emerging pattern of anopheline CP evolution is thus one of relative stasis for a majority of single-copy orthologs, juxtaposed with consistent concerted evolution of a subset of genes.

Anophelines identify hosts, oviposition sites, and other environmental cues through specialized chemosensory membrane-bound receptors. We examined three of the major gene families that encode these molecules: the odorant receptors (*ORs*), gustatory receptors (*GRs*), and variant ionotropic glutamate receptors (*IRs*). Given the rapid chemosensory gene turnover observed in many other insects, we explored whether varying host preferences of anopheline mosquitoes could be attributed to chemosensory gene gains and losses. Unexpectedly in light of the elevated genome-wide rate of gene turnover, we found that the overall size and content of the chemosensory

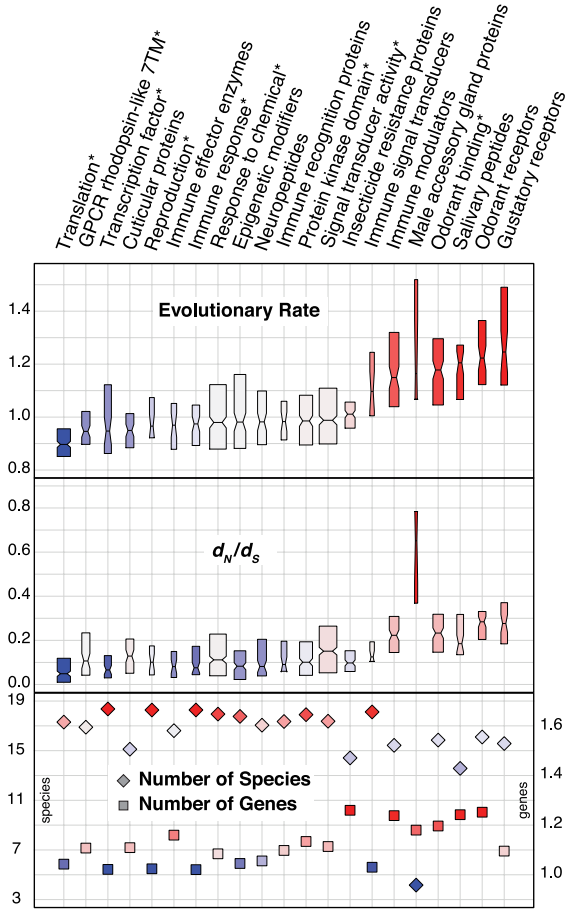


Fig. 3. Contrasting evolutionary properties of selected gene functional categories. Examined evolutionary properties of orthologous groups of genes include a measure of amino acid conservation/divergence (evolutionary rate), a measure of selective pressure (d_N/d_S), a measure of gene duplication in terms of mean gene copy-number per species (number of genes), and a measure of ortholog universality in terms of number of species with orthologs (number of species). Notched box plots show medians and extend to the first and third quartiles; their widths are proportional to the number of orthologous groups in each functional category. Functional categories derive from curated lists associated with various functions/processes as well as annotated Gene Ontology or InterPro categories (denoted by asterisks).

gene repertoire are relatively conserved across the genus. CAFE 3 (*18*) analyses estimated that the most recent common ancestor of the anophelines had approximately 60 genes in each of the *OR* and *GR* families, similar to most extant anophelines (Fig. 4C and fig. S20). Estimated gain/loss rates of *OR* and *GR* genes per million years (error-corrected $\lambda = 1.3 \times 10^{-3}$ for *OR*s and 2.0×10^{-4} for *GR*s) were much lower than the overall level of anopheline gene families. Similarly, we found almost the same number of antennae-expressed *IR*s (~20) in all anopheline genomes. Despite overall conservation in chemosensory gene numbers, we observed several examples of gene gain and loss in specific lineages. Notably, there was a net gain of at least 12 *OR*s in the common ancestor of the *An. gambiae* complex (Fig. 4C).

OR and *GR* gene repertoire stability may derive from their roles in several critical behaviors. Host preference differences are likely to be governed by a combination of functional divergence and transcriptional modulation of orthologs. This model is supported by studies of antennal transcriptomes in the major malaria vector *An. gambiae* (*35*) and comparisons between this vector and its morphologically identical sibling *An. quadriannulatus* (*36*), a very closely related species that plays no role in malaria transmission (despite vectorial competence) because it does not specialize on human hosts. Furthermore, we found that many subfamilies of *OR*s and *GR*s showed evidence of positive selection (19 of 53 *OR*s; 17 of 59 *GR*s) across the genus, suggesting potential functional divergence.

Several blood feeding–related behaviors in mosquitoes are also regulated by peptide hormones (*37*). These peptides are synthesized, processed, and released from nervous and endocrine systems and elicit their effects through binding appropriate receptors in target tissues (*38*). In total, 39 peptide hormones were identified from each of the sequenced anophelines (fig. S21). Notably, no ortholog of the well-characterized head peptide (HP) hormone of the culicine mosquito *Aedes aegypti* was identified in any of the assemblies. In *Ae. aegypti*, HP is responsible for inhibiting host-seeking behavior after a blood meal (*39*). Because anophelines broadly exhibit similar behavior (*40*), the absence of HP from the entire clade suggests they may have evolved a novel mechanism to inhibit excess blood feeding. Similarly, no ortholog of insulin growth factor 1 (*IGF1*) was identified in any anophelines even though *IGF1* orthologs have been identified in other dipterans, including *D. melanogaster* (*41*) and *Ae. aegypti* (*42*). *IGF1* is a key component of the insulin/insulin growth factor 1 signaling (IIS) cascade, which regulates processes including innate immunity, reproduction, metabolism, and life span (*43*). Nevertheless, other members of the IIS cascade are present, and four insulin-like peptides are found in a compact cluster with gene arrangements conserved across anophelines (fig. S22). This raises questions regarding the modification of IIS signaling in the absence of *IGF1* and the functional importance of this conserved genomic arrangement.

Epigenetic mechanisms affect many biological processes by modulation of chromatin structure, telomere remodeling, and transcriptional control. Of the 215 epigenetic regulatory genes in *D. melanogaster* (*44*), we identified 169 putative *An. gambiae* orthologs (table S32), which suggested the presence of mechanisms of epigenetic control in *Anopheles* and *Drosophila*. We find, however, that retrotransposition may have contributed to the functional divergence of at least one gene associated with epigenetic regulation. The ubiquitin-conjugating enzyme *E2D* (orthologous to *effete* (*45*) in *D. melanogaster*) duplicated via retrotransposition in an early anopheline ancestor, and the retrotransposed copy is maintained in a subset of anophelines. Although the entire amino acid sequence of *E2D* is perfectly conserved between *An. gambiae* and *D. melanogaster*, the retrogenes are highly divergent (Fig. 5A) and may contribute to functional diversification within the genus.

Saliva is integral to blood feeding; it impairs host hemostasis and also affects inflammation and immunity. In *An. gambiae*, the salivary proteome is estimated to contain the products of at least 75 genes, most being expressed solely in the adult female salivary glands. Comparative analyses indicate that anopheline salivary proteins are subject to strong evolutionary pressures, and these genes exhibit an accelerated pace of evolution, as well as a very high rate of gain/loss (Fig. 3 and fig. S23). Polymorphisms within *An. gambiae* populations from limited sets of salivary genes were previously found to carry signatures of positive selection (*46*). Sequence analysis across the anophelines shows that salivary genes have the highest incidence of positively selected codons among the seven gene classes (fig. S24), indicating that coevolution with vertebrate hosts is a powerful driver of natural selection in salivary proteomes. Moreover, salivary proteins also exhibit functional diversification through new gene creation. Sequence similarity, intron-exon boundaries, and secondary structure prediction point to the birth of the *SG7/SG7-2* inflammation-inhibiting (*47*) gene family from the genomic region encoding the C terminus of the 30-kD protein (Fig. 5B), a collagen-binding platelet inhibitor already present in the blood-feeding ancestor of mosquitoes and black flies (*48*). Based on phylogenetic representation, these events must have occurred before the radiation of anophelines but after separation from the culicines.

Resistance to insecticides and other xenobiotics has arisen independently in many anopheline species, fostered directly and indirectly by anthropogenic environmental modification. Metabolic resistance to insecticides is mediated by

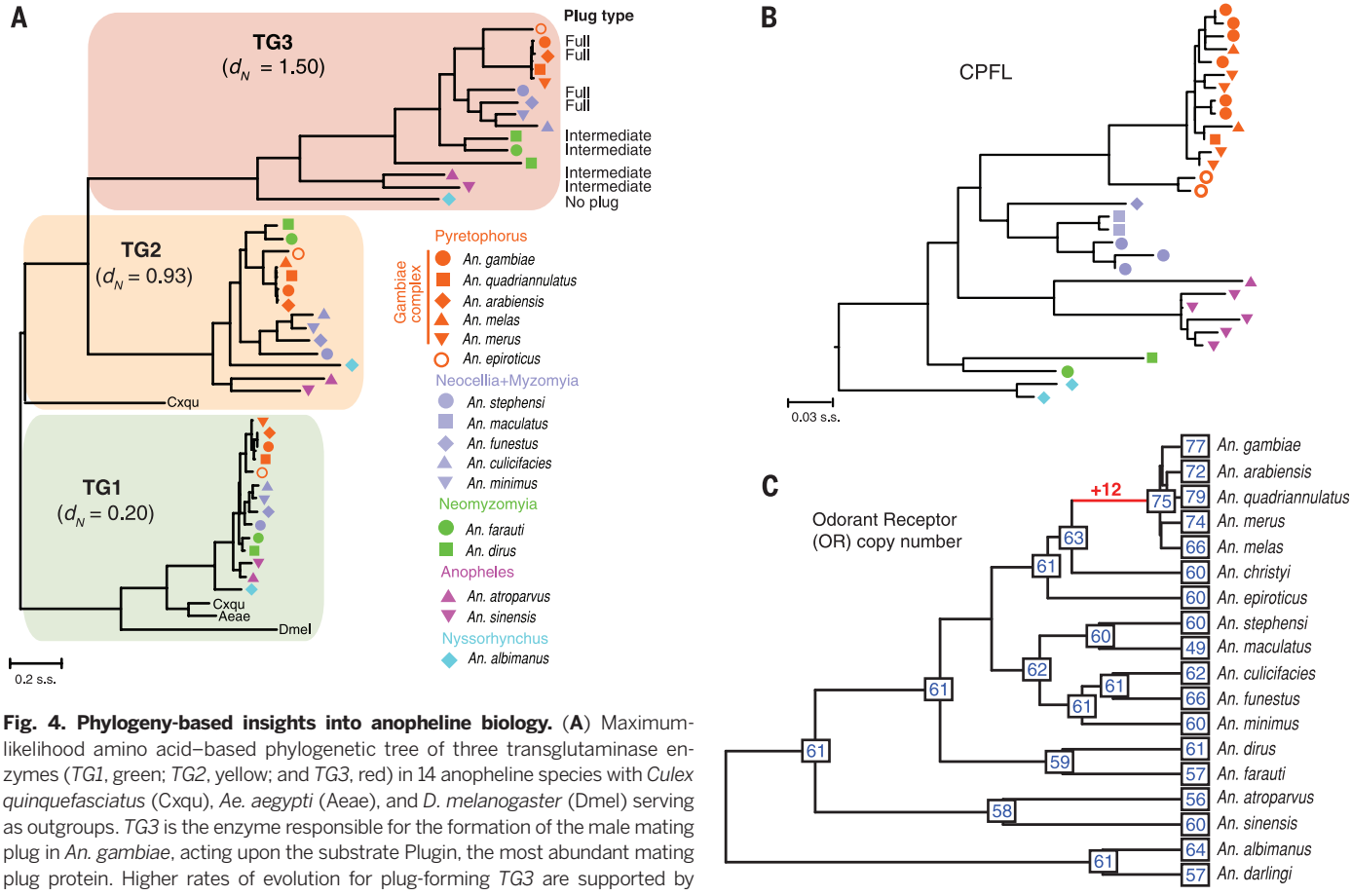


Fig. 4. Phylogeny-based insights into anopheline biology. (**A**) Maximum-likelihood amino acid–based phylogenetic tree of three transglutaminase enzymes (*TG1*, green; *TG2*, yellow; and *TG3*, red) in 14 anopheline species with *Culex quinquefasciatus* (Cxqu), *Ae. aegypti* (Aeae), and *D. melanogaster* (Dmel) serving as outgroups. *TG3* is the enzyme responsible for the formation of the male mating plug in *An. gambiae*, acting upon the substrate Plugin, the most abundant mating plug protein. Higher rates of evolution for plug-forming *TG3* are supported by elevated levels of d_N. Mating plug phenotypes are noted where known within the *TG3* clade. (**B**) Concerted evolution in CPFL cuticular proteins. Species symbols used are the same as in (A). In contrast to the *TG1/TG2/TG3* phylogeny, CPFL paralogs cluster by subgeneric clades rather than individually recapitulating the species phylogeny. Gene family size variation among species may reflect both gene gain/loss and variation in gene set completeness. (**C**) *OR* observed gene counts and inferred ancestral gene counts on an ultrametric phylogeny. At least 10 *OR* genes were gained on the branch leading to the common ancestor of the *An. gambiae* species complex, although the overall number of *OR* genes does not vary dramatically across the genus.

multiple gene families, including cytochrome P450s and glutathione *S*-transferases (GSTs), which serve to generally protect against all environment stresses, both natural and anthropogenic. We manually characterized these gene families in seven anophelines spanning the genus. Despite their large size, gene numbers (87 to 104 P450 genes, 27 to 30 GST genes) within both gene families are highly conserved across all species, although lineage-specific gene duplications and losses are often seen (tables S33 and S34). As with the *OR* and *GR* olfaction-related gene families, P450 and GST repertoires may be relatively constant due to the large number of roles they play in anopheline biology. Orthologs of genes associated with insecticide resistance either via up-regulation or coding variation [e.g., *Cyp6m2*, *Cyp6p3* (*Cyp6p9* in *An. funestus*), *Gste2*, and *Gste4*] were found in all species, which suggested that virtually all anophelines likely have genes capable of conferring insecticide resistance through similar mechanisms. Unexpectedly, one member of the P450 family (*Cyp18a1*) with a conserved role in ecdysteroid catabolism [and consequently development and metamorphosis (*49*)] appears to have been lost from the ancestor of the *An. gambiae* species complex but is found in the genome and transcriptome assemblies of other species, which indicates that the *An. gambiae* complex may have recently evolved an alternate mechanism for catabolizing ecdysone.

Susceptibility to malaria parasites is a key determinant of vectorial capacity. Dissecting the immune repertoire (*50*, *51*) (table S35) into its constituent phases reveals that classical recognition genes and genes encoding effector enzymes exhibit relatively low levels of sequence divergence. Signal transducers are more divergent in sequence but are conserved in representation across species and rarely duplicated. Cascade modulators, although also divergent, are more lineage-specific and generally have more gene duplications (Fig. 3 and fig. S25). A rare duplication of an immune signal transduction gene occurred through the retrotransposition of the signal transducer and activator of transcription *STAT2* to form the intronless *STAT1* after the divergence of *An. dirus* and *An. farauti* from the rest of the subgenus *Cellia* (Fig. 5C and table S36). It is interesting that an independent retrotransposition event appears to have created another intronless STAT gene in the *An. atroparvus* lineage. In *An. gambiae*, *STAT1* controls the expression of *STAT2* and is activated in response to bacterial challenge (*52*, *53*), and the STAT pathway has been demonstrated to mediate immunity to *Plasmodium* (*53*, *54*), so the presence of these relatively new immune signal transducers may have allowed for rewiring of regulatory networks governing immune responses in this subset of anophelines.

Conclusion

Since the discovery over a century ago by Ronald Ross and Giovanni Battista Grassi that human malaria is transmitted by a narrow range of blood-feeding female mosquitoes, the biological basis of malarial vectorial capacity has been a matter of intense interest. Inasmuch as previous successes in the local elimination of malaria have always been accomplished wholly or in part through effective vector control, an increased understanding of vector biology is crucial for continued progress against malarial disease. These 16 new reference genome assemblies provide a

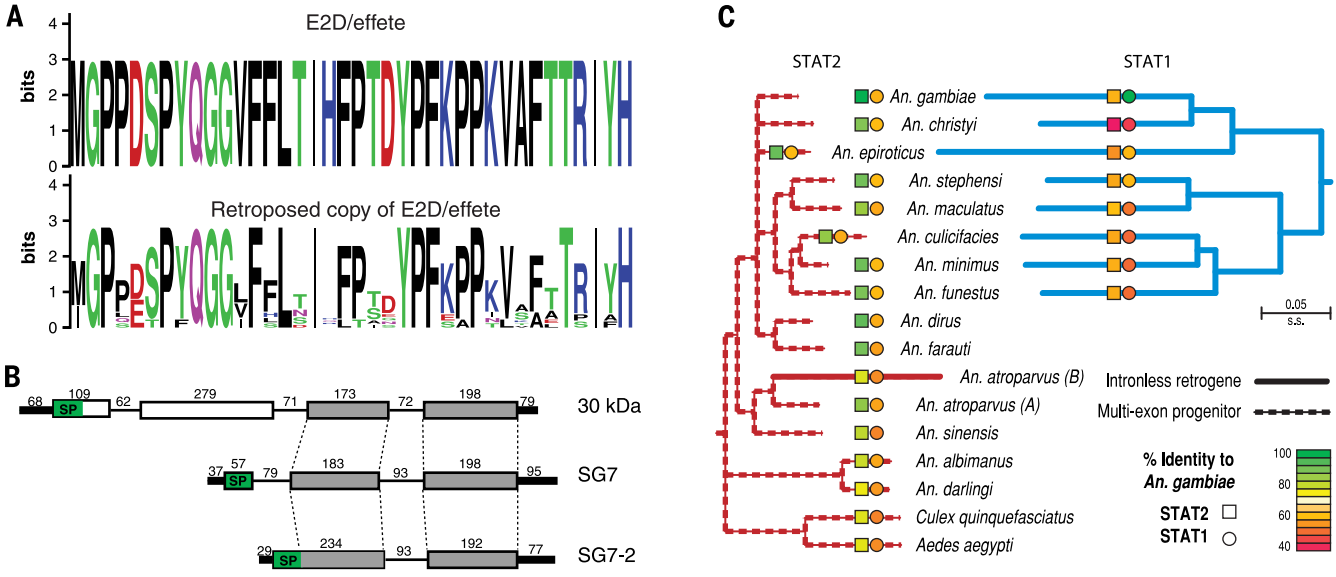


Fig. 5. Genesis of novel anopheline genes. (**A**) Retrotransposition of the *E2D/effete* gene generated a ubiquitin-conjugating enzyme at the base of the genus, which exhibits much higher sequence divergence than the original multiexon gene. WebLogo plots contrast the amino acid conservation of the original *effete* gene with the diversification of the retrotransposed copy (residues 38 to 75; species represented are *An. minimus*, *An. dirus*, *An. funestus*, *An. farauti*, *An. atroparvus*, *An. sinensis*, *An. darlingi*, and *An. albimanus*). (**B**) The *SG7* salivary protein–encoding gene was generated from the C-terminal half of the 30-kD gene. *SG7* then underwent tandem duplication and intron loss to generate another salivary protein, *SG7-2*. Numerals indicate lengths of segments in base pairs. (**C**) The origin of *STAT1*, a signal transducer and activator of transcription gene involved in immunity, occurred through a retrotransposition event in the *Cellia* ancestor after divergence from *An. dirus* and *An. farauti*. The intronless *STAT1* is much more divergent than its multiexon progenitor, *STAT2*, and has been maintained in all descendant species. An independent retrotransposition event created a retrogene copy in *An. atroparvus*, which is also more divergent than its progenitor.

foundation for additional hypothesis generation and testing to further our understanding of the diverse biological traits that determine vectorial capacity.

REFERENCES AND NOTES

1. D. P. Kwiatkowski, How malaria has affected the human genome and what human genetics can teach us about malaria. *Am. J. Hum. Genet.* **77**, 171–192 (2005). doi: 10.1086/432519; pmid: 16001361
2. A. Cohuet, C. Harris, V. Robert, D. Fontenille, Evolutionary forces on *Anopheles*: What makes a malaria vector? *Trends Parasitol.* **26**, 130–136 (2010). doi: 10.1016/j.pt.2009.12.001; pmid: 20056485
3. S. Manguin, P. Carnivale, J. Mouchet, *Biodiversity of Malaria in the World* (John Libbey Eurotext, Paris, 2008).
4. R. A. Holt *et al*., The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**, 129–149 (2002). doi: 10.1126/science.1076181; pmid: 12364791
5. O. Marinotti *et al*., The genome of *Anopheles darlingi*, the main neotropical malaria vector. *Nucleic Acids Res.* **41**, 7387–7400 (2013). doi: 10.1093/nar/gkt484; pmid: 23761445
6. D. Zhou *et al*., Genome sequence of *Anopheles sinensis* provides insight into genetics basis of mosquito competence for malaria parasites. *BMC Genomics* **15**, 42 (2014). pmid: 24438588
7. X. Jiang *et al*., Genome analysis of a major urban malaria vector mosquito, *Anopheles stephensi*. *Genome Biol.* **15**, 459 (2014). doi: 10.1186/s13059-014-0459-2; pmid: 25244985
8. D. E. Neafsey *et al*., The evolution of the *Anopheles* 16 genomes project. *Genes Genomes Genet.* **3**, 1191–1194 (2013). doi: 10.1534/g3.113.006247; pmid: 23708298
9. M. C. Fontaine *et al*., Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science* **346**, 1258524 (2014).
10. M. Moreno *et al*., Complete mtDNA genomes of *Anopheles* darlingi and an approach to anopheline divergence time. *Malar. J.* **9**, 127 (2010). doi: 10.1186/1475-2875-9-127; pmid: 20470395
11. S. Gnerre *et al*., High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1513–1518 (2011). doi: 10.1073/pnas.1017351108; pmid: 21187386
12. Materials and methods are available as supplementary materials on *Science* Online.
13. C. Holt, M. Yandell, MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**, 491 (2011). doi: 10.1186/1471-2105-12-491; pmid: 22192575
14. M. Coluzzi, A. Sabatini, A. della Torre, M. A. Di Deco, V. Petrarca, A polytene chromosome analysis of the *Anopheles gambiae* species complex. *Science* **298**, 1415–1418 (2002). doi: 10.1126/science.1077769; pmid: 12364623
15. M. Kamali *et al*., Multigene phylogenetics reveals temporal diversification of major African malaria vectors. *PLOS ONE* **9**, e93580 (2014). doi: 10.1371/journal.pone.0093580; pmid: 24705448
16. M. A. Toups, M. W. Hahn, Retrogenes reveal the direction of sex-chromosome evolution in mosquitoes. *Genetics* **186**, 763–766 (2010). doi: 10.1534/genetics.110.118794; pmid: 20660646
17. D. A. Baker, S. Russell, Role of testis-specific gene expression in sex-chromosome evolution of *Anopheles gambiae*. *Genetics* **189**, 1117–1120 (2011). doi: 10.1534/genetics.111.133157; pmid: 21890740
18. M. V. Han, G. W. C. Thomas, J. Lugo-Martinez, M. W. Hahn, Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol. Biol. Evol.* **30**, 1987–1997 (2013). doi: 10.1093/molbev/mst100; pmid: 23709260
19. M. W. Hahn, M. V. Han, S.-G. Han, Gene family evolution across 12 *Drosophila* genomes. *PLOS Genet.* **3**, e197 (2007). doi: 10.1371/journal.pgen.0030197; pmid: 17997610
20. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007). doi: 10.1093/molbev/msm088; pmid: 17483113
21. T. Dottorini *et al*., A genome-wide analysis in *Anopheles gambiae* mosquitoes reveals 46 male accessory gland genes, possible modulators of female behavior. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16215–16220 (2007). doi: 10.1073/pnas.0703904104; pmid: 17901209
22. F. Baldini, P. Gabrieli, D. W. Rogers, F. Catteruccia, Function and composition of male accessory gland secretions in *Anopheles gambiae*: A comparison with other insect vectors of infectious diseases. *Pathog. Glob. Health* **106**, 82–93 (2012). doi: 10.1179/2047773212Y.0000000016; pmid: 22943543
23. R. Assis, Q. Zhou, D. Bachtrog, Sex-biased transcriptome evolution in *Drosophila*. *Genome Biol. Evol.* **4**, 1189–1200 (2012). doi: 10.1093/gbe/evs093; pmid: 23097318
24. S. Grath, J. Parsch, Rate of amino acid substitution is influenced by the degree and conservation of male-biased transcription over 50 myr of *Drosophila* evolution. *Genome Biol. Evol.* **4**, 346–359 (2012). doi: 10.1093/gbe/evs012; pmid: 22321769
25. J. C. Perry, P. W. Harrison, J. E. Mank, The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. *Mol. Biol. Evol.* **31**, 1206–1219 (2014). doi: 10.1093/molbev/msu072; pmid: 24526011
26. D. W. Rogers *et al*., Transglutaminase-mediated semen coagulation controls sperm storage in the malaria mosquito. *PLOS Biol.* **7**, e1000272 (2009). doi: 10.1371/journal.pbio.1000272; pmid: 20027206
27. F. Baldini *et al*., The interaction between a sexually transferred steroid hormone and a female protein regulates oogenesis in the malaria mosquito *Anopheles gambiae*. *PLOS Biol.* **11**, e1001695 (2013). doi: 10.1371/journal.pbio.1001695; pmid: 24204210
28. W. R. Shaw *et al*., Mating activates the heme peroxidase HPX15 in the sperm storage organ to ensure fertility in *Anopheles gambiae*. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 5854–5859 (2014). doi: 10.1073/pnas.1401715111; pmid: 24711401
29. J. H. Willis, Structural cuticular proteins from arthropods: Annotation, nomenclature, and sequence characteristics in the genomics era. *Insect Biochem. Mol. Biol.* **40**, 189–204 (2010). doi: 10.1016/j.ibmb.2010.02.001; pmid: 20171281
30. R. S. Cornman *et al*., Annotation and analysis of a large cuticular protein family with the R&R Consensus in *Anopheles gambiae*. *BMC Genomics* **9**, 22 (2008). doi: 10.1186/1471-2164-9-22; pmid: 18205929
31. R. S. Cornman, J. H. Willis, Extensive gene amplification and concerted evolution within the CPR family of cuticular proteins in mosquitoes. *Insect Biochem. Mol. Biol.* **38**, 661–676 (2008). doi: 10.1016/j.ibmb.2008.04.001; pmid: 18510978
32. R. S. Cornman, J. H. Willis, Annotation and analysis of low-complexity protein families of *Anopheles gambiae* that are associated with cuticle. *Insect Mol. Biol.* **18**, 607–622 (2009). doi: 10.1111/j.1365-2583.2009.00902.x; pmid: 19754739
33. R. S. Cornman, Molecular evolution of *Drosophila* cuticular protein genes. *PLOS ONE* **4**, e8345 (2009). doi: 10.1371/journal.pone.0008345; pmid: 20019874
34. T. Togawa, W. A. Dunn, A. C. Emmons, J. Nagao, J. H. Willis, Developmental expression patterns of cuticular protein genes

with the R&R Consensus from *Anopheles gambiae*. *Insect Biochem. Mol. Biol.* **38**, 508–519 (2008). doi: 10.1016/j.ibmb.2007.12.008; pmid: 18405829
35. D. C. Rinker *et al*., Blood meal-induced changes to antennal transcriptome profiles reveal shifts in odor sensitivities in *Anopheles gambiae*. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8260–8265 (2013). doi: 10.1073/pnas.1302562110; pmid: 23630291
36. D. C. Rinker *et al*., Antennal transcriptome profiles of anopheline mosquitoes reveal human host olfactory specialization in *Anopheles gambiae*. *BMC Genomics* **14**, 749 (2013). doi: 10.1186/1471-2164-14-749; pmid: 24182346
37. M. Altstein, D. R. Nässel, Neuropeptide signaling in insects. *Adv. Exp. Med. Biol.* **692**, 155–165 (2010). doi: 10.1007/978-1-4419-6902-6_8; pmid: 21189678
38. J. P. Goetze, I. Hunter, S. K. Lippert, L. Bardram, J. F. Rehfeld, Processing-independent analysis of peptide hormones and prohormones in plasma. *Front. Biosci. (Landmark Ed.)* **17**, 1804–1815 (2012). doi: 10.2741/4020; pmid: 22201837
39. T. H. Stracker, S. Thompson, G. L. Grossman, M. A. Riehle, M. R. Brown, Characterization of the AeaHP gene and its expression in the mosquito *Aedes aegypti* (Diptera: Culicidae). *J. Med. Entomol.* **39**, 331–342 (2002). doi: 10.1603/0022-2585-39.2.331; pmid: 11931033
40. C. D. de Oliveira, W. P. Tadei, F. C. Abdalla, P. F. Paolucci Pimenta, O. Marinotti, Multiple blood meals in *Anopheles darlingi* (Diptera: Culicidae). *J. Vector Ecol.* **37**, 351–358 (2012). doi: 10.1111/j.1948-7134.2012.00238.x; pmid: 23181859
41. N. Okamoto *et al*., A fat body-derived IGF-like peptide regulates postfeeding growth in *Drosophila*. *Dev. Cell* **17**, 885–891 (2009). doi: 10.1016/j.devcel.2009.10.008; pmid: 20059957
42. M. A. Riehle, Y. Fan, C. Cao, M. R. Brown, Molecular characterization of insulin-like peptides in the yellow fever mosquito, *Aedes aegypti*: Expression, cellular localization, and phylogeny. *Peptides* **27**, 2547–2560 (2006). doi: 10.1016/j.peptides.2006.07.016; pmid: 16934367
43. Y. Antonova, A. J. Arik, W. Moore, M. M. Riehle, M. R. Brown, in *Insect Endocrinology*, L. I. Gilbert, Ed. (Academic Press, 2012), pp. 63–92.
44. A. Swaminathan, A. Gajan, L. A. Pile, Epigenetic regulation of transcription in *Drosophila*. *Front. Biosci. (Landmark Ed.)* **17**, 909–937 (2012). doi: 10.2741/3964; pmid: 22201781
45. F. Cipressa *et al*., Effete, a *Drosophila* chromatin-associated ubiquitin-conjugating enzyme that affects telomeric and heterochromatic position effect variegation. *Genetics* **195**, 147–158 (2013). doi: 10.1534/genetics.113.153320; pmid: 23821599
46. B. Arcà *et al*., Positive selection drives accelerated evolution of mosquito salivary genes associated with blood-feeding. *Insect Mol. Biol.* **23**, 122–131 (2014). doi: 10.1111/imb.12068; pmid: 24237399
47. H. Isawa, M. Yuda, Y. Orito, Y. Chinzei, A mosquito salivary protein inhibits activation of the plasma contact system by binding to factor XII and high molecular weight kininogen. *J. Biol. Chem.* **277**, 27651–27658 (2002). doi: 10.1074/jbc.M203505200; pmid: 12011093
48. E. Calvo *et al*., Aegyptin, a novel mosquito salivary gland protein, specifically binds to collagen and prevents its interaction with platelet glycoprotein VI, integrin alpha2beta1, and von Willebrand factor. *J. Biol. Chem.* **282**, 26928–26938 (2007). doi: 10.1074/jbc.M705669200; pmid: 17650501
49. E. Guittard *et al*., CYP18A1, a key enzyme of *Drosophila* steroid hormone inactivation, is essential for metamorphosis. *Dev. Biol.* **349**, 35–45 (2011). doi: 10.1016/j.ydbio.2010.09.023; pmid: 20932968
50. R. M. Waterhouse *et al*., Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* **316**, 1738–1743 (2007). doi: 10.1126/science.1139862; pmid: 17588928
51. L. C. Bartholomay *et al*., Pathogenomics of Culex quinquefasciatus and meta-analysis of infection responses to diverse pathogens. *Science* **330**, 88–90 (2010). doi: 10.1126/science.1193162; pmid: 20929811
52. C. Barillas-Mury, Y. S. Han, D. Seeley, F. C. Kafatos, *Anopheles gambiae* Ag-STAT, a new insect member of the STAT family, is activated in response to bacterial infection. *EMBO J.* **18**, 959–967 (1999). doi: 10.1093/emboj/18.4.959; pmid: 10022838
53. L. Gupta *et al*., The STAT pathway mediates late-phase immunity against *Plasmodium* in the mosquito *Anopheles gambiae*. *Cell Host Microbe* **5**, 498–507 (2009). doi: 10.1016/j.chom.2009.04.003; pmid: 19454353
54. A. C. Bahia *et al*., The JAK-STAT pathway controls *Plasmodium vivax* load in early stages of *Anopheles aquasalis* infection. *PLOS Negl. Trop. Dis.* **5**, e1317 (2011). doi: 10.1371/journal.pntd.0001317; pmid: 22069502

ACKNOWLEDGMENTS

All sequencing reads and genome assemblies have been submitted to the National Center for Biotechnology Information (umbrella BioProject ID, PRJNA67511). Genome and transcriptome assemblies are also available from VectorBase (https://vectorbase.org) and the Broad Institute (https://olive.broadinstitute.org/collections/anopheles.4). The authors acknowledge the NIH Eukaryotic Pathogen and Disease Vector Sequencing Project Working Group for guidance and development of this project. Sequence data generation was supported at the Broad Institute by the National Human Genome Research Institute (U54 HG003067). We thank the many members of the Broad Institute Genomics Platform and Genome Sequencing and Analysis Program who contributed to sequencing data generation and analysis.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/1258522/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S25
Tables S1 to S36
References (*55–248*)

9 July 2014; accepted 14 November 2014
Published online 27 November 2014;
10.1126/science.1258522

RESEARCH ARTICLES

STRUCTURAL BIOLOGY

Mechanistic insight from the crystal structure of mitochondrial complex I

Volker Zickermann,[1,2]*† Christophe Wirth,[3]* Hamid Nasiri,[4,5] Karin Siegmund,[1] Harald Schwalbe,[2,5] Carola Hunte,[3]† Ulrich Brandt[2,6]†

Proton-pumping complex I of the mitochondrial respiratory chain is among the largest and most complicated membrane protein complexes. The enzyme contributes substantially to oxidative energy conversion in eukaryotic cells. Its malfunctions are implicated in many hereditary and degenerative disorders. We report the x-ray structure of mitochondrial complex I at a resolution of 3.6 to 3.9 angstroms, describing in detail the central subunits that execute the bioenergetic function. A continuous axis of basic and acidic residues running centrally through the membrane arm connects the ubiquinone reduction site in the hydrophilic arm to four putative proton-pumping units. The binding position for a substrate analogous inhibitor and blockage of the predicted ubiquinone binding site provide a model for the "deactive" form of the enzyme. The proposed transition into the active form is based on a concerted structural rearrangement at the ubiquinone reduction site, providing support for a two-state stabilization-change mechanism of proton pumping.

With a molecular mass of ~1 MD, proton-pumping NADH:ubiquinone oxidoreductase (complex I) is the largest membrane protein complex of the mitochondrial respiratory chain (*1*, *2*). Complex I couples electron transfer from NADH (reduced nicotinamide adenine dinucleotide) to ubiquinone to the translocation of four protons from the mitochondrial matrix to the intermembrane space (IMS), thereby providing much of the proton-motive force for adenosine triphosphate synthesis. Dysfunction of the enzyme is the most frequent cause of mitochondrial disorders (*3*) and has been implicated in numerous neurodegenerative conditions (*4*). Complex I is also a major source of deleterious reactive oxygen species (*5*). Mitochondrial complex I consists of 14 central subunits and a large number of accessory subunits. Specific inhibitors have been essential for unraveling structure-function relationships within complex I and are a focus of biomedical research, as some of them can induce Parkinson's disease (*6*). The central subunits of complex I harboring the bioenergetic core functions are conserved from bacteria to humans. Complex I from bacteria and from mitochondria of *Yarrowia lipolytica*, a yeast genetic model for the study of eukaryotic complex I (*7*), were previously analyzed by x-ray crystallography (*8*–*10*). Yet the catalytic mechanism has remained enigmatic. Here, we present the x-ray structure of all central subunits and the largest accessory subunit of mitochondrial complex I from *Y. lipolytica*, providing insight into the molecular basis of redox-driven proton pumping.

[1]Structural Bioenergetics Group, Institute of Biochemistry II, Medical School, Goethe-University, 60438 Frankfurt am Main, Germany. [2]Cluster of Excellence Frankfurt "Macromolecular Complexes," Goethe-University, 60438 Frankfurt am Main, Germany. [3]Institute for Biochemistry and Molecular Biology, ZBMZ, BIOSS Centre for Biological Signalling Studies, University of Freiburg, 79104 Freiburg, Germany. [4]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK. [5]Institute of Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance, 60438 Frankfurt am Main, Germany. [6]Nijmegen Center for Mitochondrial Disorders, Radboud University Medical Center, 6525 GA Nijmegen, Netherlands.

*These authors contributed equally to this work. **†Corresponding author. E-mail: zickermann@med.uni-frankfurt.de (V.Z.); carola.hunte@biochemie.uni-freiburg.de (C.H.); ulrich.brandt@radboudumc.nl (U.B.)**

Overall structure

Complex I was purified and crystallized as described (*8*). Optimized cryoprotection and data collection improved x-ray diffraction in the best direction to 3.6 Å. Phases were obtained using single isomorphous replacement with anomalous scattering from highly redundant data of heavy-atom derivatives, FeS clusters, and selenomethionine. Phases were combined, averaged, and extended to the resolution of the native data set collected at 5 K. Model building with the 3.8 Å experimental electron density map was aided by the large content of α helices, close to 100 selenomethionine positions, and homology models based on the bacterial complexes (*10*, *11*). Secondary structure constraints were applied for refinement. Anisotropy correction counteracted the lower order of the crystals in the membrane plane, and the structure was refined at 3.9 Å × 3.9 Å × 3.6 Å (table S1 and fig. S1). The structure contains the majority of residues for central subunits and the largest accessory subunit (table S2). Side chains are well resolved in tightly packed subunits with high helix content, and between 50 and 89% of the residues in these subunits could be assigned. The end of the peripheral hydrophilic domain is less ordered (fig. S1 and table S1).

Mitochondrial complex I has a slightly opened L-shape with an angle of ~120° between the two arms (Fig. 1). The matrix arm protrudes into the organelle interior and is oriented perpendicular to the membrane plane and the membrane arm. The 14 central subunits forming the core of the complex are surrounded by the accessory subunits comprising 44% of the total mass. Complex I is 250 Å long and 190 Å high, with a width of 70 Å for the membrane arm and 120 Å for the matrix arm. The complex is organized in four functional modules (*1*, *8*). The distal half of the matrix arm, comprising the central 75-kDa, 51-kDa, and 24-kDa subunits, forms the N module that oxidizes NADH (nomenclature as for bovine complex I is used throughout the text). The proximal half of the peripheral arm, which comprises the central 49-kDa, 30-kDa, PSST, and TYKY subunits, is the Q module that reduces ubiquinone and docks onto the membrane arm. A chain of eight canonical FeS clusters (*2*) runs over a distance of ~100 Å through the matrix arm (Fig. 1C). The edge-to-edge distances for seven of these clusters are <14 Å, allowing fast electron tunneling (*12*). This arrangement establishes an electron transfer path from the [4Fe-4S] cluster of the 51-kDa subunit (cluster N3) to the [4Fe-4S] cluster of PSST (cluster N2). Cluster N2 reduces ubiquinone and resides ~30 Å above the membrane plane (*8*). The slightly bent membrane arm features 82 transmembrane helices (TMHs), with 64 of them contributed by central subunits (Fig. 2A). The proximal pump module (P_P) comprises central subunits ND1, ND2, ND3, ND4L, and ND6, whereas the distal pump module (P_D) contains central subunits ND4 and ND5 (Fig. 2A).

Membrane arm and proton-pumping modules

The largest membrane-embedded subunits, ND5, ND4, and ND2 (Fig. 2A), share a structurally highly similar core of 14 TMHs (fig. S2) with two repeats of five TMHs (A, TMH4–8; B, TMH9–13) in inverted topology (fig. S2). Each repeat features a discontinuous helix (TMH7a/b, TMH12a/b). Such helices are hallmarks of ion-translocating membrane proteins (*13*–*15*). Indeed, ND5, ND4, and ND2 are homologous to the Mrp Na^+/H^+ antiporter family, which suggests a role in proton pumping (*8*, *11*, *16*). ND5 has a C-terminal extension with a lateral helix, >60 Å long, lining ND5, ND4, and ND2 on the concave side of the arm close to the matrix side (Fig. 2A), thus bridging the P_P and P_D modules. The C terminus of ND5 is anchored to ND2 via a V-shaped arrangement of TMH16 and TMH17. The lateral helix was previously identified in the low-resolution analysis of mitochondrial complex I from *Y. lipolytica* (*8*) and is also present in the bacterial (*9*, *11*) and bovine complexes (*17*), in which its C terminus is anchored by one TMH only. Whether this prominent structural element is involved in energy transduction for proton pumping, as proposed previously (*8*, *9*), remains controversial (*2*).

Adjacent to ND2 are three small central subunits, ND4L, ND6, and ND3, which form two layers of straight helices crossing the membrane

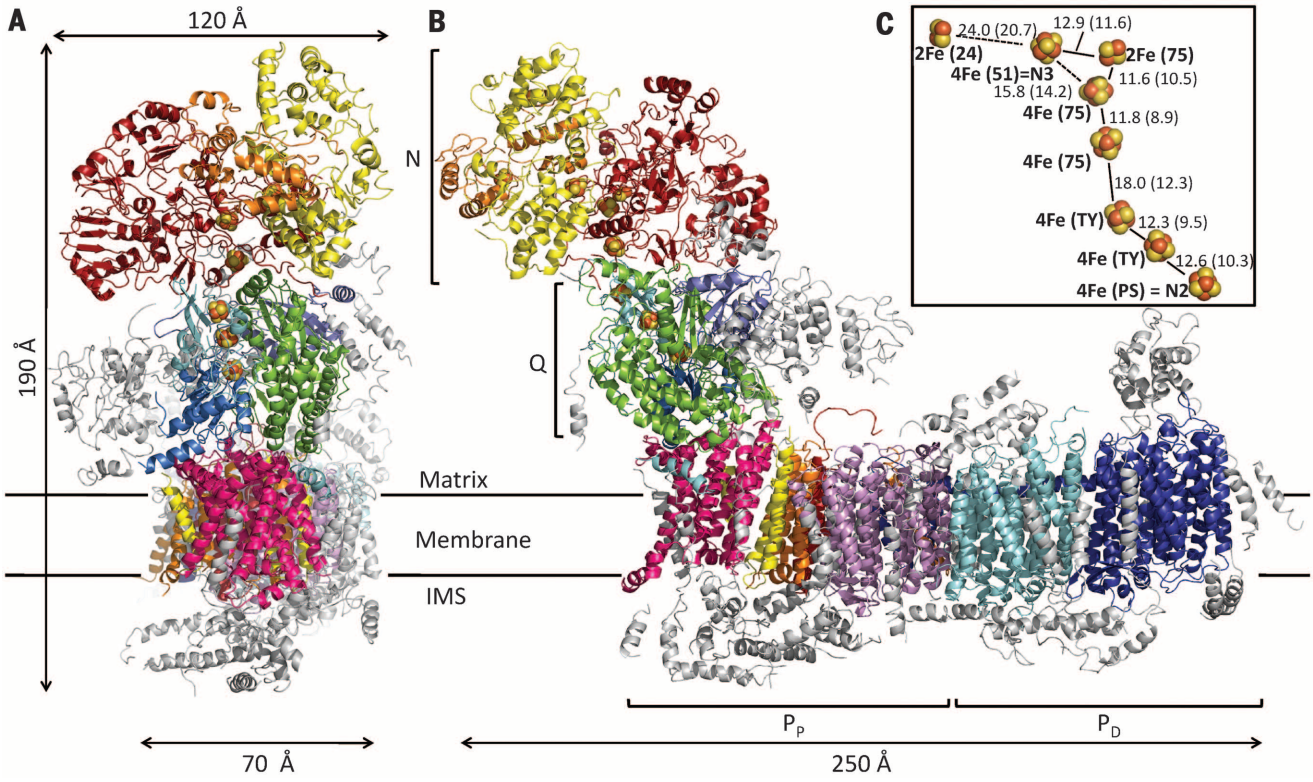


Fig. 1. Overall structure. (**A** and **B**) View from peripheral arm (A) and rotated 90° (B). N module: red, 75-kDa; yellow, 51-kDa; orange, 24-kDa. Q module: green, 49-kDa; violet, 30-kDa; blue, PSST; cyan, TYKY. P_D module: dark blue, ND5; cyan, ND4. P_P module: lilac, ND2; red, ND4L; orange, ND6; yellow, ND3; pink, ND1. Accessory subunits are depicted in gray. (**C**) Arrangement of 4Fe-4S(4Fe) and 2Fe-2S(2Fe) clusters in the peripheral arm. The coordinating subunits are denoted in brackets. Center-to-center and edge-to-edge (in brackets) distances are in angstroms.

domain (Fig. 2A and fig. S3A). Notably, $TMH3^{ND6}$ features a π bulge near the membrane center. Remote with respect to $TMH1–3^{ND6}$, $TMH4^{ND6}$ is laterally associated with the membrane arm and lines $TMH16^{ND5}$. On the intermembrane-space side, a long loop links $THM4^{ND6}$ to a short surface helix connecting to $TMH5^{ND6}$, which is surrounded by helices of ND2, ND4L, and ND3. In the bacterial complex (*9*, *11*), $TMH4^{ND6}$ sits closer to the core helices of ND6 at a position occupied in the mitochondrial complex by a single helix of an accessory subunit, tentatively identified as NUJM (*17*). $TMH1^{ND3}$ is anchored remotely in a surface cleft of ND1. ND1 at the proximal end of the membrane arm comprises eight TMHs and noticeable loop regions, including short surface helices (Fig. 2A and fig. S3B). The fold of TMH2–6 is similar to one repeat of the antiporter-like subunits. This agrees with homology predictions (*18*) and was noted in the bacterial complex (*10*). Notably, TMH5 is discontinuous in mitochondrial complex I.

The interior of the membrane arm is rich in polar and protonable residues (i.e., residues that may take up or release protons) constituting a remarkable hydrophilic central axis across all subunits (Fig. 2B). Toward the matrix and IMS, each of the antiporter-like subunits has typical arrays of titratable and polar residues constituting possible proton uptake and release pathways (Fig. 2B and fig. S4). For ND5 (Fig. 2B), Asp^{397} at the membrane center is the first in a series of polar residues (Tyr^{392}, Glu^{401}, Thr^{403}, and Ser^{402}) arranged along TMH12b to the IMS. Neighboring Lys^{396} connects this path to the central polar network with major contributions of protonable residues of TMH11 (His^{331}, His^{335}, and Lys^{339}). The likely entry point to this network is His^{251} of TMH8, to which protons could be provided from the matrix along the tip of TMH7b, TMH8, and TMH10 (Glu^{241}, Ser^{247}, Thr^{244}, and Lys^{302}). At the center of discontinuous TMH7, Lys^{226} (TMH7a) is sandwiched between the indole side chains of Trp^{143} and Trp^{235} and connects on one side to His^{251} (through Ser^{227} and Thr^{254}) and on the other side to $TMH12^{ND4}$ via protonable residues of $TMH5^{ND5}$ and $TMH6^{ND5}$ (Glu^{144}, Arg^{175}, Asp^{178}). A similar architecture is repeated in ND4 and ND2 (fig. S4). Whereas the central network and the proposed proton exit routes are consistent with suggestions for the bacterial complex, the proposed routes for proton entry differ as they were assigned between TMH5 and TMH7 in bacteria (*10*, *19*).

In summary, all three antiporter-like subunits carry the structural signature of a proton-pumping unit with two structurally distinct discontinuous helices. Consistently, TMH7a/b contributes a central Lys residue that is next to a Glu of TMH5, and TMH12a/b is part of a central polar network connecting the proton exit route along TMH12b with the likely proton entry point at loops TMH7b–8 and TMH9–10 (Fig. 2B). The dipoles of the discontinuous helices add to the polarity of the central axis, because the C termini of TMH7a and TMH12a point to the center of the membrane arm, thereby introducing evenly spaced negative partial charges along the arm. In contrast, the positive polarity of the N termini of TMH7b and TMH12b is directed away from the central axis toward the periphery of the membrane arm (Fig. 2, B and C).

In line with a pump stoichiometry of $4H^+/2e^-$ (*20*), a fourth potential proton translocation pathway could be anticipated in the small NDs or ND1. From the neighboring residues Glu^{30} and Glu^{66} of ND4L and Glu^{131} of ND2 continuing the central axis, a row of polar residues toward the IMS at the interface of ND2 and ND4L could serve as a proton exit (Fig. 2B and fig. S4). Directly above, a series of polar residues constitutes a potential fourth proton entry at the interface initially also proposed to be present in bacterial complex I (*19*). The central axis extends further toward Glu^{196} at the center of discontinuous $TMH5^{ND1}$, along Asp^{67} and Glu^{69} of ND3 and Glu^{147} of ND1. The π bulge of $TMH3^{ND6}$ provides an additional polar contribution. From Glu^{196} onward, a row of acidic residues of ND1 (Asp^{203}, Glu^{206}, Glu^{208}, and Glu^{210}) continues toward the peripheral arm. A similar arrangement was described in complex I from *T. thermophilus* (*10*) and was interpreted as a possible fourth pathway

for proton uptake, replacing the one proposed earlier (*19*). However, the structure of mitochondrial complex I suggests that the central axis of protonable residues is more likely to play a critical role in energy transmission.

Peripheral arm

The peripheral arm comprises the N module extending into the matrix and the Q module docking it onto the membrane arm. The central subunits of the Q module (the 49-kDa, PSST, TYKY, and 30-kDa subunits) are related to Ni-Fe hydrogenases (*1*, *21*). The 49-kDa subunit comprises two three-stranded antiparallel β sheets and several α helices. The longest helices form a prominent four-helix bundle inclined toward the membrane surface (Figs. 1 and 3). PSST, on the other side of the Q module, contains a central four-stranded parallel β sheet surrounded by five helices and harbors the [4Fe-4S] cluster N2. The N-terminal helix α_1^{PSST} protrudes toward the membrane surface (Figs. 1 and 3B). TYKY provides contact to the N module. Its ferredoxin-type fold holds two [4Fe-4S] clusters (Fig. 3). Its N-terminal helix α_1 docks onto ND1 and contacts the four-helix bundle of the 49-kDa subunit. The latter is also the docking site for the 30-kDa subunit, with a central five-stranded β sheet laterally surrounded by four α helices. In the N module, the 75-kDa subunit resides on the same side as PSST and coordinates one [2Fe-2S] and two [4Fe-4S] clusters (Figs. 1 and 3B). The 51-kDa subunit located above the 49-kDa subunit holds one [4Fe-4S] cluster and contains the binding sites for FMN and NADH in a Rossman-fold domain. The 24-kDa subunit coordinates the [2Fe-2S] cluster detached from the electron transfer chain leading to cluster N2. As noted previously (*8*), the Q module is rotated outward about 3° relative to the N module as compared to bacterial complex I (*10*).

The largest accessory subunit, NUEM—an ortholog of the mammalian 39-kDa subunit—flanks the Q module and docks onto PSST (fig. S5). It belongs to the family of short-chain dehydrogenases (SDRs). A characteristic Rossman fold typically found in all SDRs (*22*) is present in the N-terminal domain, with a central seven-stranded β sheet surrounded by five α helices (Fig. 3 and fig. S5). NADH phosphate (NADPH) binding to *Y. lipolytica* complex I was shown experimentally (*23*). The helices of the C-terminal domain characteristic for the subfamily of extended SDRs are in contact with the membrane arm, consistent with the reported cross-linking to ND3 (*24*). Assembly defects of *Y. lipolytica* complex I in deletion mutants (*23*) suggest a structural role, but the function of NUEM remains elusive. Some additional accessory subunits were provisionally localized in the *Y. lipolytica* complex I structure (fig. S6) on the basis of a model of supernumerary subunits of bovine complex I recently obtained by cryo–electron microscopy (*17*).

Interface between peripheral and membrane arm and ubiquinone access

Structural elements of the ND1, ND3, 49-kDa, PSST, and TYKY subunits form the interface between the membrane and the peripheral arm (Fig. 3). The matrix surface of ND1 provides the main contact area with the 49-kDa subunit and PSST. Helix α_1^{PSST} contacts surface helix $\alpha_{1\text{-}2}^{ND1}$. On the opposite side, helix α_1^{TYKY} protrudes into a groove between the four-helix bundle of the 49-kDa subunit and ND1. $TMH1^{ND3}$ and $TMH2^{ND3}$ are in contact with loop $TMH1\text{–}2^{ND1}$ including helix $\alpha_{1\text{-}2}^{ND1}$ and with loop $\beta_2\text{-}\beta_3$ of the N-terminal β sheet of the 49-kDa subunit, respectively. At a central point of the interface, contact between the two arms of complex I involves loop $TMH5\text{–}6^{ND1}$ (Glu^{206} to Gly^{221}) (Fig. 4). This loop is rich in acidic residues, protrudes into a cleft between PSST and 49-kDa subunit, and is in contact with 49-kDa loop β1-β2 (Pro^{89} to Leu^{98}) of the N-terminal β sheet, where loop $TMH1\text{–}2^{ND3}$ also approaches.

Because cluster N2 resides well above the membrane plane, an access path allowing the head group of the hydrophobic substrate ubiquinone to reach its electron donor is needed (*25*). This path is provided by a quinone exchange cavity crossing the interface region. Its opening lies between $TMH1^{ND1}$, $TMH6^{ND1}$, and amphipathic helix $\alpha_{1\text{-}2}^{ND1}$ positioned at the periphery of the matrix bilayer leaflet. The cavity extends ~30 Å toward the tip of loop $\beta1\text{-}\beta2^{49\text{-}kDa}$ (Fig. 4). The small triangular-shaped entry pore is ~7 Å wide and has a hydrophobic surface. The side chain of Ala^{54} of $\alpha_{1\text{-}2}^{ND1}$ points into the pore opening. In humans, substitution of the corresponding Ala^{52} of ND1 with the larger Thr is among the most prevalent mutations leading to Leber's hereditary optic neuropathy and interferes with ubiquinone reduction (*26*). Furthermore, substitution of Trp^{77} in PSST (lining the inner side of the narrow entry passage) by Glu abolishes ubiquinone reductase activity (*27*). The surface and immediate vicinity of the cavity exhibit a bipartite distribution of charged residues. Toward the surface of the complex, basic residues are present at the C-terminal end of $TMH1^{ND1}$ (Arg^{27}, Lys^{28}, Arg^{36}, and Arg^{37}), $TMH7^{ND1}$, and loop $TMH7\text{–}8^{ND1}$, some of which may interact with phospholipids. Deeper into the pocket toward the protein interior, the cavity is lined by acidic residues of $TMH5^{ND1}$ and adjacent loop $TMH5\text{–}6^{ND1}$ (Asp^{203}, Glu^{206}, Glu^{208}, and Glu^{210}) (Fig. 4).

Ubiquinone and inhibitor binding sites

To define the substrate and inhibitor binding sites within the Q module, we cocrystallized

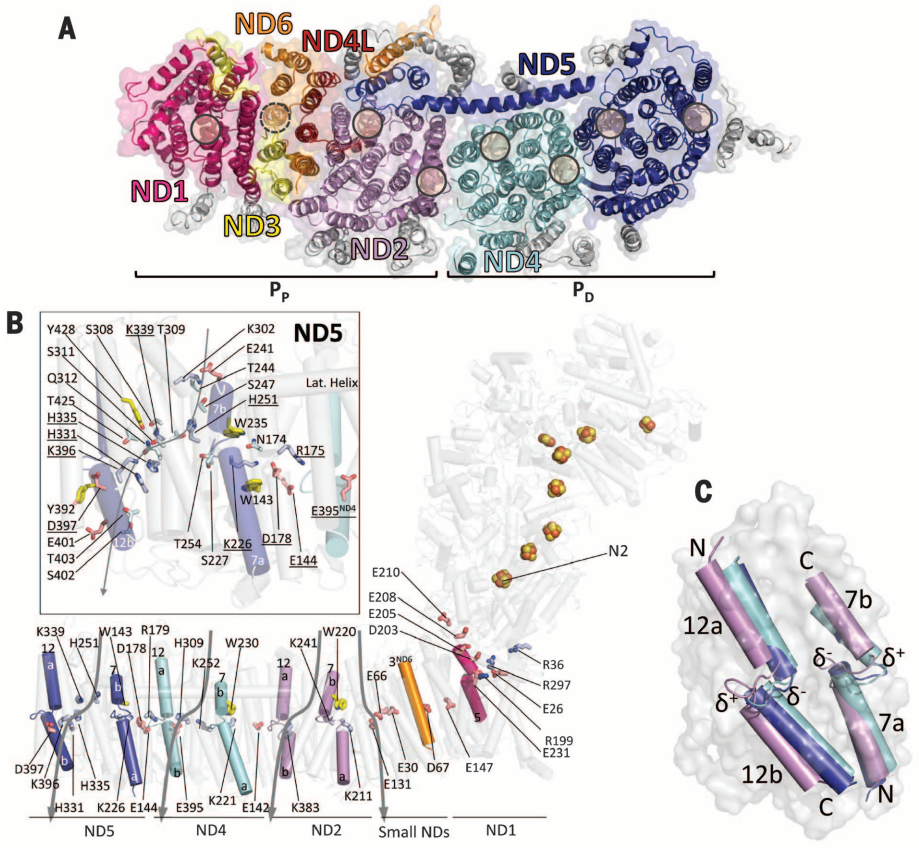


Fig. 2. Membrane arm. (**A**) Central transmembrane subunits and TMHs of accessory subunits seen from the matrix side (colors as in Fig. 1). Solid circles, discontinuous helices in ND1, ND2, ND4, and ND5; broken circle, $TMH3^{ND6}$ containing a π bulge. (**B**) Discontinuous helices and side chains of residues of the central axis. Inset shows arrays of polar and titratable residues of ND5. Residues of the central axis of protonable residues are underlined. (**C**) Discontinuous TMH7 and TMH12 after global superposition of ND2, ND4, and ND5 seen along the long axis of the membrane arm; δ^+ and δ^- are partial charges imposed by helix dipoles. Amino acid abbreviations: D, Asp; E, Glu; H, His; K, Lys; N, Asn; Q, Gln; R, Arg; S, Ser; T, Thr; W, Trp; Y, Tyr.

complex I with brominated derivatives of 2-decyl-4-quinazolinylamine (DQA; Fig. 5). DQA is a potent class I/A type inhibitor competitive to piericidin A (*28*) and shown to qualify as a true ubiquinone analog (*29*). Brominating the quinazoline scaffold (QA-1) increased the half-maximal inhibitory concentration (IC_{50}) from 17 nM for DQA to 320 nM, and replacing the side chain with a brominated phenylethylamine moiety (QA-2) hardly affected binding (IC_{50} = 23 nM) (fig. S7). Anomalous Fourier electron density maps (Fig. 5) positioned the bromine atoms close to each other in the vicinity of His^{95} of the 49-kDa subunit at the deepest point of the quinone exchange cavity (Fig. 5B). For both derivatives, the geometrical constraints imposed by anomalous signals and structure are consistent with an identical toxophore position; a quinazoline ring was modeled into the structure to visualize its likely position. The planar aromatic ring system stacks between the tip of loop β_1-$\beta_2^{49\text{-kDa}}$ and Met^{91} of PSST (Fig. 5B), placing the ring bromine at the quinazoline ring and the bromine in the tail moiety 14 Å and 16 Å away, respectively, from cluster N2. Supporting the notion that the position of the substrate analogs also reflects a binding site for ubiquinone, mutations of His^{91}, His^{95}, Val^{97}, Leu^{98}, and Arg^{99} in loop β_1-$\beta_2^{49\text{-kDa}}$ as well as substitution of Met^{91} in PSST by Lys or Glu cause drastic reduction of enzymatic activity (*21*, *27*, *30*, *31*), whereas substitution of the same Met^{91} by Cys showed a marked resistance against DQA and rotenone. Further residues (Ser^{192}, Met^{188}, and Phe^{203} of the 49-kDa subunit and Val^{88} of PSST) with a role in ubiquinone or inhibitor binding (*27*) are located nearby (Fig. 5).

Some distinct structural features in mitochondrial complex I can explain a known functional difference to the bacterial enzyme: The quinone exchange cavity does not provide access to several other residues of the 49-kDa subunit (Asp^{143}, Tyr^{144}, Val^{145}, Asp^{458}, and Val^{460}) for which biochemical evidence obtained with *Y. lipolytica* suggested a direct interaction with ubiquinone and inhibitors (*29*, *30*). Indeed, binding of ubiquinone and piericidin A next to these residues and closer to cluster N2 was modeled into the structure of the bacterial enzyme (*10*), revealing a marked difference (fig. S8) between the otherwise similar (table S3) core structures of *Y. lipolytica* and *T. thermophilus* complex I. Evidently, in the *Y. lipolytica* structure, access deeper into the protein is blocked by the tip of loop β_1-$\beta_2^{49\text{-kDa}}$ bringing His^{95} close to Tyr^{144}, whereas in the bacterial enzyme, the ubiquinone head group was located between these two residues (*10*). The different position of loop β_1-$\beta_2^{49\text{-kDa}}$ is accompanied by significantly different orientations of the adjacent acidic loop $TMH5–6^{ND1}$ and loop $TMH1–2^{ND3}$.

This observation provides a straightforward explanation for the reversible A/D transition (*32*) occurring only in complex I of some eukaryotes, including *Y. lipolytica* (*33*), and discussed as a protective mechanism against excessive oxygen radical formation (*34*). Preparations of *Y. lipolytica* complex I are always in the so-called deactive (D) form, which slowly reverts into the active (A) form upon addition of substrates (*33*). Consistent with the very low catalytic activity of the D form, the structure of *Y. lipolytica* complex I shows the ubiquinone binding site at markedly greater distance from cluster N2, as compared to the always-active structure of the bacterial enzyme, in which ubiquinone was positioned much closer to its electron donor, allowing for efficient electron transfer (*10*). Supporting this interpretation, loop $TMH1–2^{ND3}$ undergoes conformational changes during the A/D transition (*35*). Considering the remarkably high sequence conservation within the structural elements involved (fig. S9) (*10*, *30*) and because binding of ubiquinone next to cluster N2 is also supported by exhaustive mutagenesis data for the mitochondrial enzyme (*21*, *27*, *30*, *31*), we propose that in the A form, the ubiquinone binding pocket of *Y. lipolytica* complex I adopts a conformation similar to the bacterial enzyme that has no D form. We suggest that the interface region of complex I can switch between two distinct conformational states, thereby shifting the ubiquinone binding site. This switch involves a concerted movement of loops from three subunits (49-kDa, ND1, and ND3).

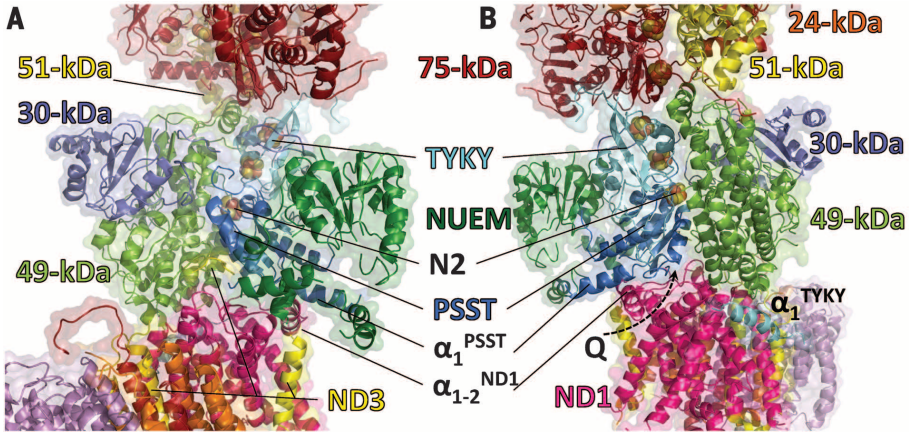


Fig. 3. Interface between peripheral and membrane arm. (**A** and **B**) Two views of the interface between central subunits of the Q module and the P_P module (colors as in Fig. 1) and accessory subunit NUEM (dark green). Dashed arrow indicates putative access for ubiquinone.

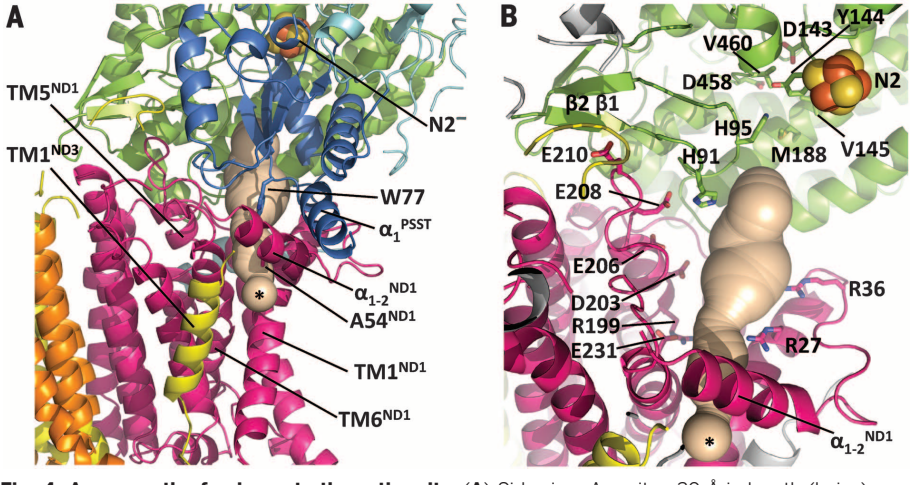


Fig. 4. Access path of quinone to the active site. (**A**) Side view. A cavity ~30 Å in length (beige) connects the active site below cluster N2, with the matrix bilayer leaflet permitting substrate access. A constricted opening (*) is located between the V-shaped arrangement of $TMH1^{ND1}$ and $TMH6^{ND1}$ and below the amphipathic helix α_{1-2}^{ND1} (colors as in Fig. 1). (**B**) View from PSST, which has been removed for clarity.

Mechanistic implications

The central question concerning the mechanism of energy conversion of complex I is how the redox energy released exclusively in the peripheral arm is transmitted to the proton pump modules of the membrane arm (Fig. 6). Several lines of evidence indicate that ubiquinone reduction plays a pivotal role in this process. The hypothetical two-state stabilization-change mechanism (*36*) proposes that stabilization of negatively charged quinone intermediates drives a conformational change, thereby transmitting energy

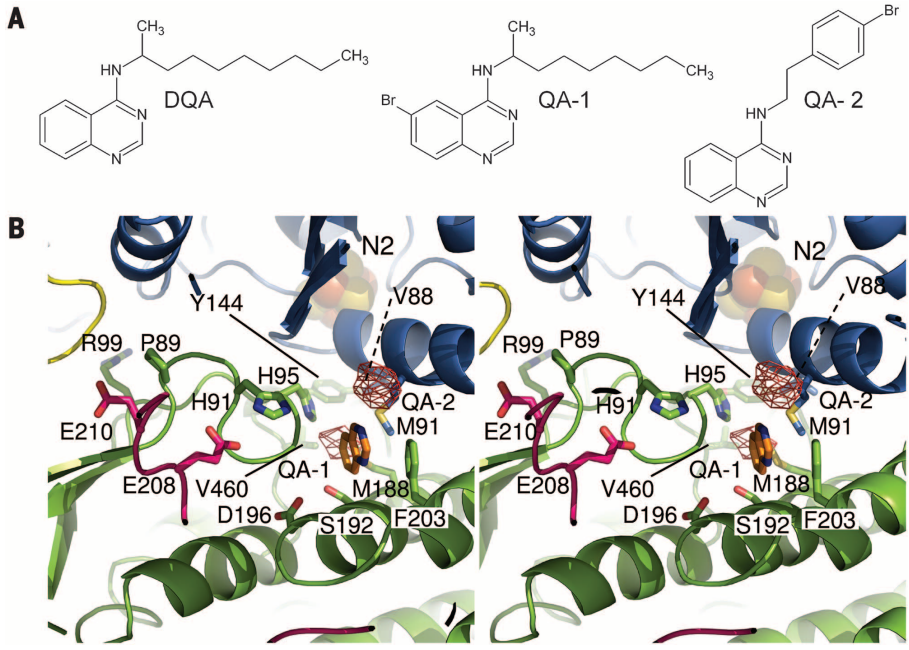


Fig. 5. Binding site of the ubiquinone analogous inhibitor DQA. (**A**) Constitutions of DQA and bromo-substituted derivatives QA-1 and QA-2. (**B**) Stereo view of inhibitor binding pocket (colors as in Fig. 1). Single peaks in the bromine anomalous Fourier electron density maps are shown (purple, QA-1, 3.8σ; red, QA-2, 4σ; superimposition of electron density maps from two separate experiments on structure). Orange, quinazoline ring modeled into the site.

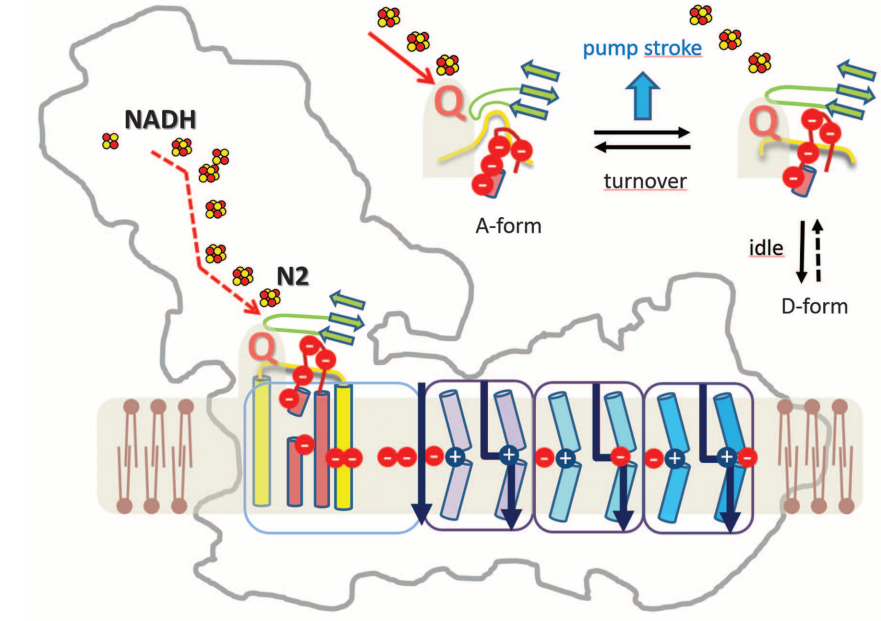


Fig. 6. Hypothetical two-state stabilization change mechanism. Electrons are transferred (red arrow) via a chain of iron-sulfur clusters from NADH to ubiquinone (Q). Within three antiporter-like subunits (violet frame) and the other ND subunits, a pattern of titratable residues defines a central axis in the membrane connected to the IMS and matrix side by putative proton translocation pathways (dark blue). Loop TMH5–6^{ND1} (red), loop β_1-$\beta_2^{49\text{-kDa}}$ (green), and the tip of loop TMH1–2^{ND3} (yellow) line the ubiquinone exchange cavity. During turnover, these loops perform a coordinated rearrangement resulting in a shift of the ubiquinone binding site and movement of the cluster of negative charges in loop TMH5–6^{ND1}, which may trigger an electrostatic pulse toward the membrane arm. Stabilization of the anionic species in the site leads to transition from E state (left) to P state (right), driving a stroke of proton pumping. The idling enzyme can convert reversibly from the active A form into the deactive D form with a structure similar to the P state.

to the membrane arm to drive proton pumping. It postulates that complex I switches between the E state, in which ubiquinone can be reduced by cluster N2, and the P state, in which it is moved away from its electron donor. This description is reminiscent of the A and D state conformations discussed above. A displacement of the acidic loop TMH5–6^{ND1} at the start of the chain of titratable residues reaching through the membrane arm seems ideally suited to transmit an "electrostatic pulse" (*37*). We therefore hypothesize that an orchestrated movement of the three loops associated with the A/D transition could also reflect the critical energy-converting steps during catalytic turnover (Fig. 6). Such a mechanism would imply that the E state corresponds essentially to the A form, whereas the P state would resemble the D form, with the notable exception that the D form cannot revert rapidly and spontaneously to the E/A state, because this is prevented by a yet unidentified structural feature not present in the bacterial enzyme. Because the A/D transition is observed only with eukaryotic enzymes, it is tempting to speculate that stabilization of the D state may involve nearby accessory subunits such as NUEM, which has been shown to take part in the associated conformational changes (*38*).

Notably, this series of events could be triggered not only by stabilization of ubisemiquinone, but also of the ubiquinol anion resulting from the second reduction step (*36*). Although so far experimental evidence for such a second pump stroke is missing, partitioning of the free energy change by making use of both electron transfer steps to ubiquinone seems to make thermodynamic and mechanistic sense, given that complex I can operate in reverse as a proton gradient–driven NAD^+ reductase.

The described mechanistic principle—charge-induced conformational changes that result in secondary electrostatic polarization of charged residues—may also be important to drive the individual proton-pumping sites of complex I. Defined localized conformational changes should ensure controlled vectorial charge translocation, whereas energy transfer between the sites could occur by electrostatic coupling. Indeed, recent large-scale molecular dynamics simulations suggested that long-range energy transmission in complex I is executed through charge-induced protonation changes of key residues (*39*).

REFERENCES AND NOTES

1. U. Brandt, *Annu. Rev. Biochem.* **75**, 69–92 (2006).
2. J. Hirst, *Annu. Rev. Biochem.* **82**, 551–575 (2013).
3. W. J. Koopman, P. H. Willems, J. A. Smeitink, *N. Engl. J. Med.* **366**, 1132–1141 (2012).
4. M. E. Breuer *et al.*, *Neurobiol. Dis.* **51**, 27–34 (2013).
5. S. Dröse, U. Brandt, *Adv. Exp. Med. Biol.* **748**, 145–169 (2012).
6. R. Betarbet *et al.*, *Nat. Neurosci.* **3**, 1301–1306 (2000).
7. S. Kerscher, S. Dröse, K. Zwicker, V. Zickermann, U. Brandt, *Biochim. Biophys. Acta* **1555**, 83–91 (2002).
8. C. Hunte, V. Zickermann, U. Brandt, *Science* **329**, 448–451 (2010).
9. R. G. Efremov, R. Baradaran, L. A. Sazanov, *Nature* **465**, 441–445 (2010).

10. R. Baradaran, J. M. Berrisford, G. S. Minhas, L. A. Sazanov, *Nature* **494**, 443–448 (2013).
11. R. G. Efremov, L. A. Sazanov, *Nature* **476**, 414–420 (2011).
12. C. C. Page, C. C. Moser, X. Chen, P. L. Dutton, *Nature* **402**, 47–52 (1999).
13. C. Hunte *et al.*, *Nature* **435**, 1197–1202 (2005).
14. E. Screpanti, C. Hunte, *J. Struct. Biol.* **159**, 261–267 (2007).
15. N. J. Hu, S. Iwata, A. D. Cameron, D. Drew, *Nature* **478**, 408–411 (2011).
16. V. K. Moparthi *et al.*, *Biochim. Biophys. Acta* **1837**, 178–185 (2014).
17. K. R. Vinothkumar, J. Zhu, J. Hirst, *Nature* **515**, 80–84 (2014).
18. B. C. Marreiros, A. P. Batista, A. M. Duarte, M. M. Pereira, *Biochim. Biophys. Acta* **1827**, 198–209 (2013).
19. R. G. Efremov, L. A. Sazanov, *Nature* **476**, 414–420 (2011).
20. A. Galkin, S. Dröse, U. Brandt, *Biochim. Biophys. Acta* **1757**, 1575–1581 (2006).
21. N. Kashani-Poor, K. Zwicker, S. Kerscher, U. Brandt, *J. Biol. Chem.* **276**, 24082–24087 (2001).
22. H. Jörnvall *et al.*, *Biochemistry* **34**, 6003–6013 (1995).
23. A. Abdrakhmanova, K. Zwicker, S. Kerscher, V. Zickermann, U. Brandt, *Biochim. Biophys. Acta* **1757**, 1676–1682 (2006).
24. M. Ciano, M. Fuszard, H. Heide, C. H. Botting, A. Galkin, *FEBS Lett.* **587**, 867–872 (2013).
25. V. Zickermann *et al.*, *J. Biol. Chem.* **278**, 29072–29078 (2003).
26. V. Zickermann, B. Barquera, M. Wikström, M. Finel, *Biochemistry* **37**, 11792–11796 (1998).
27. H. Angerer *et al.*, *Biochim. Biophys. Acta* **1817**, 1776–1784 (2012).
28. J. G. Okun, P. Lümmen, U. Brandt, *J. Biol. Chem.* **274**, 2625–2630 (1999).
29. M. A. Tocilescu *et al.*, *Biochim. Biophys. Acta* **1797**, 625–632 (2010).
30. M. A. Tocilescu, U. Fendel, K. Zwicker, S. Kerscher, U. Brandt, *J. Biol. Chem.* **282**, 29514–29520 (2007).
31. U. Fendel, M. A. Tocilescu, S. Kerscher, U. Brandt, *Biochim. Biophys. Acta* **1777**, 660–665 (2008).
32. A. B. Kotlyar, A. D. Vinogradov, *Biochim. Biophys. Acta* **1019**, 151–158 (1990).
33. E. Maklashina, A. B. Kotlyar, G. Cecchini, *Biochim. Biophys. Acta* **1606**, 95–103 (2003).
34. E. T. Chouchani *et al.*, *Nat. Med.* **19**, 753–759 (2013).
35. A. Galkin *et al.*, *J. Biol. Chem.* **283**, 20907–20913 (2008).
36. U. Brandt, *Biochim. Biophys. Acta* **1807**, 1364–1369 (2011).
37. L. Euro, G. Belevich, M. I. Verkhovsky, M. Wikström, M. Verkhovskaya, *Biochim. Biophys. Acta* **1777**, 1166–1172 (2008).
38. M. Babot *et al.*, *Biochim. Biophys. Acta* **1837**, 929–939 (2014).
39. V. R. Kaila, M. Wikström, G. Hummer, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6988–6993 (2014).

ACKNOWLEDGMENTS

Supported by the German Research Foundation (CRC 746 to C.H.; ZI 552/3-1 to V.Z.) and the Excellence Initiative of the German Federal and State Governments (EXC 115 to H.S., U.B., and V.Z.; EXC 294 BIOSS to C.H.). We thank the Swiss Light Source and European Synchrotron Radiation Facility for beamline access and staff support during visits, and A. Duchene and G. Beyer for excellent technical assistance. Coordinates and structure factors are deposited in the Protein Data Bank with accession code 4wz7.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/44/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S3
References (*40*–*58*)

11 August 2014; accepted 1 December 2014
10.1126/science.1259859

ORGANIC CHEMISTRY

Nanomole-scale high-throughput chemistry for the synthesis of complex molecules

Alexander Buitrago Santanilla,[1] Erik L. Regalado,[1] Tony Pereira,[2] Michael Shevlin,[1] Kevin Bateman,[2] Louis-Charles Campeau,[1] Jonathan Schneeweis,[3] Simon Berritt,[1] Zhi-Cai Shi,[4] Philippe Nantermet,[5] Yong Liu,[1] Roy Helmy,[1] Christopher J. Welch,[1] Petr Vachal,[6] Ian W. Davies,[1] Tim Cernak,[7]* Spencer D. Dreher[1]*

At the forefront of new synthetic endeavors, such as drug discovery or natural product synthesis, large quantities of material are rarely available and timelines are tight. A miniaturized automation platform enabling high-throughput experimentation for synthetic route scouting to identify conditions for preparative reaction scale-up would be a transformative advance. Because automated, miniaturized chemistry is difficult to carry out in the presence of solids or volatile organic solvents, most of the synthetic "toolkit" cannot be readily miniaturized. Using palladium-catalyzed cross-coupling reactions as a test case, we developed automation-friendly reactions to run in dimethyl sulfoxide at room temperature. This advance enabled us to couple the robotics used in biotechnology with emerging mass spectrometry–based high-throughput analysis techniques. More than 1500 chemistry experiments were carried out in less than a day, using as little as 0.02 milligrams of material per reaction.

High-throughput experimentation (HTE) chemistry tools have been used to aid in the discovery of new reactions (*1*–*6*) and in the scale-up optimization of known reactions (*7*–*9*), both areas where substrates for experimentation are plentiful. However, HTE is rarely used in the area where it might have the greatest impact: the synthesis of complex natural products or highly functionalized drug leads. What is needed is a tool that would allow chemists to locate successful reaction conditions (*10*) on a microgram (nanomole) scale in a high-throughput fashion without depleting precious substrate stores.

In the search for breakthrough medicines in biomedical research, the rapid preparation of new, complex molecules for biological evaluation is of paramount importance, but substrates for the synthesis of such compounds are invariably in short supply. In later stages of chemistry development, substrates are abundant, and state-of-the-art microvial (*8*, *9*) or microfluidic (*11*) HTE tools can be effective in "turning on" reactions that were otherwise unsuccessful by exploring combinations of catalysts, reagents, and other key reaction variables (*12*, *13*). Such studies require at least milligram (micromole) quantities of substrate per reaction—a prohibitively large amount in early drug discovery, where new molecules are prepared for the first time. Consequently, the set of desirable compounds designed to test a biological hypothesis is often winnowed to the much smaller subset of compounds that can be successfully synthesized using a single set of reaction conditions, with little opportunity to study and improve unsuccessful syntheses (*14*, *15*). Miniaturizing chemistry to the nanomole scale is a potential solution to this problem that has heretofore met with substantial engineering problems, such as accurately bringing together extremely small charges of materials that are often heterogeneous, effectively agitating reaction mixtures, preventing loss of volatile solvents, and incorporating general analytical approaches to assay reaction outcomes. We present the first results of a study aimed at developing general nanomole reaction screening capabilities to support the rapid synthesis of complex, highly functionalized drug leads.

Figure 1A shows highly functionalized molecules from Merck's compound collection representative of the cores of drug molecules (**1** to **8**), which in the search for new molecules with optimal biological properties would be coupled to a diversity of polar building blocks (**9** to **20**) such as those in Fig. 1B. Modern organic synthesis (and especially transition metal catalysis) is redefining the rules with which new bonds can be forged; however, it is important to recognize that many "solved" synthetic transformations are far from universal, performing well on simple model substrates yet often failing when applied to complex substrates in real-world synthesis (*16*). A recent analysis of 2149 metal-catalyzed C-N

[1]Department of Process and Analytical Chemistry, Merck Research Laboratories, Merck and Co. Inc., Rahway, NJ 07065, USA. [2]Department of Pharmacokinetics, Pharmacodynamics and Drug Metabolism, Merck Research Laboratories, Merck and Co. Inc., West Point, PA 19486, USA. [3]Department of Pharmacology, Merck Research Laboratories, Merck and Co. Inc., Kenilworth, NJ 07033, USA. [4]Department of Discovery Chemistry, Merck Research Laboratories, Merck and Co. Inc., Kenilworth, NJ 07033, USA. [5]Department of Discovery Chemistry, Merck Research Laboratories, Merck and Co. Inc., West Point, PA 19486, USA. [6]Department of Discovery Chemistry, Merck Research Laboratories, Merck and Co. Inc., Rahway, NJ 07065, USA. [7]Department of Discovery Chemistry, Merck Research Laboratories, Merck and Co. Inc., Boston, MA 02115, USA.
***Corresponding author. E-mail: timothy_cernak@merck.com (T.C.); spencer_dreher@merck.com (S.D.D.)**

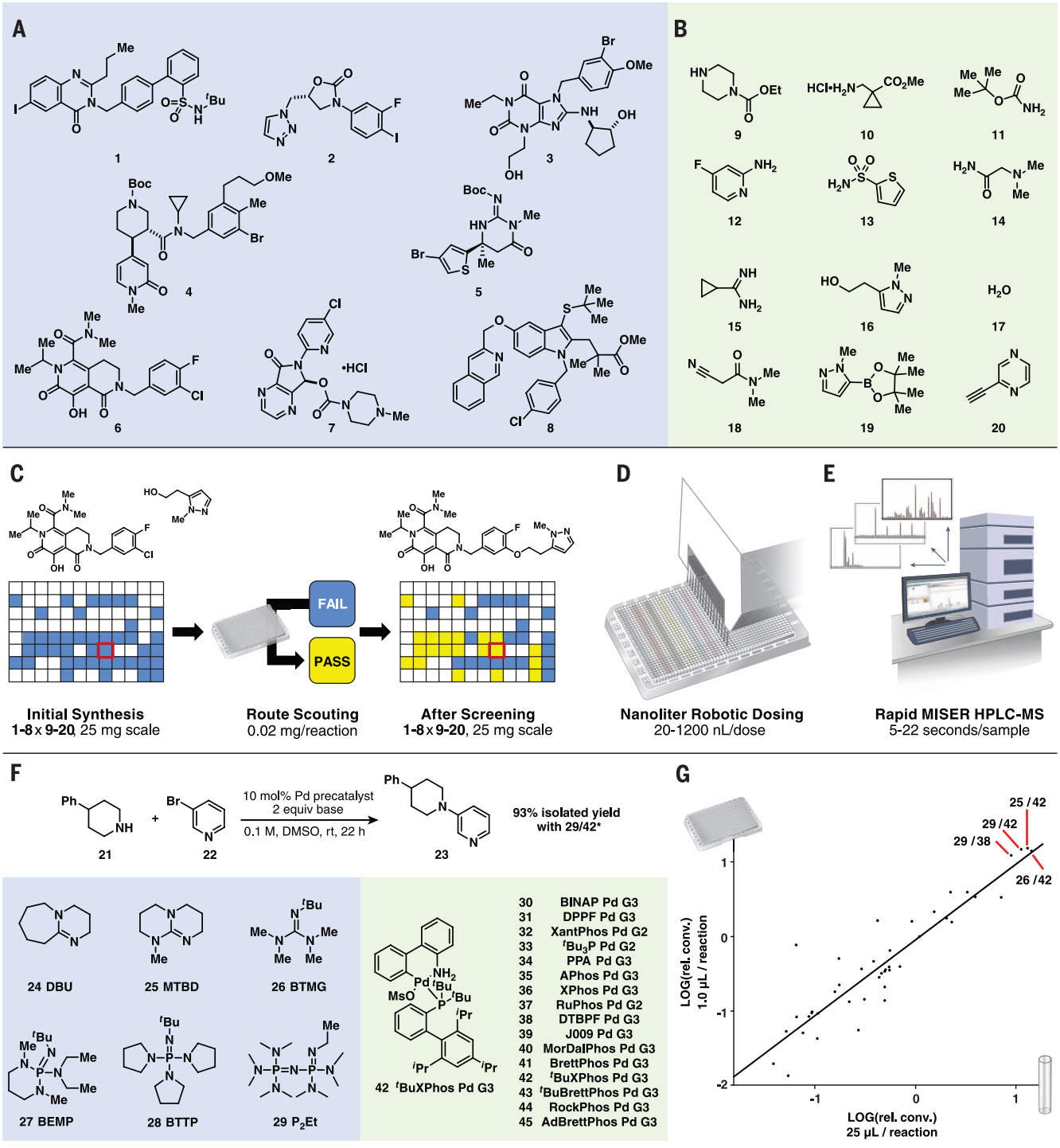


Fig. 1. Nanomole-scale reaction screening. (**A**) High-complexity electrophilic cores **1** to **8** from the Merck compound collection. (**B**) High-polarity nucleophilic building blocks **9** to **20**. (**C**) Route scouting to identify successful reaction conditions for previously failed chemistry can be accomplished using nanomole-scale reaction screening with minimal consumption of precious substrate (0.02 mg per reaction). Successful reactions run on nanoscale (0.02 mg, 50 nmol) can be repeated on a 25-mg scale (factor of 1000 scale-up). A 96-member parallel coupling library using best reaction conditions located by screening model nucleophiles (table S3) gave 54 products (white squares) and 42 failed reactions (blue squares) (table S4). After nanomole-scale screening, productive conditions were located for an additional 21 of the 96 products (yellow squares) (tables S6 and S8). (**D**) Reaction mixtures that are fully soluble in high-boiling solvents can be dosed with high-precision nanoliter robotics. Room-temperature reactions require no stirring and minimal sealing. (**E**) High-performance LC-MS MISER analysis allows for rapid label-free analysis of reactions with diverse products. (**F**) Reaction discovery: Organic superbases **24** to **29** promote room-temperature cross-coupling in DMSO with second- and third-generation biaryl palladium precatalysts **30** to **45**. (**G**) Ninety-six combinations of bases and ligands were investigated in triplicate in glass microvials with stirring and reproduced in 1536-well plastic microtiter plates without stirring.

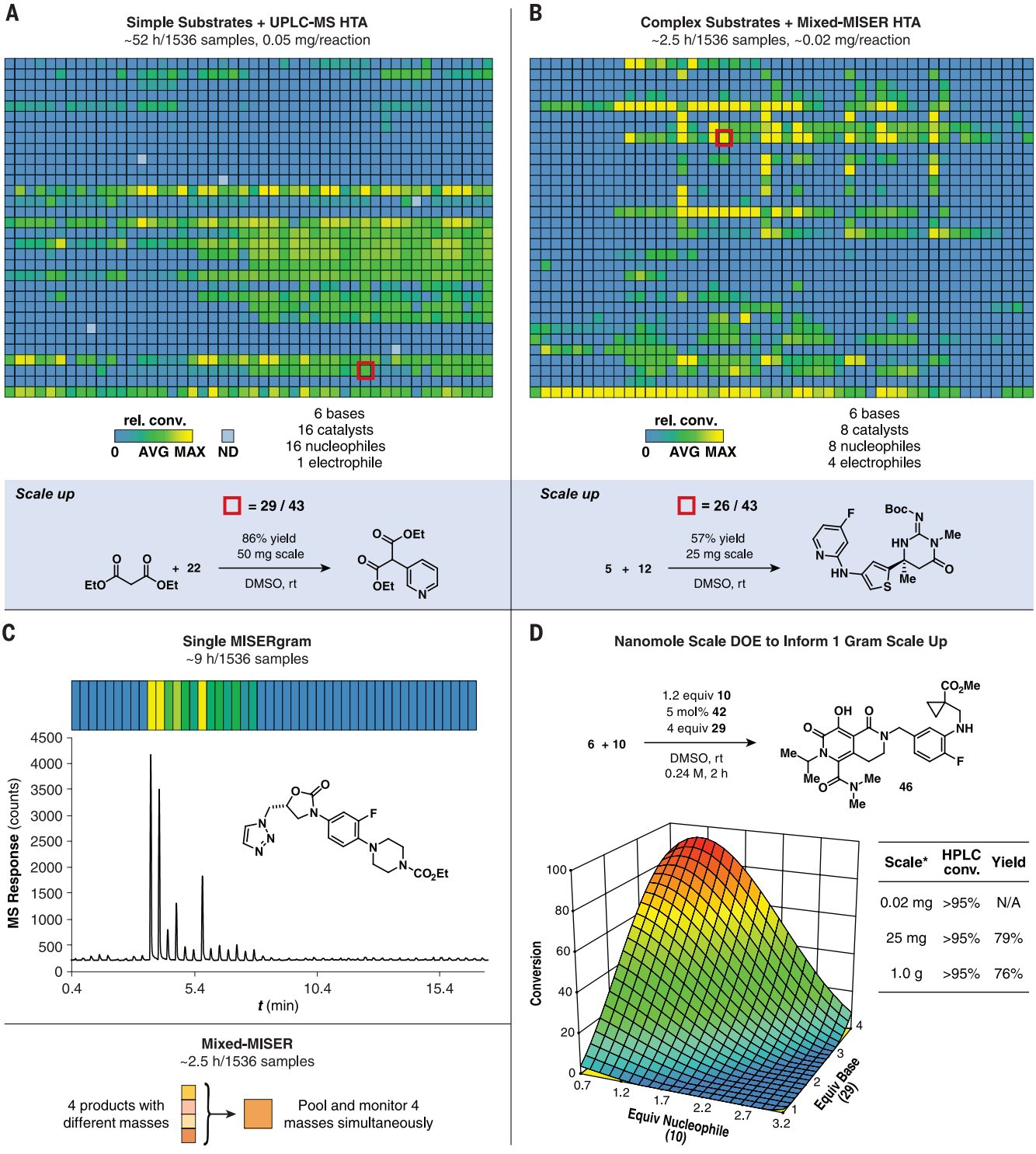


Fig. 2. A high-throughput nanomole-scale chemistry evaluation using rapid MISER and mixed-MISER LC-MS. (**A**) Heat map showing data from 1536 nanomole-scale reactions analyzed by UPLC (total analysis time ~52 hours). Sixteen diverse N/C/O/P/S nucleophiles were screened against 96 catalyst-base combinations to find room-temperature coupling conditions to 3-bromopyridine **22**. Best conditions for each nucleophile were repeated on a 50-mg scale, giving 68 to 95% isolated yields (table S3). Reaction information about best catalyst-base combination was matched to the polar nucleophile used in the 96-member parallel coupling library of **1** to **8** × **9** to **20** (Fig. 1C, initial synthesis). (**B**) Heat map showing data from 1536 nanomole-scale reactions analyzed by MISER. Thirty-two electrophile-nucleophile combinations that failed under previous best conditions were investigated under 48 catalyst-base combinations. Best conditions for each combination were repeated on a 50-μmol scale and assigned a "PASS" (Fig. 1C, after screening) if a >95% purity sample was obtained. (**C**) MISER chromatographic output (MISERgram) for 48 catalyst-base combinations that were screened with 0.02 mg of **2** per reaction, using **9** as nucleophile, and depiction of pooled MISER approach. (**D**) Nanomole-scale DOE: three-factorial, four-level DOE surface model produced by screening **6** against reagent charges of **10**, **29**, and **42**. *Conditions identified in 0.02 mg scale reactions informed practical 25-mg and 1-g scale-ups.

couplings run in the final steps of the synthesis of highly functionalized drug leads revealed that 55% of reactions failed to deliver any product at all (*17*). The missed opportunity represented by these unsuccessful syntheses haunts contemporary drug discovery, and there is a growing recognition that the tendency of polar, highly functionalized compounds to fail in catalysis may actually enrich compound sets in hydrophobic molecules that are less likely to become successful drug candidates (*15*, *16*, *18*–*20*).

Soluble superbase coupling conditions for miniaturization from micromole to nanomole scale

Using the catalysis cross-coupling space defined in Fig. 1, A and B, as a test case, this work evaluates the idea that miniaturized chemistry experimentation to inform preparative scale synthesis (Fig. 1C) could be run on the same scale as contemporary biological assays using the high-precision nanoliter robotics commonly used in biochemistry labs (Fig. 1D) in conjunction with high-throughput mass spectrometric analysis (Fig. 1E). The engineering complexities of this task could potentially be overcome by using ambient-temperature chemistry that operates in a solubilizing, low-volatility, plastic-compatible reaction medium such as dimethyl sulfoxide (DMSO) or *N*-methyl-2-pyrrolidone (NMP). We chose DMSO for the initial proof of concept because it is an environmentally friendly solvent and represents the potential to directly link chemistry with biochemical experimentation. Very few efficient ambient-temperature palladium-catalyzed cross-couplings in DMSO or other polar solvents have been reported (*21*), in part because of the strong solvent-metal coordination of such solvents and their incompatibility with the strong bases typically used in low-temperature C-N couplings (NaO^tBu, $LiN(SiMe_3)_2$, $Zn(N(SiMe_3)_2)_2$ etc.) (*22*). It seemed possible that recently reported highly hindered, electron-rich ligands (*23*) may effectively protect Pd from DMSO coordination and that non-nucleophilic organic superbases (*24*), which are soluble in most solvents and should be strongly basic enough to promote Pd C-N couplings at room temperature, could be compatible with these solvents.

Initially working in glass microvials, which represent the current validated lower limits of scalable batch miniaturization (*8*, *9*), the coupling of amine **21** with bromide **22** (2.5 μmol in 25 μl of DMSO) was evaluated in a 96-reaction array of superbases **24** to **29** with catalysts **30** to **45** [10 mole percent (mol %)] (Fig. 1F), and several highly productive room-temperature coupling conditions were identified. All solutions were homogeneous but were stirred nonetheless. This same 96-reaction array was then evaluated on a 1.0-μl scale (100 nmol of substrate) in a plastic 1536-well plate without stirring and with reagent dosing via a nanoliter liquid-handling robot (TTP LabTech Mosquito HTS) inside a nitrogen-filled glovebox, demonstrating the same hits and reaction performance (Fig. 1G). The Mosquito robot allows sequential aspiration from different wells of a source plate into a single pipette tip to create multicomponent mixtures that are dosed as single reaction drops into a 1536-well plate. This ensures proper mixing of the reaction components and obviates the need for stirring. This pipetting mechanism, in conjunction with a material-sparing source plate with minimal dead volume, ensures that almost no material is wasted, and the overage can be recovered if desired. Also, with this batch reaction approach, some degree of sample heterogeneity can be tolerated, unlike for microfluidic technologies where channel clogging can lead to systemic failure.

The coupling of bromide **22** was then evaluated with 16 nitrogen, oxygen, carbon, phosphorus, and sulfur nucleophiles under the same 96 combinations of **24** to **45** in a single miniaturized 1536-reaction experiment (Fig. 2A and table S3). In this experiment, dosing required 30 min (four components per well, 6144 total reagent doses) followed by 22 hours of reaction time, 1 hour for automated reaction quenching and sampling, and 52 hours for analysis by ultra performance liquid chromatography (UPLC). This experiment uncovered extremely mild, soluble coupling conditions for amines, alcohols, amides, sulfonamides, carbamates, amidines, aryl boronates, alkynes, and malonates, most of which have not been reported at ambient temperature with non-nucleophilic bases (table S3). Because the reactions were by design homogeneous and required no heating or cooling, translation to a larger scale was straightforward, delivering 68 to 98% yield when scaled up by a factor of 3000, to 320 μmol of **22** (50 mg), now running at 5 mol % catalyst loading and 0.2 M substrate concentration.

Implementation of nanomole screening for complex molecule synthesis

These new coupling conditions were next applied in complex synthesis to the test set of electrophiles **1** to **8** and nucleophiles **9** to **20** (Fig. 1, A and B). Initial synthesis conditions for this array were selected from the best permutation of superbase and catalyst that was previously identified for each nucleophile class when using bromide **22** (table S3). With these conditions, high-purity product samples were successfully prepared for 54 of the 96 functionality-rich substrate combinations (Fig. 1C and table S4) when run on a 500-μmol scale (~25 mg) (*25*, *26*). This chemistry was generally tolerant of polar functional groups and heterocycles; however, nearly half of the 96 reactions failed to provide any product, which is typical for application of a single set of conditions to a complex catalysis problem.

To apply nanomole-scale reaction screening to this problem, we subjected 32 different electrophile-nucleophile substrate combinations, each of which produced little or no material from the 96-reaction array, to 48 reaction conditions (8 organic superbases × 6 ligands; Fig. 2B). The substrate concentration was reduced to 0.05 M and the catalyst loading was doubled to 20 mol % so that running all 48 reactions would consume less than 1 mg of each substrate (50 nmol, ~0.02 mg per reaction) (*27*). In addition, the reaction time was reduced to 2 hours and a faster liquid chromatography–mass spectrometry (LC-MS) approach was pursued, with a goal of running and analyzing 1536 reactions in less than 24 hours. Of the available label-free techniques (*28*), MISER (multiple injections in a single experimental run) LC-MS was selected for its ease of data acquisition and analysis (*29*). Thus, 48 reactions that had the same potential product along a row of the 1536-well plate were analyzed by multiple injections (22 s per sample) in a single isocratic chromatographic run with mass detection settings observing the molecular ion for the desired product (Fig. 2C). In this way, the 1536-well plate could be analyzed in ~9 hours. To reduce the time required for analysis even further, we combined four rows of reaction samples that could give different product masses and simultaneously monitored the four desired products' *m/z* values (mass/charge ratios) in these pooled samples (Fig. 2C). By pooling four mass-encoded wells, the analysis of all 1536 reactions could be performed in ~2.5 hours. In addition, a more rigorous approach to this same analysis was pursued using a parallel two-channel LC system connected to a triple-quadrupole mass spectrometer. This instrumentation used a fast gradient to provide improved resolution and additional MS/MS structure confirmation data and required ~2.5 hours for analysis of all 1536 reactions. The MISER analyses located the expected product masses for 21 of 32 substrate combinations, and the best hits were then confirmed with UPLC-MS. These 21 hits were scaled up by a factor of 1000 (25-mg scale), and a pure product sample was obtained for 16 of these reactions (table S6).

Optimization of continuous variables (for example, the stoichiometry of reagents) is another potential path to increase product formation in HTE. Returning to the conditions originally identified for bromide **22**, six of the remaining failed reactions were evaluated under 48 combinations of catalyst loading, nucleophile charge, and base charge, again using 1 mg of substrate per 48-reaction evaluation (fig. S22). Markedly better conditions for five of the six compounds were found in this way, which allowed their isolation upon scale-up by a factor of 1000 (table S8). Hence, in these two miniaturized screening experiments, looking first at catalyst-base combinations and then at reaction stoichiometries, 21 complex new compounds could be rapidly prepared. These solutions to 21 complex synthesis problems represent the opportunity to answer biological hypotheses that would have been passed over in previous chemistry approaches because of lack of time and material.

Nanomole reaction optimization to inform gram-scale synthesis

An important consideration for this nanomole-scale approach toward discovery synthesis is the optimization of reactions for application to larger-scale synthesis. Compounds that are deemed interesting in initial biological evaluation or are synthetic intermediates in a multistep route must rapidly be prepared in larger quantities once conditions are identified. Statistical tools such

as DOE (design of experiments) are powerful tools for synthesis (*30*–*32*), yet they typically focus on performing minimal numbers of experiments, as large numbers of experiments are often resource- and time-intensive to conduct. However, because large numbers of experiments are readily feasible with this nanomole-scale chemistry platform, we were able to construct a three-factorial, four-level response surface modeling experiment to study the loading of catalyst against varying stoichiometries of base and nucleophile in the reaction of chloride **6** with amine **10**. In this DOE experiment, each condition was repeated twice, resulting in 128 total reactions with <3 mg of **6**. Indeed, a high-quality response surface model was generated with the nanomole-scale chemistry approach (Fig. 2D), which helped to define the critical charges of nucleophile and base for optimal reaction performance. The optimized conditions used 15 mol % **42** at 0.05 M concentration; by translating to more practical conditions of 5 mol % **42** and 0.24 M concentration, we obtained full conversion and a 79% isolated yield of **46** on a 25-mg scale, which was reproduced to obtain a 76% isolated yield on a 1-g scale (Fig. 2D). This result shows that advanced statistical reaction analysis, which is typically reserved for chemistry opportunities where material is plentiful, can be applied to reactions in the material-limited front lines of drug discovery or natural product synthesis.

In biomedical research, chemical synthesis should not limit access to any molecule that is designed to answer a biological question. This work demonstrates an example of how conditions for complex Pd-catalyzed C-O, C-N, and C-C cross-coupling reactions can be evolved into a powerful, substrate-focused approach to chemistry miniaturization to overcome limited access to complex products. With innovative research, other high-value modern chemistry reactions could be similarly designed into this paradigm to improve synthesis in material-limited environments by evolution of catalysts and reagents to perform in DMSO, NMP, or other high-boiling solvents at ambient temperature.

REFERENCES AND NOTES

1. M. R. Friedfeld *et al.*, *Science* **342**, 1076–1080 (2013).
2. D. A. DiRocco *et al.*, *Angew. Chem. Int. Ed.* **53**, 4802–4806 (2014).
3. D. W. Robbins, J. F. Hartwig, *Science* **333**, 1423–1427 (2011).
4. A. McNally, C. K. Prier, D. W. C. MacMillan, *Science* **334**, 1114–1117 (2011).
5. K. D. Collins, T. Gensch, F. Glorius, *Nat. Chem.* **6**, 859–871 (2014).
6. R. Moreira, M. Havranek, D. Sames, *J. Am. Chem. Soc.* **123**, 3927–3921 (2001).
7. S. M. Preshlock *et al.*, *J. Am. Chem. Soc.* **135**, 7572–7582 (2013).
8. A. Bellomo *et al.*, *Angew. Chem. Int. Ed.* **51**, 6912–6915 (2012).
9. J. R. Schmink, A. Bellomo, S. Berritt, *Aldrichim. Acta* **46**, 71–80 (2013).
10. M. Peplow, *Nature* **512**, 20–22 (2014).
11. T. Rodrigues, P. Schneider, G. Schneider, *Angew. Chem. Int. Ed.* **53**, 5750–5758 (2014).
12. S. Monfette, J. M. Blacquiere, D. E. Fogg, *Organometallics* **30**, 36–42 (2011).
13. P. M. Murray, S. N. G. Tyler, J. D. Moseley, *Org. Process Res. Dev.* **17**, 40–46 (2013).
14. S. D. Roughley, A. M. Jordan, *J. Med. Chem.* **54**, 3451–3479 (2011).
15. T. W. J. Cooper, I. B. Campbell, S. J. F. Macdonald, *Angew. Chem. Int. Ed.* **49**, 8082–8091 (2010).
16. A. Nadin, C. Hattotuwagama, I. Churcher, *Angew. Chem. Int. Ed.* **51**, 1114–1122 (2012).
17. Merck internal study of electronic notebooks.
18. M. M. Hann, G. M. Keserü, *Nat. Rev. Drug Discov.* **11**, 355–365 (2012).
19. F. Lovering, J. Bikker, C. Humblet, *J. Med. Chem.* **52**, 6752–6756 (2009).
20. H. A. Malik *et al.*, *Chem. Sci.* **5**, 2352–2361 (2014).
21. R. E. Tundel, K. W. Anderson, S. L. Buchwald, *J. Org. Chem.* **71**, 430–433 (2006).
22. D. S. Surry, S. L. Buchwald, *Chem. Sci.* **2**, 27–50 (2011).
23. N. C. Bruno, M. T. Tudge, S. L. Buchwald, *Chem. Sci.* **4**, 916–920 (2013).
24. T. Ishikawa, Y. Kondo, H. Kotsuki, T. Kumamoto, D. Margetic, K. Nagasawa, W. Nakanishi, in *Superbases for Organic Synthesis: Guanidines, Amidines, Phosphazenes and Related Organocatalysts*, T. Ishikawa, Ed. (Wiley, West Sussex, UK, ed. 1, 2009), pp. 1–326.
25. Compounds were generally purified by MS-directed purification. Isolated yields ranged from 1 to 100%, but we made no attempt to maximize the isolated yields in these reactions and instead focused on obtaining high-purity compounds as quickly as possible, which is typical in most medicinal chemistry campaigns. Some reactions showed product formation by UPLC-MS analysis but were either insufficiently pure or too low in yield for purification.
26. M. Liu *et al.*, *ACS Comb. Sci.* **14**, 51–59 (2012).
27. Some electrophiles were not fully soluble in DMSO, so NMP was used instead. Even though three stock solutions in NMP still displayed mild insolubility, the TTP Mosquito operates on positive-displacement pipetting, so viscous solutions or suspensions of small particulates are easily transferred.
28. W. Schafer, X. Bu, X. Gong, L. A. Joyce, C. J. Welch, in *Comprehensive Organic Synthesis*, C. J. Welch, Ed. (Elsevier, Oxford, ed. 2, 2014), vol. 9, pp. 28–53.
29. C. J. Welch *et al.*, *Tetrahedron Asymmetry* **21**, 1674–1681 (2010).
30. J. C. Ianni, V. Annamalai, P.-W. Phuan, M. Panda, M. C. Kozlowski, *Angew. Chem. Int. Ed.* **45**, 5502–5505 (2006).
31. S. E. Denmark, C. R. Butler, *J. Am. Chem. Soc.* **130**, 3690–3704 (2008).
32. K. C. Harper, M. S. Sigman, *Science* **333**, 1875–1878 (2011).

ACKNOWLEDGMENTS

We thank S. Krska, M. Tudge, G. Hughes, and E. Parmee for helpful discussions; M. Liu, E. Streckfuss, T. Meng, N. Pissarniski, and W. Li for assistance in purification of compounds; M. Christensen and J. Voigt for experimental assistance; and S. M. O'Brien and M. McColgan for graphic design. S.B. was supported by an NSF GOALI Grant associated with the University of Pennsylvania. Supported by the MRL Postdoctoral Research Fellows Program (A.B.S. and E.L.R.).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/49/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S32
Tables S1 to S11
References (*33*–*40*)
Data Files S1 to S5

25 July 2014; accepted 11 November 2014
Published online 20 November 2014;
10.1126/science.1259203

REPORTS

QUANTUM OPTICS

Quantum harmonic oscillator state synthesis by reservoir engineering

D. Kienzler,* H.-Y. Lo, B. Keitch, L. de Clercq, F. Leupold, F. Lindenfelser, M. Marinelli, V. Negnevitsky, J. P. Home*

The robust generation of quantum states in the presence of decoherence is a primary challenge for explorations of quantum mechanics at larger scales. Using the mechanical motion of a single trapped ion, we utilize reservoir engineering to generate squeezed, coherent, and displaced-squeezed states as steady states in the presence of noise. We verify the created state by generating two-state correlated spin-motion Rabi oscillations, resulting in high-contrast measurements. For both cooling and measurement, we use spin-oscillator couplings that provide transitions between oscillator states in an engineered Fock state basis. Our approach should facilitate studies of entanglement, quantum computation, and open-system quantum simulations in a wide range of physical systems.

Reservoir engineering is a method in which specially designed couplings between a system of interest and a zero-temperature environment can be used to generate quantum superposition states of the system as the steady state of the dissipative process, independent of the initial state of the system (*1*–*3*). Theoretical work has shown the potential for using such engineered dissipation for universal quantum computation (*4*) and in providing new routes to many-body states (*5*–*7*). Experimentally, these techniques have been used to generate entangled superposition states of qubits in atomic ensembles (*8*), trapped ions (*9*, *10*), and superconducting circuits (*11*). Theoretical proposals for quantum harmonic oscillator state synthesis by reservoir engineering extend from trapped ions (*2*, *3*) to superconducting cavities (*12*, *13*) and nanomechanics (*14*).

Here, we experimentally demonstrate the generation and stabilization of quantum harmonic oscillator states by reservoir engineering based on the original proposal of Cirac *et al.* (*1*), which we use to generate and stabilize squeezed,

Institute for Quantum Electronics, ETH Zürich, Otto-Stern-Weg 1, 8093 Zürich, Switzerland.
***Corresponding author. E-mail: daniel.kienzler@phys.ethz.ch (D.K.); jhome@phys.ethz.ch (J.P.H.)**

displaced-squeezed, and coherent states. Making use of engineered spin-motion couplings that are closely related to those used in the reservoir engineering, we introduce measurement techniques that provide simple spin population dynamics, allowing us to directly verify the coherence of the states produced and providing a measure of the fidelity with a high signal-to-noise ratio.

The energy eigenstates of the harmonic oscillator $|n\rangle$ form an equally spaced ladder connected by the action of the creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$. Alternative state ladders exist, in which each state is a superposition of energy eigenstates. These can be obtained by applying a unitary transformation, with the resulting states $|\hat{U},n\rangle \equiv \hat{U}|n\rangle$ (Fig. 1A). The transformed state ladders have their own annihilation operators $\hat{K}$, which are related to $\hat{a}$ by $\hat{K} = \hat{U}\hat{a}\hat{U}^\dagger$ (the same transformation can be performed for the creation operator). State preparation by reservoir engineering involves the choice of a suitable engineered basis for which cooling to the ground state results in the desired quantum state $|\hat{U},0\rangle$. In our experiments, we can cool in this basis by coupling the oscillator to an ancilla spin. We use an engineered spin-motion coupling Hamiltonian

$$\hat{H}_- = \hbar\Omega(\hat{K}\hat{\sigma}_+ + \hat{K}^\dagger\hat{\sigma}_-) \qquad (1)$$

where Ω is the Rabi frequency and $\hat{\sigma}_+ \equiv |\uparrow\rangle\langle\downarrow|$, $\hat{\sigma}_- \equiv |\downarrow\rangle\langle\uparrow|$ are spin flip operators. This Hamiltonian results in transitions between adjacent levels on the transformed state ladder, correlated with spin flips. The Hamiltonian dynamics are reversible, and thus cannot reduce entropy. To produce a zero-entropy pure state from a general starting state, dissipation is required, which we introduce by optical pumping of the spin. This pumps the oscillator down the engineered state ladder into the desired ground state (Fig. 1B).

We generate Gaussian oscillator states, which are related to the energetic ground state by combinations of displacements and squeezing of the wave packet (*15*). The unitary transformation is then $\hat{U} = \hat{S}(\xi)\hat{D}(\alpha)$, where $\hat{S}(\xi)$ is the squeezing operator and $\hat{D}(\alpha)$ is the displacement operator (*16*). The resulting annihilation operator in the engineered basis is $\hat{K} = e^{i\phi}(\cosh(r)\hat{a} + e^{i\phi_s}\sinh(r)\hat{a}^\dagger - \alpha)$, where $r = |\xi|$ and $\phi_s = \arg(\xi)$. The phase factor ϕ plays no role in our experiments and we set it to zero in the rest of this Report. $\hat{K}$ contains terms that are linear in the creation and annihilation operators for the energy eigenstates, meaning that the Hamiltonian $\hat{H}_-$ can be implemented by simultaneously applying a combination of the carrier ($\hat{H}_c = \hbar\Omega_c\hat{\sigma}_+ + \text{h.c.}$, where h.c. is the Hermitian conjugate), red motional sideband ($\hat{H}_{rsb} = \hbar\Omega_{rsb}\hat{\sigma}_+\hat{a} + \text{h.c.}$), and blue motional sideband ($\hat{H}_{bsb} = \hbar\Omega_{bsb}\hat{\sigma}_+\hat{a}^\dagger + \text{h.c.}$) transitions. Here Ω_c, Ω_{rsb}, and Ω_{bsb} are taken to be complex parameters, containing both the coupling strength and the phase. In our experiments, these Hamiltonians can be realized simultaneously by applying a multifrequency laser field, with each frequency component resonant with one of the transitions (*16*). The squeezing is obtained from the ratio $\Omega_{bsb}/\Omega_{rsb} = e^{i\phi_s}\tanh(r)$, and the displacement from the ratio $\Omega_c/\Omega_{rsb} = -\alpha/\cosh(r)$.

The experiments work with a single $^{40}Ca^+$ ion, which oscillates along a chosen direction with a frequency close to $\omega_z/(2\pi) = 1.9$ MHz. At the start of each experimental run, the ion is initialized by cooling all motional modes close to the Doppler limit using laser light resonant with the dipole transitions at 397 and 866 nm. All coherent manipulations (including the Hamiltonians used for reservoir engineering) make use of the quadrupole transition at 729 nm, isolating a two-state pseudospin system that we identify as $|\downarrow\rangle \equiv |L=0, J=1/2, M_J=+1/2\rangle$ and $|\uparrow\rangle \equiv |L=2, J=5/2, M_J=3/2\rangle$ (*17*). The Lamb-Dicke factor for our laser configuration is $\eta = 0.05$, which means that the experiments are well described by the Lamb-Dicke approximation (*18*). Optical pumping to $|\downarrow\rangle$ is implemented with a combination of linearly polarized light fields at 854, 397, and 866 nm (*16*). The internal electronic state of the ion is read out with state-dependent fluorescence (*16*).

The reservoir engineering is applied directly after the end of the Doppler cooling. We subsequently probe whether the state of the system has reached the dark state for the Hamiltonian $\hat{H}_-$ by optical pumping into $|\downarrow\rangle$, followed by a probe pulse using $\hat{H}_-$. Examples of data for a coherent ($\Omega_{bsb} = 0$) and a squeezed state ($\Omega_c = 0$) are shown in Fig. 2. In both cases, the state approaches a steady state that approximates a dark state of the pumping Hamiltonian. Experimentally, we have implemented two different methods of dissipative pumping. In the first (used for the coherent state data), we repeat a "cycle" involving applying $\hat{H}_-$ for a fixed duration followed by repumping of the internal state. The second method (used for the squeezed state data) involves continuous application of both $\hat{H}_-$ and the spin dissipation. We observe that the motional state pumping is faster in the continuous case. The pulsed method is easier to maintain, as it avoids AC-Stark shifts arising from the repumping laser. It also allows the use of shaped pulses to produce sideband transitions while avoiding undesired off-resonant excitation of the carrier transition (*19*).

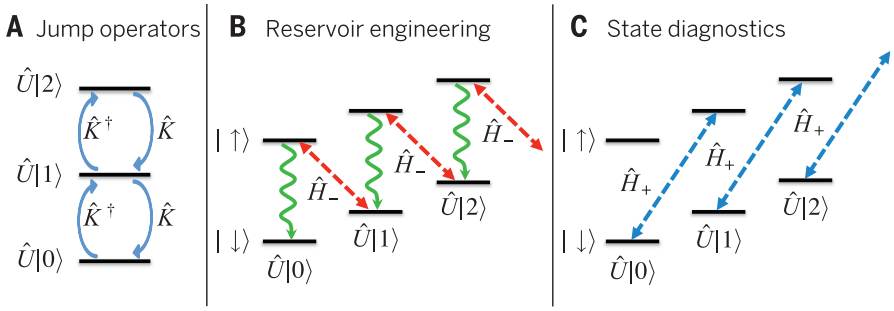


Fig. 1. Cooling and detection using an engineered state ladder. (**A**) The operators $\hat{K}$, $\hat{K}^\dagger$ are ladder operators for the set of basis states $\hat{U}|n\rangle$. (**B**) A combination of the reversible Hamiltonian $\hat{H}_-$ and internal state dissipation leads to pumping into the ground state of the engineered basis. (**C**) State measurement probes the $\hat{H}_+$ Hamiltonian, resulting in single-frequency Rabi oscillations if the system is prepared in $\hat{U}|0\rangle$.

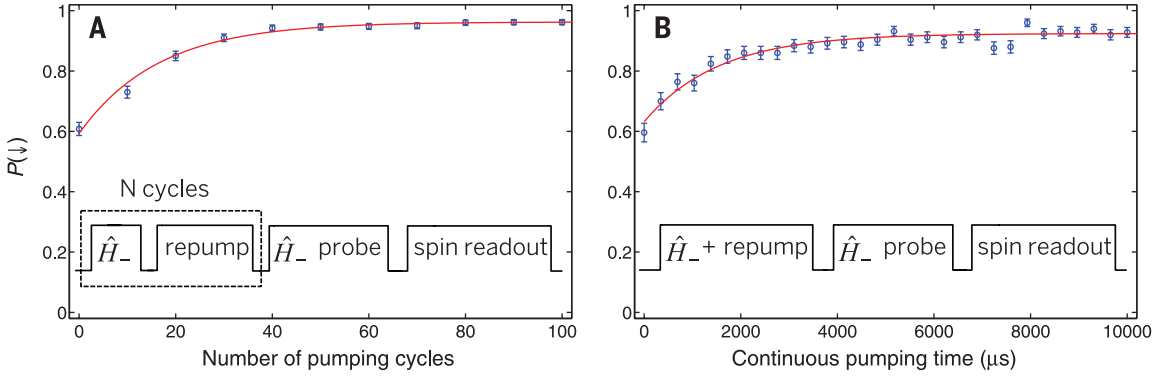


Fig. 2. Onset of the dark state of the reservoir engineering. (**A**) Pulsed preparation of a coherent state. Each pulse cycle consists of applying $\hat{H}_-$ for 25 μs and then repumping the spin. (**B**) Continuous pumping into a squeezed state, in which $\hat{H}_-$ is turned on simultaneously with optical pumping of the spin. The $\hat{H}_-$ analysis pulse lengths are 80 and 55 μs, respectively. In all figures, the data points are the mean measured spin population based on >300 repetitions of the experimental sequence, with error bars estimated from quantum projection noise.

The onset of the dark state indicates that the desired steady state has been reached. We independently verify this state using two methods. The first is a measurement of the populations of the energy eigenstates (*20*), which we perform by pumping the state to $|\downarrow\rangle$ after the end of the reservoir engineering and applying the blue sideband Hamiltonian. The population of the state $|\downarrow\rangle$ as a function of the blue sideband pulse duration t is given by

$$P(\downarrow) = \frac{1}{2}\sum_n p(n)(1 + e^{-\gamma_n t}\cos(\Omega_{n,n+1}t)) \quad (2)$$

where $p(n)$ is the probability of finding the oscillator in the nth energy eigenstate and $\Omega_{n,n+1}$ is the Rabi frequency for the transition between the $|\downarrow\rangle|n\rangle$ and $|\uparrow\rangle|n+1\rangle$ states, which in the Lamb-Dicke regime scales as $\langle n|\hat{a}|n+1\rangle = \sqrt{n+1}$ (*18*, *20*). The phenomenological decay parameter γ_n accounts for decoherence and fluctuations in the applied laser intensities (*16*, *20*). By fitting a form similar to Eq. 2 to each set of data, we obtain the probability distribution $p(n)$ (*16*). We then fit $p(n)$ using the theoretical form of the probability distributions for coherent, squeezed, and displaced-squeezed states (*21*). The data, deduced populations, and fits are shown in Fig. 3. The fit for the coherent state yields a coherent state parameter $|\alpha| = 2.00 \pm 0.01$ (error bars here and elsewhere are given as SEM). For the squeezed state, we obtain a squeezing amplitude $r = 1.45 \pm 0.03$, which for a pure state would correspond to a 12.6 ± 0.3 dB reduction in the squeezed quadrature variance. For the displaced-squeezed state, we obtain fitted parameters of $r = 0.63 \pm 0.06$, $|\alpha| = 2.2 \pm 0.2$, and $\arg(\alpha) - \phi_s/2 = 0.42 \pm 0.06$ rad. The blue sideband method does not allow us to measure the fidelity of the states, because it does not verify the phase coherence of the superposition. It is also difficult to obtain good population estimates for states with a large spread in their Fock state occupancies (*16*). To overcome these limitations, we introduce a diagnosis method that provides a Rabi frequency decomposition in the transformed state ladder that includes the

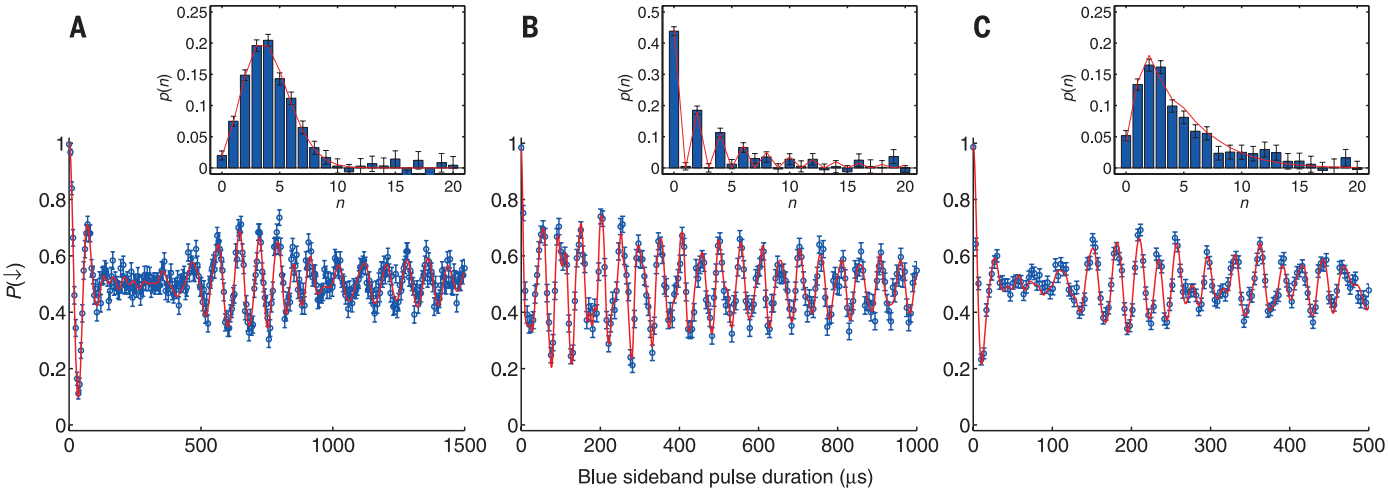


Fig. 3. Fock state analysis using a single-frequency blue-sideband probe. All data are fitted using a form similar to Eq. 2 to retrieve the Fock state populations shown in the inset bar charts. The red curves in the bar charts are fits to the populations using the relevant probability distribution $p(n)$ to determine the size of the states. Data, populations, and fitted distributions are shown for (**A**) the coherent state (with fitted $|\alpha| = 2.00 \pm 0.01$), (**B**) the squeezed vacuum state ($r = 1.45 \pm 0.03$), and (**C**) the displaced-squeezed state ($|\alpha| = 2.2 \pm 0.2$, $r = 0.63 \pm 0.06$ and $\arg(\alpha) - \phi_s/2 = 0.42 \pm 0.06$ rad).

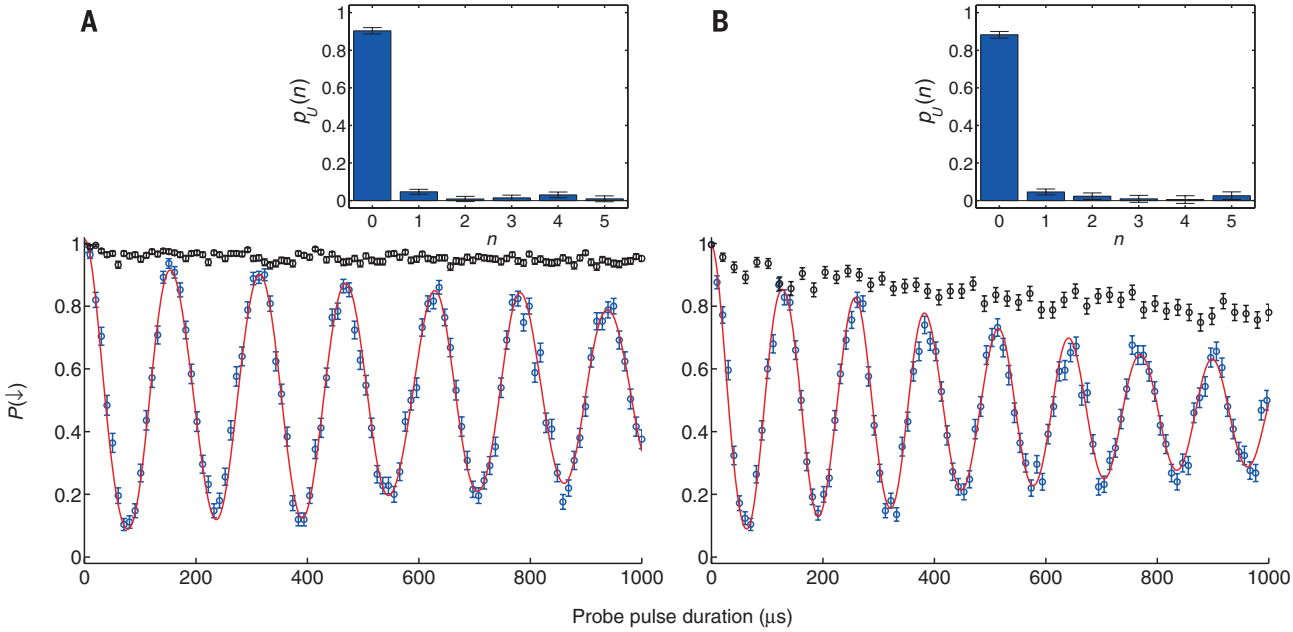


Fig. 4. Fock state analysis in the engineered basis. The data show the spin population evolution after applying $\hat{H}_+$ (blue data points) and $\hat{H}_-$ (black data points) for time t. The fitted curves for $\hat{H}_+$ are of a form similar to Eq. 2. The resulting populations are shown in the inset bar charts. Data and populations are shown for (**A**) the coherent state, with fitted $p_U(0) = 0.90 \pm 0.02$, and (**B**) the squeezed vacuum state, with $p_U(0) = 0.88 \pm 0.02$.

desired state as the ground state. Instead of driving only the blue sideband, we use the Hamiltonian

$$\hat{H}_{+} = \hbar\Omega(\hat{K}^{\dagger}\hat{\sigma}_{+} + \hat{K}\hat{\sigma}_{-}) \quad (3)$$

in which the motional state operators are conjugated with respect to $\hat{H}_{-}$ (Fig. 1C). This results in Rabi oscillations between the states $|\downarrow\rangle|\hat{U}, n\rangle$ and $|\uparrow\rangle|\hat{U}, n+1\rangle$. Because the internal states involved span a two-dimensional Hilbert space, the motional state evolution is also contracted onto two adjacent states of the engineered basis. For an arbitrary initial state, the internal state populations evolve according to Eq. 2, with the corresponding $p(n)$ being the probability of finding the ion in the nth element of the engineered basis before the application of $\hat{H}_{+}$ [we denote this as $p_U(n)$ in the figure to avoid confusion]. Data sets from this type of measurement are shown for the coherent state and for the squeezed state in Fig. 4 for the same settings as used in Figs. 2 and 3. To work in the same basis as the state engineering, we again drive combinations of the carrier and red and blue motional sidebands, but with the ratios of Rabi frequencies calibrated according to $\Omega_c/\Omega_{bsb} = -\alpha^*/\cosh(r)$ and $\Omega_{rsb}/\Omega_{bsb} = e^{-i\phi_s}\tanh(r)$ with ξ and α corresponding to the values used for the reservoir engineering (*16*).

We fit both experimental data sets with a form similar to Eq. 2, obtaining the probability of being found in the ground state of 0.90 ± 0.02 and 0.88 ± 0.02 for the coherent and squeezed states, respectively. We take these to be lower bounds on the fidelity with which these states were prepared, because these numbers include errors in the analysis pulse in addition to state-preparation errors (*16*). The $\hat{H}_{+}$ Rabi oscillations observed in our experiments involve transitions that when viewed in the energy eigenstate basis, couple Hilbert spaces that are of appreciable size. To account for 88% of the populations in oscillations between $|\hat{S}(\xi), 0\rangle$ and $|\hat{S}(\xi), 1\rangle$ for $r = 1.45$, we must include energy eigenstates up to $n = 26$. By our choice of basis, we reduce the relevant dynamics to a two-state system, greatly simplifying the resulting evolution of the spin populations and thus providing a high signal-to-noise ratio. The high fidelity with which the squeezed state is produced is a result of the robust nature of the reservoir engineering, which is insensitive to laser intensity and frequency fluctuations that are common to all frequency components of the engineered Hamiltonian. To generate the same state produced above with standard methods involving unitary evolution starting from the ground state would require simultaneously driving both second motional sidebands (*16*). We would not expect a high fidelity because these have Rabi frequencies comparable to our transition linewidth, which is broadened by magnetic field fluctuations.

This toolbox for generating, protecting, and measuring quantum harmonic oscillator states is transferrable to any physical system in which the relevant couplings can be engineered, facilitating quantum computation with continuous variables (*22*). Examples in which reservoir engineering have been proposed include superconducting circuits and nanomechanics (*12*–*14*). Reservoir engineering provides access to controlled dissipation, which can be used in quantum simulations of open quantum systems (*13*, *23*).

REFERENCES AND NOTES

1. J. I. Cirac, A. S. Parkins, R. Blatt, P. Zoller, *Phys. Rev. Lett.* **70**, 556–559 (1993).
2. J. F. Poyatos, J. I. Cirac, P. Zoller, *Phys. Rev. Lett.* **77**, 4728–4731 (1996).
3. A. R. R. Carvalho, P. Milman, R. L. de Matos Filho, L. Davidovich, *Phys. Rev. Lett.* **86**, 4988–4991 (2001).
4. F. Verstraete, M. M. Wolf, J. I. Cirac, *Nat. Phys.* **5**, 633–636 (2009).
5. S. Diehl *et al.*, *Nat. Phys.* **4**, 878–883 (2008).
6. B. Kraus *et al.*, *Phys. Rev. A* **78**, 042307 (2008).
7. F. Pastawski, L. Clemente, J. I. Cirac, *Phys. Rev. A* **83**, 012304 (2011).
8. H. Krauter *et al.*, *Phys. Rev. Lett.* **107**, 080503 (2011).
9. J. T. Barreiro *et al.*, *Nature* **470**, 486–491 (2011).
10. Y. Lin *et al.*, *Nature* **504**, 415–418 (2013).
11. S. Shankar *et al.*, *Nature* **504**, 419–422 (2013).
12. A. Sarlette, J. M. Raimond, M. Brune, P. Rouchon, *Phys. Rev. Lett.* **107**, 010402 (2011).
13. C. Navarrete-Benlloch, J. J. García-Ripoll, D. Porras, *Phys. Rev. Lett.* **113**, 193601 (2014).
14. A. Kronwald, F. Marquardt, A. A. Clerk, *Phys. Rev. A* **88**, 063833 (2013).
15. W. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH, Berlin, 2001).
16. Materials and methods are available as supplementary materials on *Science* Online.
17. C. Roos, Controlling the quantum state of trapped ions, Ph.D. thesis, Universität Innsbruck, Austria (2000).
18. D. J. Wineland *et al.*, *J. Res. Natl. Inst. Stand. Technol.* **103**, 259 (1998).
19. C. F. Roos, *New J. Phys.* **10**, 013002 (2008).
20. D. M. Meekhof, C. Monroe, B. E. King, W. M. Itano, D. J. Wineland, *Phys. Rev. Lett.* **76**, 1796–1799 (1996).
21. H. P. Yuen, *Phys. Rev. A* **13**, 2226–2243 (1976).
22. D. Gottesman, A. Kitaev, J. Preskill, *Phys. Rev. A* **64**, 012310 (2001).
23. C. J. Myatt *et al.*, *Nature* **403**, 269–273 (2000).

ACKNOWLEDGMENTS

We thank J. Alonso, A. Imamoglu, and D. Wineland for comments on the manuscript and useful discussions. We thank J. Alonso, M. Sepiol, K. Fisher, and C. Flühmann for contributions to the experimental apparatus. We acknowledge support from the Swiss National Science Foundation under grant 200021_134776 and through the National Centre of Competence in Research for Quantum Science and Technology (QSIT).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/53/suppl/DC1
Supplementary Text
Tables S1 to S3
References (*24*–*28*)

9 September 2014; accepted 24 November 2014
10.1126/science.1261033

ORGANIC CHEMISTRY

Rh-catalyzed C–C bond cleavage by transfer hydroformylation

Stephen K. Murphy,[1,2] Jung-Woo Park,[1] Faben A. Cruz,[1] Vy M. Dong[1]*

The dehydroformylation of aldehydes to generate olefins occurs during the biosynthesis of various sterols, including cholesterol in humans. Here, we implement a synthetic version that features the transfer of a formyl group and hydride from an aldehyde substrate to a strained olefin acceptor. A Rhodium (Xantphos)(benzoate) catalyst activates aldehyde carbon-hydrogen (C–H) bonds with high chemoselectivity to trigger carbon-carbon (C–C) bond cleavage and generate olefins at low loadings (0.3 to 2 mole percent) and temperatures (22° to 80°C). This mild protocol can be applied to various natural products and was used to achieve a three-step synthesis of (+)-yohimbenone. A study of the mechanism reveals that the benzoate counterion acts as a proton shuttle to enable transfer hydroformylation.

The cytochrome P450 enzymes have captured the imagination of chemists who seek to emulate their reactivity. For example, monooxygenases motivated the design of catalysts that epoxidize olefins and oxidize C–H bonds (*1*–*4*). This enzyme superfamily also includes various demethylases that break C–C bonds (*5*). In particular, lanosterol demethylase converts aldehydes to olefins by dehydroformylation during the biosynthesis of sterols in bacteria, algae, fungi, plants, and animals (*6*) (Fig. 1A). Inspired by this step in biosynthesis, we sought a transition-metal catalyst for dehydroformylations in organic synthesis.

[1]Department of Chemistry, University of California Irvine, CA 92697-2025, USA. [2]Department of Chemistry, University of Toronto, Ontario M5S 3H6, Canada.
*Corresponding author. E-mail: dongv@uci.edu

To this end, we aimed to trigger C–C bond cleavage (*7*–*11*) by chemoselective activation of aldehyde C–H bonds using Rh-catalysis (Fig. 1B). Over the past 50 years, activating aldehyde C–H bonds with Rh has been thoroughly investigated (*12*); however, the resulting acyl-Rh^{III}-hydrides have been trapped mainly by hydroacylation (*13*) or decarbonylation (*14*, *15*). This common intermediate is also implicated in hydroformylation, which is practiced on an industrial scale using synthesis gas (*16*). Thus, we needed a strategy for diverting the acyl-Rh^{III}-hydride toward dehydroformylation. To date, olefins generated by dehydroformylation have been observed in low quantities during decarbonylations (*15*, *17*, *18*). One report describes the use of stoichiometric Ru for dehydroformylation of butyraldehyde (*19*), and another uses heterogeneous Rh or Pd catalysts for transforming steroidal aldehydes

at 160 to 300°C (*20*). In contrast, an Fe-peroxo complex cleaves aldehyde C–C bonds at room temperature, but this complex must be used in stoichiometric amounts and can lead to olefin epoxidation (*21*, *22*).

Given this challenge, we designed a strategy in which dehydroformylation of an aldehyde substrate is driven by the concomitant hydroformylation of a strained olefin acceptor (Fig. 1C) (*23*, *24*). This transfer hydroformylation avoids the accumulation of CO gas, which acts as a catalyst poison in related aldehyde dehomologations. Thus, formyl group transfer should proceed under mild conditions. Brookhart's study on the linear-to-branched isomerization of aldehydes with Rh catalysis supports the feasibility of this approach (*25*). Moreover, Morimoto developed hydroformylations of monosubstituted olefins using formaldehyde as a source of CO and H_2 (*26*). Here, we report a Rh catalyst for transfer hydroformylation that operates in the 22° to 80°C temperature range, with loadings as low as 0.3 mole percent (mol %). This mild protocol for dehydroformylation can be applied to a wide range of aldehydes, including those derived from alkaloid, terpene, steroid, and macrolide natural products.

During initial studies, we obtained promising results by investigating nontraditional counterions for Rh(Xantphos) complexes (Fig. 2). The Xantphos ligand was chosen given its success in related hydroacylations, hydroformylations, and decarbonylations (*13*, *16*). Using citronellal (**1a**) and norbornadiene (**5a**) as the model substrate and acceptor, respectively, we observed that typical counterions such as BF_4^- and Cl^- yielded trace decarbonylation products, whereas a softer counterion, I^-, led to mixed dehydroformylation and decarbonylation reactivity. An increase in reactivity and selectivity was obtained by switching to organic counterions such as phenolates and sulfonamidates. The use of a benzoate counterion provided a breakthrough in efficiency. Against expectations, further tuning of the counterion revealed few trends related to acidity, Hammett parameters, or coordinating ability. This observation suggests that the counterion plays a critical role in the mechanism. 3-Methoxybenzoate provided a fivefold increase in initial rate compared with benzoate. We also identified 5-norbornene-2-carboxaldehyde (**6a**) as a stoichiometric product in each of these reactions indicating that a transfer hydroformylation mechanism operates.

The choice of olefin acceptor influences both catalyst loading and reaction temperature (Fig. 2). Because norbornadiene (**5a**) gave selectivity greater than 99:1 **2a**:**3a**, the catalyst loading could be lowered to 0.3 mol % at 80°C or 1 mol % at 60°C using this acceptor. The reaction temperature could be further reduced by using olefin acceptors that cannot chelate to Rh. For instance, norbornene (**5b**) displayed excellent reactivity at 40°C, whereas a slightly more strained acceptor, benzonorbornadiene (**5c**), provided reactivity at ambient temperature. To examine the scope of this strategy, we chose norbornadiene (**5a**) as the acceptor because it afforded the highest chemoselectivity with the lowest catalyst loadings.

This transfer hydroformylation protocol enables access to olefins from a wide range of aldehyde precursors (Fig. 3, A and B). The Diels-Alder cycloaddition was used to generate cyclohexene-4-carboxaldehdye substrates **1b** through **1d**. The trans adduct **1b** underwent dehydroformylation to yield the conjugated 1,3-diene, whereas **1c** gave a mixture of 1,3- and 1,4-dienes. The cis Diels-Alder adduct **1d** yielded the 1,3-diene exclusively, most likely as a result of a syn-selective β-hydride elimination. We reason that the observed regioselectivities are controlled by kinetics because 4-phenylbutanal (**1e**) yields the terminal olefin (**2e**) without any isomerization to the styrene derivative. In general, Lewis basic functionality, such as ethers, esters, amines, phthalimides, and indoles, were tolerated (**1f** to **1i** and **1l**). A vinylindole was derived by dehydroformylation of **1g**, which was ultimately prepared from commercial indole and acrolein. Although 4-pentenals are prototypical substrates for intramolecular olefin hydroacylation, the α-allylated aldehyde **1h** underwent chemoselective dehydroformylation to yield the conjugated diene. Disubstituted olefins enriched in the *E* stereoisomer (>20:1 *E*/*Z*) were accessed from the corresponding α-arylated aldehydes (**1i**). Substrates that do not form conjugated products

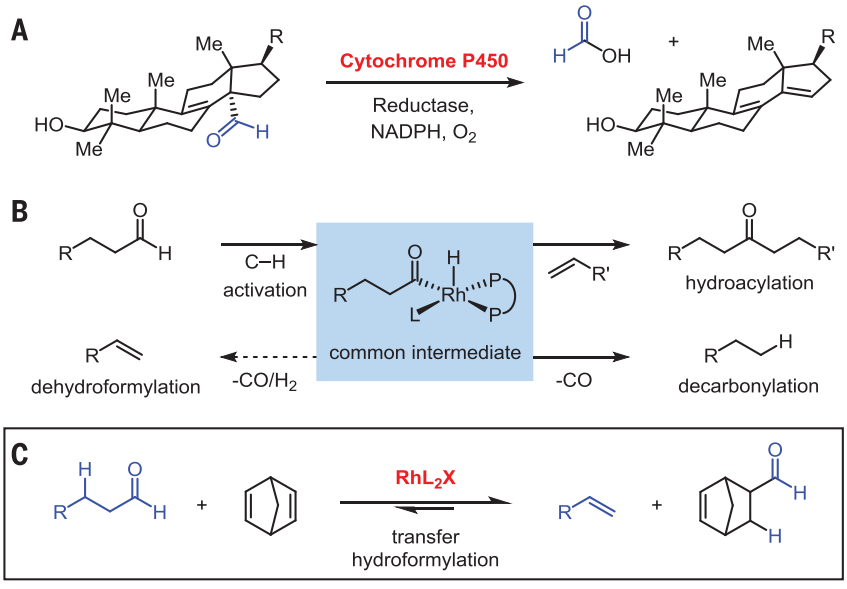


Fig. 1. Dehydroformylation in nature and organic synthesis. (**A**) Dehydroformylation during sterol biosynthesis. (**B**) Reactivity of acyl-Rh^{III}-hydrides. (**C**) Proposed transfer hydroformylation.

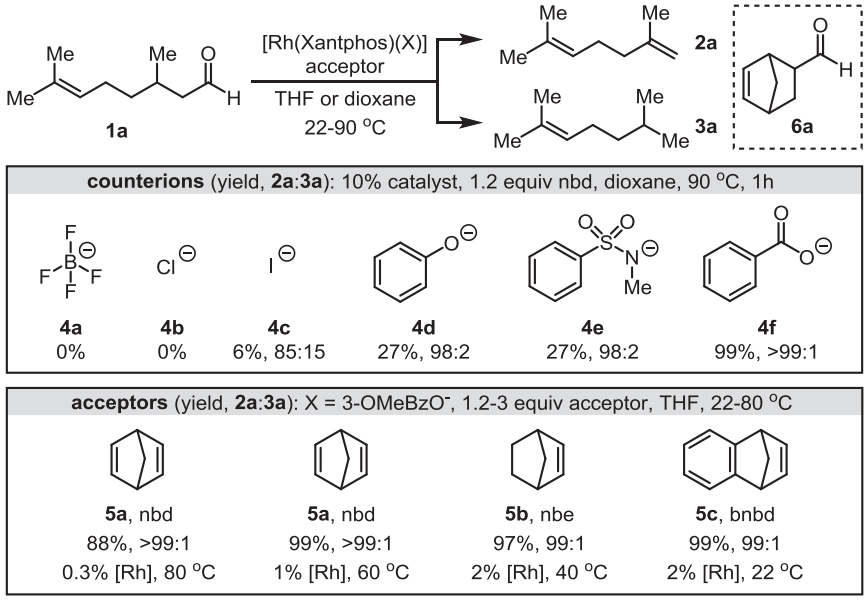


Fig. 2. Effects of counterion structure and ring strain. Yields were determined by gas chromatographic analysis of the reaction mixtures using durene as an internal standard. nbd, norbornadiene; nbe, norbornene; bnbd, benzonorbornadiene.

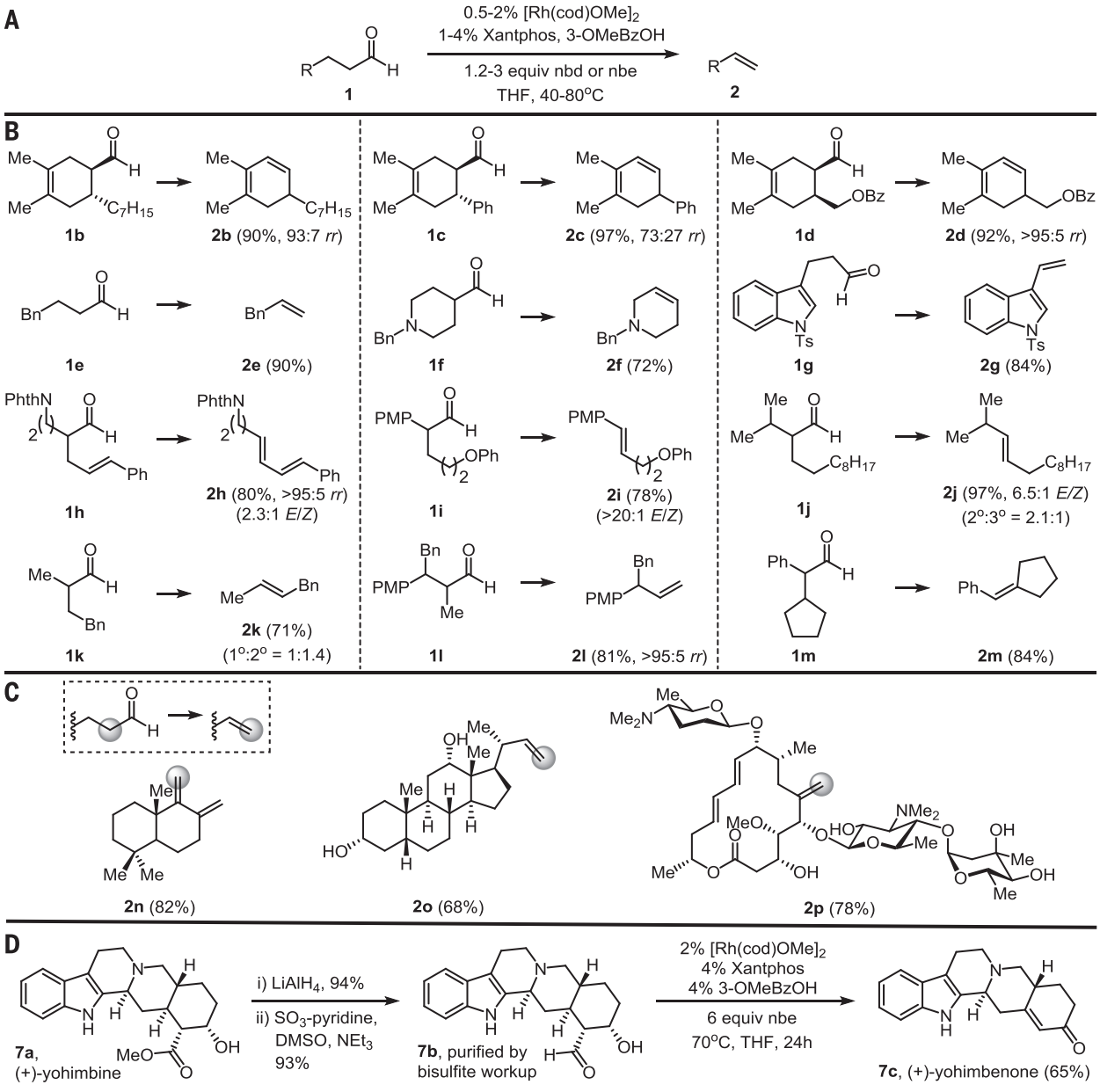


Fig. 3. Applications of dehydroformylation. (**A**) General conditions for transfer hydroformylation. (**B**) Substrate scope. (**C**) Natural product derivatization. (**D**) Three-step synthesis of (+)-yohimbenone. Yields are of isolated materials and mixtures of regioisomers where indicated. *rr* is the regioisomeric ratio; *rr* values were determined by ^{1}H NMR analysis of the reaction mixtures. The yields of **2e** and **2k** were determined by ^{1}H NMR analysis of the reaction mixtures using durene as an internal standard. See the supplementary materials for details.

upon dehydroformylation were transformed with modest regioselectivities (**1j** and **1k**); however, steric congestion favored terminal olefins over trisubstituted products (**1l**). Nonetheless, trisubstituted olefins were generated from substrates containing a single syn-β-hydrogen such as **1m**.

Next, we applied this protocol to generate structurally complex olefins from natural products (Fig. 3C). By dehydroformylation of a (+)-sclareolide derivative, we accessed a carbon-based scaffold **2n** containing an exocyclic diene adjacent to a quaternary center. This product is a key intermediate in the synthesis of several terpenes. Furthermore, (+)-sclareolide is an inexpensive and readily available precursor, whereas typical precursors such as (+)-manool and (−)-polygodial have either been discontinued by commercial suppliers or are available only in milligram quantities (*27*).

To study the chemoselectivity of dehydroformylation, we examined steroid and macrolide substrates (Fig. 3C). Deoxycholic acid derivative **2o** was prepared without protection of the hydroxyl groups, despite the potential for alcohol oxidation under Rh catalysis (*28*, *29*). Thus, activation of the aldehyde C–H bond occurred with high chemoselectivity to initiate C–C bond cleavage. Smooth dehydroformylation of the antibiotic spiramycin I to generate macrolide **2p** highlights the tolerance of this method to many functional groups, including dienes, amines, ethers, esters, and acetals. In this case, dehydroformylation introduced an exocyclic olefin that dramatically altered the topology of the macrolide.

The yohimbinoid family of indole alkaloids has often served as a testing ground for methodology (*30*). Padwa reported the de novo synthesis of racemic yohimbenone in 11 steps from methyl 3-indolylacetate (*31*). By using dehydroformylation as a key step, we prepared (+)-yohimbenone in three steps from commercially available and inexpensive (+)-yohimbine. Conversion of ester **7a** to β-hydroxy aldehyde **7b** was achieved in 87% yield by $LiAlH_4$ reduction followed by

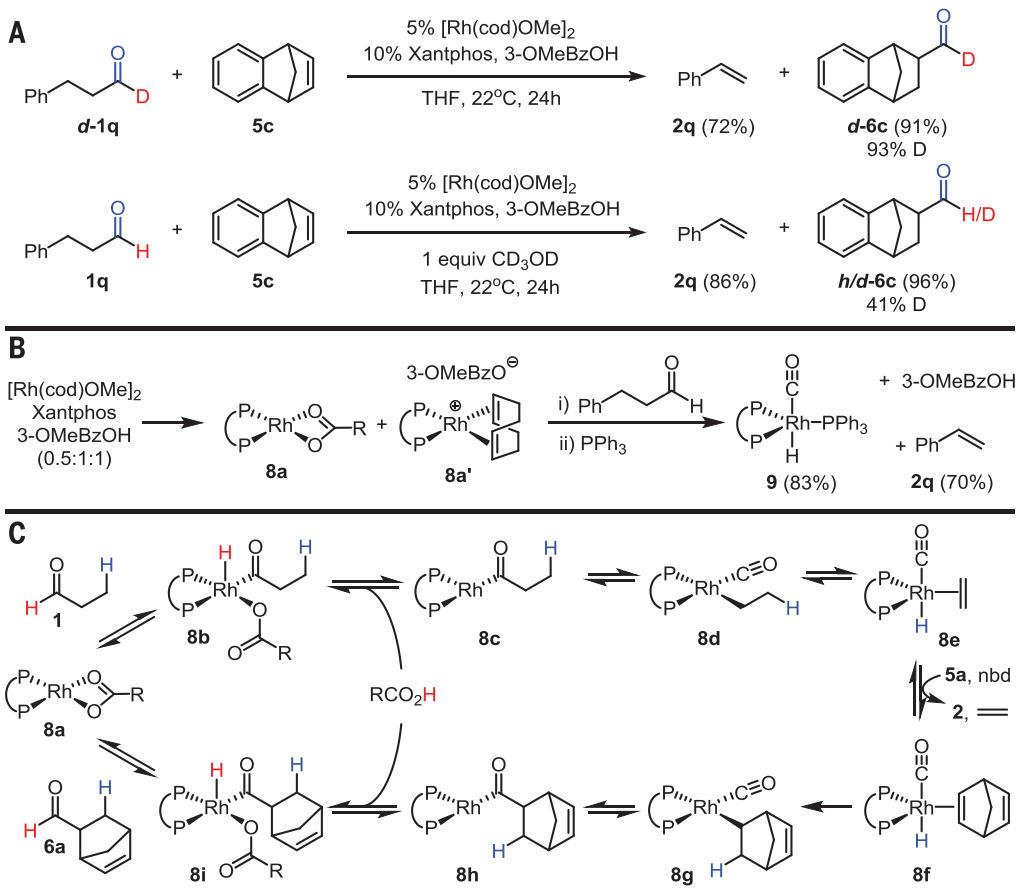


Fig. 4. Mechanistic studies. (**A**) Deuterium labeling studies. (**B**) Isolation of organometallic intermediates. (**C**) Proposed catalytic cycle.

Parikh-Doering oxidation, and the resulting aldehyde was purified by a simple workup with sodium bisulfite. This aldehyde contains both a syn- and an anti-β-hydrogen. Syn-selective dehydroformylation established the trisubstituted olefin at the ring junction. To our surprise, however, the resulting allylic alcohol underwent transfer dehydrogenation in the same pot to yield (+)-yohimbenone in 65% yield. Because dehydroformylation is faster than the allylic alcohol oxidation, either the allylic alcohol or enone product could be selectively formed by controlling the reaction temperature and stoichiometry of the strained olefin acceptor (*32*).

Through experiments designed to probe the mechanism, we obtained insight into why the counterion and strained acceptor are critical in diverting the acyl-Rh^{III}-hydride intermediate along the dehydroformylation pathway. Isotopic labeling studies revealed that the deuterium label of aldehyde ***d*-1q** was incorporated into the formyl group of the product ***d*-6c**. However, statistical scrambling occurred when protio-**1q** was subjected to transfer hydroformylation in the presence of deuterated methanol (Fig. 4A). Together, these results suggest that the aldehyde proton is transferred to the product through the intermediacy of 3-methoxybenzoic acid, which can undergo proton exchange with methanol. Experiments using stoichiometric Rh support this mechanistic scenario (Fig. 4B). Combining the Rh-source, 3-methoxybenzoic acid, and phosphine ligand resulted in an equilibrium mixture of Rh complexes **8a** and **8a′**, each with 3-methoxybenzoate counterions. Upon treatment of this mixture with hydrocinnamaldehyde (**1q**), we observed styrene (**2q**) in high yields along with the regeneration of the benzoic acid derivative (*33*). Subsequent addition of PPh_3 enabled us to identify the organometallic product, Rh-hydrido-carbonyl **9**, which is a catalyst for traditional hydroformylations (*34*). Although stoichiometric dehydroformylation takes place in the absence of the strained acceptor, our studies on the catalytic process revealed a correlation between the ring strain of the acceptor and the selectivity for dehydroformylation versus decarbonylation. Therefore, we propose that stoichiometric dehydroformylation in the absence of acceptor is thermodynamically downhill and reversible, but norbornadiene can irreversibly trap the Rh-hydrido-carbonyl intermediate to prevent decarbonylation and turn over the catalyst.

A proposed catalytic cycle for transfer hydroformylation is depicted in Fig. 4C. The neutral Rh complex **8a** activates the aldehyde C–H bond to generate acyl-Rh^{III}-hydride **8b**. The 3-methoxybenzoate counterion can then undergo reductive elimination with the hydride ligand to generate acyl-Rh^{I} **8c** and 3-methoxybenzoic acid (*35*). In contrast, most hydroacylations and decarbonylations typically employ innocent counterions such as Cl^- and BF_4^-. De-insertion of CO and subsequent β-hydride elimination forges Rh-hydrido-carbonyl **8e**. Exchange of the olefin product with norbornadiene (**5a**) generates **8f**, which irreversibly leads to the transfer hydroformylation product **6a** through similar mechanistic steps in reverse order (Fig. 4C). Thus, the ring strain of the olefin acceptor and the ability of the counterion to act as a proton shuttle by reversible redox processes afford high reactivity and selectivity.

REFERENCES AND NOTES

1. J. T. Groves, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3569–3574 (2003).
2. R. Breslow, Y. Huang, X. Zhang, J. Yang, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11156–11158 (1997).
3. M. S. Chen, M. C. White, *Science* **327**, 566–571 (2010).
4. W. Zhang, J. L. Loebach, S. R. Wilson, E. N. Jacobsen, *J. Am. Chem. Soc.* **112**, 2801–2803 (1990).
5. A. Sigel, H. Sigel, R. K. O. Sigel, Eds., *The Ubiquitous Roles of Cytochrome P450 Proteins: Metal Ions in Life Sciences*, vol. 3 (Wiley, New York, 2007).
6. G. I. Lepesheva, M. R. Waterman, *Biochim. Biophys. Acta* **1770**, 467–477 (2007).
7. A. Dermenci, J. W. Coe, G. Dong, *Org. Chem. Front.* **1**, 567–581 (2014).
8. C.-H. Jun, *Chem. Soc. Rev.* **33**, 610–618 (2004).
9. H. M. Ko, G. Dong, *Nat. Chem.* **6**, 739–744 (2014).
10. Z. Zuo *et al.*, *Science* **345**, 437–440 (2014).
11. Y. Liu *et al.*, *Adv. Synth. Catal.* **356**, 130–136 (2014).
12. M. A. Garralda, *Dalton Trans.* **2009**, 3635–3645 (2009).
13. M. C. Willis, *Chem. Rev.* **110**, 725–748 (2010).
14. J. Tsuji, K. Ohno, *Tetrahedron Lett.* **6**, 3969–3971 (1965).
15. M. Kreis, A. Palmelund, L. Bunch, R. Madsen, *Adv. Synth. Catal.* **348**, 2148–2154 (2006).

16. R. Franke, D. Selent, A. Börner, *Chem. Rev.* **112**, 5675–5732 (2012).
17. T. Iwai, T. Fujihara, Y. Tsuji, *Chem. Commun.* **2008**, 6215–6217 (2008).
18. T. Kondo, M. Akazome, Y. Tsuji, Y. Watanabe, *J. Org. Chem.* **55**, 1286–1291 (1990).
19. R. H. Prince, K. A. Raspin, *Chem. Commun.* (*London*) **1966**, 156–157 (1966).
20. C. A. McCombs, C. H. Foster, Eastman Kodak Company, Dehydroformylation of steroidal aldehydes. U.S. Patent 4272444 A, 14 August 1980.
21. D. L. Wertz, M. F. Sisemore, M. Selke, J. Driscoll, J. S. Valentine, *J. Am. Chem. Soc.* **120**, 5331–5332 (1998).
22. Y. Goto, S. Wada, I. Morishima, Y. Watanabe, *J. Inorg. Biochem.* **69**, 241–247 (1998).
23. R. Ahuja *et al.*, *Nat. Chem.* **3**, 167–171 (2011).
24. D. H. T. Phan, K. G. M. Kou, V. M. Dong, *J. Am. Chem. Soc.* **132**, 16354–16355 (2010).
25. C. P. Lenges, M. Brookhart, *Angew. Chem. Int. Ed.* **38**, 3533–3537 (1999).
26. G. Makado *et al.*, *Adv. Synth. Catal.* **352**, 299–304 (2010).
27. M. Cortés, J. A. Valderrama, M. Cuellar, V. Armstrong, M. Preite, *J. Nat. Prod.* **64**, 348–349 (2001).
28. H. Imai, T. Nishiguchi, K. Fukuzumi, *J. Org. Chem.* **39**, 1622–1627 (1974).
29. C.-H. Jun, C.-W. Huh, S.-J. Na, *Angew. Chem. Int. Ed.* **37**, 145–147 (1998).
30. T. P. Lebold *et al.*, *Nat. Chem.* **5**, 126–131 (2013).
31. C. J. Stearman, M. Wilson, A. Padwa, *J. Org. Chem.* **74**, 3491–3499 (2009).
32. For detailed experimental conditions, see the supplementary materials on *Science* Online.
33. The peaks for the regenerated benzoic acid are very broad in the nuclear magnetic resonance (NMR) spectra, possibly due to reversible reductive elimination. Analogous results were obtained using 4-fluorobenzoic acid as a surrogate, which enabled us to follow the reaction by ^{19}F NMR and confirm the results.
34. M. Kranenburg *et al.*, *Organometallics* **14**, 3081–3089 (1995).
35. M. Manger, J. Wolf, M. Teichert, D. Stalke, H. Werner, *Organometallics* **17**, 3210–3221 (1998).

ACKNOWLEDGMENTS

The authors are grateful for financial support provided by the NIH's National Institute of General Medical Sciences (GM105938). S.K.M. is grateful for a Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada. We acknowledge C. M. Rathbun for contributions during the optimization stage of reaction development.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/56/suppl/DC1
Materials and Methods
Supplementary Text
References (*36*–*54*)

15 September 2014; accepted 6 November 2014
10.1126/science.1261232

REACTION DYNAMICS

Extremely short-lived reaction resonances in Cl + HD (v = 1) → DCl + H due to chemical bond softening

Tiangang Yang,[1,2]* Jun Chen,[1]* Long Huang,[1] Tao Wang,[1] Chunlei Xiao,[1]† Zhigang Sun,[1,3]† Dongxu Dai,[1] Xueming Yang,[1,3]† Dong H. Zhang[1,3]†

The Cl + H_2 reaction is an important benchmark system in the study of chemical reaction dynamics that has always appeared to proceed via a direct abstraction mechanism, with no clear signature of reaction resonances. Here we report a high-resolution crossed–molecular beam study on the Cl + HD (v = 1, j = 0) → DCl + H reaction (where v is the vibrational quantum number and j is the rotational quantum number). Very few forward scattered products were observed. However, two distinctive peaks at collision energies of 2.4 and 4.3 kilocalories per mole for the DCl (v' = 1) product were detected in the backward scattering direction. Detailed quantum dynamics calculations on a highly accurate potential energy surface suggested that these features originate from two very short-lived dynamical resonances trapped in the peculiar H-DCl (v' = 2) vibrational adiabatic potential wells that result from chemical bond softening. We anticipate that dynamical resonances trapped in such wells exist in many reactions involving vibrationally excited molecules.

Reaction resonances are quasi-trapped quantum states in the transition state region that profoundly influence both the rate and product distribution of a chemical reaction (*1*–*3*). Since the landmark theoretical predictions of reaction resonances in the H/F + H_2 reaction in the early 1970s (*4*, *5*), extensive studies have been carried out to detect the resonances experimentally and to elucidate them theoretically. However, direct observations have proven to be extremely challenging. Through a series of crossed–molecular beam experiments (*6*–*9*), a physical picture of reaction resonances in F + H_2 (HD) beyond chemical accuracy has been established. In addition, threshold photodetachment spectroscopy has been used to probe resonances in the I + HI reaction (*10*). Recently, resonance signatures have also been detected in polyatomic reactions (*11*–*14*). Forward scattering of reaction products in crossed-beam scattering experiments can be caused by long-lived resonances. However, the presence of forward scattering does not necessarily imply that there are resonances in a chemical reaction. An intriguing question then is if and how we can probe reaction resonances in systems that show no or little forward scattering product, in which the reaction intermediate is very short lived.

Here we report a combined high-resolution crossed-beam and accurate quantum reaction dynamics study on the Cl + HD (v = 1, j = 0) → DCl + H reaction (v, vibrational quantum number; j, rotational quantum number). Our study provides very strong evidence for the existence of short-lived quantum dynamical resonances in this reaction. The Cl + H_2 system has served as one of the most important benchmark systems in the study of chemical reaction dynamics (*15*), along with the H + H_2 and F + H_2 reactions. It has also played a special role in development of the transition state theory and in the verification of kinetic isotope effects (*16*–*19*). In contrast to the F + H_2 reaction, the Cl + H_2 (v = 0) reaction was shown to be a direct abstraction with a colinear later reaction barrier (*20*–*25*).

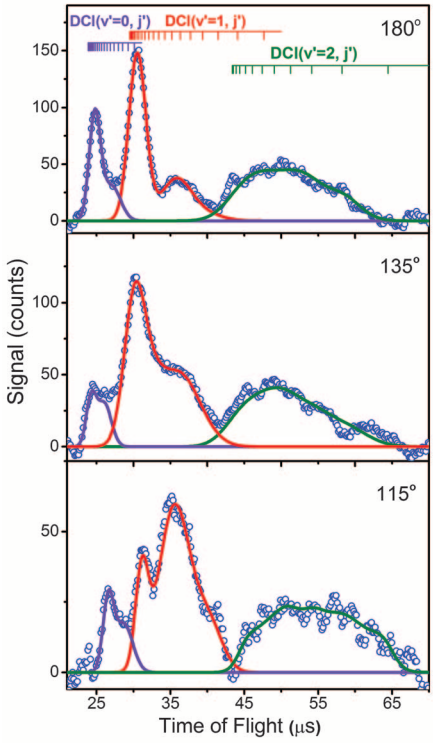


Fig. 1. Time-of-flight spectra of the H atom product from the Cl + HD (v = 1, j = 0) → DCl (v') + H reaction at the collision energy of 4.3 kcal/mol at different laboratory angles.

[1]State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, Liaoning 116023, China. [2]School of Physics and Optoelectric Engineering, Dalian University of Technology, Dalian, Liaoning 116023, China. [3]Center for Advanced Chemical Physics and 2011 Frontier Center for Quantum Science and Technology, University of Science and Technology of China, 96 Jinzhai Road, Hefei 230026, China.
*These authors contributed equally to this work. **†Corresponding author. E-mail: chunleixiao@dicp.ac.cn (C.X.); zsun@dicp.ac.cn (Z.S.); xmyang@dicp.ac.cn (X.Y.); zhangdh@dicp.ac.cn (D.H.Z.)**

No resonance features have been detected for this ground-state reaction. Kandel *et al.* investigated the Cl + HD (v = 1, j = 1,2) reaction at averaged collision energy of 1.5 kcal/mol (*26*). The observed angular distribution of the HCl product showed predominantly backward and sideways scattering, offering no clear evidence of resonances. Skouteris *et al.* observed a strong effect of the van der Waals interaction on the branching between the two product channels—HCl + D and DCl + H—in the Cl + HD (v = 0) reaction (*21*). In addition, the Cl/Cl* + H_2 (Cl/Cl*, chlorine atom in the ground/excited spin-orbit state) reaction has also served as a benchmark system for spin-orbit nonadiabatic dynamics (*27*–*29*). These extensive studies have firmly established that the Cl reaction with HD in both ground and vibrational excited states is a direct abstraction reaction, without any clear signature of reaction resonances.

In the present experiment, the Cl atom beam was generated using a home-designed double-stage–discharge molecular beam source (*30*), whereas the vibrationally excited HD (v = 1) beam was prepared using the stimulated Raman pumping (SRP) method. Because HD is not infrared-active, SRP is the only good approach to prepare a substantial amount of HD in the v = 1 state for a crossed-beam study. A high-power seeded yttrium-aluminum-garnet (YAG) laser was used as the pump beam, and a home-built high-power tunable optical parametric oscillator (OPO) as the Raman Stokes beam (*9*, *31*–*33*). Because of the frequency drift of the high-power YAG laser, a difference frequency ($\omega_{YAG} - \omega_{OPO}$) locking scheme was developed to ensure that the Raman pumping process was locked to the HD (v = 0, j = 0)–to–(v = 1, j = 0) transition during the experiment. In the crossed-beams region, ~13% of the HD molecules were pumped to the (v = 1, j = 0) level. In the experiment, the H atom products were sensitively detected using the H atom Rydberg tagging technique (*34*), and the collision energy was varied by changing the crossing angle between the Cl and HD beams. A detailed description of the experimental apparatus used in this study can be found in (*35*).

The time-of-flight (TOF) spectra of the H atom from the Cl + HD (v = 0,1) → DCl + H reaction, with the Stokes laser on and off, were measured at laboratory angles scanned in 10° intervals at collision energies of 2.4 and 4.3 kcal/mol. The H atom TOF spectra for the Cl + HD (v = 1, j = 0) reaction were obtained by subtracting the Stokes laser-off TOF spectra from the Stokes laser-on spectra (fig. S1). Figure 1 shows a few typical TOF spectra for the Cl + HD (v = 1, j = 0) → DCl(v') + H reaction at a collision energy of 4.3 kcal/mol, which exhibit clear structures corresponding to DCl products in v' = 0,1,2 vibrational states. For v' = 0,1 products (in particular, for v' = 1), bimodal features appear, indicating bimodal distribution of the DCl rotational states. In contrast, no bimodal feature was observed in the TOF spectra at collision energy E_C = 2.4 kcal/mol. The obtained TOF spectra were then converted to the center-of-mass frame to obtain the product kinetic energy distributions, from which full rovibrational state–resolved differential cross sections (DCSs) were constructed at collision energies of 2.4 and 4.3 kcal/mol (Fig. 2, A and B). The DCl products at both collision energies are predominantly backward scattered with respect to the Cl atom beam direction, but with a sideways component and even a small forward component at E_C = 4.3 kcal/mol. Overall, the reaction of Cl + HD (v = 1) seems to proceed mainly via a direct abstraction mechanism, as the Cl + HD (v = 0) reaction does. However, the observation of a small forward scattering component in the DCS at E_C = 4.3 kcal/mol merited further investigation of its origin.

Collision-energy dependence of DCSs in the backward scattering direction can provide clues to possible dynamical resonances, as previously demonstrated in the studies of the F + HD system (*9*). This is because the backward scattering mainly arises from a small number of low–angular momentum partial waves and largely retains the oscillatory structures of these partial waves as a function of collision energy that may be caused by reaction resonances. Here we designate this collision-energy–dependent DCS in the backward scattering direction as the backward scattering spectrum (BSS). To probe possible reactive resonances in this system, we have carried out a careful measurement of the BSS for the Cl + HD (v = 1) reaction, with the Stokes laser on and off. In this measurement, the microchannel plate detector was always fixed at the laboratory angle corresponding to the center-of-mass backward scattering direction at different collision energies (see table S1), and the measurement was repeated 10 times at each collision energy. The experimental DCS for DCl (v' = 1) shown in Fig. 3A was the averaged result of the 10 repeated measurements. The corresponding error bar (±1 SD of uncertainty) was determined from the scattered data points of the 10 repeated measurements (see more detailed description of error analysis in the supplementary materials). The obtained BSS for the DCl (v' = 1) product shows two clear peaks at E_C = 2.4 (peak "a") and 4.3 kcal/mol (peak "b") (Fig. 3A).

To interpret the experimental observations, we constructed a new potential energy surface

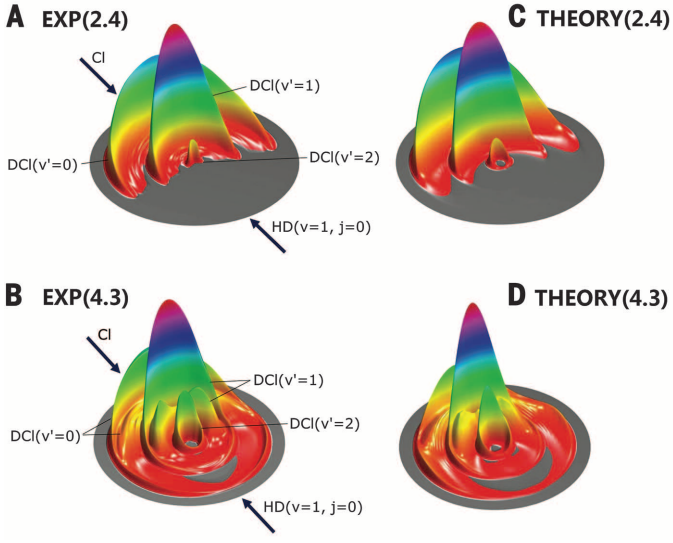


Fig. 2. Three-dimensional scattering product contour plots as a function of product velocity for the Cl + HD (v = 1) → DCl + H reaction. Product contour plots versus product velocity are shown in the center-of-mass frame for (**A** and **B**) experimental measurements and (**C** and **D**) theoretical calculations at collision energies of 2.4 kcal/mol (A and C) and 4.3 kcal/mol (B and D).

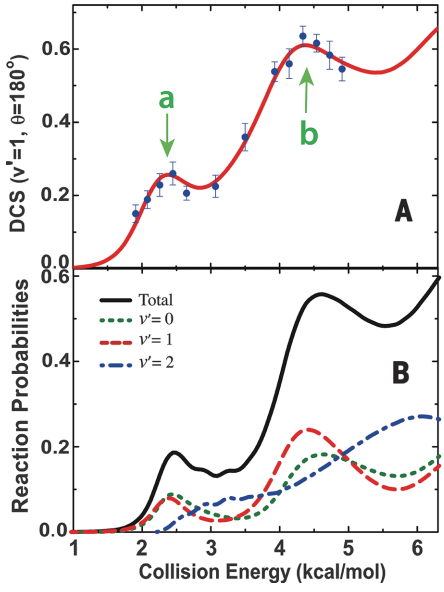


Fig. 3. Collision energy–dependent differential cross sections in the backward scattering direction and the J = 0 reaction probability as a function of collision energy. (**A**) Theoretical (solid red lines) and experimental (solid circles) DCS for the backward scattering DCl product of Cl + HD (v = 1, j = 0) over a range of collision energies. The theoretical DCS was shifted by 0.15 kcal/mol lower in energy for the comparison. The experimental DCS for DCl(v' = 1) was obtained by averaging the results of the 10 repeated measurements for each collision energy, and the corresponding error bar is ±1 SD of uncertainty (see detailed description in the supplementary materials). (**B**) Total and product state-resolved reaction probabilities with J = 0 as a function of collision energy. The resonance peaks, which correspond to peaks a and b in (A), are clearly evident.

(PES) for the Cl + HD reaction based on ~13,000 ab initio energies evaluated at the UCCSD(T)/avqz level of theory (*36*). DCSs were then calculated for the Cl + HD (v = 1, j = 0) on the PES for collision energies up to 8.0 kcal/mol using the time-dependent wave-packet method (*37*). The computed three-dimensional (3D) DCSs at E = 2.4 and 4.3 kcal/mol (Fig. 2, C and D) reproduce the corresponding experimental results shown in Fig. 2, A and B, indicating that the new PES is highly accurate in describing the dynamics of the reaction at this high total energy. We have also computed the BSS for the DCl (v' = 1) product and compared with the experimental results (Fig. 3A). Excellent agreement between theory and experiment is achieved by shifting the theoretical results lower in collision energy by 0.15 kcal/mol, most notably with the two observed peaks reproduced almost exactly. Because the estimated uncertainty of the collision energy is ~0.1 kcal/mol in this experiment, the small discrepancy of 0.15 kcal/mol in the collision energy between theory and experiment indicates that the PES is sufficiently accurate to elucidate the physical origin of the observed oscillatory structures.

The calculated total reaction probabilities for the total angular momentum J = 0 (Fig. 3B) and $J \leq 16$ (fig. S4) for the reaction exhibit clear peak structures, very similar to those in the reactions of F + H_2/HD (*7*, *8*). The peaks at higher-J partial waves clearly originate from the corresponding J = 0 peaks by progressive J shifting (figs. S4 and S5). The calculated peak positions of the total reaction probability for the J = 0 partial wave in Fig. 3B are very close to those on the measured BSS shown in Fig. 3A. These J-shifting peak structures in the total reaction probabilities of different partial waves are clear theoretical evidence of dynamical resonances in the Cl + HD (v = 1) → DCl + H reaction. We label the two resonances as resonance state a and resonance state b, corresponding to the two observed peaks in the BSS. The peaks in the J = 0 total reaction probability shown in Fig. 3B originate from the DCl (v' = 0,1) vibrational state–resolved probabilities; in contrast, the DCl (v' = 2) reaction probability shows peaks at the valley positions of the total reaction probability. Thus, the resonances enhance the DCl (v' = 0,1) probabilities but, to some extent, depress the DCl (v' = 2) probabilities.

To understand the nature of these resonances, we performed time-dependent quantum wave-packet calculations to obtain the wave functions at collision energies of 2.4 and 4.2 kcal/mol for the J = 0 partial wave. We use the node structure of the scattering wave function to determine the nature of the resonance states. The 2D contour of the scattering wave function at the collision energy of 2.4 kcal/mol (Fig. 4A1, resonance state a) shows the existence of two nodes along the DCl coordinate (correlating to the DCl product) in the H–DCl complex and one node along the reaction coordinate; the wave function for bending and the DCl coordinate (Fig. 4A2) shows no node on the bending coordinate. Note that the number of nodes in the scattering wave function provides important clues of the vibrational character of the resonance state. Here the wave function for resonance state a is the first excited state trapped in the H–DCl (v' = 2) vibrational adiabatic potential (VAP) well and can be assigned as the (v_1 = 1, v_2 = 0, v_3 = 2) or (102) state for the reaction complex (HDCl)†, where v_1 is the quantum number for the H-DCl stretching mode, v_2 for the bending mode, and v_3 for the DCl stretching mode, and † indicates that HDCl is a metastable reaction complex instead of a stable molecule. The 2D contour of the scattering wave function at the collision energy of 4.2 kcal/mol (Fig. 4A3, resonance state b) has a similar nodal structure to the (102) state for the DCl and the reaction coordinate but has two nodes on the bend mode (Fig. 4A4). This resonance state can therefore be assigned as the (122) state. Clearly, the two observed states (a and b) are quasi-bound resonance states trapped in the transition state region, with all three vibrational quantum numbers assignable. This confirms that the observed peaks in the BSS (Fig. 3A) in the Cl + HD (v = 1) → DCl + H reaction are caused by dynamical resonances.

Lifetimes of the two resonance states can also be obtained from the widths of the peaks in the total reaction probability of the J = 0 partial wave. The widths for DCl (v' = 0,1) peaks are determined to be 0.53 and 1.1 kcal/mol, corresponding to lifetimes of 28 and 14 fs for states a and b, respectively (fig. S3). This indicates that the two resonance states in the Cl + HD (v = 1) reaction are extremely short-lived in nature, in comparison with the resonance observed in the F + HD reaction, which lives longer than 100 fs. Partial waves analysis of the BSS for DCl (v' = 1) revealed that the peaks are mainly contributed by partial waves with small total angular momentum as expected, where resonance enhancement is most substantial (fig. S6).

Figure 4B shows a few VAP curves relevant to the Cl + HD (v = 0,1) → DCl + H reactions. On the ground VAP, there is a single barrier located at the product side close to the static barrier of the reaction system as shown in Fig. 4A. The interaction between H and DCl in the transition state region softens the DCl bond, manifesting not only in smaller vibrational frequency but, more importantly, also in larger anharmonicity.

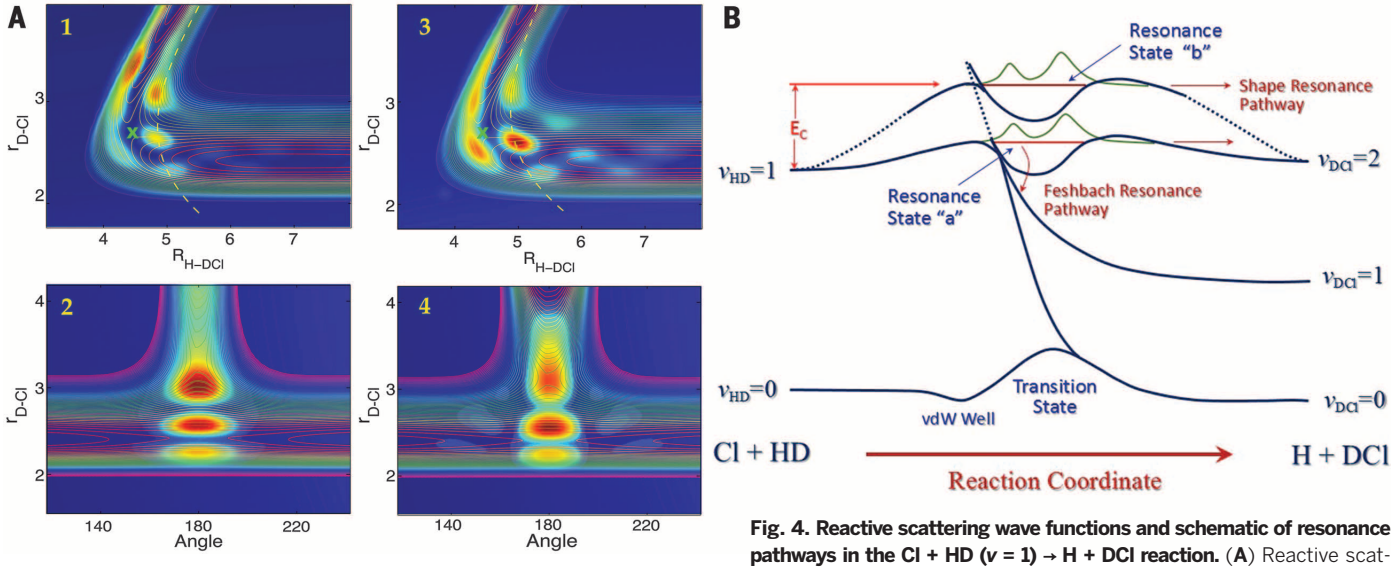


Fig. 4. Reactive scattering wave functions and schematic of resonance pathways in the Cl + HD (v = 1) → H + DCl reaction. (**A**) Reactive scattering wave functions at collision energies of E_C = 2.4 kcal/mol (panel 1) and 4.3 kcal/mol (panel 3) in the Jacobi coordinates $R_{H\text{-}DCl}$ and r_{DCl}, with the bending angle θ fixed at 0 and the transition barrier denoted as a green cross in panels 1 and 3. Reactive scattering wave functions at collision energies of E_C = 2.4 kcal/mol (panel 2) and 4.3 kcal/mol (panel 4) in the Jacobi coordinates for the bending angle θ and r_{DCl}, with the value of $R_{H\text{-}DCl}$ along the yellow dashed curves shown in panels 1 and 3, respectively. (**B**) VAPs along the reaction coordinate. The VAPs of v_{HD} = 1 and v_{DCl} = 2 correlate strongly, channeling the reactant flux and produces resonance states above the well of VAP of v_{DCl} = 2.

Mainly the latter considerably lowers the DCl energy levels for vibrationally excited states (for details, see section IV in the supplementary materials), resulting in a shallow well on the DCl (v_2 = 0, v_3 = 2)–H VAP around the peak position of the ground VAP, which supports one resonance state as shown in Fig. 4B. The resonance state b has the same origin as the state a, except on the bending excited state.

Both resonance states a and b are shape resonances on the DCl (v_3 = 2) VAP, and they enhance the overall reactivity as shown in Fig. 3B. They decay either through tunneling over the barrier to yield DCl (v' = 2) product or through coupling with the DCl (v_3 = 0,1) VAP to form DCl (v' = 0,1) products (Fig. 4B). Hence, these two resonance states are also Feshbach resonances. Because the state b is a bending excited (122) state, it produces DCl with a bimodal rotational distribution, in contrast with a Gaussian distribution from the state a (fig. S9). At E_C = 4.3 kcal/mol, the low–partial-wave contribution of the state b yields DCl with bimodal rotational distributions in the backward scattering hemisphere, whereas the contribution from high partial waves of resonance a generates broad peaks in the forward scattering hemisphere with a rotationally cold DCl product (fig. S10).

Quantum reactive scattering calculations showed that similar dynamical resonances also exist in the Cl + HD (v = 1) → HCl + D reaction channel (see supplementary materials for details). However, because of the considerably smaller cross sections of this reaction channel as compared with the DCl channel, which are due to the strong effect of the van der Waals interaction in the entrance channel as has been observed for the ground HD reaction (*21*), it is extremely difficult to perform crossed-beam experiments on this reaction channel.

From the above studies, we have demonstrated that extremely short-lived resonances in the Cl + HD (v = 1) reaction can be clearly probed with the backward scattering spectroscopy method. These resonances are supported by shallow wells on the DCl (v' = 2) VAP caused by chemical bond softening in the transition state region. Because chemical bond softening in the transition state region always occurs in reactions with energetic barriers and can result in potential wells on highly excited VAPs, we anticipate the existence of similar resonances in many other chemical reactions involving vibrationally excited reagents. Therefore, reaction dynamics studies with vibrationally excited reagents open the door to probe resonances in many direct chemical reactions.

REFERENCES AND NOTES

1. G. C. Schatz, *Science* **288**, 1599–1600 (2000).
2. F. Fernández-Alonso, R. N. Zare, *Annu. Rev. Phys. Chem.* **53**, 67–99 (2002).
3. K. Liu, *Adv. Chem. Phys.* **149**, 1–46 (2012).
4. R. D. Levine, S.-F. Wu, *Chem. Phys. Lett.* **11**, 557–561 (1971).
5. G. C. Schatz, J. M. Bowman, A. Kuppermann, *J. Chem. Phys.* **58**, 4023–4025 (1973).
6. D. M. Neumark, A. M. Wodtke, G. N. Robinson, C. C. Hayden, Y. T. Lee, *Phys. Rev. Lett.* **53**, 226–229 (1984).
7. R. T. Skodje *et al.*, *Phys. Rev. Lett.* **85**, 1206–1209 (2000).
8. M. Qiu *et al.*, *Science* **311**, 1440–1443 (2006).
9. T. Wang *et al.*, *Science* **342**, 1499–1502 (2013).
10. I. M. Waller, T. N. Kitsopoulos, D. M. Neumark, *J. Phys. Chem.* **94**, 2240–2242 (1990).
11. B. Zhang, K. Liu, *J. Chem. Phys.* **122**, 101102 (2005).
12. J. G. Zhou, J. J. Lin, K. P. Liu, *Mol. Phys.* **108**, 957–968 (2010).
13. J. Zhou, J. J. Lin, K. Liu, *J. Chem. Phys.* **121**, 813–818 (2004).
14. R. Otto *et al.*, *Science* **343**, 396–399 (2014).
15. S. C. Althorpe, D. C. Clary, *Annu. Rev. Phys. Chem.* **54**, 493–529 (2003).
16. J. Hirschfelder, H. Eyring, B. Topley, *J. Chem. Phys.* **4**, 170–177 (1936).
17. H. S. Johnston, *Gas Phase Reaction Rate Theory* (Ronald, New York, 1966).
18. K. P. Kumaran, S. S. Lim, J. V. Michael, *J. Chem. Phys.* **101**, 9487–9498 (1994).
19. J. Srinivasan, T. C. Allison, D. W. Schwenke, D. G. Truhlar, *J. Phys. Chem. A* **103**, 1487–1503 (1999).
20. M. Alagia *et al.*, *Science* **273**, 1519–1522 (1996).
21. D. Skouteris *et al.*, *Science* **286**, 1713–1716 (1999).
22. D. Skouteris *et al.*, *J. Chem. Phys.* **114**, 10662 (2001).
23. M. J. Ferguson, G. Meloni, H. Gomez, D. M. Neumark, *J. Chem. Phys.* **117**, 8181–8184 (2002).
24. D. E. Manolopoulos *et al.*, *Science* **262**, 1852–1855 (1993).
25. D. M. Neumark, *Science* **272**, 1446–1447 (1996).
26. S. A. Kandel *et al.*, *J. Chem. Phys.* **112**, 670–685 (2000).
27. M. H. Alexander, G. Capecchi, H. J. Werner, *Science* **296**, 715–718 (2002).
28. E. Garand, J. Zhou, D. E. Manolopoulos, M. H. Alexander, D. M. Neumark, *Science* **319**, 72–75 (2008).
29. X. Wang *et al.*, *Science* **322**, 573–576 (2008).
30. Z. F. Ren *et al.*, *Rev. Sci. Instrum.* **77**, 016102 (2006).
31. T. Wang, T. Yang, C. Xiao, D. Dai, X. Yang, *J. Phys. Chem. Lett.* **4**, 368–371 (2013).
32. N. Mukherjee, R. N. Zare, *J. Chem. Phys.* **135**, 184202 (2011).
33. N. Mukherjee, W. Dong, J. A. Harrison, R. N. Zare, *J. Chem. Phys.* **138**, 051101 (2013).
34. L. Schnieder, K. Seekamp-Rahn, E. Wrede, K. H. Welge, *J. Chem. Phys.* **107**, 6175–6195 (1997).
35. M. H. Qiu *et al.*, *Rev. Sci. Instrum.* **76**, 083107 (2005).
36. H. J. Werner *et al.*, MOLPRO, version 2012.1, A package of ab initio programs (2012); see www.molpro.net.
37. Z. Sun, H. Guo, D. H. Zhang, *J. Chem. Phys.* **132**, 084112 (2010).

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China, the Ministry of Science and Technology of China, and the Key Research Program of the Chinese Academy of Sciences.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/60/suppl/DC1
Supplementary Text
Figs. S1 to S13
Tables S1 and S2
References (*38*–*45*)

28 August 2014; accepted 17 November 2014
10.1126/science.1260527

BACTERIAL EVOLUTION

The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer

Sandrine Borgeaud, Lisa C. Metzger, Tiziana Scrignari, Melanie Blokesch*

Natural competence for transformation is a common mode of horizontal gene transfer and contributes to bacterial evolution. Transformation occurs through the uptake of external DNA and its integration into the genome. Here we show that the type VI secretion system (T6SS), which serves as a predatory killing device, is part of the competence regulon in the naturally transformable pathogen *Vibrio cholerae*. The T6SS-encoding gene cluster is under the positive control of the competence regulators TfoX and QstR and is induced by growth on chitinous surfaces. Live-cell imaging revealed that deliberate killing of nonimmune cells via competence-mediated induction of T6SS releases DNA and makes it accessible for horizontal gene transfer in *V. cholerae*.

Vibrio cholerae is a well-studied human pathogen that causes severe and potentially fatal diarrhea in humans. *V. cholerae* is primarily an aquatic bacterium that is often found in association with zooplankton (*1*). The molted exoskeletons of planktonic crustaceans are primarily composed of the polymer chitin. When growing on chitinous surfaces, *V. cholerae* initiates a developmental program known as natural competence (*2*, *3*), which allows the bacterium to take up free DNA from the environment (*4*) using a competence-specific DNA uptake machinery (*5*, *6*). The competence program is dependent on the regulatory protein TfoX, which is produced in the presence of chitin and chitin degradation products (*4*, *7*–*9*) (fig. S1). Natural competence is also co-regulated by carbon catabolite repression (*10*) and quorum sensing (QS) (*7*). QS requires autoinducers [cholera autoinducer 1 (CAI-1) and autoinducer 2] and a master regulator (HapR) (*11*). We recently demonstrated that only a subset of the known competence genes (e.g., *comEA* and *comEC*) (fig. S1) are co-regulated by QS and in a CAI-1–dependent manner (*12*, *13*), and we suggested that CAI-1 acts as a competence pheromone (*12*). The QS and TfoX-dependent regulator QstR links QS and TfoX activity in the induction of competence genes (*14*) (fig. S1). In this study, we demonstrate that the type VI secretion system (T6SS)

Laboratory of Molecular Microbiology, Global Health Institute, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland.
***Corresponding author. E-mail: melanie.blokesch@epfl.ch**

(*15*, *16*) of *V. cholerae* is part of the competence regulon and is expressed when the bacterium grows on chitinous surfaces. As a consequence, any nonimmune neighboring cells (*16*–*18*) are killed, and the released DNA serves as transforming material that enhances horizontal gene transfer (HGT).

To understand the extent of the TfoX-dependent competence regulon, we used an RNA sequencing (RNA-seq) approach to enable an accurate assessment of the bacterial transcriptome (*19*). We studied a variant of the pandemic *V. cholerae* O1 El Tor isolate A1552 (table S1), which carries an arabinose (ara)–inducible copy of *tfoX* on its chromosome, as the wild-type (WT) strain and grew the bacteria in the absence or presence of arabinose to simulate chitin-induced expression of TfoX (*5*, *13*). Upon the induction of *tfoX*, we observed substantial up-regulation of the three T6SS-encoding gene clusters (Fig. 1A, fig. S2, and table S2): the major or large gene cluster, which encodes the structural components of the T6SS, and two auxiliary clusters (*17*, *20*), which encode haemolysin co-regulated (Hcp) proteins. All three clusters also encode valine-glycine repeat G proteins (VgrG-1, VgrG-2, VgrG-3) and various effector-immunity modules (*17*). The T6SS machinery structurally and functionally resembles intracellular and membrane-attached phage tails (*15*, *21*) and forms a tubular structure composed of the two sheath proteins VipA and VipB (*22*). Upon contraction, the sheath propels an inner tube composed of Hcp proteins and capped by a complex between the VgrG proteins and a proline-alanine-alanine-arginine (PAAR) repeat–containing spike protein, together with the effectors, into neighboring bacterial or eukaryotic cells (*15*–*18*, *20*, *21*, *23*, *24*). The predatory population itself is protected against self-destruction by the simultaneous production of effector-compatible immunity proteins (*16*–*18*).

The regulation of the T6SS in *V. cholerae* is not well understood because the function and dynamics of the T6SS have primarily been studied in nonpandemic isolates with constitutive T6SS activity (e.g., strains V52 and 2740-80) (*17*, *20*, *21*, *23*, *24*). In strain V52, the enhancer-binding protein VasH, which is encoded by the major T6SS gene cluster (fig. S2), acts as an activator of the sigma factor RpoN (*20*). VasH has no effect on the structural genes of the major T6SS gene cluster but solely controls the auxiliary clusters and *vgrG3* (*25*). Although high osmolarity promotes Hcp secretion in strain A1552, the main T6SS gene

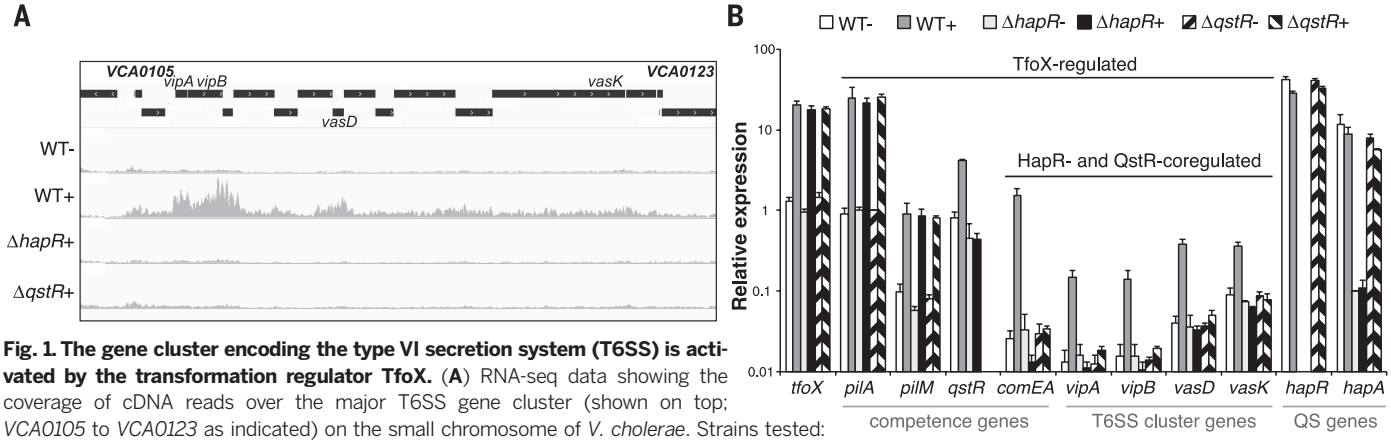


Fig. 1. The gene cluster encoding the type VI secretion system (T6SS) is activated by the transformation regulator TfoX. (**A**) RNA-seq data showing the coverage of cDNA reads over the major T6SS gene cluster (shown on top; *VCA0105* to *VCA0123* as indicated) on the small chromosome of *V. cholerae*. Strains tested: WT, *V. cholerae* strain A1552; Δ*hapR*, A1552Δ*hapR*; and Δ*qstR*, A1552Δ*qstR*. All strains contain the Tn*tfoX* transposon (*13*) on their chromosome. "-" and "+" indicate the absence and presence of *tfoX* induction, respectively. Scale of the *y* axis for each strain: 0 to 500. (**B**) Verification of the RNA-seq data by qRT-PCR. The same strains as indicated in (A) were grown in the absence (-) or presence (+) of *tfoX* induction, and the relative expression of the indicated genes was evaluated through qRT-PCR (*y* axis). The data represent the mean of three independent biological replicates (± SD, as indicated by the error bars).

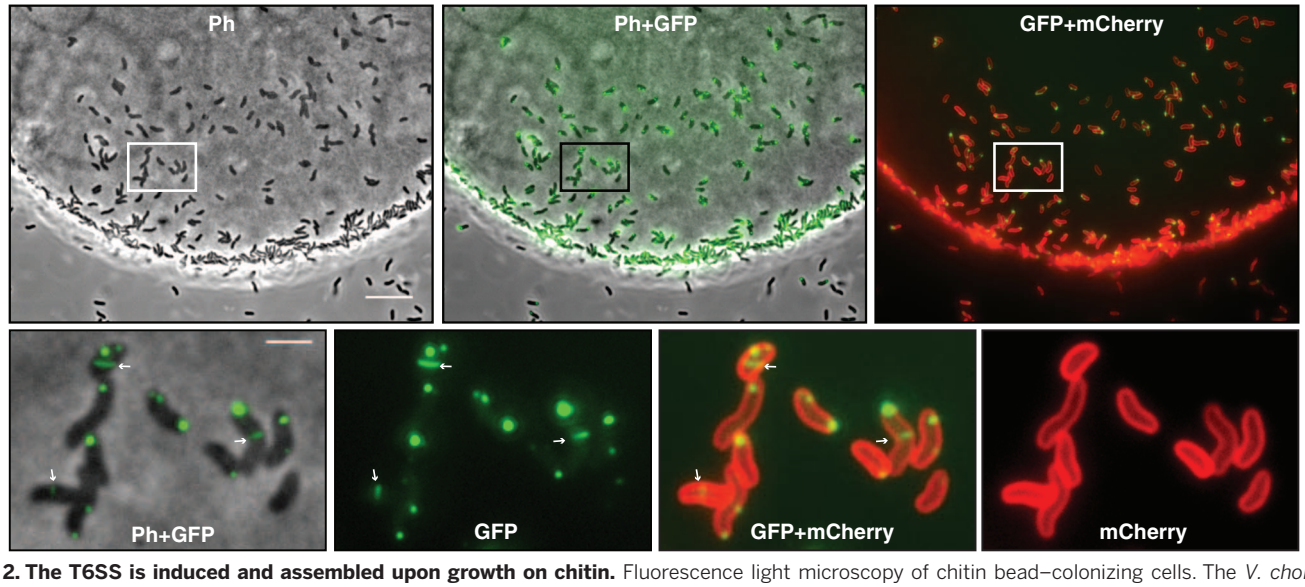


Fig. 2. The T6SS is induced and assembled upon growth on chitin. Fluorescence light microscopy of chitin bead–colonizing cells. The *V. cholerae* strain was engineered to carry translational fusions for both the competence protein ComEA-mCherry and the T6SS sheath protein VipA-sfGFP. From left to right (upper row): phase contrast image (Ph), phase contrast image overlaid with the signal from the green fluorescence channel (Ph+GFP), and merged signal from the two fluorescence channels (GFP+mCherry). Zoomed images of the boxed region are shown in the lower row. Assembled and contracted VipA-sfGFP sheath structures are indicated by the white arrows. Scale bars: 10 μm (upper row); 2 μm (lower row).

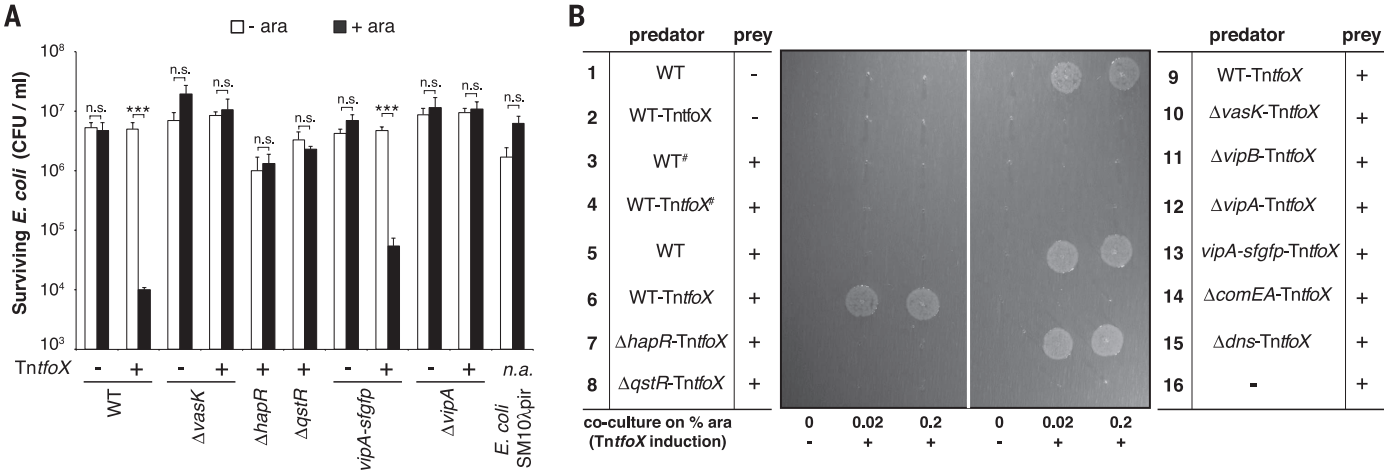


Fig. 3. TfoX-mediated expression of the T6SS leads to bacterial killing and natural transformation. (**A**) Quantification of *E. coli* TOP10 recovery [colony-forming units per milliliter (CFU/ml), as indicated on the *y* axis] after coculturing with *V. cholerae* or *E. coli* cells on plain LB agar plates (-ara) or LB agar plates supplemented with arabinose (+ara). The tested strains are indicated below the graph and contain the arabinose-inducible copy of *tfoX* (Tn*tfoX*) where indicated. The averages of three independent experiments (± SD, error bars) are shown. Asterisks denote statistically significant differences ($P < 0.001$); n.s., not significant. (**B**) TfoX-dependent expression of the T6SS enhances natural transformation. The indicated predatory strains are derivatives of *V. cholerae* O1 El Tor A1552 (Rif^R) and carry Tn*tfoX* where indicated. The predator was incubated without any prey (-) or with prey strains (+) that are derived from the environmental isolate Sa5Y (*28*) and contain a Kan^R cassette integrated into *lacZ* (Δ*lacZ*; lanes 3 and 4; T6SS-proficient) or *vipA* (Δ*vipA;* lanes 6 to 16; T6SS-defective). The strains were cocultured on LB agar plates containing 0, 0.02, or 0.2% of the *tfoX* inducer arabinose (as indicated) before the selection of transformants on LB plates containing both antibiotics (Rif+Kan). "#" denotes strains that were killed by the T6SS-proficient Sa5Y strain and are thus non-transformable. The horizontal transfer of the Kan^R cassette from the prey to the predator was confirmed for randomly picked transformants (fig. S10).

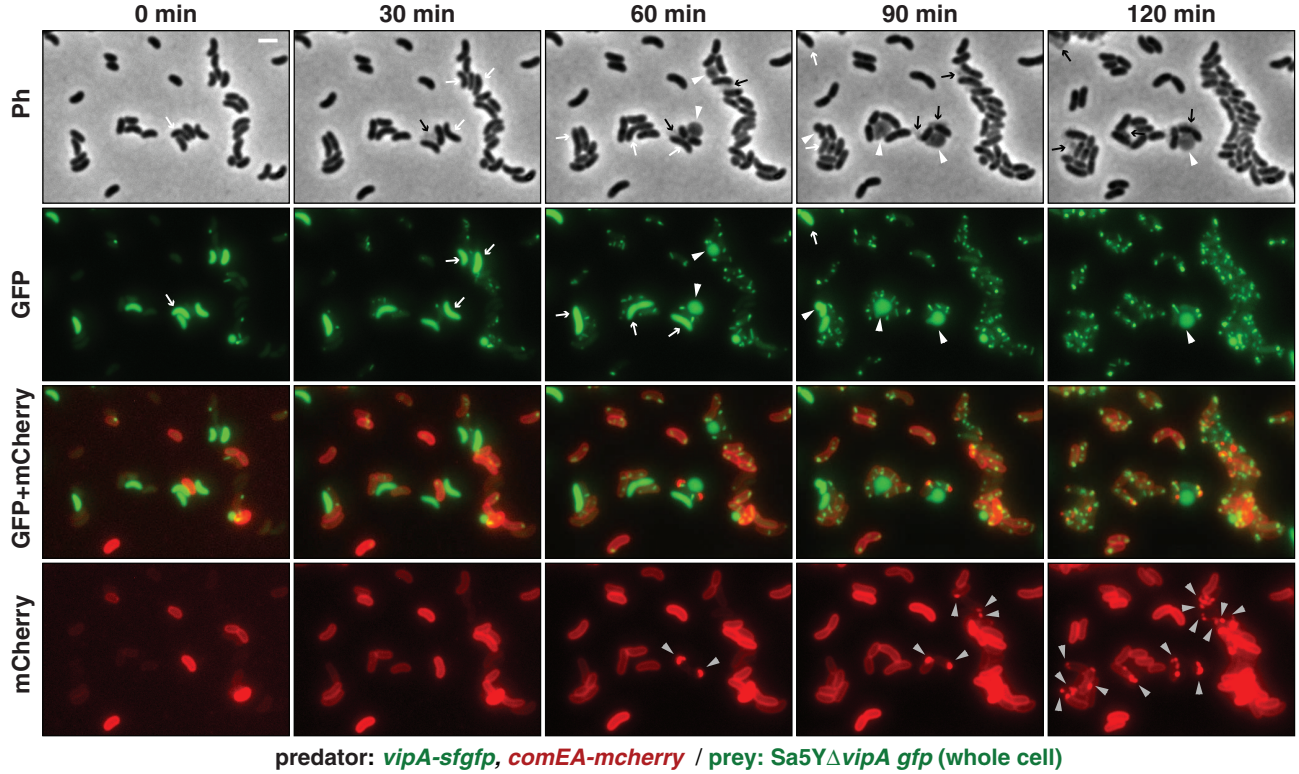


Fig. 4. TfoX-induced T6SS leads to the killing of neighboring cells followed by horizontal gene transfer. This figure shows a time-lapse microscopy image series showing the T6SS-dependent attack of nonimmune prey cells followed by their rounding and lysis and the uptake of their DNA by neighboring competent predatory cells. A *vipA-sfgfp*– and *comEA-mCherry*–carrying *V. cholerae* strain (predator) was grown to high cell density under *tfoX*-inducing conditions, mixed with the *gfp*-labeled *V. cholerae* strain Sa5YΔ*vipA gfp* (prey), and spotted onto agarose pads for imaging. The white arrows and arrowheads denote prey cells before and after attack (e.g., cell rounding), respectively, and the black arrows indicate cell debris after lysis [visualized in the phase contrast images (Ph; top row)]. DNA released by the prey lysis is taken up by the competent predator, as indicated by the gray arrowheads highlighting the redistribution of the periplasmic competence protein ComEA-mCherry (lower row). Accumulations of the VipA-sfGFP protein and assembled and contracted T6SS sheath structures are visible in the green channel (GFP). Scale bar: 2 μm (indicated in the upper left image and valid for all images).

cluster is induced by only ~1.5-fold compared with low-osmolarity conditions (*26*). Therefore, Ho *et al*. proposed that there is an additional environmental signal that triggers the transcription of the major cluster (including *vasH*) and that VasH can then activate the expression of the auxiliary clusters in an RpoN-dependent manner (*15*). On the basis of our RNA-seq data, we suggest that the competence regulatory protein TfoX initiates the transcription of the major T6SS cluster and, thus, the auxiliary clusters (Fig. 1A, fig. S2, and table S2).

TfoX regulates the majority of the competence genes that encode components of the DNA uptake machinery of *V. cholerae* (*4*–*7*, *13*). Notably, a subset of the so-far identified competence genes is co-regulated by QS in a HapR- and QstR-dependent manner (*4*, *13*, *14*) (fig. S1). We therefore repeated the RNA-seq experiment in *tfoX*-expressing *hapR*- and *qstR*-negative strains. In these mutants, TfoX-dependent up-regulation of the T6SS genes was lost (Fig. 1A, fig. S2, and tables S3 and S4). We used quantitative reverse-transcription polymerase chain reaction (qRT-PCR) to confirm the TfoX-, HapR-, and QstR-dependent regulation of the T6SS (Fig. 1B). Moreover, as the regulatory circuits often differ among strains, we tested the TfoX (and HapR)–dependent expression of the T6SS genes in five different *V. cholerae* O1 El Tor isolates (table S1) from South America, Asia, and Africa (fig. S3), which were all naturally transformable on *tfoX* induction (table S5).

The transcription and translation of native *tfoX* requires the presence of a chitinous substrate (*4*, *7*–*9*), and we confirmed that the expression of the T6SS genes was elevated after growth of *V. cholerae* on chitin flakes (*27*) (fig. S4). We also employed a chitin colonization assay (*10*) to visualize the T6SS. We generated a translational fusion between superfolder green fluorescent protein (sfGFP) and the main sheath protein VipA, as previously described (*23*), and used the *vipA-sfgfp* allele to replace the indigenous copy of *vipA*. Next we verified that VipA-sfGFP was produced under competence-inducing conditions (e.g., in a TfoX- and high cell density–dependent manner) concomitantly with the periplasmic competence protein ComEA (fig. S5; details below). In accordance with earlier studies using the T6SS hyperactive strain 2740-80 (*23*), we observed both extended and contracted sheath structures after *tfoX* induction and dynamic T6SS behavior (fig. S5). Moreover, upon growth on chitin beads, we detected VipA-sfGFP forming extended and contracted sheath structures, indicating that the full T6SS gene cluster was expressed (Fig. 2). Such chitin-induced production of T6SS sheath structures was also observed in several other *V. cholerae* O1 and non-O1 strains (fig. S6) and was dependent on the regulator QstR and on other structural components of the T6SS (fig. S7).

Having established that chitin and TfoX induce the T6SS gene cluster in *V. cholerae*, we assessed its functionality in an interspecies killing assay. Strains containing the inducible copy of *tfoX* exhibited significant killing behavior toward *Escherichia coli* in a T6SS-dependent manner when grown in the presence of the inducer (Fig. 3A and fig. S8).

To investigate whether competence-induced T6SS-mediated killing affects transformation, we cocultured the WT strain as the predator (without and with *tfoX* induction) with an environmental *V. cholerae* isolate as the prey [Sa5Y (*28*) (table S1)] in an intraspecies mixed-community assay. Comparative genomic hybridization data for strain Sa5Y (*29*) confirmed that this strain lacks the common O1 El Tor T6SS immunity genes *tsiV1*, *tsiV2*, and *tsiV3* (*30*). To avoid premature killing of the predator strain by the prey before it could reach high cell density (which is required for competence- and QS-mediated induction of the T6SS), we inactivated the T6SS of strain Sa5Y (table S1).

Natural transformants were readily obtained upon TfoX-induction in predator cells (Fig. 3B and figs. S9 and S10). These natural transformation events were fully dependent on the competence co-regulators HapR and QstR and on ComEA (Fig. 3B and fig. S9), as these proteins are also required for natural transformation when purified genomic DNA serves as the transforming material (*4*, *12*–*14*) (fig. S11). Although the activation of the T6SS system seemed to be negligible in the latter case (fig. S11), transformants were undetectable or only rarely detectable in T6SS-defective strains when grown in a mixed community with strain Sa5Y (Fig. 3B and fig. S9). A similar T6SS-dependent increase in the transformation frequency was observed when the predator and prey strains were cocultured on chitinous surfaces (without artificial *tfoX* induction) (fig. S12). We concluded that upon competence induction, *V. cholerae* induced the T6SS and thus led to the killing of neighboring nonimmune bacteria. Lysis of neighboring bacteria causes release of genomic DNA that then transforms competent predatory cells. The enhancement of HGT in *V. cholerae* by the T6SS-mediated killing of nonimmune cells resembles bacterial fratricide described for naturally competent *Streptococcus pneumoniae* (*31*). However, in contrast to fratricide, which also promotes the lysis of noncompetent sibling cells (*3*, *31*), competence-induced T6SS-mediated killing by *V. cholerae* appears to primarily target strains containing noncompatible effector-immunity modules and is contact-dependent. Earlier studies showed that the *V. cholerae* O1 El Tor strain C6706 is T6SS-silent (*30*) and unable to kill *E. coli* cells under standard laboratory conditions (*32*) (as shown in Fig. 3 for *tfoX*-uninduced conditions) owing to its inability to produce structural components of the T6SS (*32*). However, the tested strain was fully resistant against the T6SS-active strain V52, indicating that immunity is maintained even in the absence of T6SS activity (*17*, *32*).

Our final goal was to visualize prey lysis and subsequent transfer of genetic material by live-cell fluorescence microscopy imaging. Thus, we used a translational fusion between the competence protein ComEA and the fluorescent protein mCherry (*6*). We previously demonstrated that ComEA of *V. cholerae* is a periplasmic protein that is highly mobile within this compartment (*6*). Moreover, ComEA is strictly required for DNA translocation across the outer membrane of competent cells and most likely contributes to the DNA uptake process by acting as a Brownian ratchet and compacting the incoming DNA within the periplasm (*6*). We therefore combined the *comEA-mCherry* and *vipA-sfGFP* alleles and incubated the resulting predator strain (after *tfoX* induction) with a *gfp*-tagged prey strain. Under those conditions, we observed high T6SS activity in the predator cells and, as a consequence, cell rounding and lysis of the prey (Fig. 4 and fig. S13). We also observed competent bacteria in close proximity to lysed cells exhibiting the distinctive focus formation of the

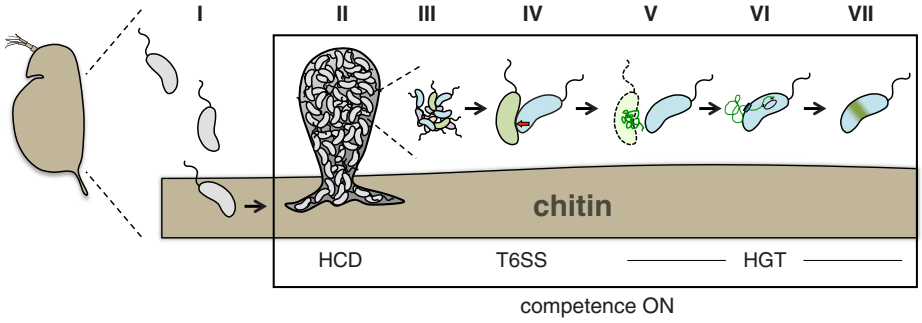


Fig. 5. Schematic summarizing the main findings of this study. In its natural environment, *V. cholerae* often colonizes the exoskeleton molts of zooplankton. Upon initial attachment (I), the bacteria form a three-dimensional biofilm (II) on the chitinous surface, thereby reaching high cell density (HCD). Within these biofilms, different *V. cholerae* strains carrying diverse and noncompatible effector-immunity modules (shown by the different colors of bacteria) most likely form mixed communities (III). Under such chitin-dependent and HCD conditions, *V. cholerae* induces its natural competence program. In this study, we showed that the T6SS (indicated by the red arrow) is part of the competence regulon and induced upon the growth of *V. cholerae* on chitin substrates. The T6SS is active against nonimmune neighboring cells (IV), leading to their lysis (V). The DNA released by the lysed cells can be taken up by a competent predator cell (VI), leading to its natural transformation and evolution in the case that new DNA material is incorporated (VII).

ComEA-mCherry protein, which is indicative of DNA translocation into the periplasm (*6*) of a predator cell (Fig. 4 and figs. S13 and S14). Similar gene-transfer events were never or only rarely observed in the T6SS-negative strain; in the presence of extracellular deoxyribonuclease; if the predator lacked the outer membrane secretin protein PilQ, which is required for efficient DNA uptake (*5*, *6*); or in the absence of any prey (figs. S14 and S15).

Our findings indicated that the T6SS of *V. cholerae* is part of the competence regulon and is induced on chitinous surfaces in a TfoX-, HapR-, and QstR-dependent manner, thereby enhancing HGT (Fig. 5). HGT plays a major role in bacterial evolution and contributes to the spread of antibiotic resistance cassettes and pathogenicity islands. Moreover, because chitinous zooplankton are thought to play an important role in cholera transmission in endemic regions (*1*), chitin-induced expression of the T6SS might also enhance the virulence potential of the pathogen due to the killing of commensal bacteria within the human gut.

REFERENCES AND NOTES

1. E. K. Lipp, A. Huq, R. R. Colwell, *Clin. Microbiol. Rev.* **15**, 757–770 (2002).
2. I. Chen, D. Dubnau, *Nat. Rev. Microbiol.* **2**, 241–249 (2004).
3. C. Johnston, B. Martin, G. Fichant, P. Polard, J. P. Claverys, *Nat. Rev. Microbiol.* **12**, 181–196 (2014).
4. K. L. Meibom, M. Blokesch, N. A. Dolganov, C.-Y. Wu, G. K. Schoolnik, *Science* **310**, 1824–1827 (2005).
5. P. Seitz, M. Blokesch, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17987–17992 (2013).
6. P. Seitz *et al.*, *PLOS Genet.* **10**, e1004066 (2014).
7. P. Seitz, M. Blokesch, *FEMS Microbiol. Rev.* **37**, 336–363 (2013).
8. S. Yamamoto *et al.*, *Mol. Microbiol.* **91**, 326–347 (2014).
9. A. B. Dalia, D. W. Lazinski, A. Camilli, *mBio* **5**, e01028-13 (2014).
10. M. Blokesch, *Environ. Microbiol.* **14**, 1898–1912 (2012).
11. W. L. Ng, B. L. Bassler, *Annu. Rev. Genet.* **43**, 197–222 (2009).
12. G. Suckow, P. Seitz, M. Blokesch, *J. Bacteriol.* **193**, 4914–4924 (2011).
13. M. Lo Scrudato, M. Blokesch, *PLOS Genet.* **8**, e1002778 (2012).
14. M. Lo Scrudato, M. Blokesch, *Nucleic Acids Res.* **41**, 3644–3658 (2013).
15. B. T. Ho, T. G. Dong, J. J. Mekalanos, *Cell Host Microbe* **15**, 9–21 (2014).
16. A. B. Russell, S. B. Peterson, J. D. Mougous, *Nat. Rev. Microbiol.* **12**, 137–148 (2014).
17. D. Unterweger *et al.*, *Nat. Commun.* **5**, 3549 (2014).
18. E. Durand, C. Cambillau, E. Cascales, L. Journet, *Trends Microbiol.* **22**, 498–507 (2014).
19. Materials and methods are available as supplementary materials on *Science* Online.
20. S. Pukatzki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1528–1533 (2006).
21. M. M. Shneider *et al.*, *Nature* **500**, 350–353 (2013).
22. G. Bönemann, A. Pietrosiuk, A. Diemand, H. Zentgraf, A. Mogk, *EMBO J.* **28**, 315–325 (2009).
23. M. Basler, M. Pilhofer, G. P. Henderson, G. J. Jensen, J. J. Mekalanos, *Nature* **483**, 182–186 (2012).
24. M. Basler, B. T. Ho, J. J. Mekalanos, *Cell* **152**, 884–894 (2013).
25. T. G. Dong, J. J. Mekalanos, *Nucleic Acids Res.* **40**, 7766–7775 (2012).
26. T. Ishikawa *et al.*, *Infect. Immun.* **80**, 575–584 (2012).
27. R. L. Marvig, M. Blokesch, *BMC Microbiol.* **10**, 155 (2010).
28. D. P. Keymer, M. C. Miller, G. K. Schoolnik, A. B. Boehm, *Appl. Environ. Microbiol.* **73**, 3705–3714 (2007).
29. M. C. Miller, D. P. Keymer, A. Avelar, A. B. Boehm, G. K. Schoolnik, *Appl. Environ. Microbiol.* **73**, 3695–3704 (2007).
30. T. G. Dong, B. T. Ho, D. R. Yoder-Himes, J. J. Mekalanos, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2623–2628 (2013).
31. J. P. Claverys, L. S. Håvarstein, *Nat. Rev. Microbiol.* **5**, 219–229 (2007).
32. S. T. Miyata, D. Unterweger, S. P. Rudko, S. Pukatzki, *PLOS Pathog.* **9**, e1003752 (2013).

ACKNOWLEDGMENTS

We thank A. Boehm, M. Miller, members of the Institut National de Recherche Biomédicale of the Democratic Republic of the Congo, and M. Lo Scrudato for providing *V. cholerae* strains. We also acknowledge the service provided by Microsynth and members of the Blokesch laboratory for scientific discussions. This work was supported by the Swiss National Science Foundation (grant 31003A_143356) and the European Research Council (grant 309064-VIR4ENV). Supporting data are provided in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/63/suppl/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S5
References (*33*–*54*)

15 August 2014; accepted 1 December 2014
10.1126/science.1260064

INNATE IMMUNITY

Dermal adipocytes protect against invasive *Staphylococcus aureus* skin infection

Ling-juan Zhang,[1] Christian F. Guerrero-Juarez,[2,3] Tissa Hata,[1] Sagar P. Bapat,[4] Raul Ramos,[2,3] Maksim V. Plikus,[2,3] Richard L. Gallo[1]*

Adipocytes have been suggested to be immunologically active, but their role in host defense is unclear. We observed rapid proliferation of preadipocytes and expansion of the dermal fat layer after infection of the skin by *Staphylococcus aureus*. Impaired adipogenesis resulted in increased infection as seen in *Zfp423*^{nur12} mice or in mice given inhibitors of peroxisome proliferator–activated receptor γ. This host defense function was mediated through the production of cathelicidin antimicrobial peptide from adipocytes because cathelicidin expression was decreased by inhibition of adipogenesis, and adipocytes from *Camp*$^{-/-}$ mice lost the capacity to inhibit bacterial growth. Together, these findings show that the production of an antimicrobial peptide by adipocytes is an important element for protection against *S. aureus* infection of the skin.

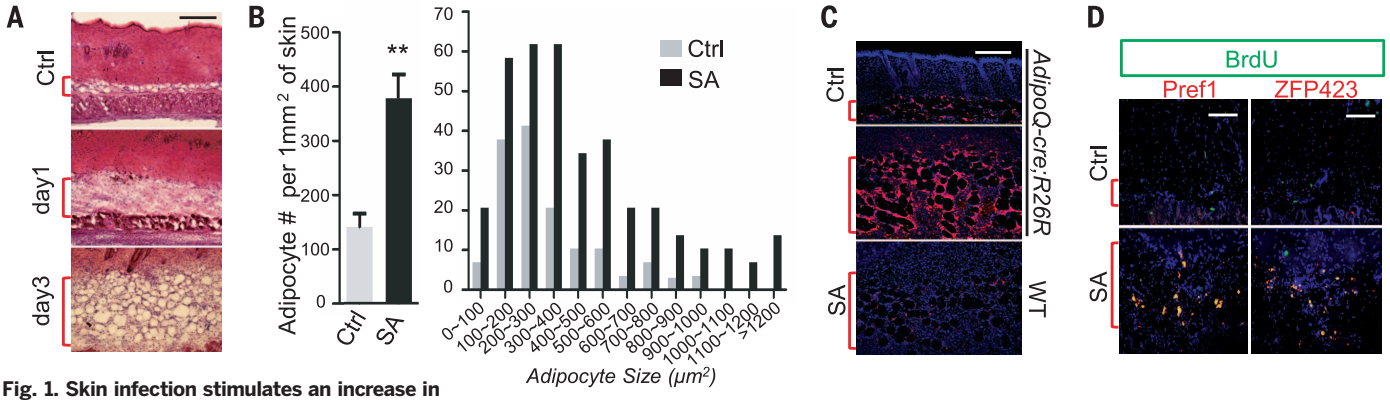
Host defense against microbial infection involves the participation of several cell types. Owing to the rapid doubling time of many microbes, immediate protection provided by local resident cells—such as epithelial cells, mast cells, and resident leukocytes—is essential to restrict the spread of infection during the lag period before recruitment of additional cells, such as neutrophils and monocytes (*1*, *2*). The production of antimicrobial peptides (AMPs) by local resident cells and recruited leukocytes is a key mechanism to limit pathogen growth (*3*–*5*).

Staphylococcus aureus is a major cause of skin and soft-tissue infections in humans, causing both local and systemic disease (*6*, *7*). We observed that a large and previously unrecognized expansion of the subcutaneous adipose layer was evident during the early response to *S. aureus* skin infection (Fig. 1A). The response to infection was confirmed with quantification of the abundance of adipocytes (Fig. 1B and fig. S1A), observations of an increase in lipid staining (fig. S1B), and increased activation of the adiponectin promoter as measured in *AdipoQcre;R26R* mice (Fig. 1C) (*8*). Adipocytes progressively increased in size after *S. aureus* infection (Fig. 1B), suggesting that the expansion of dermal adipose tissue occurs at least in part through hypertrophy of mature adipocytes. PREF1 and ZFP423 mark committed preadipocytes required for adipose tissue development and expansion (*9*–*11*). Proliferation of these preadipocytes at the site of infection was further confirmed with colocalization of PREF1 and ZFP423 with proliferation markers BrdU (Fig. 1D and fig. S1C) and Ki67 (fig. S1D). Additionally, dermal cells isolated from *S. aureus*–infected skin exhibited greater adipogenic potential than that of cells isolated from the same amount of uninfected skin, as indicated by lipid production and induction of adipocyte markers *Adipoq* and *Fabp4* in response to adipocyte differentiation medium (Fig. 1E and fig. S1E). Also supporting the conclusion that infection results in an increase of cells within the dermis with the potential to differentiate into adipocytes were observations of an increase of mRNA and protein for transcription factors driving preadipocyte differentiation, including *Cebpb*, *Pparg*, and *Cebpa* (Fig. 1F and fig. S1, D and F) (*12*, *13*). Peroxisome proliferator–activated receptor–γ (PPARγ)–positive cells at the infected sites were negative for CD11b (fig. S1G), confirming that they were not myeloid cells. To test that cell proliferation was associated with adipocyte formation, we examined BrdU

[1]Division of Dermatology, University of California, San Diego (UCSD), La Jolla, CA 92093, USA. [2]Department of Developmental and Cell Biology, Sue and Bill Gross Stem Cell Research Center, University of California, Irvine, Irvine, CA 92697, USA. [3]Center for Complex Biological Systems, University of California, Irvine, Irvine, CA 92697, USA. [4]Nomis Foundation Laboratories for Immunobiology and Microbial Pathogenesis, The Salk Institute for Biological Studies, San Diego, La Jolla, CA 92037, USA.
***Corresponding author. E-mail: rgallo@ucsd.edu**

incorporation within the nuclei of adipocytes after multiple injections of BrdU (*14*) during the first 3 days after infection. A significant increase in the number of BrdU-positive nuclei was seen within cells from *S. aureus*–infected mice stained with the adipocyte surface protein Caveolin (Fig. 1G), thus confirming that adipocytes at the site of infection were at least partially derived from proliferative precursors. Together, these data showed that infection by *S. aureus* triggers pre-adipocyte proliferation and expansion of local dermal adipocytes.

We next tested whether adipocyte activation was essential for protection against *S. aureus* infection using *Zfp423* reporter mice and *Zfp423*^{nur12}



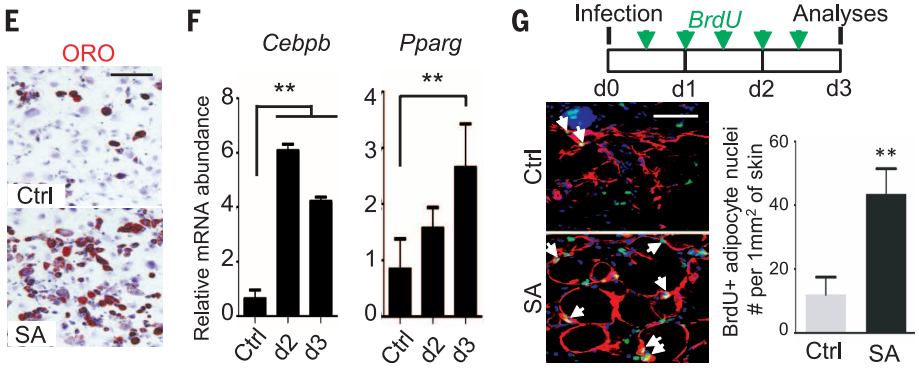
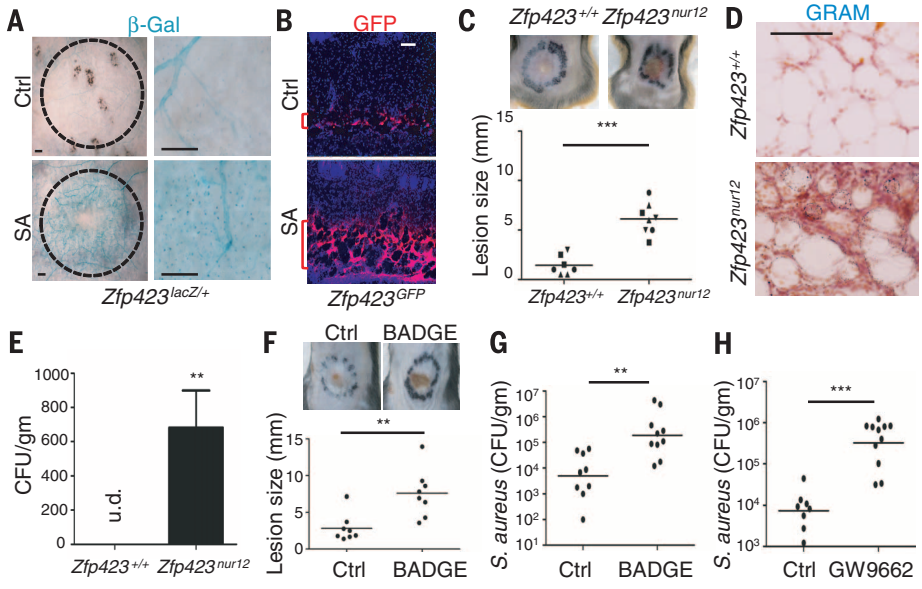


Fig. 1. Skin infection stimulates an increase in dermal adipocytes. (**A**) Hematoxylin and eosin staining of mouse skin injected with phosphate-buffered saline (PBS) control (Ctrl) or *S. aureus* (SA). Red brackets indicate subcutaneous adipose layer. Scale bars, 200 μm. (**B**) Quantification of the number and size distribution of Caveolin^{+}/Perilipin^{+} adipocytes 3 days after Ctrl or *S. aureus* injection (n = 3 to ~5 mice/group and 3 microscopy fields/mice). (**C**) Increase of adiponectin positive cells seen by staining for β-Gal (red) 3 days after *S. aureus* or Ctrl injection in AdipoQcre;R26R mice. Wild-type mice injected with *S. aureus* are shown as a staining control. Scale bars, 200 μm. Nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI) (blue). (**D**) Ctrl or *S. aureus*–infected mice (day 1) were injected with BrdU 4 hours before analysis. Proliferative preadipocytes were identified by means of yellow staining showing colocalization of PREF1 or ZFP423 (red) and BrdU (green). Scale bars, 100 μm. (**E**) Dermal cells were isolated from Ctrl or *S. aureus* (SA)–infected skin then treated with adipocyte differentiation medium for 5 days. Lipid production was shown by Oil-Red-O (ORO) staining. Scale bars, 200 μm. (**F**) Relative mRNA expression for *Cebpb* and *Pparg* (n = 3 mice/group) in skin after *S. aureus* infection. (**G**) (Top) Schematic of 3-day BrdU labeling experiments during *S. aureus* infection. (Left) Representative images for caveolin (red) and BrdU (green) staining of skin sections. Arrows indicate Caveolin^{+}BrdU^{+} adipocytes. Scale bars, 50 μm. (Right) Quantification of the number of Caveolin^{+}BrdU^{+} adipocytes (n = 3~5 mice/group and 3 microscopy fields/mice). All error bars indicate mean ± SEM; ** $P < 0.01$ (t test).

Fig. 2. Adipocytes are essential for host defense against *S. aureus* infection. (**A**) β-gal staining of the underside of *Zfp423*$^{lacZ/+}$ skin 3 days after injection with PBS (Ctrl) or *S. aureus*. Injected area is indicated by dotted line. Scale bars, 1mm (left), 200 μm (right). (**B**) GFP immunostaining (red) of skin sections from Ctrl or *S. aureus*–injected *Zfp423*^{GFP} mice. Scale bars, 100 μm. (**C**) Infected lesion size in *Zfp423*^{nur12} or *Zfp423*$^{+/+}$ mice (n = 7~8/group) and (**D**) bacteria in adipose observed by Gram staining (dark blue). Scale bars, 100 μm. (**E**) Systemic bacteremia detected 3 days after *S. aureus* skin injection in spleens isolated from *Zfp423*^{nur12} mice but not *Zfp423*$^{+/+}$ mice (n = 3 mice/group). Scale bars, 200 μm. (**F** to **H**) BADGE or GW9662 increased susceptibility to *S. aureus* as observed by (F) increased lesion size and [(G) and (H)] increased *S. aureus* (CFU) from the skin calculated with plate counting (n = 8 to ~10/group). All error bars indicate mean ± SEM. ** $P < 0.01$, *** $P < 0.001$ (t test).



mice in which adipogenesis is prominently impaired (*11*, *15*). Activation of Zfp423 during infection was confirmed by visualizing β-Gal staining on the underside of skin from infected *Zfp423*^{lacZ} reporter mice (Fig. 2A) (*16*). Immunostaining of infected *Zfp423*^{GFP} reporter mice (*17*) showed green fluorescent protein–positive (GFP$^+$) cells localized within infected dermal adipose tissue and mostly colocalized with a fibroblast marker [platelet-derived growth factor receptor–α (PDGFRα)], but not with an endothelial cell marker (CD31) (Fig. 2B and fig. S2, A and B). After *S. aureus* infection, dermal adipose tissue in *Zfp423*^{nur12} mice expanded less than in control mice (fig. S2C). Immunostaining with the adipocyte marker Perilipin (PLIN) further confirmed that adipocyte formation was reduced in the *Zfp423*^{nur12} mice compared with control (fig. S2D). Impaired adipogenesis in *Zfp423*^{nur12} mice was accompanied by increased susceptibility to *S. aureus* skin infection at the site injected with bacteria (Fig. 2, C and D) and a subsequent systemic bacteremia that was not detectable in controls (Fig. 2E). Increased susceptibility to *S. aureus* in *Zfp423*^{nur12} mice was associated with decreased activation of PDGFRα$^+$Sca1$^+$ skin preadipocytes (fig. S2E) (*18*), but not with defective leukocyte recruitment because there was no decrease in CD11b staining or neutrophil infiltration in the skin of *Zfp423*^{nur12} mice (fig. S2F).

To complement the observations made in *Zfp423*^{nur12} mice, we chemically inhibited adipogenesis by pretreating wild-type mice with bisphenol A diglycidyl ether (BADGE) or GW9662, both pharmacological inhibitors of PPARγ and acute chemical inhibitors of adipogenesis (*14*, *19*, *20*). Similarly to *Zfp423*^{nur12} mice, BADGE- or GW9662-treated mice showed decreased adipose expansion after *S. aureus* infection (fig. S3A) and became more susceptible to infection (Fig. 2, F to H, and fig. S3B). In BADGE-treated mice, PDGFRα$^+$Sca1$^+$ preadipocyte numbers appeared normal (fig. S3C), but PPARγ expression decreased (fig. S3D), suggesting as in a previous report (*14*) that inhibition of PPARγ impaired dermal adipose growth by blocking preadipocyte differentiation into adipocytes rather than by inhibiting the preadipocyte commitment process. Increased susceptibility to *S. aureus* in BADGE-treated mice was not caused by defective leukocyte recruitment (fig. S3E) nor by any impaired ability of blood neutrophils to kill *S. aureus* (fig. S3F). Because neutrophil function is a major factor in resisting *S. aureus* skin infection (*6*), this finding confirmed that the increased susceptibility to infection observed with PPARγ inhibitors was caused by their effect on adipocyte formation.

3T3-L1 cells are a well-characterized primary mouse preadipocyte line that can be induced to differentiate into mature adipocytes (*21*). To determine how adipocytes might protect against infection, a comparison of antimicrobial peptide (AMP) gene expression was done in 3T3-L1 undifferentiated preadipocytes (pAds) and differentiated adipocytes (Ads). This showed that mouse cathelicidin antimicrobial peptide (*Camp*) was strongly induced in Ads compared with pAds, whereas other AMPs belonging to the α- and β-defensin families did not show an increase in expression (Fig. 3A). *Camp* mRNA increased at day 1 after differentiation then decreased during later stages of differentiation (fig. S4A). *Camp* expression preceded expression of adipokines Resistin (*Retn*), Leptin (*Lep*), and *Fabp4*, whereas the relative expression of another AMP, *Defb14*, decreased during differentiation (fig. S4, B to H).

Cathelicidin protein peaked at day 2 to ~4 after differentiation and was abundantly detectable with Western blotting of cell extracts from differentiating pAds (fig. S4I). The size of cathelicidin detected by means of Western blot was consistent with the mCAP18 cathelicidin precursor protein (*22*). Cathelicidin was also found with Western blotting in adipocyte-conditioned medium (Fig. 3B). Secretion of cathelicidin protein preceded secretion of Resistin, Leptin, and FABP4 at days 2 to 4 after differentiation (Fig. 3B) and was similar to that of Collagen type-IV (COL4A1), which was produced 1 to 2 days after differentiation

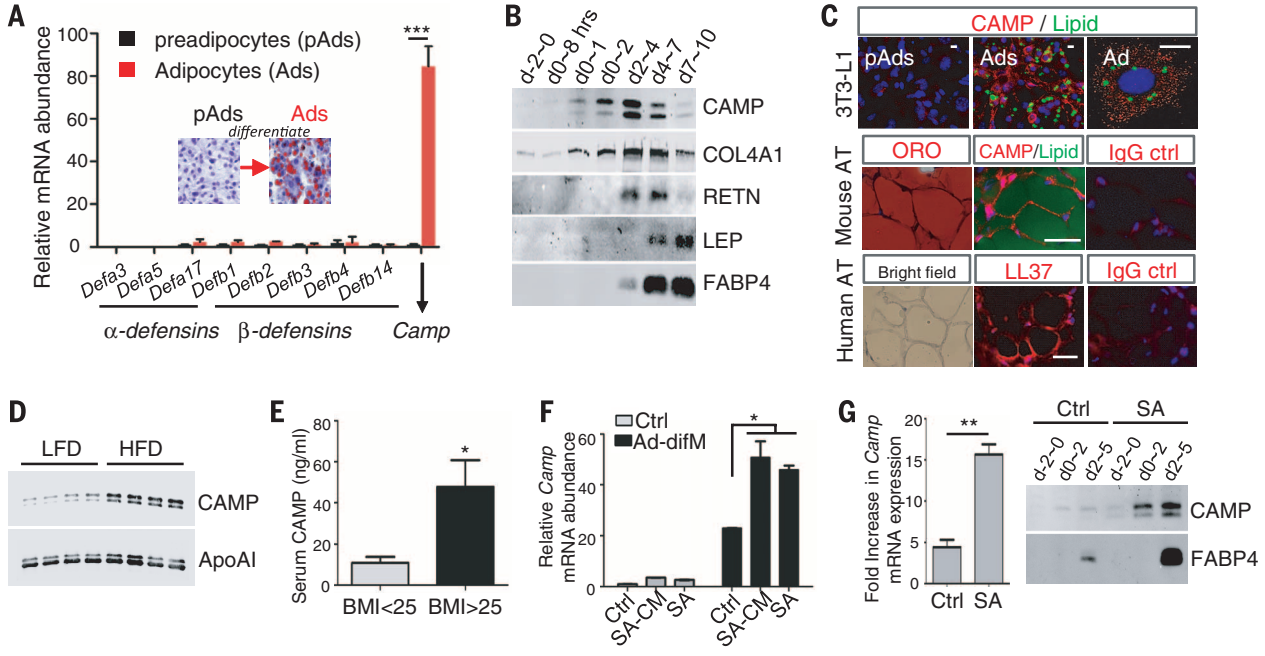


Fig. 3. Adipocytes produce cathelicidin antimicrobial peptide. (**A**) 3T3-L1 preadipocytes (pAds) were differentiated to adipocytes (Ads). Expression of mRNA for α-defensins, β-defensins, and cathelicidin (*Camp*) is shown (n = 3 cultures/group). (Inset) Oil-Red-O staining. (**B**) Western-blotting of conditioned media from differentiating 3T3-L1 cells. (**C**) Staining of 3T3-L1 cells or mouse and human subcutaneous adipose tissue (AT) by using lipid dye (Bodipy; green), mouse CAMP (red), Oil-red-O, or the human cathelicidin peptide LL37 (red) as indicated. Nuclei are counterstained with DAPI. Scale bar, 5 μm (3T3-L1) or 50 μm (adipose tissue). (**D**) C57BL6 mice fed with low-fat diet (LFD) or high-fat diet (HFD) for 4 weeks, and mouse serum was collected for Western blotting analysis by using CAMP and ApoAI antibodies. (**E**) Human serum cathelicidin measured in overweight (/BMI >25) or normal-weight human subjects (n = 10 subjects/group). (**F**) *Camp* mRNA expression levels were examined by quantitative reverse transcription PCR (RT-PCR) in 3T3L1 preadipocytes treated with or without adipocyte differentiation (Ad-difM) medium and with *S. aureus*–conditioned medium (SA-CM) or UV-killed *S. aureus* (SA). (**G**) Primary cultures of dermal cells isolated from Ctrl or *S. aureus*–injected skin were treated with Ad-difM. Relative *Camp* mRNA fold induction at day 2 compared with undifferentiated control were examined by means of quatitative (RT-PCR) (left); CAMP and FABP4 protein were detected in conditioned medium by means of Western blot (right). All error bars indicate mean ± SEM; *P < 0.05, **P < 0.01, ***P < 0.001 (t test).

(*23*). Immunostaining of 3T3-L1 cells showed cathelicidin appearing simultaneously with lipid droplets in the cytoplasm of Ads (Fig. 3C) and localized in the early endosome, but not in the lysosome (fig. S4J).

To verify that cathelicidin was also produced by human adipocytes, we evaluated cultured primary human preadipocytes with reagents directed against human cathelicidin. The *CAMP* gene was strongly induced during differentiation (fig. S5A), and Western blot showed bands consistent with the cathelicidin precursor protein hCAP18 and a shorter peptide (fig. S5B). This peptide appeared larger than LL-37, the mature peptide form of cathelicidin detected from human neutrophils (*24*), thus suggesting adipocytes may have an enzymatic processing system differing from that known in other cell types (*25*).

Cathelicidin mRNA was detected in the subcutaneous fat pad from 1- to ~2-week-old mice (fig. S6, A and B), a stage in which the tissue is undergoing early adipogenesis (*26*). *Camp* mRNA was more abundant in white adipose tissue than brown adipose tissue (fig. S6C). Cathelicidin protein immunostaining was detected in fat tissue from mice and humans (Fig. 3C) and was localized in the early endosome of adipocytes (fig. S6D). In a mouse model of diet-induced obesity, *Camp* mRNA measured with quantitative polymerase chain reaction (PCR) and cathelicidin protein levels estimated by means of Western blot (fig. S7, A and B) were more abundant in the subcutaneous fat pad from mice fed high-fat diet as compared with low-fat diet controls. High-fat diet increased both bone marrow adiposity and cathelicidin protein expression in bone marrow adipocytes (fig. S7C). Serum cathelicidin levels were also elevated in the high-fat diet group (Fig. 3D and fig. S7, D and E), similar to samples from overweight humans [body mass index (BMI) > 25] (Fig. 3E). The level of cathelicidin detected with enzyme-linked immunosorbent assay in serum (~50 ng/ml) was well below the amount of LL-37 necessary to inhibit *S. aureus* growth, which is typically above 2 to 5 mg/ml (*27*). Thus, the high local expression of cathelicidin is more likely to provide antimicrobial function than is the small increase in systemic cathelicidin found in obesity.

Cathelicidin mRNA production during 3T3-L1 differentiation was significantly enhanced when *S. aureus*–conditioned medium or ultraviolet (UV)–killed *S. aureus* was added to adipocyte differentiation medium (Fig. 3F), revealing that *S. aureus* directly enhances AMP production during adipocyte differentiation. Similar to 3T3-L1 cells, dermal cells isolated from *S. aureus*–infected skin, which were shown in Fig. 1E to exhibit high adipogenic potential, produced *Camp* mRNA and abundant cathelicidin protein preceding the production of FABP4 (Fig. 3G).

We next examined cathelicidin production during *S. aureus* infection by immunostaining infected mouse skin. Cathelicidin protein and mRNA were strongly induced in the fat layer at the site of infection (Fig. 4A and fig. S8A). Cathelicidin protein colocalized with expression of the adiponectin reporter β-Gal in AdipoQ-cre;R26R mice (fig. S8B) and COL4A1 (fig. S8, C and D), showing that cathelicidin was expressed in adipocytes in vivo during *S. aureus* infection, correlating with the increase in dermal adipocytes (Fig. 1) and its expression during early adipogenesis (Fig. 3B).

The precursor (~18 kD) and a peptide form (5 to ~10 kD) of cathelicidin were detected in differentiated 3T3-L1 adipocytes (Fig. 4B) and in the cell lysate of mouse adipose tissue (Fig. S9A). Similarly to cultured human adipocytes, the mouse adipocyte-derived cathelicidin peptide was slightly larger than mature cathelicidin-related antimicrobial peptide (CRAMP) derived from mouse neutrophils. We tested conditioned medium from differentiated 3T3-L1 adipocytes and observed that conditioned medium that contained secreted cathelicidin potently inhibited the growth of *S. aureus*, whereas conditioned medium from pAds that lacked cathelicidin did not (Fig. 4C). Brefeldin-A treatment blocked cathelicidin secretion from adipocytes and abolished the antimicrobial activity of the conditioned medium derived from adipocytes (fig. S9B). Furthermore, extracts from wild-type mouse adipose tissue inhibited growth of *S. aureus* (Fig. 4D and fig. S9C) and *Pseudomonas aeruginosa* (Fig. S9D), but adipose tissue extracts from *Camp*$^{-/-}$ mice did not possess antimicrobial properties. *Camp*$^{-/-}$ mice did not differ from wild-type in their capacity to develop metabolic syndrome after 20 weeks of high-fat diet (fig. S10).

Mice in which adipose expansion was inhibited showed reduced defense against *S. aureus* infection, and this correlated with reduced cathelicidin. BADGE treatment of 3T3-L1 during differentiation inhibited *Pparg* and *Cebpa* mRNA and blocked *Camp* (fig. S11A). The expression of cathelicidin was less in mice treated with BADGE (fig. S11B) and in *Zfp423*^{nur12} mice (Fig. 4, E and F), although the expression of cathelicidin seen in cells recruited to the site of infection was maintained. *Zfp423*^{nur12} did not inhibit neutrophil recruitment as seen by a similar increase in abundance of the neutrophil marker (*Mpo*) after infection (fig. S11D). As expected, *Camp*$^{-/-}$ mice showed increased susceptibility to infection (*3*). However, *Camp*$^{-/-}$ mice treated with BADGE or GW9662 showed

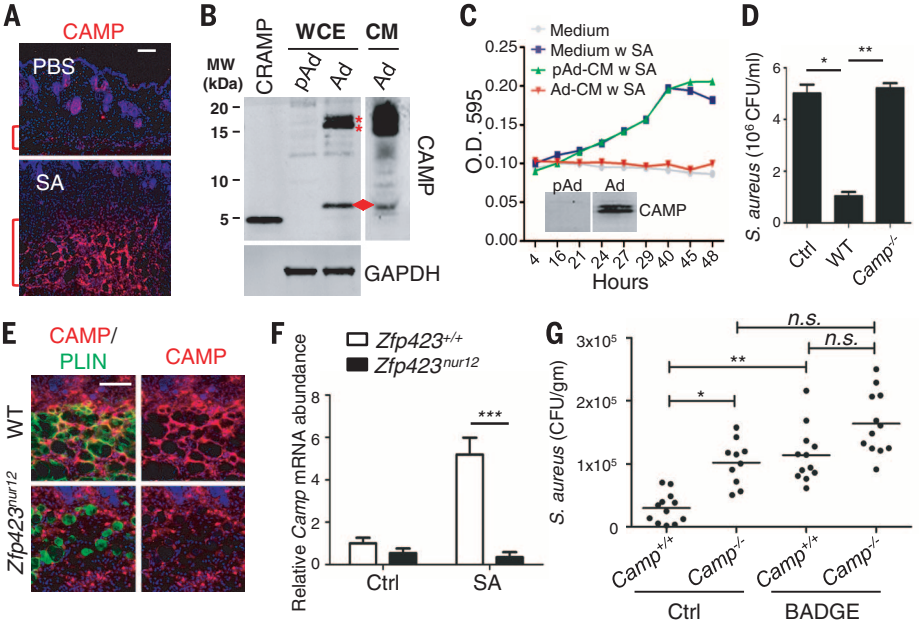


Fig. 4. Adipocytes kill bacteria by producing cathelicidin. (**A**) CAMP immunostaining (left) of adipose in mice infected with *S. aureus*. Scale bar, 100 μm. (**B**) Western-blotting of CAMP from whole-cell extracts (WCE) or conditioned medium (CM) from 3T3-L1 pAds or Ads. Red arrow indicates mature peptide form, and red asterisks indicate precursor forms of cathelicidin. Synthetic CRAMP peptide is shown as standard for the processed form of CAMP in neutrophils. (**C**) Growth curve of *S. aureus* in Ctrl medium or conditioned medium (CM) from pAd or Ad as indicated. (Inset) CAMP immunoblot in conditioned medium. (**D**) *S. aureus* growth in media alone (Ctrl) or with addition of protein extracted from subcutaneous adipose tissue from *Camp*$^{+/+}$ or *Camp*$^{-/-}$ mice (n = 3). (**E**) CAMP expression is suppressed in adipose in *Zfp423*-deficient mice. Scale bars, 100 μm. (**F**) *Camp* mRNA expression in *S. aureus*– or PBS-injected *Zfp423*$^{+/+}$ and *Zfp423*^{nur12} mice (n = 5 mice/group). (**G**) *Camp*$^{+/+}$ or *Camp*$^{-/-}$ mice were pretreated with BADGE or dimethyl sulfoxide (ctrl) before infecting with *S. aureus* and *S. aureus* survival in skin measured 3 days later (n = 7 to ~10 mice/group). All error bars indicate mean ± SEM. *P < 0.05, **P < 0.01, ***P < 0.001 (one-way analysis of variance). Red brackets indicate subcutaneous adipose layer.

no further increase in infection (Fig. 4G and fig. S12), suggesting that the effects of BADGE on cathelicidin production from adipocytes was responsible for the increase in *S. aureus* infection in vivo.

These results show that a local increase in subcutaneous adipocytes is an important host defense response against skin infection. This observation is consistent with prior observations that adipocytes secrete a variety of bioactive adipokines and cytokines that mediate immune responses after injury (*28*) and now shows that adipocytes produce an AMP that can directly kill bacteria. Local expansion of dermal fat produces cathelicidin in response to infection, but this response appears to decline as adipocytes mature. The expansion of dermal fat in response to infection may also indirectly benefit immune defense by influencing other processes such as neutrophil oxidative burst, thus further amplifying the importance of the subcutaneous preadipocyte pool in preventing infections. Defective AMP production by mature adipocytes may explain observations of elevated susceptibility to infection during obesity and insulin resistance (*29*). Cathelicidin has also been shown to be proinflammatory (*30*). Therefore, the production of cathelicidin by adipocytes may also participate in the chronic, low-level inflammation observed in obesity (*28*).

REFERENCES AND NOTES

1. F. O. Nestle, P. Di Meglio, J. Z. Qin, B. J. Nickoloff, *Nat. Rev. Immunol.* **9**, 679–691 (2009).
2. R. L. Gallo, L. V. Hooper, *Nat. Rev. Immunol.* **12**, 503–516 (2012).
3. V. Nizet *et al.*, *Nature* **414**, 454–457 (2001).
4. P. Y. Ong *et al.*, *N. Engl. J. Med.* **347**, 1151–1160 (2002).
5. K. Pütsep, G. Carlsson, H. G. Boman, M. Andersson, *Lancet* **360**, 1144–1149 (2002).
6. L. S. Miller, J. S. Cho, *Nat. Rev. Immunol.* **11**, 505–518 (2011).
7. J. M. Musser *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 225–229 (1990).
8. J. Eguchi *et al.*, *Cell Metab.* **13**, 249–259 (2011).
9. H. S. Sul, *Mol. Endocrinol.* **23**, 1717–1725 (2009).
10. C. S. Hudak *et al.*, *Cell Reports* **8**, 678–687 (2014).
11. R. K. Gupta *et al.*, *Nature* **464**, 619–623 (2010).
12. T. Tanaka, N. Yoshida, T. Kishimoto, S. Akira, *EMBO J.* **16**, 7432–7443 (1997).
13. M. I. Lefterova *et al.*, *Genes Dev.* **22**, 2941–2952 (2008).
14. E. Festa *et al.*, *Cell* **146**, 761–771 (2011).
15. W. A. Alcaraz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19424–19429 (2006).
16. L. E. Cheng, J. Zhang, R. R. Reed, *Dev. Biol.* **307**, 43–52 (2007).
17. R. K. Gupta *et al.*, *Cell Metab.* **15**, 230–239 (2012).
18. R. R. Driskell *et al.*, *Nature* **504**, 277–281 (2013).
19. O. Naveiras *et al.*, *Nature* **460**, 259–263 (2009).
20. B. A. Schmidt, V. Horsley, *Development* **140**, 1517–1527 (2013).
21. H. Green, O. Kehinde, *Cell* **5**, 19–27 (1975).
22. R. A. Dorschner *et al.*, *J. Invest. Dermatol.* **117**, 91–97 (2001).
23. T. Sillat *et al.*, *J. Cell. Mol. Med.* **16**, 1485–1495 (2012).
24. O. E. Sørensen *et al.*, *Blood* **97**, 3951–3959 (2001).
25. O. E. Sørensen *et al.*, *J. Biol. Chem.* **278**, 28540–28546 (2003).
26. N. Kawaguchi *et al.*, *Am. J. Pathol.* **160**, 1895–1903 (2002).
27. H. G. Boman, *J. Intern. Med.* **254**, 197–215 (2003).
28. A. Schäffler, J. Schölmerich, *Trends Immunol.* **31**, 228–235 (2010).
29. W. P. Cawthorn, E. L. Scheller, O. A. MacDougald, *J. Lipid Res.* **53**, 227–246 (2012).
30. K. Yamasaki *et al.*, *Nat. Med.* **13**, 975–980 (2007).

ACKNOWLEDGMENTS

This study was supported by NIH R01 AI083358, R01AI052453, and AR052728 (R.L.G.). Human serum collections were funded by The Atopic Dermatitis Research Network (HHSN272201000020C). S.P.B. was supported by the NIH (DK096828). M.V.P. is supported by the Edward Mallinckrodt Jr. Foundation Research Grant, the Dermatology Foundation Research Grant, and NIH NIAMS R01-AR067273; C.F.G.-J. is supported by the NIH MBRS-IMSD training grant (GM055246) and NSF Graduate Research Fellowship number DGE-1321846; and R.R. is supported by California Institute for Regenerative Medicine training grant (TG2-01152). The authors declare no competing financial interests. We thank T. Nakatuji for advice relating to bacterial infections and C. Aguilera for mouse technical expertise, UCSD Bio-Core facility for reagents, UCSD mouse hematology core laboratory for serum studies, and J. M. Olefsky for comments on the manuscript. Zfp423-GFP mice are available from R. Gupta under a material transfer agreement with Dana-Farber Cancer Institute. All the data reported in this manuscript are presented in the main paper and in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/67/suppl/DC1
Materials and Methods
Figs. S1 to S12
Table S1
References (*31*–*37*)

8 September 2014; accepted 21 November 2014
10.1126/science.1260972

VIRUS STRUCTURE

Structure and inhibition of EV-D68, a virus that causes respiratory illness in children

Yue Liu,[1] Ju Sheng,[1] Andrei Fokine,[1] Geng Meng,[1] Woong-Hee Shin,[1] Feng Long,[1] Richard J. Kuhn,[1] Daisuke Kihara,[1,2] Michael G. Rossmann[1]*

Enterovirus D68 (EV-D68) is a member of *Picornaviridae* and is a causative agent of recent outbreaks of respiratory illness in children in the United States. We report here the crystal structures of EV-D68 and its complex with pleconaril, a capsid-binding compound that had been developed as an anti-rhinovirus drug. The hydrophobic drug-binding pocket in viral protein 1 contained density that is consistent with a fatty acid of about 10 carbon atoms. This density could be displaced by pleconaril. We also showed that pleconaril inhibits EV-D68 at a half-maximal effective concentration of 430 nanomolar and might, therefore, be a possible drug candidate to alleviate EV-D68 outbreaks.

Picornaviruses constitute a large family of small icosahedral viruses with a single positive-stranded RNA genome and an external diameter of about 300 Å. The *Enterovirus* genus includes medically important human pathogens, such as human rhinoviruses (HRVs), polioviruses (PVs) and coxsackieviruses (CVs) (table S1) (*1*, *2*). Many of these enteroviruses (EVs) have been well characterized structurally and functionally (*3*–*10*). However, the species EV-D remains poorly characterized.

An upsurge of EV-D68 cases in the past few years has shown clusters of infections worldwide (*11*). In August 2014, an outbreak of mild to severe respiratory illnesses occurred among thousands of young children in the United States, of which 1116 cases have been confirmed to be caused by EV-D68. This virus has also been associated with occasional neurological infections (*12*). Although EV-D68 has emerged as a considerable global public health threat, there is no available vaccine or effective antiviral treatment.

The capsids of EVs consist of 60 copies of each of four different viral proteins: VP1, VP2, VP3, and VP4 (Fig. 1A). Of these, VP1 (about 300 amino acids), VP2 (about 260 amino acids), and VP3 (about 240 amino acids) each have a "jelly roll" fold arranged in the capsid with pseudo $T = 3$ icosahedral symmetry, where T represents the triangulation number (*13*). Their organization in the capsid is similar to that of the $T = 3$ RNA plant viruses, except that the three subunits related by quasi-threefold symmetry have different amino acid sequences in picornaviruses (*6*, *10*). Each roughly 70–amino–acids-long VP4 molecule forms an extended peptide on the internal surface of the capsid shell. The jelly roll fold consists of two antiparallel β sheets, which face each other to form a β barrel with a hydrophobic interior (Fig. 1B).

EVs have a deep surface depression ("canyon") circulating around each of the 12 pentameric vertices (Fig. 1A). The canyon was predicted to be the site of receptor binding, because the amino acids outside the canyon that form the external surface of the virus were more exposed and were shown to be accessible to neutralizing antibodies (Fig. 1A) (*10*, *14*). The virus thus could remain faithful to a specific receptor molecule that binds into the canyon while evading the host's immune system (*10*, *15*). The prediction that the canyon would be the site of binding to cellular receptor molecules was subsequently confirmed for numerous different EVs (*16*, *17*). All of the receptor

[1]Department of Biological Sciences, Hockmeyer Hall of Structural Biology, 240 South Martin Jischke Drive, Purdue University, West Lafayette, IN 47907, USA. [2]Department of Computer Science, 305 North University Street, Purdue University, West Lafayette, IN 47907, USA.
***Corresponding author. E-mail: mr@purdue.edu**

molecules were found to have an immunoglobulin-like structure.

A variety of small ~350-dalton hydrophobic molecules that are inhibitors of EV infection were shown to bind into the hydrophobic pocket in the center of the VP1 jelly roll (*18*) (Fig. 1B). They stabilize the virus (*19*) by filling the pocket with a well-fitting hydrophobic molecule, thereby inhibiting uncoating of the virus and the release of the genome into the infected cell. These compounds also prevent viral attachment to cells by altering the surface features of the canyon floor, where the virus attaches to a cellular receptor (*17*, *19*). The floor of the canyon is formed primarily by the GH loop of VP1, the connecting residues between β-strands G and H (*18*).

Most infectious EVs contain a small molecule or "pocket factor," probably a fatty acid in the VP1 binding pocket (*7*, *18*, *20*). Like the capsid-binding compounds, the pocket factor presumably stabilizes the virus by filling the VP1 hydrophobic pocket. Thus, the virus is stabilized while being transmitted to a new host. However, when a receptor molecule binds to the floor of the canyon, it depresses the floor (that also forms the roof of the VP1 binding pocket), which then squeezes the binding pocket, probably expelling the pocket factor (*17*). Thus, the attachment of the virus to a cell surface initiates uncoating, causing the release of the viral genome into the cell's cytosol.

In most EVs that have been investigated, these capsid-binding antiviral compounds displace the pocket factor (e,g. polioviruses, coxsackieviruses A or B, and many HRVs). The rhinoviruses HRV14 (*18*) and HRV3 (*21*) do not contain a pocket factor, and the GH loop is displaced relative to its position in other EVs that contain a pocket factor, reducing the volume of the VP1 pocket. Either neither HRV14 nor HRV3 binds a pocket factor in vivo or it was lost during the purification procedure. Much effort was made between 1985 and 2000 to design a compound that fits into the VP1 pocket and would inhibit the maximum number of rhinovirus serotypes (*22*). The final optimal structure was pleconaril (fig. S1), which not only had good efficacy but was also stable enough to maintain good bioavailability in clinical tests (*23*). However, pleconaril was not licensed, primarily because it put women using birth control drugs at risk of conception.

The EV-D68 prototype strain Fermon CA62-1 was propagated in human rhabdomyosarcoma cells at 33°C, which had previously been shown to be the optimal growth temperature for EV-D68, indicating that EV-D68 behaves much like members of the RV-A and RV-B species that are responsible for common colds (*24*). The virus was purified and crystalized as described in the supplementary materials and methods. The cubic-looking crystals had a diameter between 0.1 and 0.2 mm. After soaking in glycerol, they were flash-frozen in liquid nitrogen. X-ray diffraction data were collected at sector 14 of the Advanced Photon Source. The data extended to 2.0 Å resolution and were processed with HKL2000 (*25*). The crystal symmetry was I222, with two particles per unit cell, implying that the particles were located on a 222 symmetry position. A rotation function (*26*) differentiated between the two orthogonal possible orientations of the icosahedron. Initial phases were calculated based on the structure of HRV2 (*27*) (Protein Data Bank accession number 1FPN) after the pocket factor was removed. The phases were then extended in small steps to 2.0 Å resolution, using 15-fold averaging and solvent flattening. A model of the structure was built using the program Coot (*28*) and refined with the program CNS (*29*). The final value of R_{work} was 27.5%. R_{free} is essentially the same as R_{work} in the presence of high noncrystallographic symmetry redundancy as is the case here (table S2).

Comparison of the amino acid sequences of EV-D68 with those of other EVs shows that VP3 has a short C-terminal α helix not present in other EVs (fig. S2). The EV-D68 electron density map showed that this helix decorates the north side of the canyon in the neighboring, fivefold-related, icosahedral asymmetric unit. As a result, the canyon is narrower than in other EVs and might therefore not be large enough to accommodate immunoglobulinlike receptors (Fig. 1C).

The BC and DE loops of VP1 are structurally the most variable among known picornaviruses. The EV-D68 VP1 has two disordered regions corresponding to residues 80 to 86 (EV-D68 numbering) in the BC loop and 129 to 136 in the DE loop, both of which are near fivefold axes. These regions harbor the neutralizing immunogen sites NIm-IA and NIm-IB on HRV14, respectively (*10*) (Fig. 1A). Thus the flexible immunogenic regions around the fivefold axes might be an alternative mechanism for evading host humoral immune responses.

The electron density map of EV-D68 showed density in the VP1 pocket (Fig. 2A). The height of this pocket factor density was about three standard deviations above the mean of the noncrystallographic symmetry averaged map, as compared with about five standard deviations of most of the main-chain density. As in other EVs that have a pocket factor, the conformation of the GH loop of VP1, which defines the interface between the VP1 pocket and the floor of the canyon, is pushed into the canyon relative to the empty pocket in HRV14 (*10*, *18*) or HRV3 (*21*).

The length of the pocket factor density in the EV-D68 map corresponded to a fatty acid with an aliphatic chain of about 10 carbon atoms (Fig. 2B). Similarly, 12–carbon-atom-long pocket factors were observed for HRV16 (*9*) and HRV2 (*27*). In contrast, the well-formed pocket factors in poliovirus 1 (*6*, *7*), coxsackievirus B3 (*5*), and EV-A71 (*3*) corresponded to longer fatty acids with 18, 16, and 18 carbon atoms, respectively. Furthermore, the orientation of the pocket factor

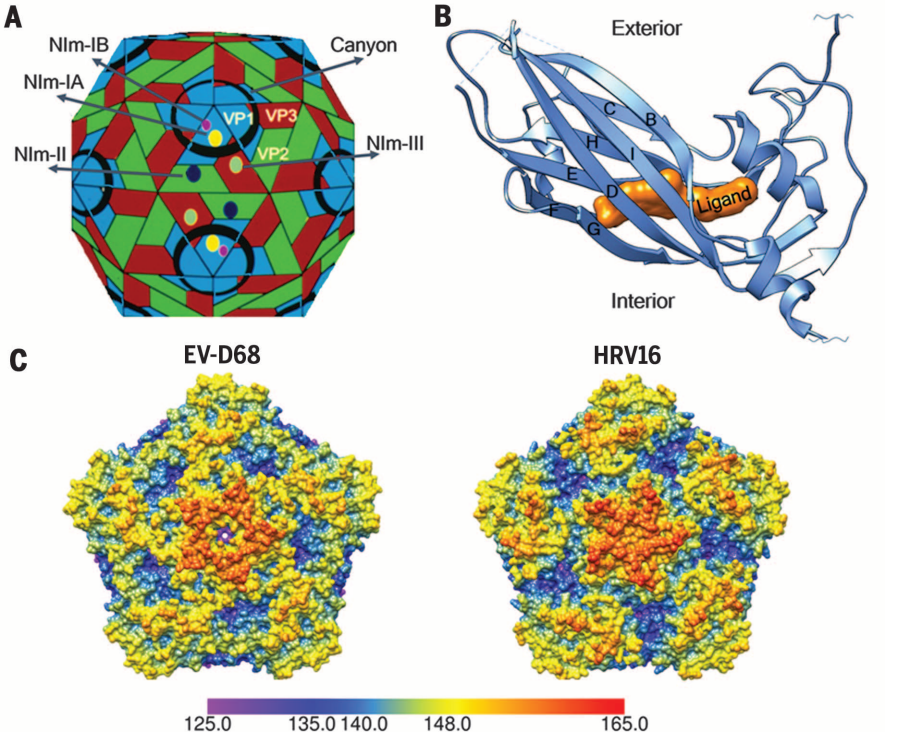


Fig. 1. The structure of EV-D68. (**A**) Diagrammatic representation of the virus. VP1, VP2, and VP3 are colored blue, green, and red, respectively. Each pentagonal icosahedral vertex is surrounded by a canyon colored black. The epitopes for neutralizing antibodies for the homologous HRV14 are labeled as NIm-I, NIm-II, and NIm-III. The NIm-IA and NIm-IB sites are disordered in EV-D68. (**B**) The VP1 jelly roll in EV-D68 shown as a ribbon diagram. If the β strands along the polypeptide are identified sequentially as A, B, C,….I, then one of the sheets is composed of the antiparallel strands BIDG and the other by the antiparallel strands CHEF. (**C**) Surface features of EV-D68 compared to HRV16. The canyon of EV-D68 is shallower and narrower than the canyon of HRV16.

tails in EV-D68, HRV16, and HRV2 is slightly different than that of the long pocket factors of EV-A71, poliovirus 1, and coxsackievirus B3 (Fig. 2C). The orientation is controlled in part by VP1 residue 184, which is Leu or Ile for the viruses with small, short pocket factors and Val for viruses with longer pocket factors (fig. S2). The larger Leu or Ile residue pushes the pocket factor sideways. The similarity between the properties of the pocket factor in EV-D68 and that in the HRVs might partially explain why these viruses are less stable.

The anti–EV-D68 activity of two capsid-binding compounds, pirodavir and BTA-188 (fig. S1), that had significant anti-rhinovirus activity was compared with pleconaril, using plaque reduction assays in HeLa cells (Fig. 3, A and B). The half-maximal effective concentration (EC_{50}) value of these two compounds was found to be comparable to previous results using cytopathic effect inhibition assays against EV-D68 (*30*). However, pleconaril was found to be more potent against EV-D68 than pirodavir and BTA-188 (Fig. 3A). The inhibitory effect of pleconaril is similar against EV-D68, HRV16, and HRV14, but better than against EV-A71. It is therefore noteworthy that pleconaril was an effective inhibitor in extensive clinical tests for the treatment of common colds (*23*, *31*). Furthermore, fluorescence-based thermal stability assays indicated that when EV-D68 was incubated with pleconaril at either 10 or 50 μg/ml, 4°C higher temperatures were required to release the RNA genome than when no pleconaril was present (Fig. 3C and materials and methods in the supplementary materials). Thus, pleconaril stabilizes EV-D68 capsids, preventing the virus from uncoating during viral entry.

The structure of EV-D68 was also determined to 2.3 Å resolution when co-crystalized with pleconaril (fig. S3). The crystallographic procedure was the same as for the native structure determination (materials and methods in the supplementary materials). The electron density of the ligand inside the hydrophobic pocket was of the same height but much longer than that of the pocket factor in the native EV-D68 structure (Fig. 4A) and could be easily fitted with the structure of pleconaril (Fig. 4B). This demonstrated that the native pocket factor was replaced by pleconaril. However, the densities of the ligand and of the pocket factor are slightly lower than that of the polypeptide main chain. In part this is to be expected, because the ligand or fatty acid pocket factor is composed primarily of only carbon atoms as opposed to the heavier combination of carbon and nitrogen plus oxygen atoms of the main chain. In part, the slightly lower density of the ligand might be due to some variations of conformation within the VP1 binding pocket.

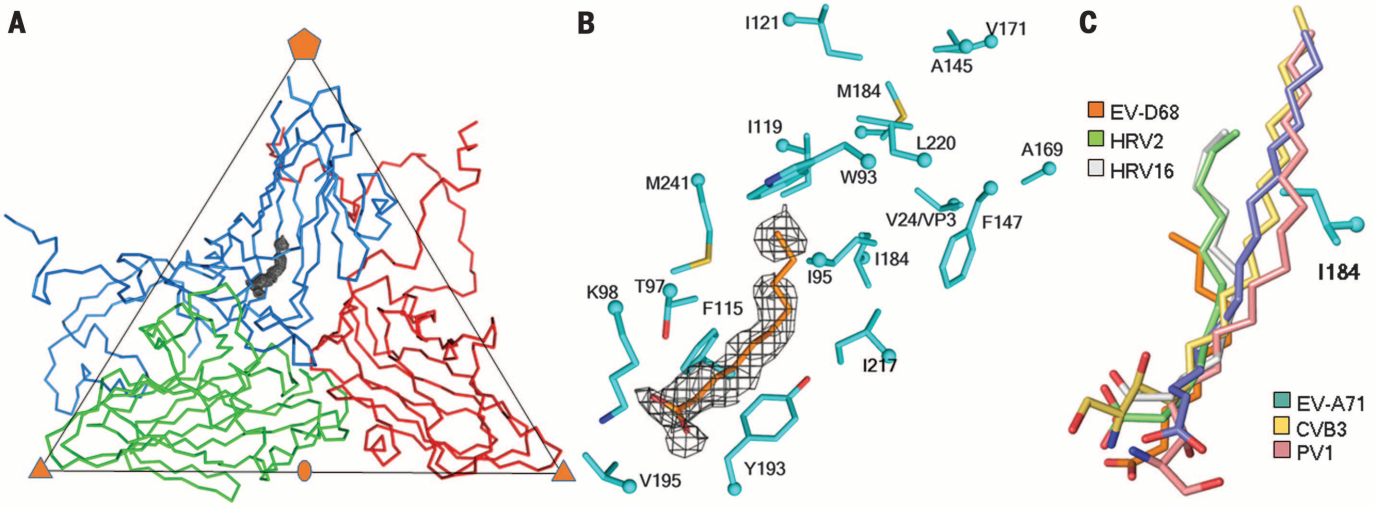


Fig. 2. The pocket factor. (**A**) One icosahedral asymmetric unit of the EV-D68 structure, showing the Cα atom backbone for VP1, VP2, and VP3 in marine blue, green, and red, respectively. The pocket factor electron density outline is shown in gray. (**B**) Enlargement of the pocket factor density with a fitted putative C10 fatty acid. Shown also are the amino acids lining the pocket. (**C**) Comparison of the putative pocket factor structures in six known EV structures. Shown also is the residue I184 in EV-D68 that alters the direction of the shorter pocket factors in some of the viruses as compared to others with longer pocket factors and a smaller residue at the equivalent position of 184. Atoms at the head of the pocket factors and of the VP1 residue side chains are colored red (oxygen), dark blue (nitrogen), and yellow (sulfur).

A

Anti-enterovirus activity of three compounds

Virus	EC_{50} (μM)		
	Pleconaril	Pirodavir	BTA-188
EV-D68	0.43±0.02*	1.61±0.30*	>4.93*
EV-D68	-†	2.44±0.54‡	3.34±0.08‡
HRV16	0.57§	0.016±0.034‡	0.008±0.002‡
HRV14	0.16§	0.027±0.043‡	0.068±0.033‡
EV-A71	>262‖	5.42±0.49‡	0.082±0.139‡

*Data were reported in this work
†No reported data were available
‡According to Barnard et al. (2004) *(30)*
§According to Ledford et al. (2004) *(32)*
‖According to Tijsma et al. (2014) *(33)*

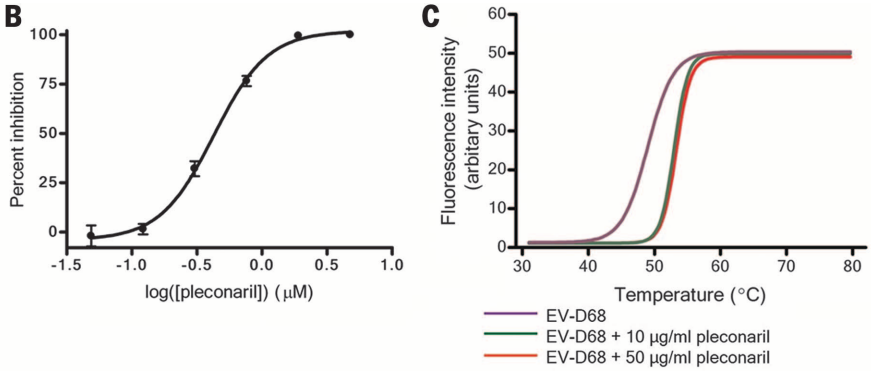


Fig. 3. Anti-EV activity of three capsid-binding compounds. (**A**) List of EC_{50} values (*30*, *32*, *33*). Each experiment was performed at least three times. (**B**) Plot of the percent of plaque reduction of pleconaril as a function of the log of its concentration. Error bars indicate standard deviations. (**C**) Release of EV-D68 genome upon increase in temperature as monitored by Sybr green II. The experiments were done in triplicate. For each data point, the ratio between mean fluorescence intensity and standard deviation is at least 7.6. Shown are curves fitted with a sigmoidal function for the native virus and for the virus after incubation with pleconaril.

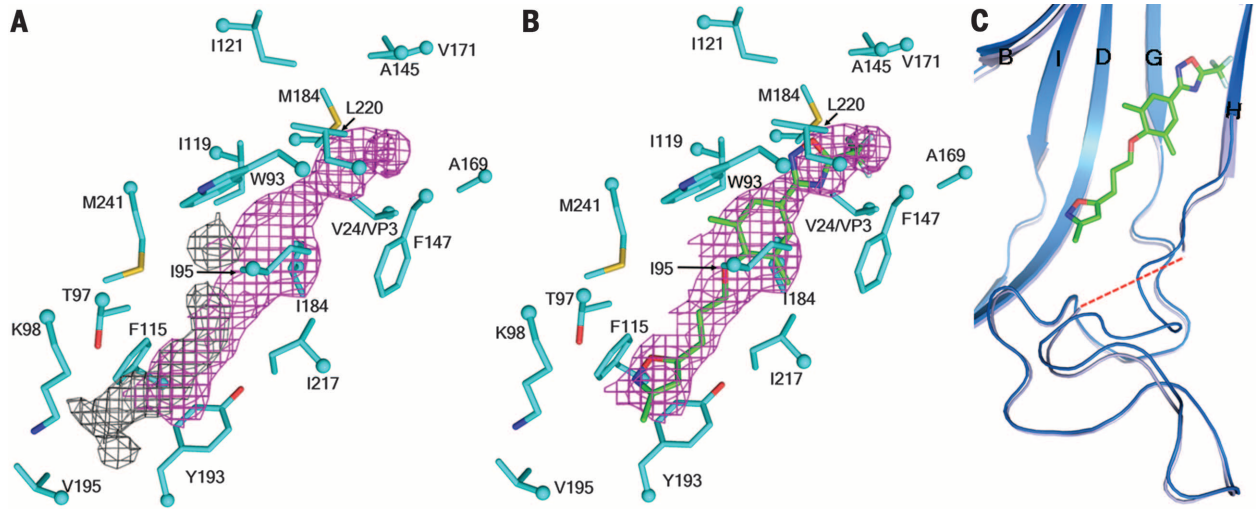


Fig. 4. Structure of pleconaril bound into the VP1 pocket of EV-D68. (**A**) Pocket factor density (gray) compared to the pleconaril density (magenta). (**B**) Pleconaril (green) fitted to density in the structure of the complex. (**C**) Conformational change of the VP1 GH loop as a consequence of the presence of pleconaril. The native and complex structures are shown in marine blue and baby blue, respectively. Oxygen, nitrogen, sulfur, and fluorine atoms are shown in red, dark blue, yellow, and light green, respectively.

Part of the VP1 GH loop (residues 212 to 215) that forms the entrance to the VP1 pocket had become less ordered in the EV-D68-pleconaril complex. The Cα atom of residue 211 had moved 1.2 Å toward the inside of the pocket relative to that of the native structure, possibly blocking the entrance to the pocket once pleconaril had entered (Fig. 4C). Thus, the dynamics of the GH loop might be a consideration for future structure-based design of EV-D68 capsid-binding inhibitors.

A comparison of the EV-D68–pleconaril, HRV14-pleconaril, and HRV16-pleconaril structures showed a similar binding mode for pleconaril in the VP1 pocket of these three viruses (fig. S4 and table S3). This may explain why pleconaril is similarly effective against these three EVs. To investigate why pleconaril is more effective against EV-D68 than pirodavir or BTA-188, we performed in silico docking experiments. The presence of a trifluoromethyl-substituted oxadiazole moiety in pleconaril, rather than a more hydrophilic group in either pirodavir (ethyl carboxylate group) or BTA-188 (O-ethyloxime group) at structurally equivalent positions (fig. S5), probably contributes to more favorable interactions of pleconaril with the hydrophobic residues deep inside the VP1 pocket of EV-D68.

Our results show that the structure of EV-D68 has considerable similarities to those of the well-studied HRVs for which pleconaril was specifically designed. We also show that pleconaril replaces the pocket factor and is a potent inhibitor of EV-D68, with an EC_{50} value of 430 nM. The size and location of the pocket factor lodged in the VP1 pocket are similar to those found in other HRVs and different from those of the pocket factors found in poliovirus 1 and EV-A71. This correlates with the observation that pleconaril is far more active when the natural pocket factor is short, as in the HRVs and in EV-D68. Furthermore, sequence alignment of 188 EV-D68 strains found between 1962 and 2013 indicates that residues in VP1 that interact with pleconaril, as identified from the complex structure, are completely conserved, with one exception. Therefore, pleconaril is likely to inhibit not only the prototype strain examined here but also many other strains. In view of the previous extensive clinical trials that have established its safety, pleconaril would be a possible drug candidate to alleviate EV-D68 outbreaks.

REFERENCES AND NOTES

1. M. J. Adams, A. M. King, E. B. Carstens, *Arch. Virol.* **158**, 2023–2030 (2013).
2. N. J. Knowles *et al.*, in *Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses*, A. M. Q. King, M. J. Adams, E. B. Carstens, E. J. Lefkowitz, Eds. (Elsevier, San Diego, CA, 2012), pp. 855–880.
3. P. Plevka, R. Perera, J. Cardosa, R. J. Kuhn, M. G. Rossmann, *Science* **336**, 1274 (2012).
4. X. Wang *et al.*, *Nat. Struct. Mol. Biol.* **19**, 424–429 (2012).
5. J. K. Muckelbauer *et al.*, *Structure* **3**, 653–667 (1995).
6. J. M. Hogle, M. Chow, D. J. Filman, *Science* **229**, 1358–1365 (1985).
7. D. J. Filman *et al.*, *EMBO J.* **8**, 1567–1579 (1989).
8. K. N. Lentz *et al.*, *Structure* **5**, 961–978 (1997).
9. M. A. Oliveira *et al.*, *Structure* **1**, 51–68 (1993).
10. M. G. Rossmann *et al.*, *Nature* **317**, 145–153 (1985).
11. R. Tokarz *et al.*, *J. Gen. Virol.* **93**, 1952–1958 (2012).
12. J. D. Kreuter *et al.*, *Arch. Pathol. Lab. Med.* **135**, 793–796 (2011).
13. D. L. Caspar, A. Klug, *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24 (1962).
14. B. Sherry, A. G. Mosser, R. J. Colonno, R. R. Rueckert, *J. Virol.* **57**, 246–257 (1986).
15. M. G. Rossmann, *J. Biol. Chem.* **264**, 14587–14590 (1989).
16. N. H. Olson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 507–511 (1993).
17. M. G. Rossmann, Y. He, R. J. Kuhn, *Trends Microbiol.* **10**, 324–331 (2002).
18. T. J. Smith *et al.*, *Science* **233**, 1286–1293 (1986).
19. M. G. Rossmann, *Protein Sci.* **3**, 1712–1725 (1994).
20. M. Smyth, T. Pettitt, A. Symonds, J. Martin, *Arch. Virol.* **148**, 1225–1233 (2003).
21. R. Zhao *et al.*, *Structure* **4**, 1205–1220 (1996).
22. J. M. Rogers, G. D. Diana, M. A. McKinlay, *Adv. Exp. Med. Biol.* **458**, 69–76 (1999).
23. F. G. Hayden *et al.*, *Clin. Infect. Dis.* **36**, 1523–1532 (2003).
24. M. S. Oberste *et al.*, *J. Gen. Virol.* **85**, 2577–2584 (2004).
25. Z. Otwinowski, W. Minor, *Methods Enzymol.* **276**, 307–326 (1997).
26. L. Tong, M. G. Rossmann, *Methods Enzymol.* **276**, 594–611 (1997).
27. N. Verdaguer, D. Blaas, I. Fita, *J. Mol. Biol.* **300**, 1179–1194 (2000).
28. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
29. A. T. Brünger *et al.*, *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
30. D. L. Barnard *et al.*, *Antimicrob. Agents Chemother.* **48**, 1766–1772 (2004).
31. D. C. Pevear *et al.*, *Antimicrob. Agents Chemother.* **49**, 4492–4499 (2005).
32. R. M. Ledford *et al.*, *J. Virol.* **78**, 3663–3674 (2004).
33. A. Tijsma *et al.*, *Antimicrob. Agents Chemother.* **58**, 6990–6992 (2014).

ACKNOWLEDGMENTS

We thank M. Steven Oberste of the Centers for Disease Control and Prevention for supplying us with the prototype strain of EV-D68; M. A. McKinlay of the Task Force for Global Health for helpful discussion and suggestions; S. Kelly for help with the manuscript preparation; and V. Srajer, R. Henning, and the other staff of the Advanced Photon Source BioCARS beamline 14 for help with x-ray diffraction data collection. Use of BioCARS sector 14 was supported by the National Institutes of Health, National Center for Research Resources (NIH/NCRR) grant RR007707. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract DE-AC02-006CH11357. This study was supported by NIH grant award AI11219 to M.G.R. Coordinates for native EV-D68 and EV-D68–pleconaril structures were deposited with the Protein Data Bank with accession numbers 4WM8 and 4WM7, respectively.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/71/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S3
References (*34*–*49*)

1 October 2014; accepted 25 November 2014
10.1126/science.1261962

PROTEIN SYNTHESIS

Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains

Peter S. Shen,[1] Joseph Park,[2] Yidan Qin,[8,9] Xueming Li,[7]* Krishna Parsawar,[10] Matthew H. Larson,[3,4,5,6] James Cox,[1,10] Yifan Cheng,[7] Alan M. Lambowitz,[8,9] Jonathan S. Weissman,[3,4,5,6]† Onn Brandman,[2]† Adam Frost[1,7]†

In Eukarya, stalled translation induces 40S dissociation and recruitment of the ribosome quality control complex (RQC) to the 60S subunit, which mediates nascent chain degradation. Here we report cryo–electron microscopy structures revealing that the RQC components Rqc2p (YPL009C/Tae2) and Ltn1p (YMR247C/Rkr1) bind to the 60S subunit at sites exposed after 40S dissociation, placing the Ltn1p RING (Really Interesting New Gene) domain near the exit channel and Rqc2p over the P-site transfer RNA (tRNA). We further demonstrate that Rqc2p recruits alanine- and threonine-charged tRNA to the A site and directs the elongation of nascent chains independently of mRNA or 40S subunits. Our work uncovers an unexpected mechanism of protein synthesis, in which a protein—not an mRNA—determines tRNA recruitment and the tagging of nascent chains with carboxy-terminal Ala and Thr extensions ("CAT tails").

Despite the processivity of protein synthesis, faulty messages or defective ribosomes can result in translational stalling and incomplete nascent chains. In Eukarya, this leads to recruitment of the ribosome quality control complex (RQC), which mediates the ubiquitylation and degradation of incompletely synthesized nascent chains (*1*–*4*). The molecular components of the RQC include the AAA adenosine triphosphatase Cdc48p and its ubiquitin-binding cofactors, the RING-domain E3 ligase Ltn1p, and two proteins of unknown function, Rqc1p and Rqc2p. We set out to determine the mechanism(s) by which relatively rare (*5*) proteins such as Ltn1p, Rqc1p, and Rqc2p recognize and rescue stalled 60*S* ribosome nascent chain complexes, which are vastly outnumbered by ribosomes translating normally or in stages of assembly.

To reduce structural heterogeneity and enrich for complexes still occupied by stalled nascent chains, we immunoprecipitated Rqc1p-bound RQC assemblies from *Saccharomyces cerevisiae* strains lacking the C-terminal RING domain of Ltn1p, which prevents substrate ubiquitylation and Cdc48 recruitment (*1*). Three-dimensional (3D) classification of Ltn1ΔRING particles revealed 60*S* ribosomes with nascent chains in the exit tunnel and extraribosomal densities (Fig. 1). These extraribosomal features were resolved between 5 and 14Å and proved to be either Tif6p or RQC components as characterized below (Fig. 1 and figs. S1 to S7). Tif6p was not observed bound to the same 60*S* particles bound by RQC factors (figs. S1 to S3). We repeated the purification, imaging, and 3D classification from *rqc2*Δ cells and computed difference maps. This analysis did not reveal density attributable to Rqc1p but did identify Rqc2p as a transfer RNA (tRNA)–binding protein that occupies the 40*S* binding surface and Ltn1p as the elongated molecule that meets Rqc2p at the sarcin-ricin loop (SRL) (Figs. 1 and 2 and figs. S1 to S5). Comparison of the 60*S*-bound Ltn1p with reconstructions of isolated Ltn1p suggests that the N terminus of Ltn1p engages the SRL with Rqc2p and that the middle region—which contains long HEAT/Armadillo repeats that adopt an elongated superhelical structure—reaches around the 60*S* (*6*). This conformation probably positions the C-terminal RING domain near the exit tunnel to ubiquitylate stalled nascent chains [figs. S5 and S6 and (*7*)].

A refined reconstruction of the Rqc2p-occupied class demonstrated that Rqc2p makes extensive contacts with an approximately P-site positioned (~P-site) tRNA (Figs. 1 and 2 and fig. S7). Rqc2p has a long coiled coil that makes direct contact

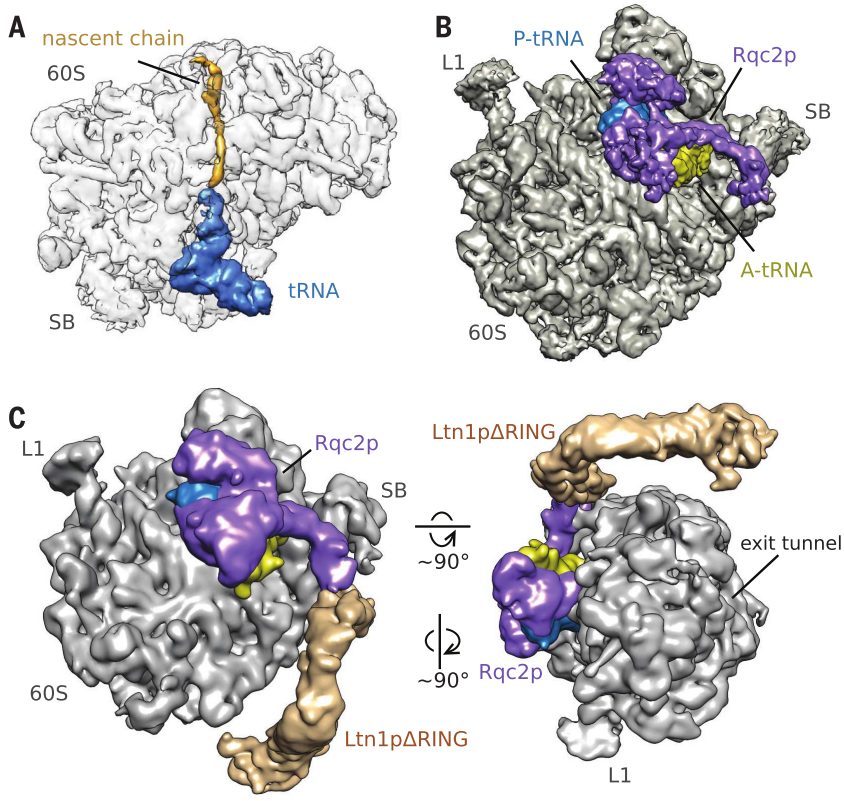


Fig. 1. Cryo-EM reconstructions of peptidyl-tRNA-60S ribosomes bound by the RQC components Rqc2p and Ltn1p. (**A**) A peptidyl-tRNA-60S complex isolated by immunoprecipitation of Rqc1p. The ribosome density is transparent to visualize the nascent chain. (**B**) Rqc2p (purple) and an ~A-site tRNA (yellow) bound to peptidyl-tRNA-60S complexes. Landmarks are indicated (L1, L1 stalk; SB, P-stalk base). (**C**) Ltn1p (tan) bound to Rqc2p-peptidyl-tRNA-60S complexes (B).

[1]Department of Biochemistry, University of Utah, UT 84112, USA. [2]Department of Biochemistry, Stanford University, Palo Alto, CA 94305, USA. [3]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA. [4]Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA. [5]California Institute for Quantitative Biomedical Research, University of California, San Francisco, San Francisco, CA 94158, USA. [6]Center for RNA Systems Biology, University of California, San Francisco, San Francisco, CA 94158, USA. [7]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA. [8]Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712, USA. [9]Department of Molecular Biosciences, University of Texas at Austin, Austin, TX 78712, USA. [10]Mass Spectrometry and Proteomics Core Facility, University of Utah, UT 84112, USA. *Present address: School of Life Sciences, Tsinghua University, Beijing 100084, China. †Corresponding author. E-mail: jonathan.weissman@ucsf.edu (J.S.W.), onn@stanford.edu (O.B.), adam.frost@ucsf.edu (A.F.)

with the SRL and the 60*S* P-stalk base (Fig. 2A). This structure also revealed Rqc2p binding to an ~A-site tRNA, whose 3′-CCA tail is within the peptidyl transferase center of the 60*S* (Fig. 2B and fig. S7). This observation was unexpected, because A-site tRNA interactions with the large ribosomal subunit are typically unstable and require mRNA templates and elongation factors (*8*). Rqc2p's interactions with the ~A-site tRNA appeared to involve recognition between the anticodon loop and a globular N-terminal domain, as well as D-loop and T-loop interactions along Rqc2p's coiled coil (Figs. 2 and 3).

To determine whether Rqc2p binds specific tRNA molecules, we extracted total RNA after RQC purification from strains with intact *RQC2* versus *rqc2*Δ strains. Deep sequencing by a new method using a thermostable group II intron reverse transcriptase (*9*) revealed that the presence of Rqc2p leads to an ~10-fold enrichment of $tRNA^{Ala(AGC)}$ and $tRNA^{Thr(AGT)}$ in the RQC (Fig. 3A). In complexes isolated from strains with intact *RQC2*, Ala(AGC) and Thr(AGT) are the most abundant tRNA molecules, even though they are less abundant than a number of other tRNAs in yeast (*10*).

Our structure suggested that Rqc2p's specificity for these tRNAs is due in part to direct interactions between Rqc2p and positions 32 to 36 of the anticodon loop, some of which are edited in the mature tRNA (Fig. 3). Adenosine 34 in the anticodon of both $tRNA^{Ala(AGC)}$ and $tRNA^{Thr(AGT)}$ is deaminated to inosine (*11*–*13*), leading to a diagnostic guanosine upon reverse transcription (*13*, *14*) (Fig. 3, B and C). Further analysis of the sequencing data revealed that cytosine 32 in $tRNA^{Thr(AGT)}$ is also deaminated to uracil in ~70% of the Rqc2p-enriched reads [Fig. 3 and (*15*)]. Together with the structure, this suggests that Rqc2p binds to the D-, T-, and anticodon loop of the ~A-site tRNA, and that recognition of the 32-UUIGY-36 edited motif accounts for Rqc2's specificity for these two tRNAs (Fig. 3, C and D). The pyrimidine at position 36 could explain the discrimination between the otherwise similar anticodon loops that harbor purines at base 36.

While assessing why Rqc2p evolved to bind these specific tRNA molecules, we considered these observations: First, our structural and biochemical data indicate that Rqc2p binds the 60*S* subunit after a stalled ribosome dissociates [fig. S6 (*1*, *2*)]. Second, stalled nascent chains accumulate as higher-molecular-weight species in the presence of Rqc2p than in its absence [Fig. 4A, also seen in Fig. 3E of (*1*)]. Finally, amino acid addition to a nascent chain can be mediated by the large ribosomal subunit in vitro even when decoupled from an mRNA template and the small subunit (*16*). Together, these facts led us to hypothesize that Rqc2p may promote the extension of stalled nascent chains with alanine and threonine residues in an elongation reaction that is mRNA- and 40*S*-free. This hypothesis makes specific predictions. First, the Rqc2p-dependent increase in the molecular weight of the nascent chain should occur from the C terminus exclusively. Second, the C-terminal extension should consist entirely of alanine and threonine residues that start immediately at the stalling sequence. Finally, the alanine and threonine extension should not have a defined sequence.

To test these predictions, we expressed a series of reporters containing a stalling sequence [tracts of up to 12 consecutive arginine codons, including pairs of the difficult-to-decode CGA codon (*17*)], inserted between the coding regions of green fluorescent protein (GFP) and red fluorescent protein (Fig. 4A). Null mutations in RQC components or inhibition of the proteasome led to the accumulation of nascent chain fragments that are normally degraded in wild-type cells (Fig. 4A) (*1*–*4*, *18*). Furthermore, *ltn1*Δ and *rqc2*Δ cells have different phenotypes: Expression of the stalling reporter in *ltn1*Δ led to the formation and accumulation of higher-molecular-weight species that resolve as a smear ~1.5 to 5 kD above the expected position of GFP (Fig. 4A). GFP mass-shifted products are observable in *rqc1*Δ*ltn1*Δ

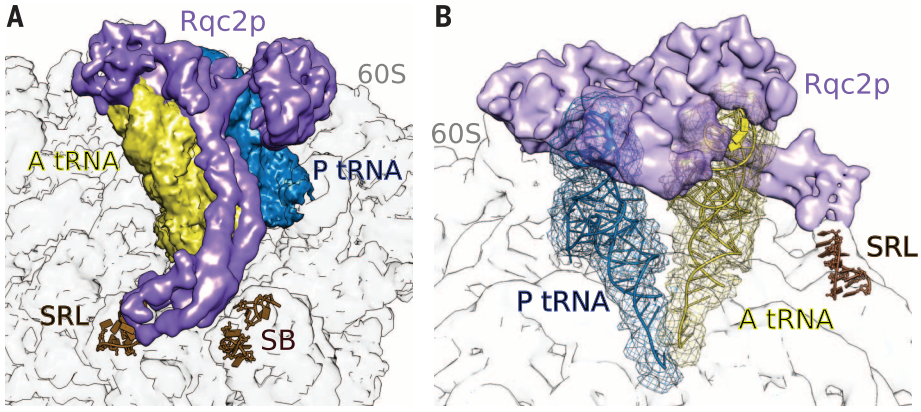


Fig. 2. Rqc2p binding to the 60*S* ribosome and ~P-site, and ~A-site tRNAs. (**A**) Rqc2p contacts ~P- and ~A-site tRNAs, the SRL, and P-stalk base ribosomal RNA (SB). (**B**) Rigid body fitting of tRNAs structures (ribbons) into EM densities (mesh).

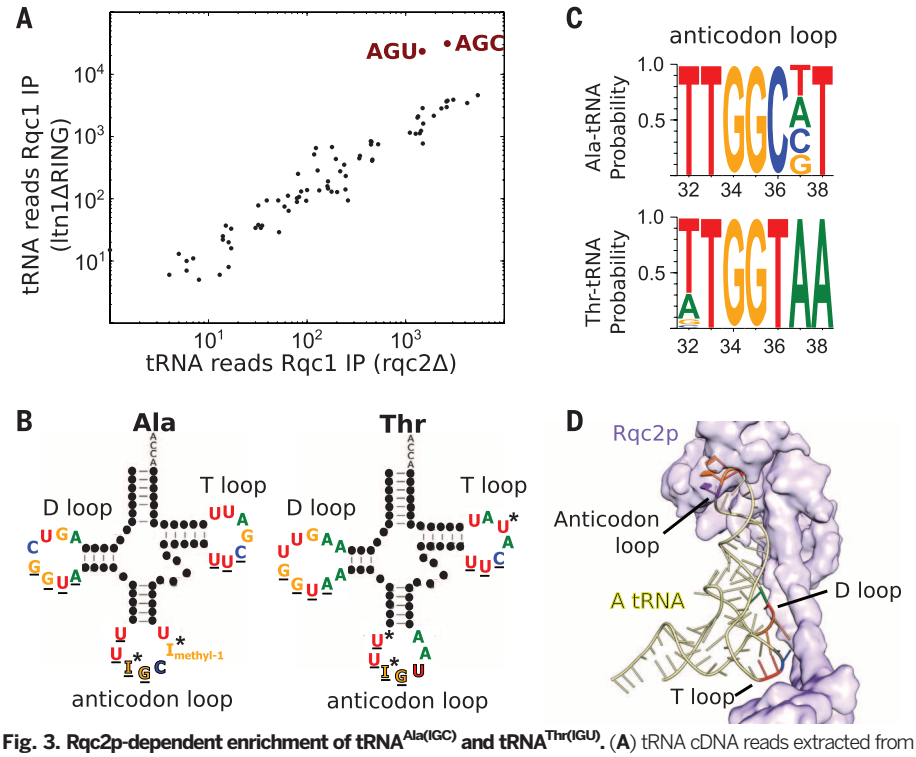


Fig. 3. Rqc2p-dependent enrichment of $tRNA^{Ala(IGC)}$ and $tRNA^{Thr(IGU)}$. (**A**) tRNA cDNA reads extracted from purified RQC particles and summed per unique anticodon, with versus without Rqc2p. (**B**) Secondary structures of $tRNA^{Ala(IGC)}$ and $tRNA^{Thr(IGU)}$. Identical nucleotides are underlined. Edited nucleotides are indicated with asterisks (*24*, *25*). (**C**) Weblogo representation of cDNA sequencing reads related to shared sequences found in anticodon loops (positions 32 to 38) of mature $tRNA^{Ala(IGC)}$ and $tRNA^{Thr(IGU)}$ (*26*). (**D**) ~A-tRNA contacts with Rqc2p at the D-, T-, and anticodon loops. Identical nucleotides between $tRNA^{Ala(IGC)}$ and $tRNA^{Thr(IGU)}$ are colored as in (B) (A, green; U, red; C, blue; G, orange) and pyrimidine, purple. Anticodon nucleotides are indicated as slabs.

double mutants, less prominent but still observable in *rqc1*Δ single mutants, but absent in all *rqc2*Δ single and double mutants (Fig. 4A). Thus, Rqc2p is necessary for the production of these higher-molecular-weight GFP species.

We probed the location of the extra mass along the GFP by inserting a tobacco etch virus (TEV) protease cleavage site upstream of the stalling tract (Fig. 4B). GFP resolved as a single band of the expected size with TEV treatment, indicating that the extra mass is located at or after the stall sequence. To pinpoint the location of the extra mass along the GFP, we moved the TEV cleavage site after the R12 stalling sequence. This created a mass-shifted GFP that was insensitive to TEV treatment, suggesting that the post-R12 TEV cleavage site was not synthesized. One possible model is that a translational frameshift occurs near the R12 sequence, which causes the mRNA to be mistranslated until the next out-of-frame stop codon. We falsified this model in two ways. First, we detected an Rqc2p-dependent GFP mass shift using a shorter R4 reporter in which multiple STOP codons were engineered in the +1 and +2 frames following the polyarginine tract (fig. S8). Second, we detected the Rqc2p-dependent GFP mass shift in a construct encoding a hammerhead ribozyme. The ribozyme cleaves the coding sequence of the GFP mRNA, leaving a truncated non-stop mRNA that causes a stall during translation of its final codon [fig. S8 (*19*)]. Thus, the GFP mass shift is located at or after the stall sequence but cannot be explained by mRNA translation past the stalling tract in any frame.

In order to determine the composition of the GFP mass-shifted products, we performed total amino acid analysis of immunopurified GFP from strains expressing the stalling reporter. Purified GFP from *ltn1*Δ (Fig. 4C) or *rqc1*Δ strains (fig. S9) is enriched in alanine and threonine as compared to purified GFP from double mutants with *rqc2*Δ which do not produce extended GFP. We then used Edman degradation to sequence TEV release fragments after purification of the stalled GFP reporter from the *ltn1*Δ strain. The first three codons in the R12 sequence are CGG-CGA-CGA, and Edman degradation suggested that the ribosome stalls at the first pair of the challenging-to-decode CGA codons (fig. S10). Following the encoded arginine residues, rising levels of alanine and threonine were detected at the C terminus (fig. S10). We further characterized these fragments by mass spectrometry and detected diverse poly-Ala and poly-Thr species ranging from 5 to 19 residues, with no defined sequence (table S1). Together, these observations demonstrate that Rqc2p directs the elongation of stalled nascent chains with nontemplated carboxy-terminal Ala and Thr extensions, or "CAT tails."

Earlier work (*1*) revealed that the accumulation of stalled nascent chains (e.g., by deletion of *LTN1*) led to a robust heat shock response that is fully dependent on Rqc2p, although the mechanism by which Rqc2p enabled this stress response was unclear. We hypothesized that CAT tails may be required for activation of heat shock factor 1 (Hsf1p). To isolate the effect of CAT tails in this context, we sought an *rqc2* allele that could not support CAT tail synthesis but could still bind to the 60*S* and facilitate Ltn1p-dependent ubiquitylation of the nascent chains. Rqc2p belongs to the conserved NFACT family of nucleic acid–binding proteins (*20*), and the N-terminal NFACT-N domain of Rqc2p is 22% identical to the NFACT-N domain of the *Staphylococcus aureus* protein Fbp (PDB:3DOA). Based on sequence and predicted secondary structure conservation, we fit this structure into a portion of the cryo–electron microscopy (cryo-EM) density ascribed to Rqc2p (figs. S11 and S12). This modeling exercise predicts that Rqc2p's NFACT-N domain recognizes features of both the P- and A-site tRNA molecules and that conserved residues D9, D98, and R99, which have been hypothesized to play roles in nucleic acid–binding or –modifying reactions (*20*), may contact the ~A-site tRNA (*20*) (fig. S12). An Rqc2p variant in which these residues were mutated to alanine ($rqc2_{aaa}$) rescued 60*S* recognition and the clearance of the stalling reporter almost as effectively as wild-type Rqc2p but did not support CAT tail synthesis (Fig. 4D and fig. S12). This CAT tail–deficient $rqc2_{aaa}$ allele also failed to rescue Hsf1p transcriptional activation (Fig. 4E), indicating that CAT tails may promote Hsf1p activation.

Integrating our observations, we propose the model schematized in fig. S13. Ribosome stalling leads to dissociation of the 60*S* and 40*S* subunits, followed by recognition of the peptidyl-tRNA-60*S* species by Rqc2p and Ltn1p. Ltn1p ubiquitylates the stalled nascent chain, and this

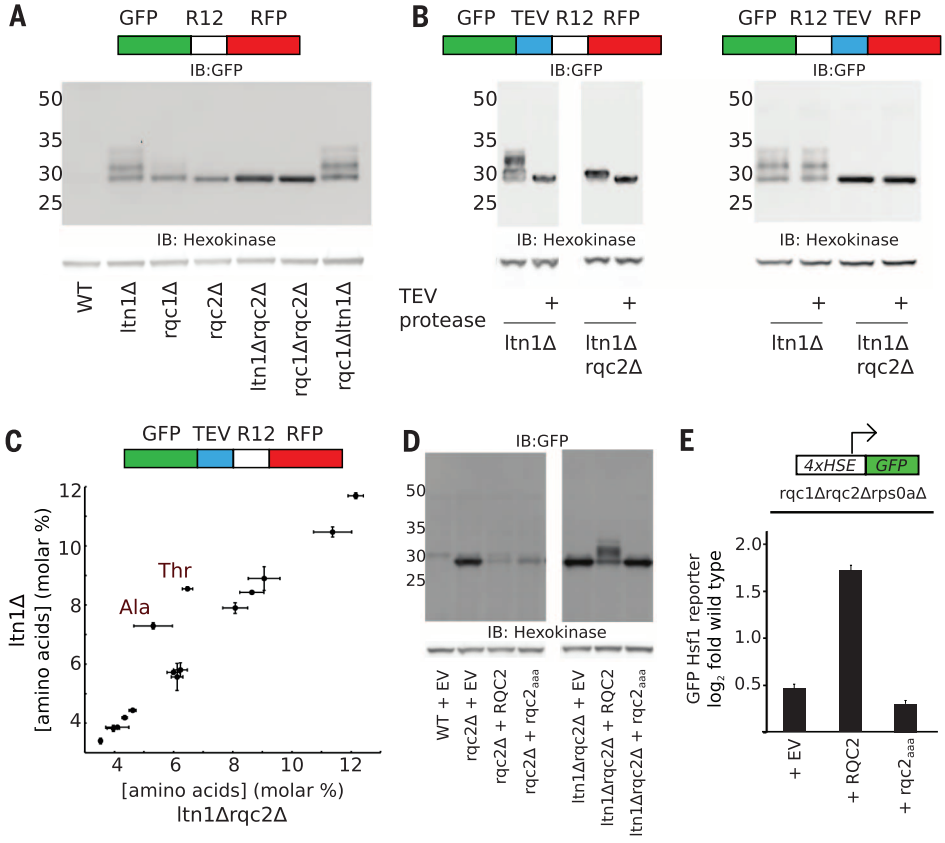


Fig. 4. Rqc2p-dependent formation of CAT tails. (**A**, **B**, and **D**) Immunoblots of stalling reporters in RQC deletion strains. (**C**) Total amino acid analysis of immunoprecipitated GFP expressed in *ltn1*Δ and *ltn1*Δ*rqc2*Δ strains, *n* = 3 independent immunoprecipitations. (**E**) Triplicate GFP levels measured with a flow cytometer and normalized to a wild-type control. EV, empty vector. All error bars are standard deviations.

leads to Cdc48 recruitment for extraction and degradation of the incomplete translation product. Rqc2p, through specific binding to Ala(IGC) and Thr(IGU) tRNAs, directs the template-free and 40*S*-free elongation of the incomplete translation product with CAT tails. CAT tails induce a heat shock response through a mechanism that is yet to be determined.

Hypomorphic mutations in the mammalian homolog of *LTN1* cause neurodegeneration in mice (*21*). Similarly, mice with mutations in a central nervous system–specific isoform of $tRNA^{Arg}$ and GTPBP2, a homolog of yeast Hbs1 which works with PELOTA/Dom34 to dissociate stalled 80S ribosomes, suffer from neurodegeneration (*22*). These observations reveal the consequences that ribosome stalls impose on the cellular economy. Eubacteria rescue stalled ribosomes with the transfer-messenger RNA (tmRNA)–SmpB system, which appends nascent chains with a unique C-terminal tag that targets the incomplete protein product for proteolysis (*23*). The mechanisms used by eukaryotes, which lack tmRNA, to recognize and rescue stalled ribosomes and their incomplete translation products have been unclear. The RQC—and Rqc2p's CAT tail tagging mechanism in particular—bear both similarities and contrasts to the tmRNA *trans*-translation system. The evolutionary convergence upon distinct mechanisms for extending incomplete nascent chains at the C terminus argues for their importance in maintaining proteostasis. One advantage of tagging stalled chains is that it may distinguish them from normal translation products and facilitate their removal from the protein pool. An alternate, not mutually exclusive, possibility is that the extension serves to test the functional integrity of large ribosomal subunits, so that the cell can detect and dispose of defective large subunits that induce stalling.

REFERENCES AND NOTES

1. O. Brandman *et al.*, *Cell* **151**, 1042–1054 (2012).
2. Q. Defenouillère *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 10.1073/pnas.1221724110 (2013).
3. R. Verma, R. S. Oania, N. J. Kolawa, R. J. Deshaies, *eLife* **2**, e00308 (2013).
4. S. Shao, K. von der Malsburg, R. S. Hegde, *Mol. Cell* **50**, 637–648 (2013).
5. G.-W. Li, D. Burkhardt, C. Gross, J. S. Weissman, *Cell* **157**, 624–635 (2014).
6. D. Lyumkis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1702–1707 (2013).
7. S. Shao, R. S. Hegde, *Mol. Cell* **55**, 880–890 (2014).
8. R. Lill, J. M. Robertson, W. Wintermeyer, *Biochemistry* **25**, 3245–3255 (1986).
9. G. E. Katibah *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 12025–12030 (2014).
10. D. Chu, D. J. Barnes, T. von der Haar, *Nucleic Acids Res.* **39**, 6705–6714 (2011).
11. P. F. Agris, F. A. P. Vendeix, W. D. Graham, *J. Mol. Biol.* **366**, 1–13 (2007).
12. F. H. C. Crick, *J. Mol. Biol.* **19**, 548–555 (1966).
13. A. P. Gerber, W. Keller, *Science* **286**, 1146–1149 (1999).
14. E. Delannoy *et al.*, *Plant Cell* **21**, 2058–2071 (2009).
15. M. A. T. Rubio, F. L. Ragone, K. W. Gaston, M. Ibba, J. D. Alfonzo, *J. Biol. Chem.* **281**, 115–120 (2006).
16. R. E. Monro, *Nature* **223**, 903–905 (1969).
17. D. P. Letzring, K. M. Dean, E. J. Grayhack, *RNA* **16**, 2516–2528 (2010).
18. S. Ito-Harashima, K. Kuroha, T. Tatematsu, T. Inada, *Genes Dev.* **21**, 519–524 (2007).
19. K. Kobayashi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 17575–17579 (2010).
20. A. M. Burroughs, L. Aravind, *RNA Biol.* **11**, 360–372 (2014).
21. J. Chu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2097–2103 (2009).
22. R. Ishimura *et al.*, *Science* **345**, 455–459 (2014).
23. S. D. Moore, R. T. Sauer, *Annu. Rev. Biochem.* **76**, 101–124 (2007).
24. A. Gerber, H. Grosjean, T. Melcher, W. Keller, *EMBO J.* **17**, 4780–4789 (1998).
25. K. W. Gaston *et al.*, *Nucleic Acids Res.* **35**, 6740–6749 (2007).
26. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, *Genome Res.* **14**, 1188–1190 (2004).

ACKNOWLEDGMENTS

Electron microscopy was performed at the University of Utah and the University of California. We thank D. Belnap (University of Utah) and M. Braunfeld (University of California, San Francisco) for supervision of the electron microscopes; A. Orendt and the Utah Center for High Performance Computing and the NSF Extreme Science and Engineering Discovery Environment consortium for computational support; D. Sidote (University of Texas at Austin) for help processing RNA-seq data; and D. Herschlag and P. Harbury for helpful comments. Amino acid analysis was performed by J. Shulze at the University of California, Davis Proteomics Core. Edman sequencing was performed at Stanford University's Protein and Nucleic Acid Facility by D. Winant. This work was supported by the Searle Scholars Program (A.F.); Stanford University (O.B.); NIH grants 1DP2GM110772-01 (A.F.), GM37949, and GM37951 (A.M.L.); the Center for RNA Systems Biology grants P50 GM102706 (J.S.W.) and U01 GM098254 (J.S.W.); and the Howard Hughes Medical Institute (J.S.W.). The authors declare no competing financial interests. The cryo-EM structures have been deposited at the Electron Microscopy Data Bank (accession codes 2811, 2812, 6169, 6170, 6171, 6172, 6176, and 6201).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/75/suppl/DC1
Materials and Methods
Figs. S1 to S13
Table S1
References (*27*–*41*)

7 August 2014; accepted 14 November 2014
10.1126/science.1259724

CANCER ETIOLOGY

Variation in cancer risk among tissues can be explained by the number of stem cell divisions

Cristian Tomasetti[1]* and Bert Vogelstein[2]*

Some tissue types give rise to human cancers millions of times more often than other tissue types. Although this has been recognized for more than a century, it has never been explained. Here, we show that the lifetime risk of cancers of many different types is strongly correlated (0.81) with the total number of divisions of the normal self-renewing cells maintaining that tissue's homeostasis. These results suggest that only a third of the variation in cancer risk among tissues is attributable to environmental factors or inherited predispositions. The majority is due to "bad luck," that is, random mutations arising during DNA replication in normal, noncancerous stem cells. This is important not only for understanding the disease but also for designing strategies to limit the mortality it causes.

Extreme variation in cancer incidence across different tissues is well known; for example, the lifetime risk of being diagnosed with cancer is 6.9% for lung, 1.08% for thyroid, 0.6% for brain and the rest of the nervous system, 0.003% for pelvic bone and 0.00072% for laryngeal cartilage (*1*–*3*). Some of these differences are associated with well-known risk factors such as smoking, alcohol use, ultraviolet light, or human papilloma virus (HPV) (*4*, *5*), but this applies only to specific populations exposed to potent mutagens or viruses. And such exposures cannot explain why cancer risk in tissues within the alimentary tract can differ by as much as a factor of 24 [esophagus (0.51%), large intestine (4.82%), small intestine (0.20%), and stomach (0.86%)] (*3*). Moreover, cancers of the small intestinal epithelium are three times less common than brain tumors (*3*), even though small intestinal epithelial cells are exposed to much higher levels of environmental mutagens than are cells within the brain, which are protected by the blood-brain barrier.

Another well-studied contributor to cancer is inherited genetic variation. However, only 5 to 10% of cancers have a heritable component (*6*–*8*), and even when hereditary factors in predisposed individuals can be identified, the way in which these factors contribute to differences in cancer incidences among different organs is obscure. For example, the same, inherited mutant *APC* gene is responsible for both the predisposition to colorectal and small intestinal cancers

[1]Division of Biostatistics and Bioinformatics, Department of Oncology, Sidney Kimmel Cancer Center, Johns Hopkins University School of Medicine and Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 550 North Broadway, Baltimore, MD 21205, USA. [2]Ludwig Center for Cancer Genetics and Therapeutics and Howard Hughes Medical Institute, Johns Hopkins Kimmel Cancer Center, 1650 Orleans Street, Baltimore, MD 21205, USA.
***Corresponding author. E-mail: ctomasetti@jhu.edu (C.T.); vogelbe@jhmi.edu (B.V.)**

in familial adenomatous polyposis (FAP) syndrome patients, yet cancers occur much more commonly in the large intestine than in the small intestine of these individuals.

If hereditary and environmental factors cannot fully explain the differences in organ-specific cancer risk, how else can these differences be explained? Here, we consider a third factor: the stochastic effects associated with the lifetime number of stem cell divisions within each tissue. In cancer epidemiology, the term "environmental" is generally used to denote anything not hereditary, and the stochastic processes involved in the development and homeostasis of tissues are grouped with external environmental influences in an uninformative way. We show here that the stochastic effects of DNA replication can be numerically estimated and distinguished from external environmental factors. Moreover, we show that these stochastic influences are in fact the major contributors to cancer overall, often more important than either hereditary or external environmental factors.

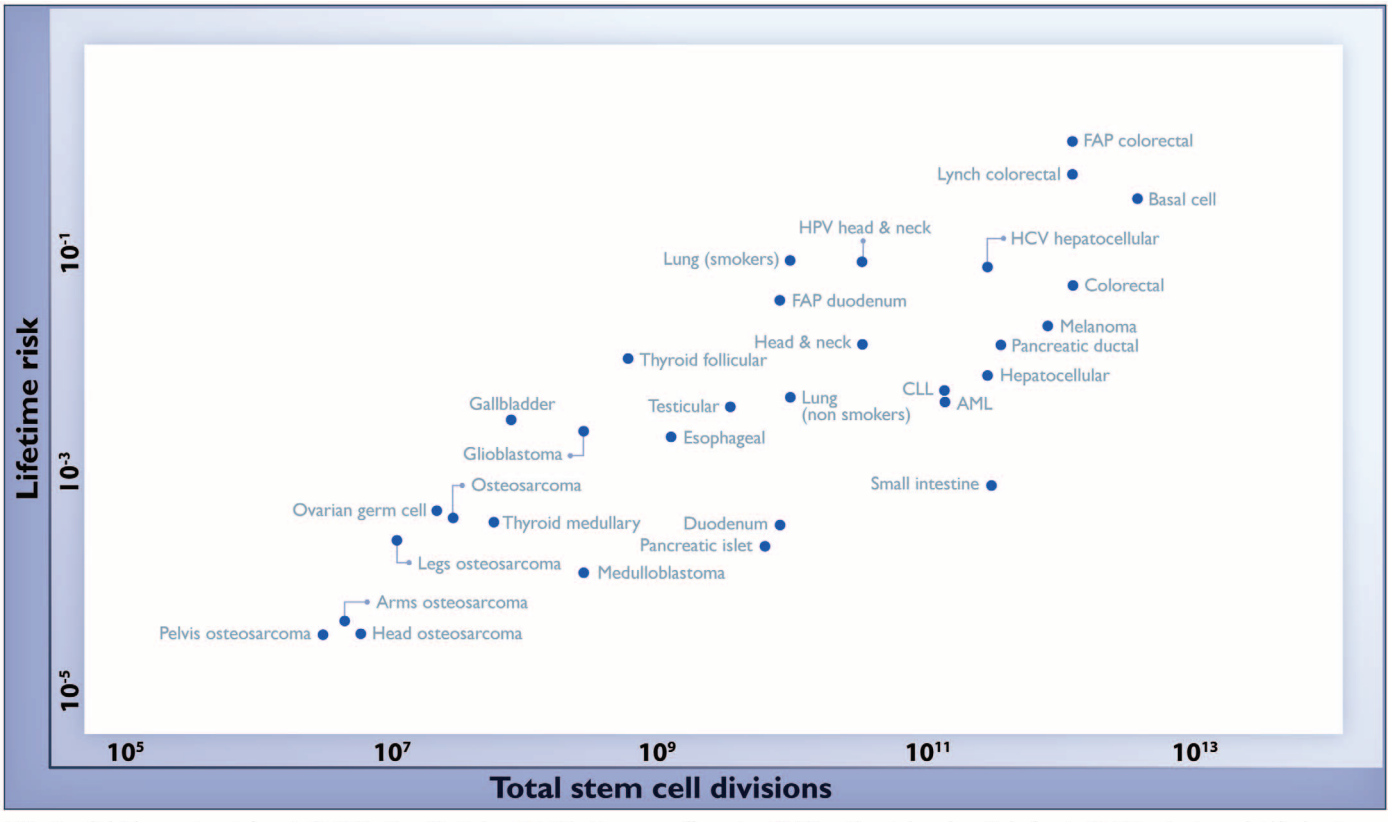
That cancer is largely the result of acquired genetic and epigenetic changes is based on the somatic mutation theory of cancer (*9*–*13*) and has been solidified by genome-wide analyses (*14*–*16*). The idea that the number of cells in a tissue and their cumulative number of divisions may be related to cancer risk, making them more vulnerable to carcinogenic factors, has been proposed but is controversial (*17*–*19*). Other insightful ideas relating to the nature of the factors underlying neoplasia are reviewed in (*20*–*22*).

The concept underlying the current work is that many genomic changes occur simply by chance during DNA replication rather than as a result of carcinogenic factors. Since the endogenous mutation rate of all human cell types appears to be nearly identical (*23*, *24*), this concept predicts that there should be a strong, quantitative correlation between the lifetime number of divisions among a particular class of cells within each organ (stem cells) and the lifetime risk of cancer arising in that organ.

To test this prediction, we attempted to identify tissues in which the number and dynamics of stem cells have been described. Most cells in tissues are partially or fully differentiated cells that are typically short-lived and unlikely to be able to initiate a tumor. Only the stem cells—those that can self-renew and are responsible for the development and maintenance of the tissue's architecture—have this capacity. Stem cells often make up a small proportion of the total number of cells in a tissue and, until recently, their nature, number, and hierarchical division patterns were not known (*25*–*28*). Tissues were not included in our analysis if the requisite parameters were not found in the literature or if their estimation was difficult to derive.

Through an extensive literature search, we identified 31 tissue types in which stem cells had been quantitatively assessed (see the supplementary materials). We then plotted the total number of stem cell divisions during the average lifetime of a human on the x axis and the lifetime risk for cancer of that tissue type on the y axis (Fig. 1) (table S1). The lifetime risk in the United States for all included cancer types has been evaluated in detail, such as in the Surveillance, Epidemiology, and End Results (SEER) database (*3*). The correlation between these two very different parameters—number of stem cell divisions and lifetime risk—was striking, with a highly positive correlation (Spearman's rho = 0.81; $P < 3.5 \times 10^{-8}$) (Fig. 1). Pearson's linear correlation 0.804 [0.63 to 0.90; 95% confidence interval (CI)] was equivalently significant ($P < 5.15 \times 10^{-8}$). One of the most impressive features of this correlation was that it extended across five orders of magnitude, thereby applying to cancers with enormous differences in incidence. No other environmental or inherited factors are known to be correlated in this way across tumor types. Moreover, these correlations were extremely robust; when the parameters used to construct Fig. 1 were varied over a broad range of plausible values, the tight correlation remained intact (see the supplementary materials).

A linear correlation equal to 0.804 suggests that 65% (39% to 81%; 95% CI) of the differences in cancer risk among different tissues can be explained by the total number of stem cell divisions in those tissues. Thus, the stochastic effects of



FAP = Familial Adenomatous Polyposis ◆ HCV = Hepatitis C virus ◆ HPV = Human papillomavirus ◆ CLL = Chronic lymphocytic leukemia ◆ AML = Acute myeloid leukemia

Fig. 1. The relationship between the number of stem cell divisions in the lifetime of a given tissue and the lifetime risk of cancer in that tissue. Values are from table S1, the derivation of which is discussed in the supplementary materials.

DNA replication appear to be the major contributor to cancer in humans.

We next attempted to distinguish the effects of this stochastic, replicative component from other causative factors—that is, those due to the external environment and inherited mutations. For this purpose, we defined an "extra risk score" (ERS) as the product of the lifetime risk and the total number of stem cell divisions ($\log_{10}$ values). Machine learning methods were employed to classify tumors based only on this score (see the supplementary materials). With the number of clusters set equal to two, the tumors were classified in an unsupervised manner into one cluster with high ERS (9 tumor types) and another with low ERS (22 tumor types) (Fig. 2).

The ERS provides a test of the approach described in this work. If the ERS for a tissue type is high—that is, if there is a high cancer risk of that tissue type relative to its number of stem cell divisions—then one would expect that environmental or inherited factors would play a relatively more important role in that cancer's risk (see the supplementary materials for a detailed explanation). It was therefore notable that the tumors with relatively high ERS were those with known links to specific environmental or hereditary risk factors (Fig. 2, blue cluster). We refer to the tumors with relatively high ERS as D-tumors (D for deterministic; blue cluster in Fig. 2) because deterministic factors such as environmental mutagens or hereditary predispositions strongly affect their risk. We refer to tumors with relatively low ERS as R-tumors (R for replicative; green cluster in Fig. 2) because stochastic factors, presumably related to errors during DNA replication, most strongly appear to affect their risk.

The incorporation of a replicative component as a third, quantitative determinant of cancer risk forces rethinking of our notions of cancer causation. The contribution of the classic determinants (external environment and heredity) to R-tumors is minimal (Fig. 1). Even for D-tumors, however, replicative effects are essential, and environmental and hereditary effects simply add to them. For example, patients with FAP are ~30 times as likely to develop colorectal cancer than duodenal cancer (Fig. 1). Our data suggest that this is because there are ~150 times as many stem cell divisions in the colon as in the duodenum. The lifetime risk of colorectal cancer would be very low, even in the presence of an underlying *APC* gene mutation, if colonic epithelial stem cells were not constantly dividing. A related point is that mice with inherited *APC* mutations display the opposite pattern: Small intestinal tumors are more common than large intestinal tumors. Our analysis provides a plausible explanation for this striking difference between mice and men; namely, in mice the small intestine undergoes more stem cell divisions than the large intestine (see the supplementary materials for the estimates). Another example is provided by melanocytes and basal epidermal cells of the skin, which are both exposed to the same carcinogen (ultraviolet light) at the identical dose, yet melanomas are much less common than basal cell carcinomas. Our data suggest that this difference is attributable to the fact that basal epidermal cells undergo a higher number of divisions than melanocytes (see the supplementary materials for the estimates). The total number of stem cells in an organ and their proliferation rate may of course be influenced by genetic and environmental factors such as those that affect height or weight.

In formal terms, our analyses show only that there is some stochastic factor related to stem cell division that seems to play a major role in cancer risk. This situation is analogous to that of the classic studies of Nordling and of Armitage and Doll (*10*, *29*). These investigators showed that the relationship between age and the incidence of cancer was exponential, suggesting that many cellular changes, or stages, were required for carcinogenesis. On the basis of research since that time, these events are now interpreted as somatic mutations. Similarly, we interpret the stochastic factor underlying the importance of stem cell divisions to be somatic mutations. This interpretation is buttressed by the large number of somatic mutations known to exist in cancer cells (*14*–*16*, *30*).

Our analysis shows that stochastic effects associated with DNA replication contribute in a substantial way to human cancer incidence in the United States. These results could have important public health implications. One of the most promising avenues for reducing cancer deaths is through prevention. How successful can such approaches be? The maximum fraction of tumors that are preventable through primary prevention (such as vaccines against infectious agents or altered lifestyles) may be evaluated from their ERS. For nonhereditary D-tumors, this fraction is high and primary prevention could make a major impact (*31*). Secondary prevention, obtainable in principle through early detection, could further reduce nonhereditary D-tumor–related deaths and is also instrumental for reducing hereditary D-tumor–related deaths. For R-tumors, primary prevention measures are not likely to be very effective, and secondary prevention should be the major focus.

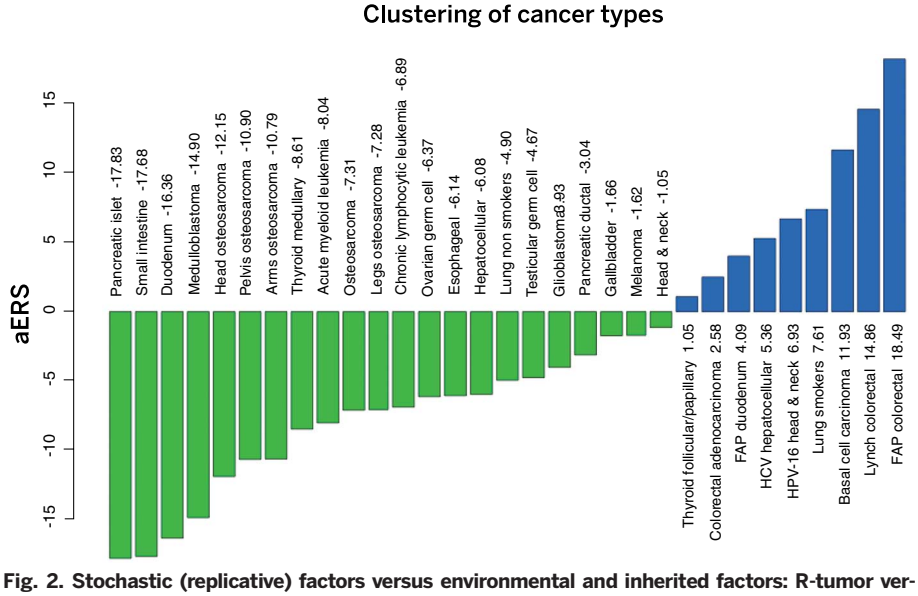


Fig. 2. Stochastic (replicative) factors versus environmental and inherited factors: R-tumor versus D-tumor classification. The adjusted ERS (aERS) is indicated next to the name of each cancer type. R-tumors (green) have negative aERS and appear to be mainly due to stochastic effects associated with DNA replication of the tissues' stem cells, whereas D-tumors (blue) have positive aERS. Importantly, although the aERS was calculated without any knowledge of the influence of environmental or inherited factors, tumors with high aERS proved to be precisely those known to be associated with these factors. For details of the derivation of aERS, see the supplementary materials.

REFERENCES AND NOTES

1. P. M. Dubal *et al.*, *Laryngoscope* **124**, 1877–1881 (2014).
2. N. Jaffe, *Pediatric and Adolescent Osteosarcoma* (Springer, New York, 2009).
3. National Cancer Institute, Surveillance, Epidemiology, and End Results Program; www.seer.cancer.gov.
4. G. Danaei *et al.*, *Lancet* **366**, 1784–1793 (2005).
5. Centers for Disease Control and Prevention; www.cdc.gov.
6. E. R. Fearon, *Science* **278**, 1043–1050 (1997).
7. P. Lichtenstein *et al.*, *N. Engl. J. Med.* **343**, 78–85 (2000).
8. American Cancer Society; www.cancer.org.
9. P. Armitage, R. Doll, *Br. J. Cancer* **11**, 161–169 (1957).
10. P. Armitage, R. Doll, *Br. J. Cancer* **8**, 1–12 (1954).
11. T. Boveri, *Zur Frage der Entstehung Maligner Tumoren* (G. Fischer, Jena, Germany, 1914).
12. E. R. Fearon, B. Vogelstein, *Cell* **61**, 759–767 (1990).
13. A. G. Knudson Jr., *Proc. Natl. Acad. Sci. U.S.A.* **68**, 820–823 (1971).
14. L. A. Garraway, E. S. Lander, *Cell* **153**, 17–37 (2013).
15. M. R. Stratton, P. J. Campbell, P. A. Futreal, *Nature* **458**, 719–724 (2009).
16. B. Vogelstein *et al.*, *Science* **339**, 1546–1558 (2013).
17. D. Albanes, M. Winick, *J. Natl. Cancer Inst.* **80**, 772–775 (1988).
18. L. Tomatis; International Agency for Research on Cancer, *Environ. Health Perspect.* **101** (suppl. 5), 149–151 (1993).
19. J. M. Ward, H. Uno, Y. Kurata, C. M. Weghorst, J. J. Jang, *Environ. Health Perspect.* **101** (suppl. 5), 125–135 (1993).
20. S. Sell, *Environ. Health Perspect.* **101** (suppl. 5), 15–26 (1993).
21. S. A. Frank, in *Dynamics of Cancer: Incidence, Inheritance, and Evolution* (Princeton Univ. Press, Princeton, NJ, 2007), chap. 4.
22. S. G. Baker, A. Cappuccio, J. D. Potter, *J. Clin. Oncol.* **28**, 3215–3218 (2010).
23. M. Lynch, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 961–968 (2010).
24. C. Tomasetti, B. Vogelstein, G. Parmigiani, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1999–2004 (2013).

25. C. Blanpain, E. Fuchs, *Nat. Rev. Mol. Cell Biol.* **10**, 207–217 (2009).
26. C. Booth, C. S. Potten, *J. Clin. Invest.* **105**, 1493–1499 (2000).
27. T. Reya, S. J. Morrison, M. F. Clarke, I. L. Weissman, *Nature* **414**, 105–111 (2001).
28. H. J. Snippert *et al.*, *Cell* **143**, 134–144 (2010).
29. C. O. Nordling, *Br. J. Cancer* **7**, 68–72 (1953).
30. C. Kandoth *et al.*, *Nature* **502**, 333–339 (2013).
31. G. A. Colditz, K. Y. Wolin, S. Gehlert, *Sci. Transl. Med.* **4**, 127rv4 (2012).

ACKNOWLEDGMENTS

We thank E. Cook for artwork. This work was supported by the The Virginia and D. K. Ludwig Fund for Cancer Research, The Lustgarten Foundation for Pancreatic Cancer Research, The Sol Goldman Center for Pancreatic Cancer Research, and NIH grants P30-CA006973, R37-CA43460, RO1-CA57345, and P50-CA62924. Authors' contributions: C.T. formulated the hypothesis. C.T. and B.V. designed the research. C.T. provided mathematical and statistical analysis. C.T. and B.V. performed research. C.T. and B.V. wrote the paper.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/78/suppl/DC1
Materials and Methods
Fig. S1
Table S1
References (*32*–*146*)

4 September 2014; accepted 20 November 2014
10.1126/science.1260825

MUTAGENESIS

Smoking is associated with mosaic loss of chromosome Y

Jan P. Dumanski,[1,2]* Chiara Rasi,[1,2] Mikael Lönn,[3] Hanna Davies,[1,2] Martin Ingelsson,[4] Vilmantas Giedraitis,[4] Lars Lannfelt,[4] Patrik K. E. Magnusson,[5] Cecilia M. Lindgren,[6,7] Andrew P. Morris,[6,8] David Cesarini,[9] Magnus Johannesson,[10] Eva Tiensuu Janson,[11] Lars Lind,[11] Nancy L. Pedersen,[5] Erik Ingelsson,[2,11] Lars A. Forsberg[1,2]*

Tobacco smoking is a risk factor for numerous disorders, including cancers affecting organs outside the respiratory tract. Epidemiological data suggest that smoking is a greater risk factor for these cancers in males compared with females. This observation, together with the fact that males have a higher incidence of and mortality from most non–sex-specific cancers, remains unexplained. Loss of chromosome Y (LOY) in blood cells is associated with increased risk of nonhematological tumors. We demonstrate here that smoking is associated with LOY in blood cells in three independent cohorts [TwinGene: odds ratio (OR) = 4.3, 95% confidence interval (CI) = 2.8 to 6.7; Uppsala Longitudinal Study of Adult Men: OR = 2.4, 95% CI = 1.6 to 3.6; and Prospective Investigation of the Vasculature in Uppsala Seniors: OR = 3.5, 95% CI = 1.4 to 8.4] encompassing a total of 6014 men. The data also suggest that smoking has a transient and dose-dependent mutagenic effect on LOY status. The finding that smoking induces LOY thus links a preventable risk factor with the most common acquired human mutation.

Tobacco smoking killed ~100 million people during the 20th century and is projected to kill ~1 billion people during the current century, assuming that the current frequency of smoking is retained (*1*, *2*). Lung cancer is the prime cause of cancer-associated death in relation to smoking. However, smoking is also a risk factor for tumors outside the respiratory tract, and these are more common in males than females [hazard ratio in males: 2.2, 95% confidence interval (CI) = 1.7 to 2.8; in females: 1.7, 95% CI = 1.4–2.1] (*2*). Moreover, males have a higher incidence and mortality from most non–sex-specific cancers, disregarding smoking status, and this fact is largely unexplained by known risk factors (*3*, *4*). A recent analysis of noncancerous blood cells revealed that a male-specific chromosomal aberration, acquired mosaic loss of chromosome Y (LOY), is associated with an increased risk of nonhematological tumors among aging males (*5*).

Here, we analyzed possible causes of LOY by studying 6014 men from three independent prospective cohorts—TwinGene, n = 4373 (*6*, *7*); Uppsala Longitudinal Study of Adult Men (ULSAM), n = 1153 (*8*); and Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS), n = 488 (*9*)—from which comprehensive epidemiological records are available (tables S2 to S4). We included the following environmental, lifestyle, and clinical factors in the analyses: smoking, age, hypertension, exercise habits, diabetes, body mass index, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, education level, and alcohol intake. We also included genotyping quality as a confounder in the regression analyses, to adjust for possible influence of experimental noise. Similar definitions of factors were used in all cohorts, as outlined in tables S2 to S5 and described in detail in the materials and methods section of the supplementary materials. Estimation of LOY was based on single-nucleotide polymorphism (SNP)–array data using the 2.5MHumanOmni and HumanOmniExpress beadchips in the ULSAM and PIVUS/TwinGene studies, respectively (fig. S1). The estimation of the degree of mosaicism and scoring of LOY was undertaken using the continuous median logR ratio (mLRR-Y) estimate, calculated from SNP-array data as the median of the logR ratio of all SNP probes within the male-specific part of chromosome Y (MSY), as described previously (*5*). An mLRR-Y estimate close to zero indicates a normal chromosome Y state, whereas more negative mLRR-Y values denote an increasing level of blood cells with LOY. To facilitate comparisons between the three cohorts, we corrected the mLRR-Y values for all participants, using cohort-specific correction constants, as explained in the supplementary materials (figs. S1 and S2).

LOY was by far the most common postzygotic mutation found in the three cohorts. The age range at sampling in ULSAM and PIVUS was 70.7 to 83.6 years and 69.8 to 70.7 years, respectively, and we found LOY in 12.6% of ULSAM participants and 15.6% of PIVUS participants (figs. S3 and S4). The age range at sampling in TwinGene was 48 to 93 years, and the frequency of LOY in the entire cohort was 7.5% (fig. S5).

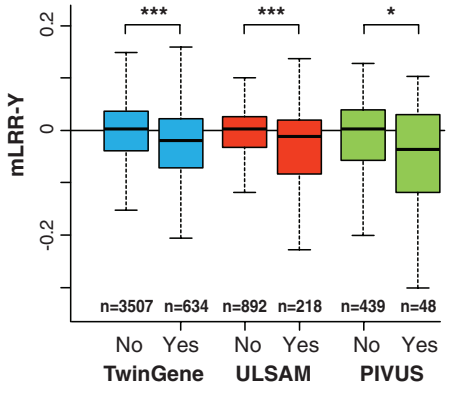


Fig. 1. The association between smoking status and the level of LOY (i.e., mLRR-Y) in three independent cohorts. In all cohorts, these unadjusted analyses indicate that the current smokers (Yes) (table S5) had a significantly higher degree of mosaic LOY in blood, compared with noncurrent smokers (No), composed of never-smokers and previous smokers. ***P < 0.001; *P < 0.05 (Kolmogorov-Smirnov tests: TwinGene, D = 0.15, $P = 1.131 \times 10^{-11}$; ULSAM, D = 0.15, P = 0.0006; PIVUS, D = 0.23, P = 0.0203). The definitions used for LOY scoring and the entire ranges of mLRR-Y data observed in each cohort are shown in figs. S3 to S5.

[1]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [2]Science for Life Laboratory, Uppsala University, Uppsala, Sweden. [3]Södertörn University, School of Life Sciences, Biology, Huddinge, Sweden. [4]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden. [5]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [6]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [7]Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, USA. [8]Department of Biostatistics, University of Liverpool, Liverpool, UK. [9]Center for Experimental Social Science, New York University, New York, NY 10012, USA. [10]Department of Economics, Stockholm School of Economics, Stockholm, Sweden. [11]Department of Medical Sciences, Uppsala University, Uppsala, Sweden.
***Corresponding Author. E-mail: jan.dumanski@igp.uu.se (J.P.D.); lars.forsberg@igp.uu.se (L.A.F.).**

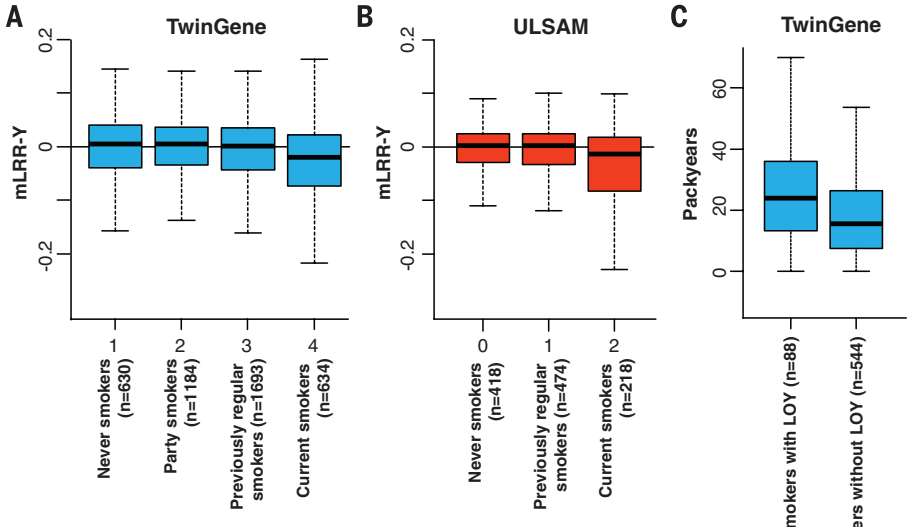
However, in TwinGene participants aged 70 years or older, 15.4% had LOY, which is similar to the LOY frequency in the other cohorts in the same age range. In TwinGene participants younger than 70 years, only 4.1% were affected by LOY. Thresholds for LOY scoring, at the lower 99% confidence limit of the distributions of experimental mLRR-Y variation (TwinGene = −0.1324, ULSAM = −0.1024, and PIVUS = −0.1182), were used for frequency calculations, as explained in the supplementary materials and figs. S3 to S5. At this degree of mosaicism, ~10% of the analyzed nucleated blood cells from each sample are affected by LOY. Mosaic LOY in blood detected with SNP arrays was validated using whole-genome next-generation sequencing (NGS) in 100 random participants in the ULSAM cohort. There was 100% concordance in LOY scoring between results from SNP array and NGS data (*5*).

We found a strong association between smoking and LOY status in the three independent cohorts. Current smokers had a significantly higher degree of LOY mosaicism, compared with nonsmokers and past smokers, in unadjusted models (Fig. 1) and in the multivariable models adjusting for the above-mentioned potential confounders (table S1, analysis of covariance: TwinGene $F_{(1,1666)}$ = 45.4, $P = 2.225 \times 10^{-11}$; ULSAM $F_{(1,968)}$ = 17.6, $P = 2.984 \times 10^{-5}$; PIVUS $F_{(1,385)}$ = 9.1, P = 0.0028). Apart from smoking, the only other factor significantly associated with LOY was age, which was observed in TwinGene, with higher degree of LOY in older participants. The age range in ULSAM and PIVUS was narrower (see above), which may explain why age had no effect on LOY in these regression models (tables S1 and S5). To assess the risk of LOY in blood cells of smokers, we used logistic regression adjusting for the same confounders as in table S1. Participants were LOY-scored as 1 or 0 before these analyses based on the continuous mLRR-Y estimate, using the same threshold as for the estimations of LOY frequencies—i.e., the lower 99% CIs of the experimentally induced mLRR-Y variation (figs. S3 to S5). The adjusted odds ratio (OR) estimates from logistic regressions were highest in TwinGene (OR = 4.3, 95% CI = 2.8 to 6.7), followed by PIVUS (OR = 3.5, 95% CI = 1.4 to 8.4) and ULSAM (OR = 2.4, 95% CI = 1.6 to 3.6). The corresponding unadjusted ORs are given in table S6. Based on these calculations, we estimate that current smokers in the studied cohorts had a 2.4 to 4.3 times greater risk of displaying LOY compared with nonsmokers. Furthermore, among the current smokers in the large TwinGene cohort, we found a strong dose-response effect with more LOY in heavy smokers—i.e., smokers with LOY had been smoking significantly more pack-years compared to smokers without LOY (Fig. 2C).

Our results suggest that the association between smoking and LOY is valid for current smokers only (Fig. 2, A and B). Previous epidemiological studies showed that smoking cessation at any age is associated with dramatically reduced death rates. For smokers who quit at 25 to 34 years of age, survival was nearly identical with those who had never smoked (*2*, *10*). We analyzed the level of LOY after smoking cessation in the ULSAM and the TwinGene cohorts by using LOY data in past regular smokers (table S5). No difference in LOY frequency between never-smokers and previously regular smokers was found (Fig. 2, A and B). One possible explanation for these results is that the previous smokers with LOY died off faster than the rest of the cohort. Another, and perhaps more likely, explanation is that LOY is induced and sustained by smoking and that LOY is a dynamic and reversible process.

Whether the LOY induced by smoking plays a direct role in cancer is unclear. One hypothesis is that smoking is clastogenic—that is, it induces many chromosomal abnormalities, including an incidental loss of the Y chromosome. In this scenario, LOY would be a neutral passenger mutation and a reporter of a general tendency of chromosome missegregation in mitosis, which is enhanced by smoking and associated with risk for cancer and mortality. A second hypothesis is that LOY in blood cells is a causative factor in cancer development, possibly through effects on tumor immunosurveillance (*11*). We conducted a preliminary test of this hypothesis by investigating possible functional consequences of LOY in sorted blood cells from three ULSAM survivors scored with LOY and still alive at the age of 91 years. We sorted cells from three compartments (granulocytes, CD4+ T-lymphocytes, and CD19+ B-lymphocytes) in subjects that displayed LOY in earlier serial analyses of whole blood, performed at four time points during two decades (fig. S6). The three subjects were free from cancer diagnoses at the time of blood collection at 91 years. The data from these experiments are tantalizing because (i) the percentage of cells with LOY differed between different compartments of the hematopoietic system, and (ii) ULSAM-1412 suggests that LOY might be an oligo-clonal process, because cells derived from myeloid and some (but not all) lymphoid progenitors display LOY. These preliminary results support the second hypothesis. If LOY were a phenotypically neutral passenger mutation, one would expect that LOY cells would be randomly distributed within all components of the hematological system. It was recently shown that LOY status of blood cells is associated with a higher risk for all-cause mortality as well as a higher risk for nonhematological cancers and that it can be considered as a biomarker of male carcinogenesis (*5*). We hypothesized that a disrupted tumor-immunosurveillance in LOY-affected cells could help explain the connection between LOY status of noncancerous blood cells and risk for tumors in other tissues (*5*, *11*). These results also support the second hypothesis and the increasingly recognized view that chromosome Y carries many vital functions in biological processes beyond sex determination and sperm production (*12*–*16*).

Our results are consistent with a previously described dynamic nature of expanding-contracting noncancerous cell clones in blood affected with



P-values adjusted for multiple testing

	Group 1	Group 2	Group 3
Group 2	0.9828	-	
Group 3	0.1322	0.1260	-
Group 4	<0.0001	<0.0001	<0.0001

P-values adjusted for multiple testing

	Group 0	Group 1
Group 1	0.1791	-
Group 2	<0.0001	0.0011

Fig. 2. Differences in degree of LOY between different smoking categories within TwinGene (A) and ULSAM (B), as defined in table S5. In TwinGene (A), there was a significant difference between four smoking categories [analysis of variance (ANOVA): $F_{(3,4137)}$ = 22.2, $P = 3.028 \times 10^{-14}$], and results from post hoc analysis adjusting P values for multiple testing using a Tukey post hoc test are displayed. In ULSAM (B), there was also a significant difference between three smoking categories (ANOVA; $F_{(2,1107)}$ = 12.2, $P = 5.812 \times 10^{-6}$), and post hoc analysis is shown. In both cohorts, the current smokers had a significantly higher degree of LOY compared with all other categories. The average degree of LOY in the previously regular smokers was not significantly different from the average degree of LOY in the never-smokers in both cohorts. (**C**) A dose-response effect within current smokers in TwinGene, with men smoking the most pack-years also being associated with higher degree of LOY, as defined in fig. S5 (Kolmogorov-Smirnov test: D = 0.2244, P = 0.0010).

mosaic genetic aberrations—i.e., it appears that the relative frequency of cells from a cell clone can first increase and then decrease later in life (*5*, *17*, *18*). In the present analyses, LOY was detected in ≥10% of blood cells from about 15% of elderly males in three cohorts (figs. S3 to S5). The cell clones with LOY were likely detectable in our analyses because they are enriched due to an increased proliferative potential as a consequence of LOY, which is in agreement with chromosome Y containing tumor suppressor genes. Recent analysis of >8200 tumor-normal pairs suggest that two genes (*ZFY* and *UTY*, from the male-specific part of Y) have properties of tumor suppressors (*19*). Interestingly, both genes have homologs on chromosome X and escape X inactivation (*19*, *20*). Moreover, other analyses of various tumor collections show that chromosome Y is lost from numerous types of tumors in frequencies ranging from 15 to 80% of cases (*21*–*24*). Thus, counting both LOY in noncancerous blood clones and in transformed tumor cells, nullisomy Y is among the most common, if not the most common, human mutation. The results presented here suggest that this aneuploidy, affecting 1.6% of the genome, is likely induced by smoking.

In conclusion, we show that LOY is more common in current smokers compared with non-current smokers in three cohorts (Fig. 1 and table S1), that the effect from smoking on LOY is dose dependent, and that this effect appears to be transient, as it disappears after smoking cessation (Fig. 2). Epidemiological observations suggest that smoking could be a greater risk factor for cancer outside the respiratory tract in males compared with females (*2*, *4*, *10*). Moreover, males have a higher incidence and mortality from most sex-unspecific cancers (*3*, *4*). The molecular mechanisms behind these observations are not well understood, but LOY, being a male-specific, smoking-induced risk factor, could provide a missing link and help explain these sex differences.

REFERENCES AND NOTES

1. P. Jha, *Natl. Rev.* **9**, 655–664 (2009).
2. P. Jha *et al.*, *N. Engl. J. Med.* **368**, 341–350 (2013).
3. M. B. Cook, K. A. McGlynn, S. S. Devesa, N. D. Freedman, W. F. Anderson, *Cancer Epidemiol. Biomarkers Prev.* **20**, 1629–1637 (2011).
4. G. Edgren, L. Liang, H. O. Adami, E. T. Chang, *Eur. J. Epidemiol.* **27**, 187–196 (2012).
5. L. A. Forsberg *et al.*, *Nat. Genet.* **46**, 624–628 (2014).
6. P. Lichtenstein *et al.*, *Twin Res. Hum. Genet.* **9**, 875–882 (2006).
7. P. K. Magnusson *et al.*, *Twin Res. Hum. Genet.* **16**, 317–329 (2013).
8. H. Hedstrand, *Ups. J. Med. Sci. Suppl.* **19**, 1–61 (1975).
9. L. Lind, N. Fors, J. Hall, K. Marttala, A. Stenborg, *Arterioscler. Thromb. Vasc. Biol.* **25**, 2368–2375 (2005).
10. M. J. Thun *et al.*, *N. Engl. J. Med.* **368**, 351–364 (2013).
11. G. P. Dunn, A. T. Bruce, H. Ikeda, L. J. Old, R. D. Schreiber, *Nat. Immunol.* **3**, 991–998 (2002).
12. B. Lemos, L. O. Araripe, D. L. Hartl, *Science* **319**, 91–93 (2008).
13. B. Lemos, A. T. Branco, D. L. Hartl, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15826–15831 (2010).
14. A. G. Clark, *Nature* **508**, 463–465 (2014).
15. D. W. Bellott *et al.*, *Nature* **508**, 494–499 (2014).
16. D. Cortez *et al.*, *Nature* **508**, 488–493 (2014).
17. L. A. Forsberg *et al.*, *Am. J. Hum. Genet.* **90**, 217–228 (2012).
18. L. A. Forsberg, D. Absher, J. P. Dumanski, *J. Med. Genet.* **50**, 1–10 (2013).
19. T. Davoli *et al.*, *Cell* **155**, 948–962 (2013).
20. A. Schneider-Gädicke, P. Beer-Romero, L. G. Brown, R. Nussbaum, D. C. Page, *Cell* **57**, 1247–1258 (1989).
21. L. J. Zhang, E. S. Shin, Z. X. Yu, S. B. Li, *Chin. Med. J. (Engl.)* **120**, 2002–2005 (2007).
22. N. O. Bianchi, *Mutat. Res.* **682**, 21–27 (2009).
23. L. C. Silva Veiga, N. A. Bérgamo, P. P. Reis, L. P. Kowalski, S. R. Rogatto, *Braz. J. Med. Biol. Res.* **45**, 172–178 (2012).
24. P. H. Duijf, N. Schultz, R. Benezra, *Int. J. Cancer* **132**, 2316–2326 (2013).

ACKNOWLEDGMENTS

We thank K. Lindblad-Toh, R. M. Myers, C.-H. Heldin, D. H. Ledbetter, G. J. B. van Ommen, and U. Landegren for critical evaluation of the manuscript. This study was sponsored by the Swedish Cancer Society, the Swedish Research Council, the Swedish Heart-Lung Foundation and Science for Life Laboratory Uppsala to J.P.D. and the Olle Enqvist Byggmästare Foundation to L.A.F. Genotyping and next-generation sequencing were performed by the SNP&SEQ Technology Platform in Uppsala, Sweden, and supported by Wellcome Trust Grants WT098017, WT064890, and WT090532, Uppsala University, Uppsala University Hospital, the Swedish Research Council, and the Swedish Heart-Lung Foundation. The SNP&SEQ Technology Platform is part of Science for Life Laboratory at Uppsala University and supported as a national infrastructure by the Swedish Research Council. C.M.L. is a Wellcome Trust Research Career Development Fellow (086596/Z/08/Z). A.P.M. is a Wellcome Trust Senior Research Fellow in Basic Biomedical Science. A.P.M. acknowledges funding from the Wellcome Trust under awards WT064890, WT090532, and WT098017. TwinGene was supported by the Swedish Research Council (M-2005-1112), GenomEUtwin (EU/QLRT-2001-01254 and QLG2-CT-2002-01254), NIH DK U01-066134, the Swedish Foundation for Strategic Research (SSF), and the Heart and Lung foundation no. 20070481. J.P.D. and L.A.F. are cofounders and shareholders in Cray Innovation AB, as well as co-inventors on Patent Application No. PCT/EP2014/071448, protecting the commercial applications of LOY for the assessment of cancer risk. Genetic variants detected in this study are available at the Database of Genomic Structural Variation (dbVar) under accession code nstd92 for ULSAM and PIVUS cohorts and accession code nstd104 for the TwinGene cohort.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/81/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S6
References (*25*–*27*)

4 July 2014; accepted 24 November 2014
Published online 4 December 2014;
10.1126/science.1262092

AGING

Lysosomal signaling molecules regulate longevity in *Caenorhabditis elegans*

Andrew Folick,[1] Holly D. Oakley,[2,3] Yong Yu,[2,3] Eric H. Armstrong,[4] Manju Kumari,[5]* Lucas Sanor,[2] David D. Moore,[1,6] Eric A. Ortlund,[4] Rudolf Zechner,[5] Meng C. Wang[1,2,3]†

Lysosomes are crucial cellular organelles for human health that function in digestion and recycling of extracellular and intracellular macromolecules. We describe a signaling role for lysosomes that affects aging. In the worm *Caenorhabditis elegans*, the lysosomal acid lipase LIPL-4 triggered nuclear translocalization of a lysosomal lipid chaperone LBP-8, which promoted longevity by activating the nuclear hormone receptors NHR-49 and NHR-80. We used high-throughput metabolomic analysis to identify several lipids in which abundance was increased in worms constitutively overexpressing LIPL-4. Among them, oleoylethanolamide directly bound to LBP-8 and NHR-80 proteins, activated transcription of target genes of NHR-49 and NHR-80, and promoted longevity in *C. elegans*. These findings reveal a lysosome-to-nucleus signaling pathway that promotes longevity and suggest a function of lysosomes as signaling organelles in metazoans.

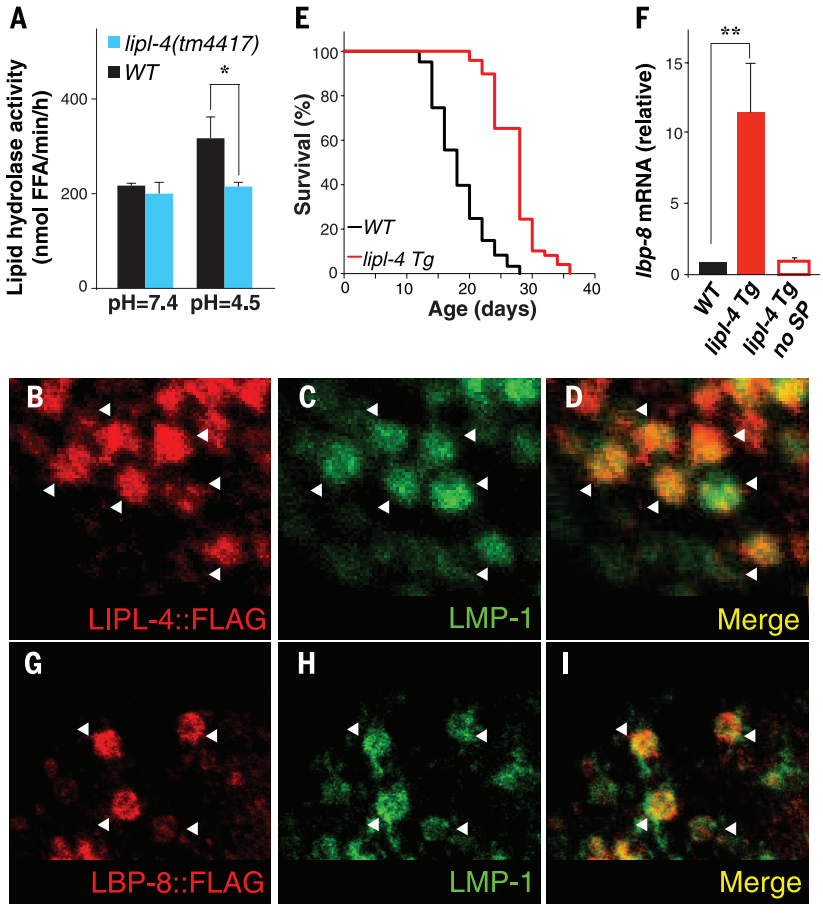
Lysosomes contain acid hydrolytic enzymes, digesting macromolecules taken up by endocytosis and recycling dysfunctional cellular components during autophagy (*1*). Lysosomal deficiency is associated with human diseases. For example, loss of human lysosomal acid lipase, *LIPA*, results in severe systemic metabolic malfunction known as infantile Wolman disease (*2*). Here, we explored how lysosomes might generate signaling molecules that regulate aging by influencing nuclear transcription.

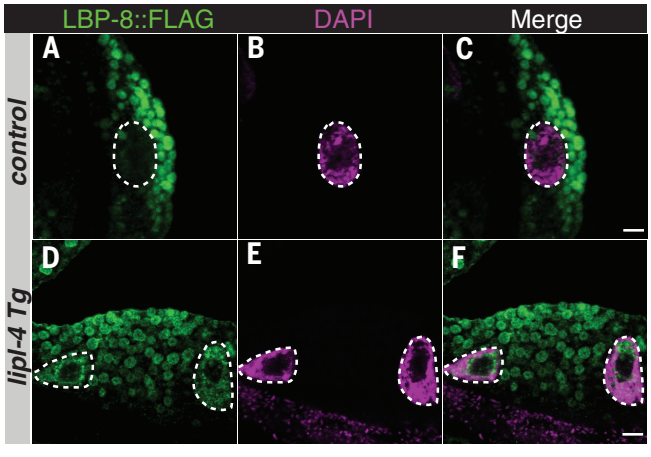
[1]Program in Developmental Biology, Baylor College of Medicine, Houston, TX 77030, USA. [2]Huffington Center on Aging, Baylor College of Medicine, Houston, TX 77030, USA. [3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. [4]Department of Biochemistry, Discovery and Developmental Therapeutics, Winship Cancer Institute, Emory University School of Medicine, Atlanta, GA 30322, USA [5]Institute of Molecular Biosciences, University of Graz, Graz, A-8010, Austria. [6]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA.
*Present address: Division of Endocrinology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA.
†Corresponding author. E-mail: wmeng@bcm.edu

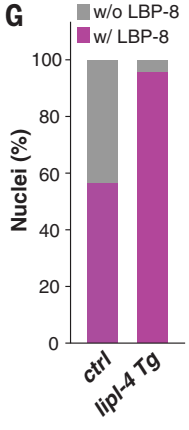
We analyzed a *Caenorhabditis elegans* longevity-promoting lipase, LIPL-4, which has sequence and functional similarities with human *LIPA* (fig. S1). Lipid hydrolase activity was decreased in *lipl-4(tm4417)* loss-of-function mutants at pH 4.5 but not at pH 7.4 (Fig. 1A). FLAG-tagged LIPL-4 protein was localized to intestinal lysosomes (Fig. 1, B to D, and fig. S2). Increased *lipl-4* expression is associated with longevity (*3*). A transgenic strain (*lipl-4 Tg*) that constitutively expressed *lipl-4* in the intestine had 55% mean life-span increase compared with wild-type (WT) animals (Fig. 1E and table S1) and delayed age-related decline of physical activity (fig. S3A). Constitutive expression of LIPL-4 without the signal peptide (*lipl-4 Tg no SP*), which was not targeted to the lysosome, caused little extension of life span (fig. S4 and table S1), which suggests that the lysosomal activity of LIPL-4 is essential for its longevity effect.

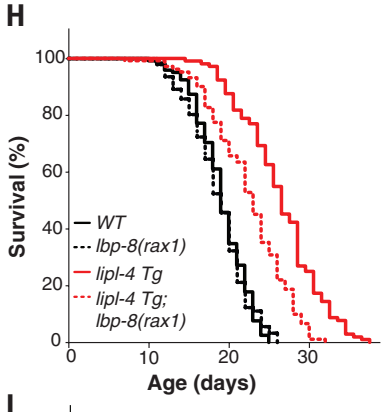
To elucidate whether lipid signals are affected by the LIPL-4 lipase, we examined fatty acid–binding proteins (FABPs) that are intracellular lipid chaperones shuttling lipid molecules between cellular

Fig. 1. Lysosomal lipid chaperone is increased in long-lived worms. (**A**) The amount of free fatty acids (FFAs) liberated from [^{3}H]triolein is significantly decreased in the *lipl-4(tm4417)* loss-of-function mutant compared with wild-type (WT) at pH = 4.5 but not at pH = 7.4. *$P < 0.05$, Student's *t* test. (**B** to **D**) Adult worms (*raxEx20[ges-1p::lipl-4::3xFLAG]*) were stained with antibodies against FLAG and LMP-1. LIPL-4 colocalizes with LMP-1, an established protein marker of lysosomes (*18*). Scale bar, 10 μm. (**E**) Mean life span is increased 55% in *lipl-4 Tg* worms (*raxIs3[ges-1p::lipl-4::sl2gfp]*) compared with WT. $P < 0.0001$, log-rank test. (**F**) *lbp-8* mRNA amounts were increased in *lipl4 Tg* compared with WT but not in the transgenic strain overexpressing *lipl-4* that lacks the signal peptide for lysosomal expression (*lipl-4 Tg no SP*). Error bars represent standard deviation (SD). **$P < 0.001$, Student's *t* test. (**G** to **I**) Adult worms (*raxEx31[lbp-8p::lbp-8::3xFLAG]*) were stained with antibodies against FLAG and LMP-1. LBP-8 colocalizes with LMP-1 in intestinal lysosomes. Scale bar, 10 μm.









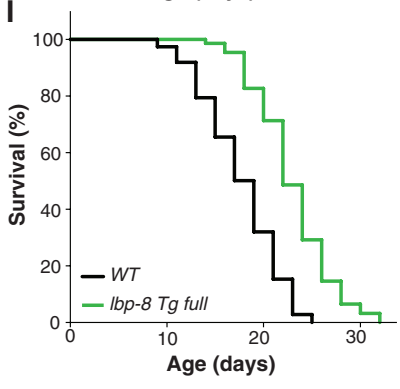
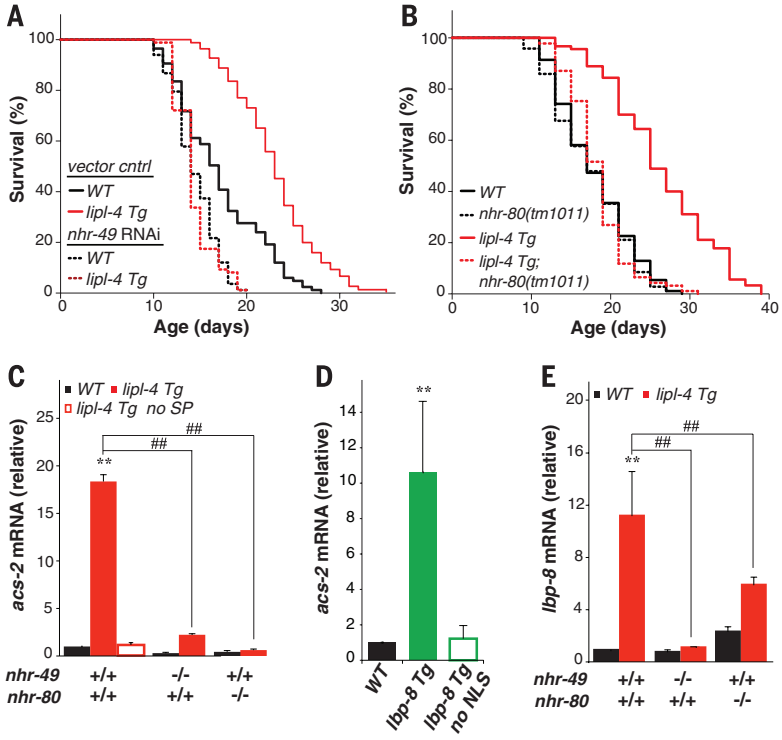


Fig. 2. Lysosomal lipid chaperone promotes longevity. (**A** to **F**) Adult worms expressing LBP-8::3xFLAG were stained with FLAG-specific antibody and 4′,6-diamidino-2-phenylindole (DAPI). More LBP-8 nuclear staining in *lipl-4 Tg* versus control worms. Scale bar, 10 μm. (**G**) Quantification of the percentage of nuclei positive for LBP-8::3xFLAG staining. $n = 100$. (**H**) The *lbp-8(rax1)* loss-of-function mutation reduces mean life-span extension in *lipl-4 Tg* by 46% ($P < 0.0001$) but has no effect on the life span of WT ($P > 0.05$). Log-rank test. (**I**) Mean life span is increased 30% in *lbp-8 Tg* (*raxIs4[lbp-8p::lbp-8::sl2gfp]*) compared with WT. $P < 0.0001$, log-rank test.



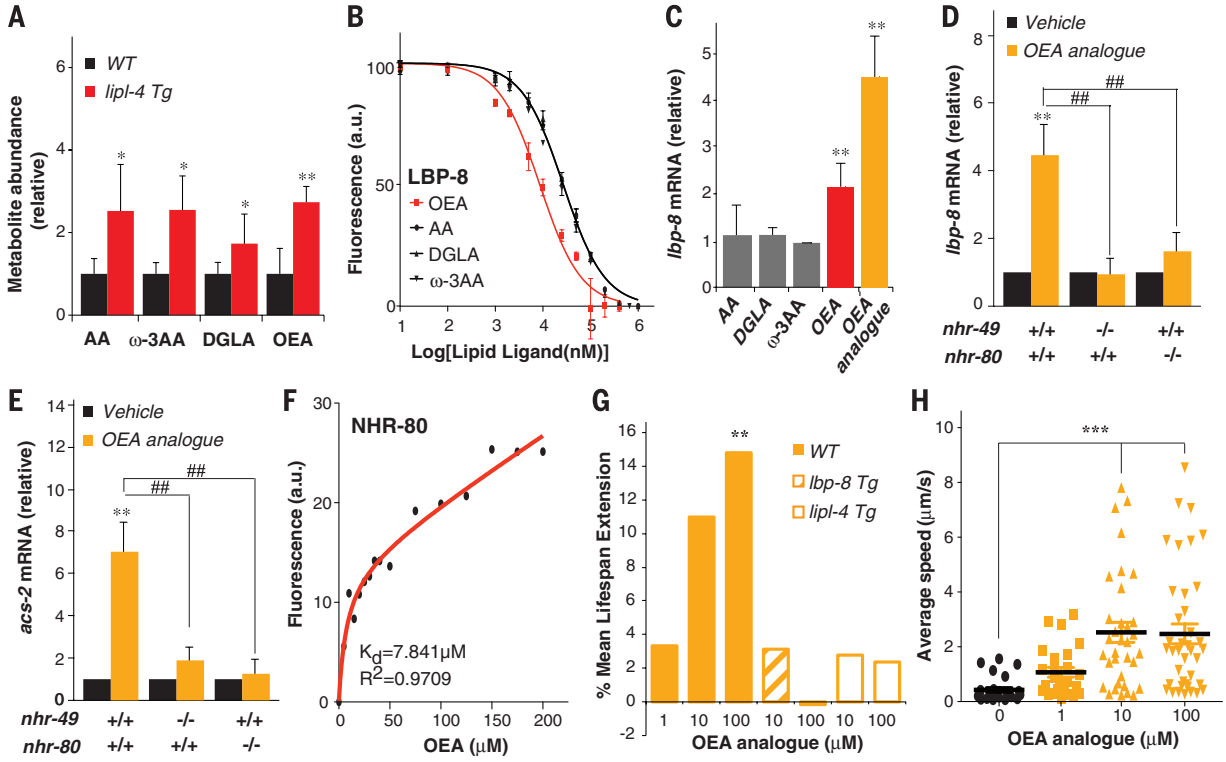


Fig. 3. Nuclear receptors act in lysosomal longevity signaling. (**A** and **B**) With adult-only RNAi inactivation of *nhr-49* or with the *nhr-80(tm1011)* loss-of-function mutation, the mean life spans of *lipl-4 Tg* and WT are not significantly different. $P > 0.5$, log-rank test. (**C**) Increased mRNA amount of NHR-49 target gene *acs-2* is suppressed by the *nhr-49(nr2041)* or *nhr-80(tm1011)* loss-of-function mutation in *lipl-4 Tg*, and is absent in *lipl-4 Tg no SP*. (**D**) *acs-2* mRNA amount is increased in *lbp-8 Tg* but not in the transgenic strain expressing *lbp-8* without the N-terminal NLS (*lbp-8 Tg no NLS*). (**E**) Increased mRNA amount of *lbp-8* in *lipl-4 Tg* is suppressed by the *nhr-49(nr2041)* or *nhr-80(tm1011)* mutant. Error bars represent SD. $^{**}P < 0.01$ by Student's *t* test; $^{\#\#}P < 0.01$ by two-way analysis of variance (ANOVA).

Fig. 4. OEA activates nuclear receptors and promotes longevity. (**A**) Increased levels of arachidonic acid (AA), ω-3 arachidonic acid (ω-3 AA), dihomo-γ-linoleic acid (DGLA), and OEA in *lipl-4 Tg* compared with WT. $^{*}P < 0.05$, $^{**}P < 0.001$, Welch's *t* test. (**B**) Decreased fluorescence derived from binding of 1-anilinonaphthalene-8-sulfonic acid (1,8-ANS) to LBP-8 by increasing OEA, AA, DGLA, and ω-3 AA competition. OEA has higher binding affinity than the other lipids by a factor of 3. (**C**) *lbp-8* mRNA amounts are increased in WT supplemented with OEA or the OEA analog, but not AA, DGLA, or ω-3 AA, $^{**}P < 0.01$, Student's *t test*. (**D** and **E**) mRNA amounts of *lbp-8* and *acs-2* are increased by OEA analog supplementation in WT, but not in the *nhr-49(nr2041)* or *nhr-80(tm1011)* mutant. $^{**}P < 0.01$, Student's *t* test. $^{\#\#}P < 0.01$, two-way ANOVA. (**F**) The intrinsic fluorescence intensity of GST-NHR-80 fusion proteins is decreased with increasing concentration of OEA [dissociation constant (K_d) of the binding reaction, 7.8 ± 4.1 μM]. R^2 is the coefficient of determination. (**G**) Supplementation of the OEA analog increases mean life span in WT, but not in *lipl-4 Tg* or *lbp-8 Tg*. $^{**}P < 0.01$, log-rank test. (**H**) Mean locomotion velocity is increased in WT treated with the OEA analog at 18 days of age. $^{***}P < 0.0001$, Student's *t* test. Error bars, SD. Error bars represent SD in (A) to (E), and in (H), standard error of the mean.

compartments for different functions (*4*, *5*). Of the nine *C. elegans* FABP homologs, only mRNA amounts of *lbp-8* were increased in *lipl-4 Tg* animals, but not in the *lipl-4 Tg no SP* strain (Fig. 1F). A green fluorescent protein (GFP) reporter strain showed that *lbp-8* was exclusively expressed in the intestine (fig. S5A). Both FLAG- and mCherry-tagged LBP-8 proteins were predominantly localized to intestinal lysosomes (Fig. 1, G to I, and fig. S5, B to J).

We also detected partial nuclear localization of LBP-8 in the intestine, which was enhanced in *lipl-4 Tg* animals (Fig. 2, A to G, and fig. S6, A to F). LBP-8 contains an N-terminal nuclear localization signal (NLS) (fig. S6G) and was present in both cytoplasmic and nuclear fractions of total worm lysate (fig. S6H). Both RNA interference (RNAi)–mediated depletion of LBP-8 and a newly isolated deletion mutant, *lbp-8(rax1)*, suppressed the life-span extension in *lipl-4 Tg* animals without affecting WT life span (Fig. 2H, fig. S7, and table S1). Thus, LBP-8 appears to be required for LIPL-4 lysosomal activity to confer longevity.

We found that a transgenic strain (*lbp-8 Tg)* constitutively expressing *lbp-8* had a 30% increase in mean life span compared with WT animals (Fig. 2I and table S1) and improved maintenance of physical activity in old age (fig. S3B). However, a transgenic strain that constitutively expresses LBP-8 lacking NLS (*lbp-8 Tg no NLS)* was excluded from nuclei and showed little or no life-span extension (fig. S8 and table S1). Thus, LBP-8 may function as a lysosomal lipid chaperone transducing lipid signals to the nucleus.

To test whether lysosomal signals might influence nuclear transcription, we screened several transcription factors implicated in longevity regulation (*6*–*11*). Nuclear hormone receptors *nhr-49* and *nhr-80*, previously demonstrated to physically interact (*10*), were both required for *lipl-4*– and *lbp-8*–mediated longevity. RNAi-mediated inactivation of *nhr-49* in adult worms shortened the life span of WT worms but also completely suppressed longevity extension in *lipl-4 Tg* and *lbp-8 Tg* worms (Fig. 3A, fig. S9A and table S1). The loss-of-function mutation *nhr-80(tm1011)* abrogated longevity extension without affecting the life span of WT worms (Fig. 3B, fig. S9B, and table S1). Neither *nhr-49* nor *nhr-80* is required for dietary restriction–induced longevity (*6*, *12*), which suggests that the LIPL-4–mediated longevity mechanism may act independently of dietary restriction. Concordantly, the longevity extensions by *lipl-4 Tg* and *eat-2(ad1116)*, a genetic model of dietary restriction in *C. elegans* (*13*), were additive (fig. S10).

acs-2 encodes an acyl-CoA synthetase required for mitochondrial β-oxidation and is a target gene of NHR-49 (*11*). *acs-2* transcription was increased more than 15-fold in *lipl-4 Tg* animals; this effect was dependent on *nhr-49* and *nhr-80,* and absent in the *lipl-4 Tg no SP* strain (Fig. 3C). Transcription of *acs-2* was also increased more than 10-fold in *lbp-8 Tg* but not in *lbp-8 Tg no NLS* animals (Fig. 3D). Thus, LIPL-4–induced activation of NHR-49 and NHR-80 can be reproduced by nuclear action of LBP-8. Transcriptional increase of *lbp-8* by *lipl-4 Tg* was in turn mediated by NHR-49 and NHR-80 (Fig. 3E).

To identify lipid molecules that might function in this lysosome-to-nucleus lipid signaling, we performed high-throughput metabolomic profiling analyses on WT and *lipl-4 Tg* worms. Among 352 metabolites detected, 71 had significantly altered abundance in *lipl-4 Tg* animals (table S2). Long-chain fatty acids and their derivatives are likely binding partners of FABPs (*4*). Thus, we focused our analysis on three C20 fatty acids—arachidonic acid, ω-3 arachidonic acid, and dihomo-γ-linolenic acid—and oleoylethanolamide (OEA), an *N*-acylethanolamine fatty acid derivative (Fig. 4A and table S2). In fluorescence-based binding assays, all four lipids bound to LBP-8, and the binding affinity of OEA for LBP-8 was 3 times that of the fatty acids (Fig. 4B).

Next, we tested the effects of the four lipids on transcription when directly applied to WT adult worms. We also used an OEA analog, KDS-5104, that is more resistant to hydrolysis than OEA (*14*). Only OEA and its analog were sufficient to increase the transcription of *lbp-8* in WT worms, and the analog exerted a stronger effect (Fig. 4C). After 3 hours of treatment with the OEA analog, transcription of *lbp-8* and *acs-2* was increased more than 4- and 7-fold above the control levels, respectively (Fig. 4, D and E). This effect was abrogated in the *nhr-49(nr2041)* or *nhr-80(tm1011)* mutant (Fig. 4, D and E). Thus, accumulation of OEA in response to LIPL-4 may act to promote transcription via NHR-49/NHR-80.

To test whether OEA directly binds to NHR-49 or NHR-80, or both, we measured intrinsic fluorescence changes of glutathione *S*-transferase (GST)–NHR fusion proteins in the presence of OEA. OEA binding significantly decreased the fluorescence intensity of the NHR-80 fusion protein in a dose-dependent manner [equilibrium dissociation constant (K_d) of 7.8 μM] (Fig. 4F). In a differential protease-sensitivity assay, chymotrypsin digestion of [^{35}S]NHR-80 in the presence of the OEA analog resulted in protease-resistant fragments of approximately 45 and 35 kD (fig. S11), which indicated direct binding between NHR-80 and the OEA analog. However, no binding was detected between NHR-49 and OEA or the OEA analog (fig. S12). Thus, NHR-80 appears to act as a direct nuclear receptor of OEA and NHR-49 may function as a cofactor of NHR-80.

N-Acylphosphatidylethanolamine–specific phospholipase D (NAPE-PLD) mediates OEA synthesis (*15*). In *C. elegans*, *nape-1* and *nape-2* encode NAPE-PLD (*16*). The *nape-1(tm3860)* loss-of-function mutation suppressed the life-span extension in *lipl-4 Tg* and *lbp-8 Tg* by half (fig. S13 and table S1). Additionally, a loss-of-function mutant *lipl-4 (tm4417)* reduced the longevity of *lbp-8 Tg* by 68% (fig. S14), which supports the possibility that LIPL-4 activity promotes the generation of longevity-promoting OEA carried by LBP-8.

Direct treatment of WT worms with the OEA analog prolonged life span (Fig. 4G and table S1) and improved physical activity maintenance in aged animals (Fig. 4H). In contrast, neither *lipl-4 Tg* nor *lbp-8 Tg* life span was affected by OEA analog supplementation (Fig. 4G and table S1), which suggests that OEA may promote longevity by the same mechanism as occurs in *lipl-4 Tg* and *lbp-8 Tg* animals. Note that OEA supplementation decreased life span in the *nhr-80(tm1011)* mutant (fig. S15 and table S1), which indicated that OEA requires NHR-80 to promote longevity, and it can have detrimental effects in the absence of NHR-80. Thus, OEA may act as a lipid messenger to transduce lysosome-to-nucleus signaling in promoting longevity.

Overall, our studies suggest that bioactive lipid messengers and lipid chaperones link lysosomal activity and nuclear transcription to promote longevity. All the components of this lysosome-to-nucleus signaling pathway are well conserved in mammals. Notably, mammalian peroxisome proliferator–activated receptor α is activated by OEA (*17*), whereas NHR-80 is homologous to mammalian HNF4, which suggests that different nuclear receptors bind the same ligands despite divergent ligand-binding domains. Considering that FABPs are quite promiscuous in ligand binding (*4*), there may be other lipid molecules binding to LBP-8 and functioning in this longevity pathway.

REFERENCES AND NOTES

1. C. Settembre, A. Fraldi, D. L. Medina, A. Ballabio, *Nat. Rev. Mol. Cell Biol.* **14**, 283–296 (2013).
2. A. D. Patrick, B. D. Lake, *Nature* **222**, 1067–1068 (1969).
3. M. C. Wang, E. J. O'Rourke, G. Ruvkun, *Science* **322**, 957–960 (2008).
4. M. Furuhashi, G. S. Hotamisligil, *Nat. Rev. Drug Discov.* **7**, 489–503 (2008).
5. J. Storch, B. Corsico, *Annu. Rev. Nutr.* **28**, 73–95 (2008).
6. J. Goudeau *et al.*, *PLOS Biol.* **9**, e1000599 (2011).
7. C. Kenyon, J. Chang, E. Gensch, A. Rudner, R. Tabtiang, *Nature* **366**, 461–464 (1993).
8. S. Ogg *et al.*, *Nature* **389**, 994–999 (1997).
9. S. H. Panowski, S. Wolff, H. Aguilaniu, J. Durieux, A. Dillin, *Nature* **447**, 550–555 (2007).
10. P. P. Pathare, A. Lin, K. E. Bornfeldt, S. Taubert, M. R. Van Gilst, *PLOS Genet.* **8**, e1002645 (2012).
11. M. R. Van Gilst, H. Hadjivassiliou, A. Jolly, K. R. Yamamoto, *PLOS Biol.* **3**, e53 (2005).
12. B. N. Heestand *et al.*, *PLOS Genet.* **9**, e1003651 (2013).
13. B. Lakowski, S. Hekimi, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13091–13096 (1998).
14. G. Astarita *et al.*, *J. Pharmacol. Exp. Ther.* **318**, 563–570 (2006).
15. J. Fu *et al.*, *J. Biol. Chem.* **282**, 1518–1528 (2007).
16. M. Lucanic *et al.*, *Nature* **473**, 226–229 (2011).
17. J. Fu *et al.*, *Nature* **425**, 90–93 (2003).
18. G. Hadwiger, S. Dour, S. Arur, P. Fox, M. L. Nonet, *PLOS ONE* **5**, e10161 (2010).

ACKNOWLEDGMENTS

We thank H. Y. Mak and A. Antebi for providing strains; A. Dervisefendic and H. Jen for experimental support; N. Timchenko, J. Wang, Z. Yu, and D. Chow for instrumental support; H. Dierick, C. Herman, H. Zheng, F. Xia, S. Rosenberg, and H. Zoghbi for critical reading of the manuscript; P. P. Metoyer for scientific editing. Supported by NIH grants T32GM008602 (E.H.A.), R01DK095750 (E.A.O.), T32HD055200 (A.F.), F30AG046043 (A.F.), R00AG034988 (M.C.W.), and R01AG045183 (M.C.W.); Ellison New Scholar Award (M.C.W.); Welch Chair in Science (Q-0022) (D.D.M.); European Research Council Advanced Grant (R.Z.); and Fondation Leducq (R.Z.). Requests for materials should be addressed to M.C.W.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/83/suppl/DC1
Materials and Methods
Figs. S1 to S16
Tables S1 and S2
References (*19*–*23*)

16 July 2014; accepted 13 November 2014
10.1126/science.1258857

Magnetic Beads

MagSi-Direct provides a convenient and efficient way to attach virtually any biomolecule to nanoscopic paramagnetic beads. It is particularly suited for exploiting interactions between a biomolecule and its binding partners. With MagSi-Direct, you can start with a molecule of interest, attach magnetic beads to it, and then use a magnet to isolate that molecule from any desired reaction mixture, together with any other molecules, complexes, or even intact cells to which the starting molecule has bound. In effect, the starting molecule is converted into a magnetic affinity reagent, or "bait," which can be used to purify, isolate, or characterize the partners with which it interacts. MagSi-Direct beads attach to the starting molecule via strong coordinate bonds between the surface of the bead and any electron donating group on your molecule (carboxyl, amide, 1°, 2°, 3° amine, hydroxyl, phosphate, halogen, etc). Molecular orientation is non-specific and random.

AMS Biotechnology
For info: +44-(0)-1235-828200
www.amsbio.com/magnetic-beads.aspx

Ultrasafe High-Pressure Reactor

PressureSyn is a state-of-the-art, 125 mL working volume high-pressure reactor that uniquely combines outstanding performance, ease of use, and the highest level of operational safety for users. PressureSyn reactors provide an ideal tool for stirred, or non-stirred, high-pressure applications including hydrogenations, carbonylations, catalyst screening, and polymerizations.
The standard safety features of PressureSyn include a bursting disk and a pressure relief valve. PressureSyn reactors feature a unique bracket and key operated locking system ensuring easy assembly. However, this novel clamping arrangement also prevents the clasp from being disassembled while the reactor is still under pressure. Each reactors locking collar has a unique key which ensures only that key can be used to open that specific individual reactor. With customer safety paramount, each individual PressureSyn reactor is tested to 170 bar, witnessed and certified by Zurich Insurance, and is rated for use up to a maximum pressure of 100 bar and temperature of 200°C.

Asynt
For info: +44-(0)-1638-781709
www.asynt.com/product/pressuresyn

Paraffin Section Flotation Bath

A new paraffin section flotation bath makes the manipulation and placement of paraffin wax sections onto glass slides both simple and efficient. With its user-friendly design and optimized performance, it builds on the success of Electrothermal's original appliance and continues the brands legacy of robust, safe, and efficient laboratory equipment. The Paraffin Section Flotation Bath benefits from a PTFE black interior, which provides high contrast for easy viewing of samples. Virtually maintenance-free and simple to clean, it delivers ultimate ease of use and reliability. The bath also has improved heat-up rates (raising 2.25 L of water to a nominal temperature of 45°C in just 40 minutes), and enhanced temperature stability of ± 0.5°C. These improved levels of control also prevent overheating and lower the risk of damaging samples. Users can not only prepare slides for viewing in a shorter timeframe, but can have upmost confidence in the safe handling of their biological materials.

Bibby Scientific
For info: +44-(0)-1785-812121
www.bibby-scientific.com

High-Performance Liquid Chromatography System

A new high-performance liquid chromatography (HPLC) system has been designed from the ground up to provide new levels of performance, productivity, and usability when used as a standalone system or with the latest mass spectrometers. The monolithic case contains an instrument that combines the ruggedness of an integrated system with the flexibility and serviceability of a modular system. Vanquish stands about 25% lower than comparable modular stacks for safety and convenience in the laboratory. Central to the Vanquish concept is the new family of Accucore Vanquish UHPLC columns, specially designed to optimize performance. The new columns feature 1.5 µm solid core particles utilizing Core Enhanced Technology to take full advantage of the Vanquish system's 1,500 bar (22,000 psi) maximum pump pressure and flow rate up to 5 mL/min for ultrashort diffusion path lengths and highly efficient separations.

Thermo Fisher Scientific
For info: 800-556-2323
www.thermoscientific.com/vanquish

Automated Electrophoresis System

The new LabChip GX Touch and GXII Touch electrophoresis systems automate conventional gel electrophoresis, eliminating the need for manual sample preparation. The LabChip GX Touch and GXII Touch reduce hands-on time and simplify data generation via a user friendly operator touch screen. The LabChip platform provides the data and throughput necessary to perform effective nucleic acid quantification and protein characterization. The LabChip GX system can enhance genomics research process by precisely quantifying nucleic acids. The GXII Touch system helps to accelerate the development of biotherapeutic drugs by automating the protein characterization process. The LabChip GXII Touch system's rapid reporting of multiple proteins' critical quality attributes can support customer needs throughout the product life cycle. The LabChip Touch platform's top differentiator is its ability to perform electrophoresis in microfluidic channels. The LabChip GX Touch system provides rapid, quantifiable data for pre- and post-polymerase chain reaction sample analysis, which is critical for next generation sequencing and genomics workflows.

PerkinElmer
For info: 800-762-4000
www.perkinelmer.com/labchipsystems

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/products/newproducts.dtl for more information.

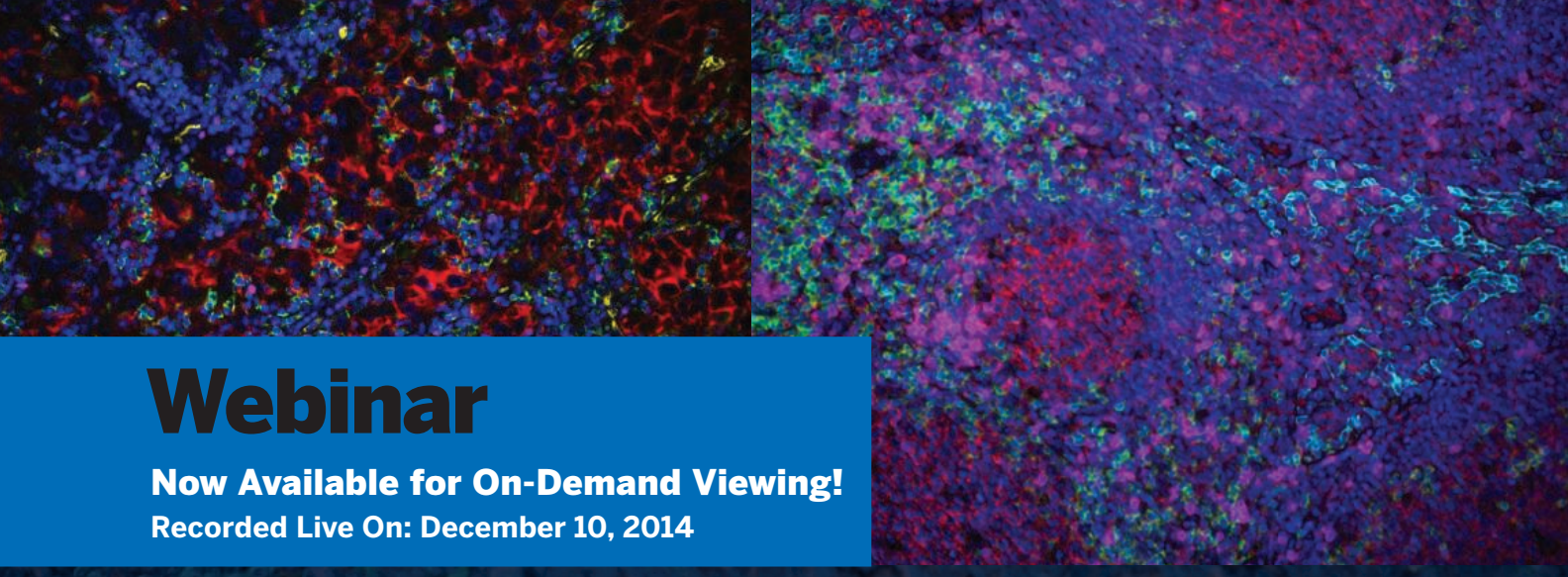
Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

This is the start of something big.

Science Advances | AAAS

SIGNIFICANT RESEARCH, GLOBAL IMPACT

Introducing *Science Advances* – the new, online-only, open-access journal from *Science* and AAAS. Find out how you can be among the first authors published at **scienceadvances.org.**



Webinar

Now Available for On-Demand Viewing!
Recorded Live On: December 10, 2014

Untangling the tumor microenvironment

Illuminating the complex interactions and functions of immune cells

During the webinar viewers will learn about:

- Cutting-edge research targeting the B-cell receptor signaling pathway that has recently demonstrated therapeutic promise
- Methods that can provide in depth information on cancer phenotypes, including simultaneous immunohistochemistry of multiple biomarkers, multiplexed imaging, single cell quantitative analysis, and automated phenotyping
- How host-tumor interaction analysis in breast cancers could form the basis for assays to guide therapy and monitor response.

Certain antibody therapies have demonstrated the potential for directing a patient's own immune system against tumors. Further advances in this area will depend upon a detailed understanding of the tumor microenvironment and characterization of the location and status of immune cells and their interaction with tumor cells. This will require methods that provide phenotyping of immune and cancer cells combined with information about their spatial relationship in tumor regions. Additionally, a deeper understanding of the signaling cascades active in immune recognition of cancers is crucial. During this webinar, we will discuss the bringing together of multiplexed fluorescent immunohistochemistry, advanced microscopy techniques, and bioinformatics, and how these are now enabling new insights into cancer biology and immunology.

Speakers

Scott J. Rodig, M.D., Ph.D.
Dana-Farber Cancer Institute
Boston, MA

Edward C. Stack, Ph.D.
PerkinElmer
Hopkinton, MA

View Now! webinar.sciencemag.org

Webinar sponsored by

PerkinElmer
For the Better

Brought to you by the *Science*/AAAS Custom Publishing Office

Science
AAAS

@SciMagWebinars

Learn more and conduct your job search the easy way.

- Search thousands of job postings
- Create job alerts based on your criteria
- Get career advice from our Career Forum experts
- Download career advice articles and webinars
- Complete an individual development plan at "myIDP"

Target your job search using relevant resources on **ScienceCareers.org**.

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

HUMAAAN PROTEOME

Interested in the human proteome? Order a free copy of this poster to discover how leading scientists from the Human Protein Atlas project have mapped the complete human proteome using Triple A Polyclonals. Today we provide over 17,000 antibodies, which target more than 75% of the human protein coding genes, to researchers worldwide.

Triple A Polyclonals offer advanced levels of specificity, reproducibility and versatility.

Learn more today and request your free poster copy at atlasantibodies.com/proteome-poster

ATLAS ANTIBODIES
Totally human

Science Careers Advertising

For full advertising details, go to ScienceCareers.org and click For Employers, or call one of our representatives.

Tracy Holmes
Worldwide Associate Director
Science Careers
Phone: +44 (0) 1223 326525

THE AMERICAS
E-mail: advertise@sciencecareers.org
Fax: 202 289 6742

Tina Burks
Phone: 202 326 6577

Nancy Toema
Phone: 202 326 6578

Marci Gallun
Sales Administrator
Phone: 202 326 6582

Online Job Posting Questions
Phone: 202 312 6375

EUROPE / INDIA / AUSTRALIA / NEW ZEALAND / REST OF WORLD
E-mail: ads@science-int.co.uk
Fax: +44 (0) 1223 326532

Axel Gesatzki
Phone: +44 (0) 1223 326529

Sarah Lelarge
Phone: +44 (0) 1223 326527

Kelly Grace
Phone: +44 (0) 1223 326528

JAPAN
Katsuyoshi Fukamizu (Tokyo)
E-mail: kfukamizu@aaas.org
Phone: +81 3 3219 5777

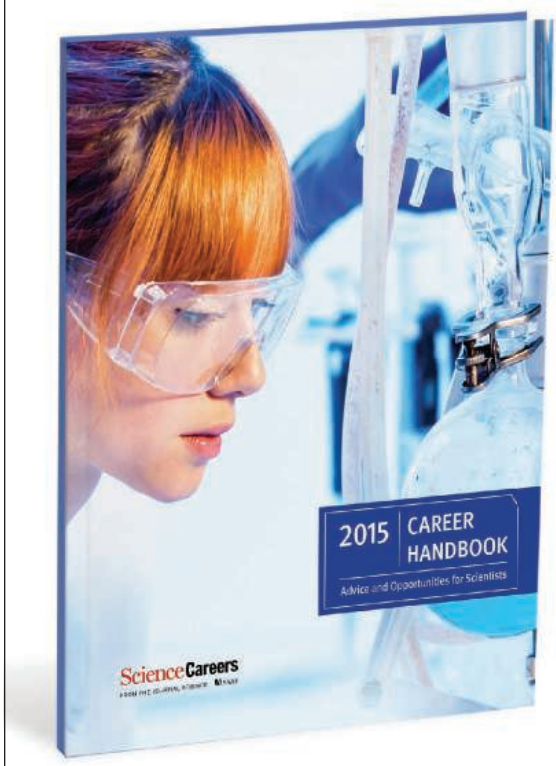
Hiroyuki Mashiki (Kyoto)
E-mail: hmashiki@aaas.org
Phone: +81 75 823 1109

CHINA / KOREA / SINGAPORE / TAIWAN / THAILAND
Ruolei Wu
Phone: +86 186 0082 9345
E-mail: rwu@aaas.org

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

ScienceCareers.org

Recruit and Promote Your Brand All Year Long



Book now to secure your listing.

Ad materials due January 14.

NEW COVER POSITIONS AVAILABLE

Benefits to your company

- Opportunity to brand your organization to scientists beyond your normal reach.
- Print bonus distribution of 3,000 copies to career fairs and meetings around the globe.
- Digital copy e-mailed to 100,000 scientists including all *Science* Careers registrants.
- Your logo included in two Career Path newsletters and online on the booklet landing page

To book your ad, contact:

advertise@sciencecareers.org

THE AMERICAS
202 326 6582

EUROPE/ROW
+44 (0) 1223 326500

JAPAN
+81 3 3219 5777

CHINA / KOREA / SINGAPORE / TAIWAN
+86 186 0082 9345

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

SCIENCECAREERS.ORG

Shandong University of Technology Recruits for High-end Talents Globally

Founded in 1956 and located in the cradle of Ancient Qi Culture-Zibo City of Shandong Province, Shandong University of Technology is a provincially prioritized university with a focus on science and engineering. They cover eight fields of learning, such as engineering, science, economics, management, literature, law, history and pedagogy. It also holds the name of Shandong Provincial Featured University of Applied Talents Cultivation.

I.Disciplines of Recruitment

Mechanical Engineering, Agricultural Engineering, Chemical Engineering and Technology, Mechanics, Materials Science and Engineering, Power Engineering and Engineering Thermophysics, Instruments Science and Technology, Food Science and Engineering, Electric Engineering, Electronic Science and technology, Information and Communication Engineering, Control Science and Engineering, Computer Science and Technology, Civil Engineering, Surveying and Mapping technology, Transportation Engineering, Architecture, Mining Engineering, Textile Engineering, Biology, Economics, Management Science and Engineering, Mathematics, Physics, Law, Literature and other disciplines.

II.Requirements and Treatment

(1). for the leading personnel who holds the post of professor or equivalent position at renowned university from home and abroad, the university will provide a 50,000-1,200,000 RMBs annual salary plus 1,000,000 settling-in allowance and housing of 273m2 . Start-up funding will be offered according to the research work.
(2). for the academic leader who achieved outstanding performance with high academic influence in certain area, the university will provide a 400,000-600,000 RMBs annual salary plus 500,000 settling-in allowance and housing of 248m2 , a reasonable research funds will be offered as well.
(3). for excellent young talents with doctorate, the university will provide 300,000-500,000 RMBs settling-in and housing allowance. Start-up funding will be 50,000-100,000 RMBs for science and engineering, and 30,000-50,000 RMBs for humanities and social science. An associate professor treatment will be offered within 3 years after the talents work in the university.

III.The university will provide support if the talents accord with condition of applying National Thousand Talent Program or Taishan Scholar Program and the applicants will enjoy the same treatment.

IV.Contact Information

Tel:+86-533-2782311/2781219
E-mail: rshch@sdut.edu.cn

Joint Faculty Positions at the Center for Quantitative Biology and Peking-Tsinghua Center for Life Sciences

北京大学 PEKING UNIVERSITY

JOINT FACULTY POSITIONS

The Center for Quantitative Biology (CQB) and Peking-Tsinghua Center for Life Sciences (CLS) at Peking University jointly open for applications for faculty positions at all ranks.

We seek for creative individuals in all areas of quantitative biology, with emphases on (but not limited to):
(1) Systems biology;
(2) Synthetic biology;
(3) Computational biology/Bioinformatics;
(4) Disease mechanism and drug design from system biology perspective;
(5) Development of quantitative methods and technology.

CQB (http://cqb.pku.edu.cn/) is dedicated to research and education at the interface between the traditionally more quantitative disciplines (such as mathematics, physical sciences, engineering, computer science) and the biological sciences. CLS (http://www.cls.edu.cn/) is a center of excellence to support and nurture creative research of long lasting impact.

Application materials (with cover letter, summary of research interests, CV and less than 5 representative publications, all in a single PDF file) should be sent to Ms. Wei Xiao (gsmkyb@pku.edu.cn), to whom you should also ask your references to send in their recommendation letters.

西南交通大学 Southwest Jiaotong University, P.R.China Anticipates Your Working Application

Southwest Jiaotong University (SWJTU), founded in 1896, situates itself in Chengdu, the provincial capital of Sichuan. It is a national key multidisciplinary "211" and "985 Feature" Projects university directly under the jurisdiction of the Ministry of Education, featuring engineering and a comprehensive range of study programs and research disciplines spreading across more than 20 faculties and institutes/centers. Boasting a complete Bachelor-Master-Doctor education system with more than 2,600 members of academic staff, our school also owns 2 first-level national key disciplines, 2 supplementary first-level national key disciplines (in their establishment), 15 first-level doctoral programs, 43 first-level master programs, 5 key undergraduate programs, 10 post-doctoral stations and more than 40 key laboratories at national and provincial levels.

Our university is currently implementing the strategy of "developing and strengthening the university by introducing and cultivating talents". Therefore, we sincerely look forward to your working application.

More information available at http://www.swjtu.edu.cn/

I. Positions and Requirements

A.High-level Leading Talents

It is required that candidates be listed in national top talents programs such as ***Program of Global Experts, Top Talents of National Special Support Program, "Chang Jiang Scholars", China National Funds for Distinguished Young Scientists and National Award for Distinguished Teacher.***

Candidates are supposed to be no more than 60 years old. The limitation could be extended in the most-needed areas of disciplinary development.

Candidates who work in high-level universities/institutes and reach the above requirements are supposed to be no more than 45 years old.

B. Young Leading Scholars

Candidates are supposed to be listed in or qualified to apply for the following programs:

- ***National Thousand Young Talents Program***
- ***The Top Young Talents of National Special Support Program(Program for Supporting Top Young Talents)***
- ***Science Foundation for the Excellent Youth Scholars***

Candidates should have good team spirit and leadership, outstanding academic achievements, broad academic vision and international cooperation experience and have the potential of being a leading academic researcher.

C. Excellent Young Academic Backbones

Candidates under 40 years old are expected to graduate from high-level universities/institutes either in China or other countries. Those who are professors, associate professors and other equivalent talents from high-level universities/institutes overseas could be employed as professors and associate professors as well.

D. Excellent Doctors and Post Doctoral Fellows

Candidates under 35 years old are supposed to be excellent academic researchers from high-level universities either in China or other countries.

II. Treatments

The candidates will be provided with competitive salaries and welfares that include settling-in allowance, subsidy of rental residence, start-up funds of scientific research, assistance in establishing scientific platform and research group as well as international-level training and promotion . As for outstanding returnees, we can offer further or specific treatments that can be discussed personally.

III . Contact us:

Contacts: Ye ZENG & Yinchuan LI Telephone number: 86-28 66366202 Email: *talent@swjtu.edu.cn*

Address: Human Resources Department of SWJTU, the western park of high-tech zone, Chengdu, Sichuan, P.R.China, 611756

http://www.swjtu.edu.cn/

徐州医学院 XUZHOU MEDICAL COLLEGE 江苏省骨髓干细胞重点实验室 Key Laboratory of Bone Marrow Stem Cells, Jiangsu Province

Tenure-Track positions at the Assistant/Associate/Full Professor level in Key Laboratory of Bone Marrow Stem Cells, Jiangsu Province, Xuzhou Medical College

The Key Laboratory of Bone Marrow Stem Cells (KLBMSC), Jiangsu Province, at Xuzhou Medical College, Xuzhou, China (http://www.xzmc.edu.cn) is seeking applications for a full-time tenure-track position at academic ranks of Assistant, Associate or Full Professor level. Highly motivated and talented individuals with demonstrated accomplishments in, but not limited to, the following fields are encouraged to apply: hematology, hematopoietic and/or leukemia stem cells, graft-versus host disease (GVHD), endothelial progenitor and/or endothelial cells.

Applicants must have a PhD and/or MD degree and postdoctoral experience in relevant field. Meanwhile, applicants must have strong research capabilities as demonstrated by publications in high-impact journals and can use cutting-edge approaches involving transplantation mice model, knock-out mice, flow cytometry, immunofluorescence and confocal microscopy to conduct the following-related studies: 1) GVHD, 2) endothelial progenitor/endothelial cells and 3) leukemia stem cells.

KLBMSC, headed by Prof. Kailin Xu, was approved in 2014 with major research areas being GVHD, endothelial progenitor cells and leukemia stem cells. Currently, KLBMSC has been established a series of excellent platforms including animal model platform (conditional knock-out mice), pathology platform, cell culture and protein platform, etc. Meanwhile, researchers from KLBMS have been chairing 19 National Natural Science Foundation of China (NSFC) and 20 provincial foundations in recent three years.

The College and KLBMSC provide a very supportive research environment with excellent resources conducive to developing a successful research program. Qualified candidates will receive a comprehensive package including competitive salary, medical insurance and other benefits.

Interested individuals should submit a detailed letter of expression of interest or research proposal, CV, list of publications, and three reference letters to: rsc@xzmc.edu.cn and lihmd@163.com. For further information, please contact Professor Kailin Xu, Director, Key Laboratory of Bone Marrow Stem Cells, Jiangsu Province, at lihmd@163.com or Mr. Yuan-Dong Li, Department of HR, Xuzhou Medical College, at rsc@xzmc.edu.cn.

UNI BASEL

ETH zürich

Assistant Professor (Tenure Track) of Engineering of Synthetic Systems

The Department of Chemistry of the Faculty of Science at the University of Basel (www.chemie.unibas.ch) and the Department of Biosystems Science and Engineering (www.bsse.ethz.ch) at ETH Zurich invite applications and nominations for the above-mentioned assistant professorship. The position will be jointly hosted by both institutions in the framework of the National Center of Competence in Research (NCCR) "Molecular Systems Engineering" (www.nccr-mse.ch). The location of both Departments in Basel, the heart of the BioValley area, provides excellent opportunities for collaborations within this strong life science research community at the academic, clinical, chemical, and pharmaceutical level. Collaborations within the NCCR "Molecular Systems Biology" are essential and collaborations in other important networks, such as SystemsX.ch (www.systemsx.ch), are strongly encouraged.

The successful candidate should have a strong experimental background in the engineering of in vitro and/or in vivo (typically bacterial) systems with advanced novel functionalities such as compartmentalized reaction cascades, communication and transport across compartmental boundaries, cellular orthogonalization strategies, or applied chassis design. Ideally, the candidate should have a strong background in disciplines such as chemistry and (bio)molecular engineering. The assistant professor will be a member of the Faculties of both host departments. The ideal candidate should have demonstrated success in obtaining support for independent research projects and a strong publication record reflecting innovative, interdisciplinary, and collaborative approaches to engineering synthetic systems.

The successful candidate will be expected to teach undergraduate level courses (German or English) and graduate level courses (English) in Chemistry, Biotechnology, and in the newly to be established courses in Molecular Systems Engineering.

This assistant professorship has been established to promote the careers of younger scientists. The University of Basel has an established program for career progression allowing promotion to associate and full professor upon successful evaluation.

Please submit your application together with a curriculum vitae, a list of publications, and a statement on future teaching and research activities **no later than 28 February 2015 to Prof. Dr. Jörg Schibler, University of Basel, Dean of the Faculty of Sciences, Klingelbergstrasse 50, 4056 Basel, dekanat-philnat@unibas.ch**. When applying electronically, do only send one PDF file. The University of Basel is an equal opportunity and family friendly employer and is further responsive to the needs of dual career couples. We specifically encourage women to apply.

For further information, please contact: Prof. Dr. Thomas Ward, Head, Department of Chemistry, E-mail: thomas.ward@unibas.ch.

UNI BASEL

ETH zürich

Assistant Professor (Tenure Track) of the Synthesis of Functional Modules

The Department of Chemistry of the Faculty of Science at the University of Basel (www.chemie.unibas.ch) invites applications and nominations for the above-mentioned assistant professorship. The department is located in Basel, the heart of the BioValley area providing excellent opportunities for collaborations within this strong life science research community at the academic, clinical, chemical and pharmaceutical level. The assistant professor will also be a member of the Basel-based Department of Biosystems Science and Engineering (www.bsse.ethz.ch) at ETH Zurich and of the recently granted National Center of Competence in Research (NCCR) "Molecular Systems Engineering" (www.nccr-mse.ch). Collaborations within the NCCR are essential, and collaborations in other important networks are strongly encouraged.

The successful candidate should have a strong background in the synthesis and exploitation of complex molecular systems. Areas of particular interest include: molecular prosthetics, non-equilibrium systems, hierarchical self-assembly, cross-regulation, systems chemistry, self-repair etc. Having a strong background in chemistry, the future colleague is expected to interact with biologists with the ultimate goal of assembling molecular systems whose functionality approaches that of cells.

Furthermore, candidates should have demonstrated success in obtaining support for independent research projects and a strong publication record reflecting innovative, interdisciplinary, and collaborative approaches.

The selected candidate will be expected to participate in teaching Chemistry (all levels of the BSc, MSc and PhD programs), and in the newly to be established courses in Molecular Systems Engineering.

This assistant professorship has been established to promote the careers of younger scientists. The University of Basel has an established program for career progression allowing promotion to associate and full professor upon successful evaluation.

Please submit your application together with a curriculum vitae, a list of publications, a statement on future teaching and research activities **no later than 28 February 2015 to Prof. Dr. Jörg Schibler, University of Basel, Dean of the Faculty of Sciences, Klingelbergstrasse 50, 4056 Basel, dekanat-philnat@unibas.ch**. When applying electronically, do only send one PDF file. The University of Basel is an equal opportunity and family friendly employer and is further responsive to the needs of dual career couples. We specifically encourage women to apply.

For further information, please contact: Prof. Dr. Thomas Ward, Head, Department of Chemistry, E-mail: thomas.ward@unibas.ch.

ETH zürich

UNI BASEL

Assistant Professor (Tenure Track) of Biosystems Engineering

The Department of Biosystems Science and Engineering (www.bsse.ethz.ch) at ETH Zurich invites applications for the above-mentioned assistant professorship. The department is located in Basel, the heart of the BioValley area providing excellent opportunities for collaborations within this strong life science research community at the academic, clinical and pharmaceutical level. Its main research thrust is the understanding and engineering of biological systems for medical or chemical purposes, which generates numerous opportunities for interactions. The assistant professor will also be a member of the Department of Chemistry at the University of Basel (www.chemie.unibas.ch) and of the newly established National Center of Competence in Research (NCCR) "Molecular Systems Engineering" (www.nccr-mse.ch).

The successful candidate is experimentally working on the engineering of biomolecular systems for the design of complex molecular factories or/and the control of cellular systems. Ideally, he or she should have a strong background in disciplines like molecular engineering, molecular cell biology, genetics, chemistry, and be interested in engineering synthetic and/or cellular systems. Furthermore, candidates should have demonstrated success in obtaining support for independent research projects and a strong publication record reflecting innovative, interdisciplinary, and collaborative approaches to identify and engineer important biomolecular systems. Collaborations within the NCCR "Molecular Systems Engineering", are essential and collaborations in other important networks, such as SystemsX.ch "The Swiss Initiative in Systems Biology" (www.systemsx.ch), are encouraged.

The educational goal of the Department of Biosystems Science and Engineering is to teach students at both the undergraduate and graduate levels by integrating expertise and knowledge from biologists, chemists, physicists, engineers, computer scientists and mathematicians, along with industrial collaborators. The new assistant professor will be expected to teach undergraduate level courses (German or English) and graduate level courses (English) in Biotechnology and in the newly to be established courses in Molecular Systems Engineering.

This assistant professorship has been established to promote the careers of younger scientists. The initial appointment is for four years with the possibility of renewal for an additional two-year period and promotion to a permanent position.

Please apply online at www.facultyaffairs.ethz.ch

Applications should include a curriculum vitae, a list of publications, and a statement of future research and teaching interests. The letter of application should be addressed **to the President of ETH Zurich. The closing date for applications is 28 February 2015**. ETH Zurich is an equal opportunity and family friendly employer and is further responsive to the needs of dual career couples. We specifically encourage women to apply.

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

International Search for Academic Positions of Assistant Professor or above

The University of Macau is a leading higher education institution in Macao, with English as its working language. In recent years, the University has made great progress in various areas, with increasing international recognition of its excellence in teaching, research and community service. The Times Higher Education World University Rankings 2014/2015 ranks the University among the world's top 300 universities. With the beautiful new campus (20 times larger than the old one) becoming fully operational, the launch of Asia's largest residential college system, the establishment of new faculties, and the increasing numbers of students and faculty members recruited from around the world, UM possesses great potential and provides exciting new possibilities for growth and development.

We plan to develop a strong team of top-notch scholars to help us realize our vision. Applications are therefore invited from those with excellent academic achievements in the following disciplines:

* Business and Management
* Education
* Law
* Liberal Arts and Humanities
* Social Sciences
* Mathematics, Sciences and Engineering
* Health Sciences
* Chinese Medicines

Remuneration and appointment rank offered will be competitive and commensurate with the successful applicants' academic qualification, current position and professional experience. The current local maximum income tax rate is 12% but is effectively around 5% - 7% after various discretionary exemptions.

For details about the above open positions and related information, please visit the following websites:

Job vacancy website: **http://www.umac.mo/vacancy**
University website: **http://www.umac.mo**
Macao government website: **http://www.gov.mo**

Further particulars about the job openings are available at **http://www.umac.mo/vacancy**. Kindly apply online through the E-application system. Applications will be accepted until the positions are filled. Applicants may consider their applications not successful if they were not invited for an interview within 3 months of application.

Human Resources Office
University of Macau,
Av. da Universidade, Taipa, Macau, China
Website: https://isw.umac.mo/recruitment;
Email: vacancy@umac.mo
Tel: +853 8822 8578; Fax: +853 8822 2412

The effective position and salary index are subject to the Personnel Statute of the University of Macau in force. The University of Macau reserves the right not to appoint a candidate. Applicants with less qualification and experience can be offered lower positions under special circumstances.

****Personal data provided by applicants will be kept confidential and used for recruitment purpose only****
Under the equal condition of qualifications and experience, priority will be given to the Macao Permanent Residents

University of Macau - An ideal place to pursue your career
http://www.umac.mo

ASSISTANT/ASSOCIATE PROFESSOR

Integrative Ecoimmunology

The Department of Biological Sciences at the University of Alabama invites applications for a tenure-track faculty position at the rank of Assistant/Associate Professor in Integrative Ecoimmunology. We seek applicants whose research focuses on mechanisms driving changes in host physiology and behavior in response to parasites or pathogens under ecologically relevant contexts. Applicants should use broadly integrative approaches including, but not limited to, molecular, cellular, and physiological methods to characterize host immune responses, hormonal modulation of immune responses, disease susceptibility in a social context, parasite/pathogen manipulation of host behavior and physiology, and life history evolution. Candidates that utilize field-based approaches and/or laboratory-based investigations using model or non-model systems are encouraged to apply. The successful candidate will be expected to establish an extramurally funded and internationally recognized research program in Ecoimmunology. Teaching responsibilities will include basic undergraduate courses in biology, immunology, and graduate courses in the successful candidate's area of expertise.

Candidates must have a Ph.D. in the Biological Sciences or related field, postdoctoral (or equivalent job) experience, demonstrated excellence in research, and a commitment to excellence in teaching and the training of undergraduate and graduate students. Queries regarding additional details should be addressed to the chair of the search committee: Dr. Ryan L. Earley at **rlearley@as.ua.edu**

To apply, go to **https://facultyjobs.ua.edu**, complete the online application (Job #0809244), and upload: (1) an application letter with a list of three to five references (including contact information); (2) CV; (3) statement of research interests and goals; and (4) statement of teaching interests and philosophy. Letters of reference will be requested by the search committee as appropriate. Consideration of applications will begin 1 December 2014, and will continue until the position is filled. Prior to hiring, the final candidate will be required to pass a pre-employment background investigation. The anticipated start date is August 16, 2015.

Additional information about the Department of Biological Sciences and this available position can be found on our website at **http://bsc.ua.edu**. Applications from women and members of traditionally under-represented groups in Biology are especially encouraged.

The University of Alabama is an equal-opportunity educational institution/employer.

THE UNIVERSITY OF ALABAMA®
THE CAPSTONE OF HIGHER EDUCATION

Perelman School of Medicine
University of Pennsylvania

Assistant Professor
Otorhinolaryngology: Head & Neck Surgery

The Department of Otorhinolaryngology: Head and Neck Surgery at the **Perelman School of Medicine** at the **University of Pennsylvania** seeks candidates for an Assistant Professor position in the tenure track.

The successful applicant will have experience in the field of airway epithelial biology with a focus on live cell imaging. Responsibilities include real time recordings using fluorescent dyes of intra and extracellular pH changes, NO production, calcium flux, airway surface liquid hydration, chloride flux as well as in vitro recordings of mucociliary transport and ciliary beat frequency. Additionally, experiments investigating interactions between respiratory pathogens and airway epithelium are expected to be performed. The successful applicant will be expected to obtain extramural research support, contribute to collaborative research efforts and participate in the instruction of medical and graduate students. Applicants must have an Ph.D. and/or M.D. degree and have demonstrated excellent qualifications in education and research.

We seek candidates who embrace and reflect diversity in the broadest sense. The University of Pennsylvania is an EOE. Minorities/Women/Individuals with disabilities/ Protected Veterans are encouraged to apply.

Apply online at: **https://www.med.upenn.edu/apps/faculty_ad/index.php/g329/d3778**

Rochester Institute of Technology

Director of Carlson Center for Imaging Science | *Requisition Number -1504BR*

RIT seeks a visionary and dynamic leader to serve as the Director of the Chester F. Carlson Center for Imaging Science, an interdisciplinary and experiential research and education center dedicated to advancing the technical frontiers of imaging science.

Required Minimum Qualifications - The successful candidate must have:

- An earned doctorate, with significant research experience related to imaging science
- Demonstrated strong management, leadership, communication, and advocacy skills
- The ability to facilitate a broad research portfolio that encompasses the many dimensions of imaging science, fostering and encouraging collaboration with other academic disciplines

An appropriate candidate may be appointed at the level of full professor with tenure.

Required Minimum Education Level: PhD

Required Application Documents:

- Curriculum Vitae or Resume
- Cover Letter
- List of References

HOW TO APPLY: Apply online at **http://apptrkr.com/555968** search openings, then Keyword search 1504BR.

R·I·T

RIT provides reasonable accommodation to applicants with disabilities, veterans or wounded warriors where appropriate. If you need reasonable accommodation for any part of the application and hiring process, please contact the Human Resources office at 585-475-2424 or email your request to Careers@rit.edu. Determinations on requests for reasonable accommodation will be made on a case-by-case basis.

Associate Professor/Professor – Virology

Cornell is a community of scholars, known for intellectual rigor and engaged in deep and broad research, teaching tomorrow's thought leaders to think otherwise, care for others, and create and disseminate knowledge with a public purpose.

The Department of Microbiology and Immunology at the Cornell University Ithaca Campus (**http://www.vet.cornell.edu/microbiology/**) seeks applicants for a tenure-track position at the level of Associate Professor/Professor.

Applicants should have a Ph.D., D.V.M., M.D., or equivalent degree and an established research program with vigorous extramural funding. We are interested in individuals whose primary focus is in virology or viral pathogenesis and who have developed an exceptional career in science. Areas of research interest include human and/or animal viral pathogens and their interaction with hosts. Candidates who can leverage the department's strength in studies of intracellular pathogens and in immunology would be particularly welcome.

Applications should be submitted to **https://academicjobsonline.org/ajo/jobs/5200**, where further details can be found. The committee will begin evaluating completed applications on January 15, 2015 until a suitable applicant is identified. Administrative questions can be e-mailed to **mi-facultysearch@cornell.edu**, or you can contact the Chair of the Search Committee, Dr. Colin Parrish, at **crp3@cornell.edu**.

Cornell University is an innovative Ivy League university and a great place to work. Our inclusive community of scholars, students and staff impart an uncommon sense of larger purpose and contribute creative ideas to further the university's mission of teaching, discovery and engagement. Located in Ithaca, NY, Cornell's far-flung global presence includes the medical college's campuses on the Upper East Side of Manhattan and in Doha, Qatar, as well as the new CornellNYC Tech campus to be built on Roosevelt Island in the heart of New York City.

Diversity and Inclusion are a part of Cornell University's heritage. We're an employer and educator recognized for valuing AA/EEO, Protected Veterans, and Individuals with Disabilities.

Institut Pasteur

Group Leader Positions in Quantitative Biology and Modeling

The Institut Pasteur (Paris, France) announces an international call for group leader positions in the field of quantitative biology and modeling in developmental systems. Successful applicants will be integrated into the cutting edge interdisciplinary environment offered by an internationally renowned institute combining fundamental and translational research, in an attractive location in central Paris, in close proximity to other major research centers. Candidates with both an experimental and theoretical background, using quantitative approaches and willing to develop multidisciplinary projects related to developmental processes are encouraged to apply.

Successful **junior candidates** [1] will be appointed with a permanent position, and as head of a group of 6 people. These groups will be created for a period of 5 years and may thereafter compete for a full research group.

Successful **mid-career and senior candidates** will be appointed with a permanent position, and as head of a research group of 8 to 12 people. The groups will be created for 10 years (mid-term evaluation at 5 years) with the possibility of renewal.
Candidates should send their formal applications by E-mail to the Director of Scientific Evaluation, Prof. Alain Israël, at the Institut Pasteur (**qubimo@pasteur.fr**).

Application deadline: February 6, 2015

Short-listed candidates will be contacted for interview.

Applicants should provide the following (in order) in a single pdf file:

1. A brief introductory letter of motivation, including the name of the proposed group. Candidates are encouraged to contact the head of the Search Committee Francois Schweisguth (fschweis@pasteur.fr) or the head of the Department of Developmental & Stem Cell Biology Shahragim Tajbakhsh for queries (**shaht@pasteur.fr**).
2. A Curriculum Vitae and a full publication list.
3. A description of past and present research activities (up to 6 pages with 1.5 spacing; Times 11 or Arial 10 font size).
4. The proposed research project (up 6 pages with 1.5 spacing; Times 11 or Arial 10 font size).

Junior candidates [1] should also provide:

5. The names of 3 scientists from whom letters of recommendation can be sought, together with the names of scientists with a potential conflict of interest from whom evaluations should not be requested.

[1] Institut Pasteur is an equal opportunity employer. Junior group leaders should be less than 8 years after PhD at the time of submission (Dec 31, 2014). Women are eligible up to 11 years after their PhD if they have one child and up to 14 years after their PhD if they have two or more children.

भारतीय विज्ञान संस्थान

THE INFOSYS CHAIR VISITING PROFESSORSHIPS

Indian Institute of Science, Bangalore, India

The Infosys Foundation has endowed Visiting Professorships in Mathematics and Physics at the Indian Institute of Science (IISc). The purpose of these Professorships is to bring distinguished scientists in the broad areas of Physical and Mathematical Sciences to the IISc campus for collaborative research interactions.

Visiting Professors are expected to spend a total of one year at IISc; visits can be spread out over a period of 4 years. The remuneration package for these Visiting Professorships is internationally competitive.

If you are interested in this Professorship, please contact a faculty member at IISc who will then take the process forward. If you would like to nominate a distinguished Physicist or Mathematician for this Professorship, please send a nomination to the Chairman, Division of Physical and Mathematical Sciences, Indian Institute of Science by e-mail (**ichair@admin.iisc.ernet.in**).

DUKE NUS
GRADUATE MEDICAL SCHOOL SINGAPORE

FACULTY POSITIONS
ALL RANKS
CANCER AND STEM CELL BIOLOGY

The Duke-NUS Graduate Medical School Singapore (Duke-NUS) invites applications from outstanding physician-scientists and scientists for tenure track positions with a focus in the area of cancer and/or stem cell biology. We seek internationally competitive applicants, especially those with expertise in proteomics, stem cells, cancer, genomics, and/or Precision or Translational Medicine. Scientists with existing programs that are translational in nature are especially encouraged to apply.

Duke-NUS is a global partnership between Duke University in the United States and the National University of Singapore. Our faculty are expected to establish extramurally funded translational or basic science research programs and to participate in medical education. Applicants should have a PhD, MD, or equivalent and a record of outstanding promise or achievement. New recruits will work alongside existing faculty and trainees with strong programs in cancer signaling, genomics, bioinformatics, cell biology, and stem cell biology. Opportunities also exist for translational studies in collaboration with world-class clinical services including the Singapore General Hospital and National Cancer Centre Singapore, as well as other Research Institutes and Universities across the island.

Faculty positions at Duke-NUS include full salary support and generous start-up funding, to be supplemented by competitive grant awards.

For full consideration, interested candidates should submit, by 1 March, 2015, a cover letter, curriculum vitae, a summary of research accomplishments and outline of future plans.

Assistant Professor candidates should also arrange for three letters of reference to be forwarded to:

S. Tiong ONG, Chair,
Faculty Search Committee
Program in Cancer and Stem Cell Biology
Duke-NUS Graduate Medical School Singapore
8 College Road, Singapore 169857
email: hr@duke-nus.edu.sg

Partner in Academic Medicine
SingHealth

Learn more and don't hit the pavement for your job search.

- Search thousands of job postings
- Create job alerts based on your criteria
- Get career advice from our Career Forum experts
- Download career advice articles and webinars
- Complete an individual development plan at "myIDP"

Target your job search using relevant resources on **ScienceCareers.org**.

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

POSITIONS OPEN

FACULTY POSITION in Musculoskeletal Sciences and Engineering

The Division of Musculoskeletal Sciences, Department of Orthopaedics and Rehabilitation at The Pennsylvania State University College of Medicine is seeking applications for a faculty Position in Musculoskeletal Sciences or Engineering. This tenure-track or tenured appointment will be at the rank of **ASSISTANT, ASSOCIATE**, or **FULL PROFESSOR** and in the area of musculoskeletal sciences or engineering. This position includes a highly competitive salary and significant startup funding.

The successful candidate should have a Ph.D. in a biological science or engineering field relevant to musculoskeletal science and a strong record of research accomplishment. Applicants for Assistant Professor should demonstrate a strong potential for obtaining extramural research funding while applicants for Associate Professor or Professor should have an active extramurally funded research program.

This is a unique opportunity to join a well-established, highly interactive research group consisting of engineers, material, clinical, and basic scientists focusing on musculoskeletal research. This appointment will include a primary appointment in the College of Medicine and a secondary appointment in an appropriate department at Penn State's University Park campus.

The Penn State College of Medicine is located in Hershey, Pennsylvania and offers a highly desirable lifestyle and an affordable cost of living in close proximity to many metropolitan areas including Baltimore, Washington, D.C., Philadelphia, and New York City.

Applications will be accepted until the position is filled. Applicants should submit their curriculum vitae and the names of at least three references. Apply to **job #54895** at **website: http://apptrkr.com/553763.**

Campus Security Crime Statistics: For more about safety at Penn State, and to review the Annual Security Report which contains information about crime statistics and other safety and security matters, please go to **website: http://www.police.psu.edu/clery/**, which will also provide you with detail on how to request a hard copy of the Annual Security Report.

Penn State is an Equal Opportunity/Affirmative Action Employer, and is committed to providing employment opportunities to minorities, women, veterans, disabled individuals, and other protected groups.

USDA AGRICULTURAL RESEARCH SERVICE

The USDA, Agricultural Research Service, Plant Genetics Research Unit in Columbia, Missouri, is seeking a **RESEARCH GENETICIST** (Plants), at GS-12/13 grade level (Announcement Number ARS-D15E-0019A) salary range $69,497 - $107,434. The research focus of the position is to provide solutions to problems limiting maize productivity and utilization. The applicant is expected to develop molecular genetic resources for improving genetic analysis in maize, characterization, and manipulation of genetic resources for maize through the application of molecular, quantitative, and statistical genetics, and bioinformatics-based strategies, and to increase our understanding of the genetic basis of traits of importance to maize productivity. The scientist is expected to provide leadership within teams on aspects of computational genetics, quantitative genetics, genomic, and bioinformatics approaches and resources. Research will include collaboration with other agency scientists and with scientists in academia and private industry as needed to accomplish objectives. United States citizenship is required. For information on the research program and/or position, contact **M. J. Oliver**, research leader, at **telephone: 573-882-9645** or at **e-mail: mel.oliver@ars.usda.gov**. For further information, and complete application instructions, go to **website: https://www.usajobs.gov/GetJob/ViewDetails/389187800.** Applications must be received by January 16, 2015. *USDA/ARS is an Equal Opportunity Provider and Employer.*

POSITIONS OPEN

The new Center for Research on Environmental Disease, College of Medicine, University of Kentucky is seeking candidates for several faculty positions at the **ASSISTANT/ASSOCIATE/PROFESSOR** (tenure-track and tenured) levels with 12-month appointment. Applicants are sought in the broad area of environmental disease. Applicants should submit their curriculum vitae to **Dr. Xianglin Shi**, professor and director, Center for Research on Environmental Disease to **e-mail: xshi5@email.uky.edu**. Review of applications will begin immediately and continue until the positions are filled.

By **Mary Poffenroth**

My liberation through science

I always knew I would go to college—which, when I think about it, is a little strange, because I come from a family in which no one had ever gone to college. Born to a teen mom, I was adopted at birth by my maternal grandmother, a high school graduate, and her second husband, an eighth-grade dropout who died of alcohol-induced cirrhosis when I was a year old. My adoptive parents were the Poffenroths and so I, too, am a Poffenroth, the youngest of five siblings and the only one to graduate from high school.

Everything I knew about college came from the movies. On my very first day, overwhelmed with the sights and sounds of being on a real college campus, I wandered toward the comforting aroma of barbecue. Mesmerized by all the activity, I fell into a 6-foot construction trench that was inches from the source of the smell: a fire pit where a fraternity was grilling meat to recruit new pledges.

It wasn't the last hole I'd find myself in during college, but all the others were less literal. As the months turned into semesters, I got a classic case of too much fun and too little class time. At the end of my second year, my major professor sat me down and said, "You are never going to graduate unless something big changes."

Teen pregnancy, alcoholism, trouble with the law—my family had done it all. I wanted to break free from old Poffenroth habits—which, I realized, I had been slipping toward, unwittingly. But it wasn't until I entered my graduate program in conservation ecology that I felt I was on the right path. My conservation work gave me something bigger than myself to strive for. It exposed me to passionate and dedicated people. I had positive role models to inspire me, but I struggled to fit the new me into my old family.

Over dinners I wanted to discuss Emerson and Dawkins, while my family wanted to discuss soap operas and game shows. One Thanksgiving, I decided to invite over a few school friends who could not be with their faraway families. As we sat for dinner, my family gravitated toward our large, mahogany table that still bore the scars of one of my mom's china-breaking tantrums. My college friends sat at a rented table. I lingered, hoping the matter would be decided for me, but there were two open chairs after everyone else was seated, one at each table. I knew my choice was consequential. I chose the place I felt most comfortable, where I could most be myself: nestled with my new family of college friends.

"My liberation through science has brought me joy, frustration, excitement, and adventure."

My liberation through science has brought me joy, frustration, excitement, and adventure. I have been shot at while doing fieldwork in the remote Warner Mountains of California (although the shooters were actually aiming for quail). I've had wine with former Vice President Al Gore and dinner with nature activist Paul Watson. I have eaten my weight in olives along the coast of Spain while writing a book. I've stepped over cobras in Marrakech, Morocco, and looked out over the Adriatic Sea during a cyclone in Croatia. Science made all that possible.

Today, my job is to get people excited about science. I specialize in teaching nonmajor courses, in megalecture classes with several hundred students per section. My class may be one of few opportunities they have in their lives to learn about such vital topics as global climate change, sustainability, and human biology. In a single year, I can reach nearly 2000 students.

There are benefits to being your family's first college graduate. You can get out of doing things you don't want to do by saying you have to study. Even when you're wrong, they assume you're right. Today, 17 years after I fell in that hole, I feel less like the family outcast and more like an explorer who carved a new path. I've redefined, for the better, what it means to be a Poffenroth. ■

Mary Poffenroth is a full-time lecturer at San Jose State University in California. For more on life and careers, visit www.sciencecareers.org. Send your story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER