An activist uses science to fight animal research *p. 366*

A battle of principles in the e-cigarettes debate *p. 375*

Counting molecular garbage chutes in intact neurons *p. 439*

Science

$10
23 JANUARY 2015
sciencemag.org

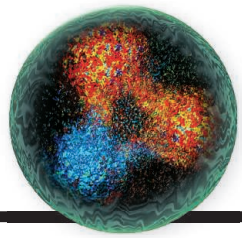
AAAS

Catching a comet

Rosetta follows a relic of the early solar system toward the Sun

pp. 358 & 387

CONTENTS



363
The inner life of the proton

23 JANUARY 2015 • VOLUME 347 • ISSUE 6220

NEWS

IN BRIEF

356 Roundup of the week's news

IN DEPTH

358 **COMET CLOSE-UP REVEALS A WORLD OF SURPRISES**
Diverse landscape of 67P suggests an unexpectedly complicated origin
By E. Hand
▸ CATCHING A COMET SECTION P. 387

359 **SURVEYS REVEAL STATE OF AFGHAN POPULATION**
From fertility to mortality, new data will help guide aid efforts *By M. Hvistendahl*

361 **JAPAN'S NUCLEAR RENAISSANCE DOGGED BY WASTE CHALLENGE**
Nuclear reprocessing is years late; repository is in limbo *By D. Normile*

362 **MALARIA MAY ACCELERATE AGING IN BIRDS**
Infection erodes telomeres and life spans—hinting that there are no benign parasites *By G. Vogel*
▸ REPORT P. 436

FEATURES

363 **PROBING THE PROTON**
A newly upgraded accelerator explores the seething maelstrom at the heart of matter *By A. Cho*

366 **THE INSURGENT**
Animal rights activist Justin Goodman is using science's own tools to combat animal research *By D. Grimm*

INSIGHTS

PERSPECTIVES

370 **PLAY IT AGAIN, SAM**
Lasers shine new light on the drying of Mars while reviving the mystery of methane
By K. Zahnle
▸ REPORTS PP. 412 & 415

371 **FLEXIBILITY FOR SPECIFICITY**
Human T cells show remarkable differentiation plasticity in responding to pathogens
By M. M. Davis
▸ RESEARCH ARTICLE P. 400

373 **UNDERSTANDING ARTEMISININ RESISTANCE**
Genetic studies provide clues to how malaria parasites become drug resistant *By C. H. Sibley*
▸ REPORTS PP. 428 & 431

375 **SMOKE AND FIRE OVER E-CIGARETTES**
As nations adopt regulatory measures for e-cigarettes, it is imperative to understand how approaches to risk, cost-benefit, and trade-offs have shaped interpretations of evidence
By A. L. Fairchild and R. Bayer

377 **BRAZIL'S SOY MORATORIUM**
Supply-chain governance is needed to avoid deforestation *By H. K. Gibbs* et al.

BOOKS *ET AL.*

379 **THE ALTRUISTIC BRAIN**
By D. W. Pfaff, reviewed by F. B. M. de Waal

380 **THE NEW MATH**
By C. J. Phillips, reviewed by J. Kilpatrick

LETTERS

382 **CARNIVORE COEXISTENCE: VALUE THE WILDERNESS**
By J. J. Gilroy et al.

382 **CARNIVORE COEXISTENCE: AMERICA'S RECOVERY**
By M. E. Gompper et al.

383 **CARNIVORE COEXISTENCE: TROPHIC CASCADES**
By T. M. Newsome and W. J. Ripple

383 **FERNS TO FULFILLMENT**
By L. Reynolds

DEPARTMENTS

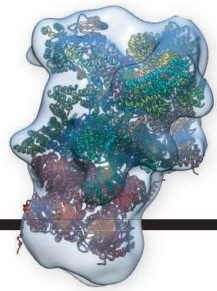
355 **EDITORIAL**
Rethink the Nicaragua Canal
By Jorge A. Huete-Perez et al.

450 **WORKING LIFE**
Teaming up against tsunamis
By Elisabeth Pain

Science Staff 354
New Products 443
Science Careers 444

Published by AAAS

CONTENTS



439 Counting proteasomes one at a time

23 JANUARY 2015 • VOLUME 347 • ISSUE 6220

SPECIAL SECTION

Catching a comet

INTRODUCTION

387 Rosetta begins its comet tale *M. G. G. T. Taylor* et al.

RESEARCH ARTICLE ABSTRACTS

On the nucleus structure and activity of comet 67P/Churyumov-Gerasimenko *H. Sierks* et al.

The morphological diversity of comet 67P/Churyumov-Gerasimenko *N. Thomas* et al.

Dust measurements in the coma of comet 67P/Churyumov-Gerasimenko inbound to the Sun *A. Rotundi* et al.

REPORT ABSTRACTS

67P/Churyumov-Gerasimenko, a Jupiter family comet with a high D/H ratio *K. Altwegg* et al.

The organic-rich surface of comet 67P/Churyumov-Gerasimenko as seen by VIRTIS/Rosetta *F. Capaccioni* et al.

Time variability and heterogeneity in the coma of 67P/Churyumov-Gerasimenko *M. Hässig* et al.

Birth of a comet magnetosphere: A spring of water ions *H. Nilsson* et al.

Subsurface properties and early activity of comet 67P/Churyumov-Gerasimenko *S. Gulkis* et al.

SEE ALSO ▸ NEWS STORY P. 358 ▸ PODCAST ▸ SLIDESHOW ▸ sciencemag.org/site/special/rosetta

RESEARCH

IN BRIEF

384 From *Science* and other journals

REVIEW

393 SEA LEVEL CHANGE
Inherited landscapes and sea level change *S. Cloetingh and B. U. Haq*
REVIEW SUMMARY; FOR FULL TEXT: dx.doi.org/10.1126/science.1258375

RESEARCH ARTICLES

394 PROTEOMICS
Tissue-based map of the human proteome *M. Uhlén* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: dx.doi.org/10.1126/science.1260419

395 HUMAN EVOLUTION
Human-like hand use in *Australopithecus africanus* *M. M. Skinner* et al.

400 T CELL IMMUNITY
Functional heterogeneity of human memory $CD4^+$ T cell clones primed by pathogens or vaccines *S. Becattini* et al.
▸ PERSPECTIVE P. 371

REPORTS

406 ASTROPHYSICS
The exceptionally powerful TeV γ-ray emitters in the Large Magellanic Cloud *The H.E.S.S. Collaboration*

MARS ATMOSPHERE

412 The imprint of atmospheric evolution in the D/H of Hesperian clay minerals on Mars *P. R. Mahaffy* et al.

415 Mars methane detection and variability at Gale crater *C. R. Webster* et al.
▸ PERSPECTIVE P. 370

418 PLANETARY SCIENCE
Shock compression of stishovite and melting of silica at planetary interior conditions *M. Millot* et al.

420 MAGNETIC MATERIALS
Tilt engineering of spontaneous polarization and magnetization above 300 K in a bulk layered perovskite *M. J. Pitcher* et al.

425 MATERIALS CHEMISTRY
Composition-matched molecular "solders" for semiconductors *D. S. Dolzhnikov* et al.

DRUG RESISTANCE

428 K13-propeller mutations confer artemisinin resistance in *Plasmodium falciparum* clinical isolates *J. Straimer* et al.

431 Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance *S. Mok* et al.
▸ PERSPECTIVE P. 373

436 CHRONIC INFECTION
Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds *M. Asghar* et al.
▸ NEWS STORY P. 362

439 PROTEASOMES
A molecular census of 26*S* proteasomes in intact neurons *S. Asano* et al.

ON THE COVER

View of a cliff and gravel field on the small lobe of comet 67P/Churyumov-Gerasimenko from a distance of 8 kilometers to the surface, as seen by Rosetta/OSIRIS. See pages 358 and 387 and sciencemag.org/site/special/rosetta. *Image: ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA*

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. *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

EDITORIAL

Rethink the Nicaragua Canal

Jorge A. Huete-Perez is vice-president of the Academy of Sciences of Nicaragua and a professor at the University of Central America, Managua, Nicaragua. E-mail: jorgehuete@uca-cbm.org

Axel Meyer is a profesor in the Department of Biology at the University of Konstanz, Konstanz, Germany. E-mail: Axel.Meyer@uni-konstanz.de

Pedro J. Alvarez is the George R. Brown Professor and Chair of Civil and Environmental Engineering at Rice University, Houston, Texas, USA. E-mail: alvarez@rice.edu

At the end of 2014, construction began on the Grand Canal in Nicaragua, a project shrouded in secrecy since its inception 2.5 years ago. The Nicaraguan government showed scant evidence of having accounted for the impact on the environment and on local residents, or of having adequately consulted the public in selecting the final 278-km route. Such disregard should be alarming to everyone. Projects of this magnitude warrant dialogue among all stakeholders. As construction is projected to span 5 years, there is still time to reconsider it and convene independent assessments and meetings that are transparent, inclusive, and respectful of different perspectives, to guide the project toward the best outcome.

Declared an income generator by the Nicaraguan government, the Grand Canal is owned by a Chinese company (HKND). It will enter Nicaragua from the Pacific through the Brito River, cross agricultural land and forests, and traverse Lake Cocibolca (the largest tropical lake of the Americas). On the Caribbean side, it will cross pasturelands, forests, natural reserves, wetlands, and indigenous communities, entering the Caribbean close to the Punta Gorda River.

At an international multidisciplinary workshop convened in Managua in November 2014, experts from the Academy of Sciences of Nicaragua (ASN) and the InterAmerican Network of Academies of Sciences, with support from the International Council for Science in Latin America and the Caribbean, explored technical and scientific aspects of the canal and shared their scientific concerns with the public, HKND, and the government.* The final communiqué highlighted the importance of environmental impact assessments (EIAs) on biodiversity and regional water resources, including Lake Cocibolca, which supplies drinking water to the surrounding population. Particularly disquieting is how the initial construction and recurrent dredging of the lake channel might affect the trophic dynamics, water quality, and sedimentation of local rivers, along with implications for domestic water supplies, crop irrigation, and fisheries. Also planned is an artificial lake to serve the canal and lock system. The potential impact of "Lake Atlanta" and the canal on forests, wetlands, and coastal ecosystems has been grossly underexplored. Species cohesiveness and migration are also at risk. Although a private EIA was commissioned by HKND—undermining its independence—it will not be completed before April 2015, diminishing its influence.

"...there is still time...to guide the project toward the best outcome."

Most countries follow rigorous environmental guidelines before approving large infrastructure projects. The Nicaraguan government's decision to grant the canal concession without public consultation, feasibility studies, or EIAs departs from such practices and increases the project's risks. Without independent EIAs, unintended and irreversible impacts on wildlife and indigenous populations are more likely. Social unrest is growing along the canal route. Government and security forces are responding by harassing, repressing, and jailing opposition protesters and leaders, increasing fear among the populace.

It is incumbent upon scientists, human rights advocates, nongovernmental organizations, and wildlife protection organizations to share knowledge, voice concerns, provide guidance, and demand a greater role for science in the design and construction of this massive project. There is tremendous need for the international scientific community to join ASN and global organizations such as the Association for Tropical Biology and Conservation and the International Society of Limnology to analyze design plans of the canal and its subprojects for safety, social responsibility, and sustainability; make recommendations to protect the region's water resources and biodiversity; and draft statements urging the Nicaraguan government to halt construction until studies can be performed and evaluated by experts. In this matter of great urgency and importance, this is an opportunity to exercise scientific leadership, raise awareness, and contribute to averting a potential environmental disaster.

– Jorge A. Huete-Perez, Axel Meyer, Pedro J. Alvarez

*www.interacademies.net/File.aspx?id=26001.

10.1126/science.aaa6998

PHOTO: GUIZIOU FRANCK/HEMIS/CORBIS

NEWS

> "I don't know if it is all [man's fault] but the majority is. ... It is man who continuously slaps down nature."
>
> **Pope Francis** on 15 January, responding to questions about his views on humankind's role in global climate change.

A controlled burn of surface oil from the 2010 *Deepwater Horizon* oil spill in the Gulf of Mexico.

IN BRIEF

Judge's ruling: BP Gulf spill totaled 3.19 million barrels of oil

After hearing "voluminous, dense, highly technical, and conflicting" testimony from geoscientists, a federal judge on 15 January ruled that petroleum giant BP spilled 3.19 million barrels of oil into the Gulf of Mexico during the 2010 *Deepwater Horizon* disaster. The U.S. government had argued for 4.2 million barrels; BP 2.45 million. The ruling means BP faces a maximum fine of $13.7 billion, or $4300 per barrel, for violations of the Clean Water Act; the final amount will be set in a proceeding that begins 20 January. Under the 2012 RESTORE Act, 80% of the fine will be dedicated to Gulf Coast restoration, recovery, and research. "There is no way to know with precision how much oil discharged. ... There was no meter counting off each barrel of oil," wrote District Court Judge Carl Barbier of the Eastern District of Louisiana in New Orleans. Six experts testified that as few as 2.4 million barrels of oil, and as many as 6 million, escaped during the 86-day ordeal; BP and the government agreed that 810,000 barrels were captured. http://scim.ag/_BPruling

AROUND THE WORLD

First light for planet hunter

PARANAL, CHILE | A group of European institutions announced first light on the Next-Generation Transit Survey (NGTS), a new planet-hunting instrument in Chile, last week. NGTS, like NASA's Kepler satellite, hunts for exoplanets by measuring dips in brightness as a planet crosses in front of its star. From the duration and frequency of the dips, scientists can deduce an exoplanet's size and orbit. Previous transit surveys by ground-based telescopes have only been able to detect Jupiter-sized or bigger planets due to the blurring effect of Earth's atmosphere. Kepler, by focusing on one patch of sky for years, can spot Earth-sized or smaller planets with long orbits. NGTS aims to fill the gap between those rocky planets and gas giants, focusing on planets between two and eight times the diameter of Earth—super-Earths to exo-Neptunes. http://scim.ag/NGTSopen

The Next-Generation Transit Survey observatory.

Long-lost Mars lander found

PASADENA, CALIFORNIA | At last, scientists know the fate of the Beagle 2, a British-built probe sent to the Red Planet in late 2003 as part of the European Space Agency's (ESA's) Mars Express mission. When it couldn't establish contact with the probe, ESA declared the probe lost in 2004. But last week, scientists examining three images taken by NASA's Mars Reconnaissance Orbiter over several years (and at different sun angles) with its High Resolution Imaging Science Experiment camera, announced they had spotted signs of the errant spacecraft on the planet's surface. Although mission scientists originally feared Beagle 2 had crash-landed, the images show that it landed successfully enough to partially deploy its solar panels—but the partial deployment wasn't enough to transmit data.

Lawsuit over mine approval

SYDNEY, AUSTRALIA | Australian environmentalists are taking the nation's environment minister to court over his approval last July of Indian firm Adani's AU$16.5 billion Carmichael coal mine in Queensland. Lawyers for the Mackay

PHOTOS: (TOP TO BOTTOM) PATRICK SEMANSKY/AP PHOTO; ESO/R. WESSON

Plan to save monarch butterflies backfires

It started with the best of intentions: Well-meaning gardeners across the United States, upon learning that monarch butterflies were losing the milkweed they depend on because of the spread of herbicide-resistant crops, began planting milkweed in their own gardens. But inadvertently, they may be endangering the monarchs' iconic migration to Mexico by planting the wrong species of milkweed, scientists reported last week in the *Proceedings of the Royal Society B*. Many gardeners planted widely available tropical milkweed *Asclepias curassavica*, which doesn't die back in the winter like native milkweed does, offering the monarchs a year-round place to lay their eggs—so that many monarchs don't bother making the trip to Mexico at all. Worse, tropical milkweed hosts the protozoan parasite *Ophryocystis elektroscirrha*, which caterpillars ingest along with their milkweed meals; when they hatch, the butterflies are covered in spores and are much weaker than their healthy counterparts. If an infected monarch does try to migrate, it will probably die long before it arrives in central Mexico. **http://scim.ag/monarchfail**

To help out the monarchs, planting the right milkweed matters.

PHOTOS: (TOP TO BOTTOM) © WILLIAM WARNER/ALAMY; © PALLAVA BAGLA

Conservation Group will argue that the minister, Greg Hunt, failed in his duty to properly consider the impact that greenhouse gas emissions from the burned coal will have on global warming and therefore on the Great Barrier Reef. Hearings begin 3 February. Over the mine's 60-year life, mining and transporting the coal is expected to produce more than 200 million tonnes of CO_2; the burned coal will also emit 130 million tonnes yearly. A representative for Hunt said the minister made his decision in accordance with federal environmental protection legislation. The case is one of three lodged by conservation groups against the mine.

Horizon 2020 cuts unveiled

STRASBOURG, FRANCE | A controversial plan to divert research funds to economic stimulus became more concrete on 13 January, as European Commission President Jean-Claude Juncker unveiled legislation to implement the shift. It would deduct €2.7 billion over 5.5 years from Horizon 2020, the commission's main research program, set to invest about €80 billion between 2014 and 2020. The largest cut—€350 million—would come from the European Institute of Innovation and Technology (EIT) in Budapest, which funds industry-academia collaborations. The cut amounts to 13% of EIT's budget. The European Research Council loses €221 million, mostly in 2016 and 2017. The move has drawn muted opposition. "I'm surprised that there isn't a louder outcry and no clearer opposition from the scientific community," Hans-Olaf Henkel, a member of the European Parliament, told *Science*|*Business*. **http://scim.ag/Hor2020cuts**

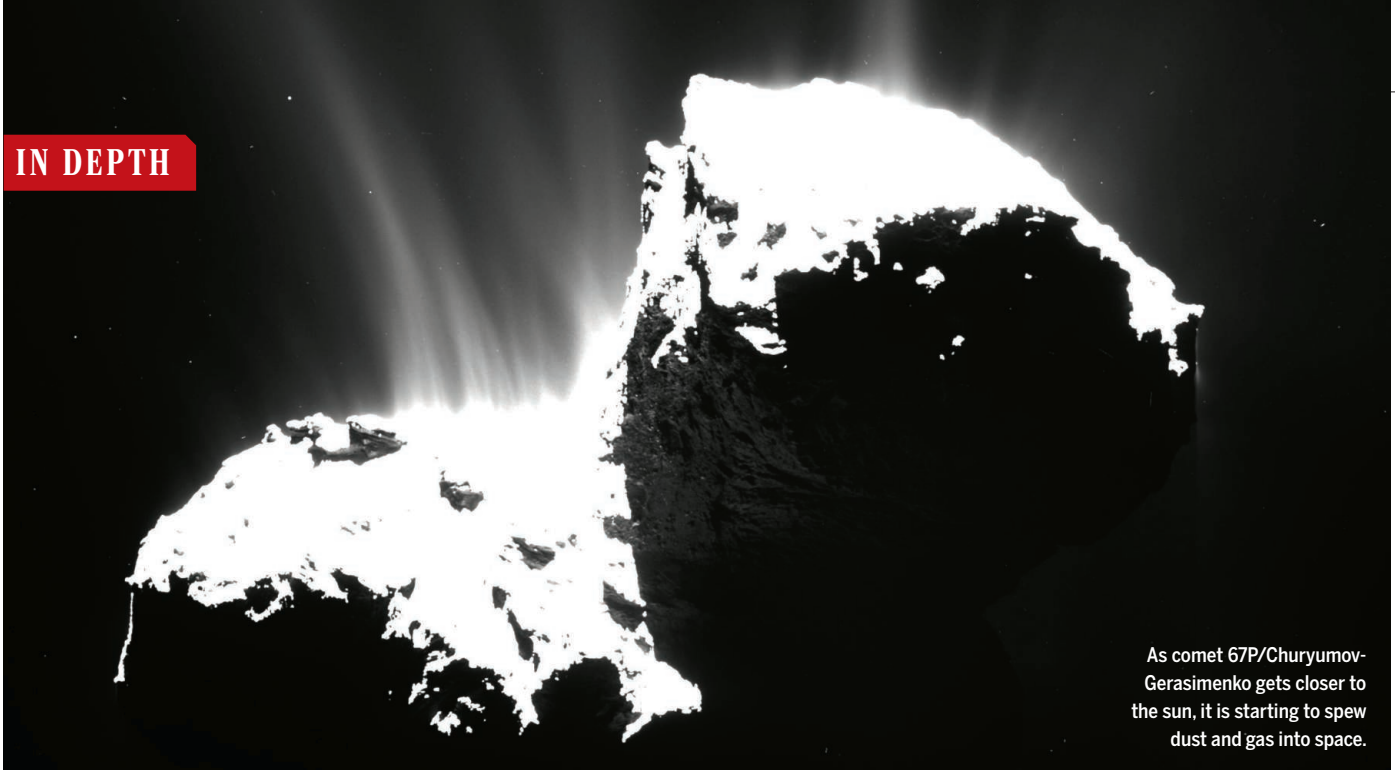
Squeeze put on flush scientists

BETHESDA, MARYLAND | In the latest example of U.S. biomedical research budget stretching, the National Institute of General Medical Sciences (NIGMS) is imposing a strict one-grant limit on scientists who already have plentiful no-strings support. The rule, which will take effect in January 2016, will apply to researchers who have at least $400,000 per year in research funding not tied to a specific project (not including salary or overhead costs). The new limit "will enable NIGMS to fund additional labs, increasing the likelihood of making significant scientific advances," NIGMS says. NIGMS Deputy Director Judith Greenberg estimates there are at least 22 Howard Hughes Medical Institute investigators who hold two or three NIGMS research grants. Limiting them to a single grant could free up $6 million for 25 to 30 grants. **http://scim.ag/onegrant**

NEWSMAKERS

'Eye in the sky' to lead space org

Veteran space engineer **Alur Seelin Kiran Kumar** has taken the reins of the Indian Space Research Organisation (ISRO), an organization riding high after its Mangalyaan probe reached Mars last September on the nation's first try. Known as "India's eye in the sky," Kumar, 62, has overseen development of India's remote sensing systems, including the Bhaskara-1 Earth observation satellite in 1979 and a constellation of high-resolution imaging satellites now in orbit. As director of the Space Applications Centre in Ahmedabad, Kumar headed the team that designed three of Mangalyaan's five scientific payloads, including the Mars Color Camera. Kumar will oversee an ambitious slate of missions at ISRO, including the launch this spring of India's prototype space shuttle and Chandrayaan-2, India's second moon foray featuring a lander and a rover and slated for liftoff in 2017.



As comet 67P/Churyumov-Gerasimenko gets closer to the sun, it is starting to spew dust and gas into space.

PHOTO: ESA/ROSETTA/MPS FOR OSIRIS TEAM MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

PLANETARY SCIENCE

Comet close-up reveals a world of surprises

Diverse landscape of 67P suggests an unexpectedly complicated origin

By **Eric Hand**

A boring lump of ice and dust it's not. Comet 67P/Churyumov-Gerasimenko—already the best explored comet ever—turns out to be pocked with pits, incised by cracks and cliffs, and decorated with ripples and flows of dust: all signs of an extraordinarily active place.

Five months after the European Space Agency's (ESA's) Rosetta spacecraft arrived at 67P—and 2 months after the spacecraft dropped the Philae lander to the surface—the mission team this week publishes a suite of papers in *Science* that detail a surprising diversity of features on the 4-kilometer-long duck-shaped comet.

The main science camera, OSIRIS (Optical, Spectroscopic and Infrared Remote Imaging System), cataloged much of that diversity. Many of the intriguing landforms it has found testify to the power of the sun, which heats up 67P during part of every orbit, igniting jets of gas and dust that resculpt the surface of the comet. Other discoveries could be primordial, dating from the comet's formation more than 4.5 billion years ago.

The sun alone can't have created all the layers, topography, and chemical diversity that scientists see, says Nicolas Thomas, an OSIRIS co-investigator and an experimental physicist at the University of Bern. "I find it hard to believe at the present time that [67P] was originally homogeneous," he says. "You've got to produce all of these diverse morphologies on the surface with that one energy source—I find this tough." He and others say the complexity of the comet today suggests that the comet-forming regions of the early solar system were more turbulent and chemically diverse than theorists have thought.

The diversity occurs at scales both small and big. At the big end, the team has divided the comet into 19 distinct regions. Some are buried in dust, while others contain brittle and rocky-seeming terrain. One puzzling area called Aten, on the "wing" of the duck-shaped comet, is a depression that is oddly free of dust. It, at least, is unlikely to be primordial, Thomas says. He speculates that the hollow formed during recent passes by the sun. Sharp changes in temperature may have cracked and weakened the surface, and rising subsurface gas pressures may have blown off chunks of the comet in one or more catastrophic shedding events.

Cliff faces also record the ongoing transformation of the comet, says Holger Sierks, principal investigator for OSIRIS at the Max Planck Institute for Solar System Research in Göttingen, Germany. Boulders at the base of cliffs indicate that material can fall from them, exposing fresh ice and dust for the sun to erode. Pits a few hundred meters across, which appear all over the comet, are less important sites of material loss, at the moment at least. Many of them are clogged with several meters of insulating dust, and scientists are eager to see if they become unplugged as the comet heats up.

The pits contained some surprises. In one, OSIRIS revealed a delta of material that had flowed out, like ooze from a baking-soda-and-vinegar volcano. Thomas says the flow is a sign that pockets of pressure can build up so high in the icy interior that fluid mixes of material occasionally erupt.

In the walls of other pits, OSIRIS has spotted what could be features dating back to the comet's formation: what the team calls "goosebumps" or "dinosaur eggs," nodules about 3 meters across that could represent the fundamental chunks of material that coalesced into 67P. "It's a window to the early days of the solar system, to the building blocks when no Earth was present," Sierks says.

Other clues about the comet's origins came from the spacecraft's chemical sensors. Scanning the surface, for instance, a spectrometer detected an absorption feature associated with complex organic molecules that could include carboxylic acids—precursors to amino acids. Finding organic molecules on 67P is not a surprise—scientists have spotted them in the halos of other comets, and NASA's Stardust mission found actual amino acids in comet dust it

sampled in 2004—but the molecules Rosetta has detected are more complex than those seen on other comets.

These organics form only in extreme cold, when ultraviolet light and cosmic rays strike various types of ice-coated dust. That suggests they formed when the comet did, instead of evolving later, says the spectrometer's principal investigator, Fabrizio Capaccioni of the National Institute for Astrophysics in Rome. The molecules could indicate that 67P formed farther out in the solar system than comets that lack them. "This comet is more pristine than others," he says.

Another spectrometer, looking just above the comet's surface, found a diversity of gases emanating from it. In some places, the comet is emitting mostly water. In others, emissions are dominated by carbon dioxide and carbon monoxide, which sublime more quickly than water ice when the comet is warmed by the sun. The gas concentrations don't always vary as the sunlight falling on the comet changes—evidence that they may be due to heterogeneous concentrations of different ices.

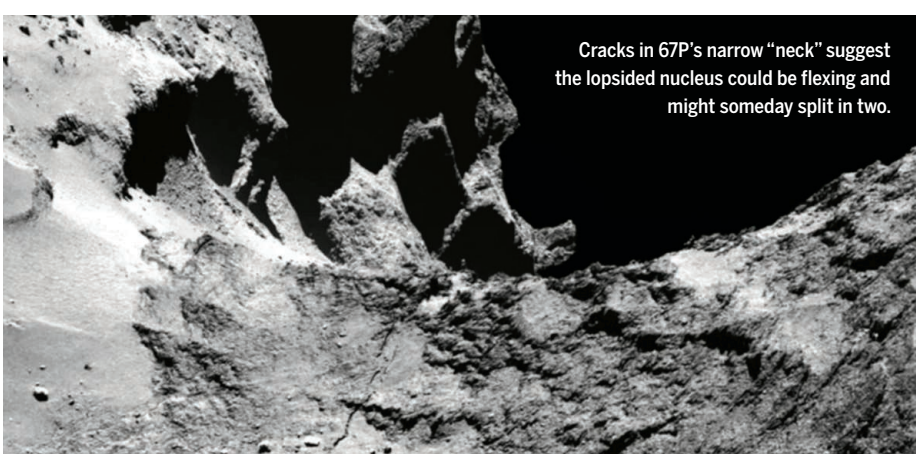
The comet is still relatively quiet. Its jets are expected to become about a hundred times as active by the time 67P reaches its closest point to the sun in August. Pits will become geysers, cliff faces will tumble, and perhaps a more catastrophic shedding event awaits. "The fun thing is going to be seeing these changes," Thomas says. Watching them will help investigators understand how much of the evolution can be explained by the sun and how much by underlying heterogeneity that was emplaced at the comet's birth. "The answer, always, is: It's probably both," says Jessica Sunshine, a planetary scientist at the University of Maryland, College Park, who is not on the mission.

Rosetta's envoy to the comet itself, the Philae lander, has not contributed to current papers. But the lander team is preparing a set of manuscripts for publication. During the 57 hours the lander spent on the surface before its batteries died, cameras on the sides took images, but mission scientists do not think that the lander's drill successfully extracted a sample and fed it to the two chemical analysis ovens. However, one of the ovens performed a gas measurement as the lander first bounced off the surface, and it may have sucked in and analyzed some dust as well. "There is a possibility that maybe we collected something," says Hermann Böhnhardt, a lead lander scientist at the Max Planck Institute for Solar System Research.

Philae, which came to rest on its side in a shadowed area, still has not been found. In December, Rosetta swooped in to 18 kilometers above the surface to scrutinize a 500-by-300–meter area where the lander is presumed to be. Combing through the images by hand, the OSIRIS team found many bright spots, but none that could be distinguished from boulders. Böhnhardt hopes that the OSIRIS team will now allow the wider Rosetta and Philae teams to study the images. "Looking by eyes, this is certainly the first approach," he says. "It is just not as good as reading the pixel values and trying to do something with computers."

Knowing just where Philae ended up will help its science team assess their data and allow engineers to evaluate how precarious the lander's situation is. The lander team is lobbying for ESA to schedule another close reconnaissance pass by the orbiter—but this becomes more risky as the comet wakes up and its jets of gas and dust become stronger.

Another option is simply to wait. As the comet's seasons change and more light falls on the lander's solar panels, Philae could just wake up. Böhnhardt says there's a good chance that its computer will boot up in May and that its radio will regain contact with the orbiter by June. "It can shift earlier, it could also be never," he says. On 16 January, the lander steering committee agreed to resume science planning for later this year—a sign of confidence that Philae's mission is not yet done. ■



Cracks in 67P's narrow "neck" suggest the lopsided nucleus could be flexing and might someday split in two.

PHOTO: OSIRIS TEAM MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

DEMOGRAPHY

Surveys reveal state of Afghan population

From fertility to mortality, new data will help guide aid efforts

By **Mara Hvistendahl**

Last winter, thousands of surveyors fanned out across Kabul to learn the state of the population after decades of turmoil. Key findings from that door-to-door exercise, part of a nationwide effort called the Socio-Demographic and Economic Survey (SDES), were released last week. They show that Afghanistan, where U.S.-led forces formally ended combat missions last month, faces steep challenges as it seeks to rebuild: Newborns are dying in droves, and fully 64% of adults in Kabul had no meaningful income in the year before the survey.

Scholars and aid workers welcome the hard data if not the bleak findings, because Afghanistan was once a demographic mystery. "All the reports of mortality for decades had basically been opinion," says David Peters, a public health researcher at the Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland. Now, the SDES, which will run until 2018, and a slew of other surveys carried out by Afghan agencies in collaboration with overseas scientists are revealing a wealth of information about everything from fertility to mortality. The results will help government officials and aid agencies forecast population growth, assess the labor market, plan for education and other services, and analyze health programs. Such data are "critical," Peters says, "to know if your programs are working."

Merely starting the SDES was a fraught process. In the 2001 Bonn Agreement, which followed the U.S. invasion of Afghanistan and laid out steps for establishing a transition government, parties called for a full census: a simultaneous tally of households in all 34 Afghan provinces. (Afghanistan's last attempt at a census had been in 1979, on the eve of the Soviet invasion.) "Donors were pouring in aid for reconstruction," notes Mercedita Tia, a statistician with the United Nations Population Fund (UNFPA) in Kabul, and they needed to track results. But by 2010, with the Taliban dominating

sampled in 2004—but the molecules Rosetta has detected are more complex than those seen on other comets.

These organics form only in extreme cold, when ultraviolet light and cosmic rays strike various types of ice-coated dust. That suggests they formed when the comet did, instead of evolving later, says the spectrometer's principal investigator, Fabrizio Capaccioni of the National Institute for Astrophysics in Rome. The molecules could indicate that 67P formed farther out in the solar system than comets that lack them. "This comet is more pristine than others," he says.

Another spectrometer, looking just above the comet's surface, found a diversity of gases emanating from it. In some places, the comet is emitting mostly water. In others, emissions are dominated by carbon dioxide and carbon monoxide, which sublime more quickly than water ice when the comet is warmed by the sun. The gas concentrations don't always vary as the sunlight falling on the comet changes—evidence that they may be due to heterogeneous concentrations of different ices.

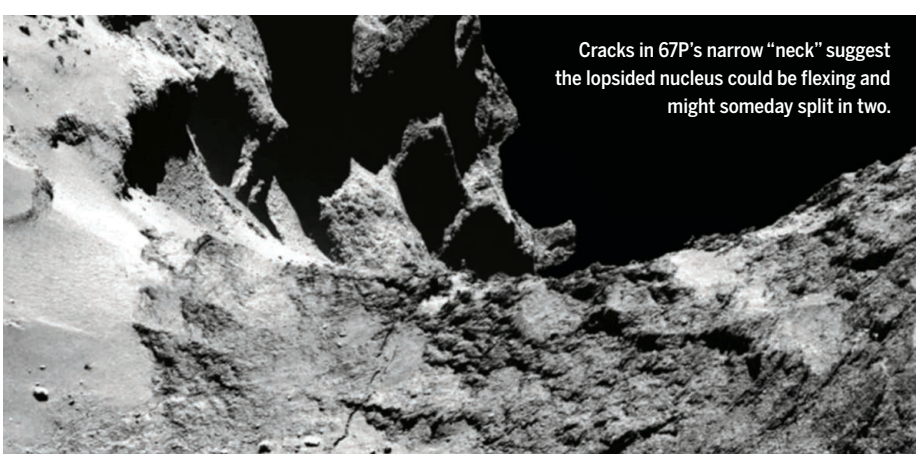
The comet is still relatively quiet. Its jets are expected to become about a hundred times as active by the time 67P reaches its closest point to the sun in August. Pits will become geysers, cliff faces will tumble, and perhaps a more catastrophic shedding event awaits. "The fun thing is going to be seeing these changes," Thomas says. Watching them will help investigators understand how much of the evolution can be explained by the sun and how much by underlying heterogeneity that was emplaced at the comet's birth. "The answer, always, is: It's probably both," says Jessica Sunshine, a planetary scientist at the University of Maryland, College Park, who is not on the mission.

Rosetta's envoy to the comet itself, the Philae lander, has not contributed to current papers. But the lander team is preparing a set of manuscripts for publication. During the 57 hours the lander spent on the surface before its batteries died, cameras on the sides took images, but mission scientists do not think that the lander's drill successfully extracted a sample and fed it to the two chemical analysis ovens. However, one of the ovens performed a gas measurement as the lander first bounced off the surface, and it may have sucked in and analyzed some dust as well. "There is a possibility that maybe we collected something," says Hermann Böhnhardt, a lead lander scientist at the Max Planck Institute for Solar System Research.

Philae, which came to rest on its side in a shadowed area, still has not been found. In December, Rosetta swooped in to 18 kilometers above the surface to scrutinize a 500-by-300–meter area where the lander is presumed to be. Combing through the images by hand, the OSIRIS team found many bright spots, but none that could be distinguished from boulders. Böhnhardt hopes that the OSIRIS team will now allow the wider Rosetta and Philae teams to study the images. "Looking by eyes, this is certainly the first approach," he says. "It is just not as good as reading the pixel values and trying to do something with computers."

Knowing just where Philae ended up will help its science team assess their data and allow engineers to evaluate how precarious the lander's situation is. The lander team is lobbying for ESA to schedule another close reconnaissance pass by the orbiter—but this becomes more risky as the comet wakes up and its jets of gas and dust become stronger.

Another option is simply to wait. As the comet's seasons change and more light falls on the lander's solar panels, Philae could just wake up. Böhnhardt says there's a good chance that its computer will boot up in May and that its radio will regain contact with the orbiter by June. "It can shift earlier, it could also be never," he says. On 16 January, the lander steering committee agreed to resume science planning for later this year—a sign of confidence that Philae's mission is not yet done. ■



Cracks in 67P's narrow "neck" suggest the lopsided nucleus could be flexing and might someday split in two.

PHOTO: OSIRIS TEAM MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

DEMOGRAPHY

Surveys reveal state of Afghan population

From fertility to mortality, new data will help guide aid efforts

By **Mara Hvistendahl**

Last winter, thousands of surveyors fanned out across Kabul to learn the state of the population after decades of turmoil. Key findings from that door-to-door exercise, part of a nationwide effort called the Socio-Demographic and Economic Survey (SDES), were released last week. They show that Afghanistan, where U.S.-led forces formally ended combat missions last month, faces steep challenges as it seeks to rebuild: Newborns are dying in droves, and fully 64% of adults in Kabul had no meaningful income in the year before the survey.

Scholars and aid workers welcome the hard data if not the bleak findings, because Afghanistan was once a demographic mystery. "All the reports of mortality for decades had basically been opinion," says David Peters, a public health researcher at the Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland. Now, the SDES, which will run until 2018, and a slew of other surveys carried out by Afghan agencies in collaboration with overseas scientists are revealing a wealth of information about everything from fertility to mortality. The results will help government officials and aid agencies forecast population growth, assess the labor market, plan for education and other services, and analyze health programs. Such data are "critical," Peters says, "to know if your programs are working."

Merely starting the SDES was a fraught process. In the 2001 Bonn Agreement, which followed the U.S. invasion of Afghanistan and laid out steps for establishing a transition government, parties called for a full census: a simultaneous tally of households in all 34 Afghan provinces. (Afghanistan's last attempt at a census had been in 1979, on the eve of the Soviet invasion.) "Donors were pouring in aid for reconstruction," notes Mercedita Tia, a statistician with the United Nations Population Fund (UNFPA) in Kabul, and they needed to track results. But by 2010, with the Taliban dominating

Surveyors for the Socio-Demographic and Economic Survey at a training session in Bamiyan province.

PHOTO: © LORENZO TUGNOLI/UNFPA AFGHANISTAN

southern Afghanistan, international experts recommended scrapping the census in favor of the SDES: a staggered, province-by-province survey of households in all but the most dangerous districts.

The SDES, which is being carried out by the Afghan Central Statistics Organization (CSO) with assistance from UNFPA, uses an unusual method. In every second household, surveyors collect detailed information on age, gender, education, literacy, employment, fertility, mortality, birth registration, and other indicators. In the others, they simply record the name of the head of household and the number of members by sex. That arrangement cuts costs while still giving reliable results at the district level, Tia says. Even so, the SDES is immensely time-consuming. In rural areas, surveyors have to set out by donkey or motorbike, and for months at a time snow makes parts of the mountainous country inaccessible.

Earlier, more focused surveys gave glimpses of demographic and health conditions in parts of Afghanistan. In 2002, Linda Bartlett, then an epidemiologist and maternal health specialist with the U.S. Centers for Disease Control and Prevention in Atlanta, set out on horseback with a team of health researchers to assess maternal and infant mortality in rural Afghanistan. Bartlett hired Afghan women as surveyors, but a Taliban restriction held they could only travel escorted by their husbands or male relatives. The team canvassed 13,000 households, asking family members for information on an estimated 85,000 women. They later reconstructed how many maternal deaths had occurred from 1999 to 2002.

Their findings were grim: The four districts surveyed showed an average of 1600 maternal deaths per 100,000 live births—a rate that would now be among the highest in the world. In one district in northern Afghanistan, Bartlett says, "there was no access to any kind of prenatal care, obstetric care, or postnatal care, and there was no family planning available." Traditional practices included pushing on a woman's belly during delivery.

In the years that followed, the Afghan government made maternal and child health a priority, vastly expanding access to health care services. The number of midwives shot up from 467 in 2002 to more than 4600 today, according to UNFPA. And the 2010 Afghanistan Mortality Survey (AMS) of 24,000 households suggested that such efforts were paying off. It uncovered a steep drop in maternal mortality, compared with previous estimates: 327 deaths per 100,000 live births. But some researchers skewered the finding as the product of underreporting, possibly reflecting Afghan reluctance to discuss women who died. "It was pretty concerning for a lot of us, because we felt that it falsely represents what's going on," Bartlett says.

The finding "may be on the low side," says Fred Arnold, a demographer at ICF International in Fairfax, Virginia, which provided technical assistance for the AMS. But it jibes with figures from neighboring countries, he notes. Earlier surveys in Afghanistan drew largely from areas with higher rates of maternal deaths, says Arindam Das of the Indian Institute of Health Management Research in Jaipur, which also assisted on the AMS. The survey is not nationally representative: Since 2008, volatile conditions have made it virtually impossible to collect data from areas like Kandahar, a Taliban stronghold in southern Afghanistan covered in Bartlett's survey. (On two occasions, the Taliban kidnapped AMS surveyors, despite the team avoiding rural areas in Kandahar and two other provinces.) Still, Das says, they covered a larger area than did previous surveys.

Other studies have underscored how hard it will be to get a clear picture of demographic trends. In 2006, for example, Peters and his Hopkins colleague Gilbert Burnham set out to estimate mortality for children under 5. Available estimates had been derived from models relying on 1970s data and hadn't been updated since 1993. Surveying more than 8200 households in 29 of 34 Afghan provinces, they found signs pointing to significant underreporting of girl births—and possibly girl deaths—meaning that any future child mortality figures would need to be adjusted. "People don't like to talk about girls, period," Peters says.

The SDES has surveyed six provinces, says Mohammad Sami Nabi, director of the CSO's field operations department, and data are still limited. Work in the other 28 provinces is projected to wrap up in 2018.

Although survey costs in Afghanistan are skyrocketing with worsening security conditions—organizations now have to shell out for helicopters to circumvent roadblocks—plans for still more ambitious surveys are now being rolled out. In March, Arnold says, fieldwork is scheduled to begin on Afghanistan's inaugural Demographic and Health Survey, part of a worldwide program that collects detailed population and health data for developing countries.

Meanwhile, SDES investigators are learning all they can while minimizing risk. Because ethnicity is an explosive issue in the fractured country, surveyors aren't including it in their questions. Even so, the findings are potentially charged. In 2012, survey results from Bamiyan province, home to a persecuted Shia Muslim minority, showed a smaller than anticipated population. Misconstruing the figure as a sign of prejudice, hundreds of residents took to the streets, shouting "death to the enemies of Bamiyan," reported Pajhwok Afghan News. The incident prompted UNFPA to step up publicity campaigns for the survey, Tia says: "We have learned a lot." ■

NUCLEAR POWER

Japan's nuclear renaissance dogged by waste challenge

Nuclear reprocessing is years late; repository is in limbo

By **Dennis Normile**, *in Tokyo*

Japan's battered nuclear industry is finally getting back on its feet, raising anew a vexed question: what to do with the waste. Nearly 4 years after the Tohoku earthquake and tsunami sparked meltdowns at the Fukushima Daiichi Nuclear Power Station and led to the shutdown of all 48 of the country's reactors, the central government is expected this spring to approve the restart of two nuclear reactors in Kagoshima Prefecture. But critics charge that uncertainties and wishful thinking are clouding prospects for dealing with the waste.

"Japan is totally at a loss about how to resolve the high-level waste disposal problem," says antinuclear crusader Hideyuki Ban of the Citizens' Nuclear Information Center here. "The waste issue should be addressed before restarting the reactors." Prime Minister Shinzo Abe's administration has acknowledged the challenge. The energy plan of April 2014, which reversed the previous government's no-nukes policy, noted that delays, accidents, and cover-ups in developing technologies for handling radioactive waste "have generated a feeling of distrust among the people."

When Japan commissioned its first nuclear power plant in 1966, it envisioned reprocessing spent fuel to extract uranium and plutonium for reuse in reactors and to reduce the amount of radioactive waste that would need to be entombed underground. In theory, reprocessing could help ensure a supply of fuel and make it easier to cope with the remaining waste. But a half-century later, none of the reprocessing technologies are ready. Japan now has 17,000 tons of spent nuclear fuel submerged in cooling pools at reactor sites across the country, and a planned reprocessing plant has suffered years of delays. Reprocessing turns out to be the answer to "a very complicated equation," says Hajimu Yamana, a reprocessing expert at Kyoto University.

Japan Nuclear Fuel Ltd. (JNFL), owned by 10 utility companies, broke ground on a reprocessing facility in Rokkasho, a village at the northern tip of Honshu, in 1993 and planned to complete it 4 years later. Two decades later, it is still a work in progress. The new facility will use the Plutonium Uranium Redox EXtraction (PUREX) method, pioneered in the Manhattan Project, to separate plutonium and uranium from spent fuel. The process greatly reduces the volume of waste but leaves a highly radioactive residue that amounts to about 3% of the original volume. JNFL opted to design its own method for encasing these leftovers in glass, a process known as vitrification. Yamana says Japan chose an experimental liquid-fed ceramic melter vitrification process that promised to be more durable and provide a higher throughput than the metallic melter-based techniques then in commercial use. But when scaled up, the chemistry involved in mixing the aqueous nitric acid solution containing the waste into the molten glass, he says, proved "extremely complex."

JNFL announced in early 2013 that it had finally worked out the kinks. But tougher seismic and safety requirements that followed the Fukushima disaster led to further delay. The price tag for the Rokkasho facility will end up at about $18.8 billion, according to JFNL, nearly triple the original $6.5 billion budget. The new target date for finally completing the plant is March 2016, after which it will take several years to ramp up capacity to process a planned 800 tons of waste per year. Yamana says this amount would gradually put a dent in the accumulated waste if Japan reduces its reliance on nuclear power in accordance with the latest energy plan.

Critics question whether Japan should even be in the business of reprocessing. The original plan was to use the separated uranium in conventional reactors and the plutonium in so-called fast reactors, once thought to be a next-generation nuclear power source. But Japan's experimental Monju fast reactor in Tsuruga on the Sea of Japan has been largely shuttered since it was commissioned 20 years ago. Most other countries have abandoned fast reactor projects because of technical difficulties, high costs, and concerns about safeguarding plutonium, though China and India are still studying the technology. The 2014 energy strategy calls for using Monju to "burn" spent fuel to reduce the volume of high-level waste.

But that won't eliminate the waste issue. Despite more than a decade of trying, the government has been unable to interest any locality in hosting a deep underground repository. In September 2012, the Science Council of Japan recommended abandoning the idea, arguing that it is impractical in seismically active Japan to identify a geological formation that would remain stable for millennia. The council suggested instead that Japan develop secure facilities for temporarily storing nuclear waste until further research hits on a better solution and a public consensus can be reached on an overall nuclear energy policy.

That could take a while. ■

Recycling radioactive remains

Japan's reprocessing takes spent fuel rods and separates low-level waste for disposal, high-level waste for vitrification, and purifies uranium and plutonium for reuse.

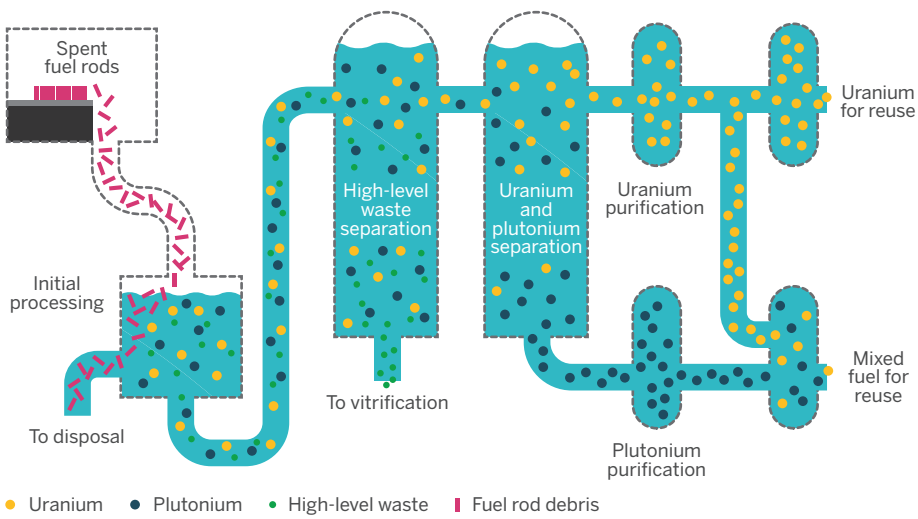


IMAGE: JAPAN NUCLEAR FUEL LTD./ADAPTED BY G. GRULLON/*SCIENCE*

BIOLOGY

Malaria may accelerate aging in birds

Infection erodes telomeres and life spans—hinting that there are no benign parasites

By **Gretchen Vogel**

Malaria is a scourge of humankind, but many birds seem to shrug it off. Although they are chronically infected with malaria parasites, their behavior seems unaffected, and they mostly reproduce and raise young just as well as noninfected birds. That was a puzzle not just for ornithologists but also for evolutionary biologists, who have long theorized that parasites inevitably take a toll on fitness.

The birds' healthy appearance turns out to be deceiving, however. Drawing on data from a 3-decade study of great reed warblers in southern Sweden, researchers report on page 436 that long-term infection with malaria significantly shortened the birds' lives. The analysis also revealed a possible explanation: The blood cells of infected birds had shorter telomeres, stretches of DNA that cap the ends of chromosomes and protect them during cell division. In many species, shorter telomeres are associated with aging and shorter life span.

Malaria doesn't affect the great reed warbler's care for its chicks, but does cut its lifetime number of offspring.

The work "is an almost deceptively simple study that puts the nail in the coffin of the idea of benign parasites," says Marlene Zuk, an evolutionary biologist at the University of Minnesota, Twin Cities. And although it's not clear exactly how malaria infection influences telomere length, the study raises the question of whether other mild, but chronic, infections have similar hidden costs in different animals, even humans.

Evolutionary biologists have been arguing for more than a century about the long-term costs of nonlethal diseases to fitness and reproductive success. It is famously difficult, however, to tease out the effects of disease from all the other variables that affect a wild animal's ability to reproduce. The problem is compounded by small sample sizes in most studies of wild populations, says evolutionary ecologist Ben Sheldon of the University of Oxford in the United Kingdom.

In 1983, two biology students at Lund University in Sweden, Dennis Hasselquist and Staffan Bensch, began collecting data on great reed warblers nesting at nearby Lake Kvismaren. Dozens of the brown-and-white birds with a loud, raspy song visit the lake each summer to breed, and most come back to the nesting site where they were born. Bensch and Hasselquist, now both professors at Lund, still spend summers at the lake, and with colleagues keep track of which warblers return, who mates with whom, and how many offspring each bird raises. The scientists also take blood samples from hatchlings and adults.

Roughly 40% of the birds that nest at the site are infected with one or more species of malaria parasites picked up in their winter habitat, which ranges across sub-Saharan Africa. After an initial acute phase of the disease, in which birds are anemic, lethargic, and lose their appetite, survivors recover but often remain infected, with a low level of parasites in their blood. Many seem to do just fine nonetheless. In a 17-year study, for example, the research team found no significant differences in annual breeding success and survival to the next year of infected and uninfected birds.

Then came the telomere clue. Muhammad Asghar, a Ph.D. student in Bensch's group, recently measured their lengths in the blood samples using a PCR technique pioneered for birds by Pat Monaghan's group at the University of Glasgow. "Some amazing patterns started to emerge," Bensch says: Infected birds had significantly shorter telomeres, and those with more intense infections showed more dramatic telomere loss.

That could have been because birds with shorter telomeres were more susceptible to malaria. But in an earlier study, the group had deliberately infected captive warblers. Examining the birds' stored blood, the researchers found that the telomeres shrank in a direct response to the infection. "The combination of comparative and experimental results makes for a very strong case," says Robert Fleischer, an evolutionary geneticist at the Smithsonian Institution's National Zoological Park in Washington, D.C.

Infection shortened life spans as well as telomeres—by about a year, the lake records showed. That stole breeding opportunities. On average, uninfected birds lived 2.5 years and raised more than eight offspring to fledglings. Infected birds lived an average of 1.6 years and raised just four offspring. "It's a different kind of cost" than researchers expected, Hasselquist says.

It isn't clear if the shorter telomeres are a result of the extra red blood cells infected birds have to make—telomeres typically shrink with every cell division—or whether they are a more general consequence of the body's efforts to fight the infection. But the authors and others suggest it's worth looking more closely at chronic infections in birds and elsewhere, perhaps even in people with asymptomatic malaria. "Maybe we underestimate what 'mild' disease is," Hasselquist says. "Like herpesvirus. We have it, and we don't think there are any problems. But the potential is there at least" that it could carry some hidden long-term costs. ■

PHOTO: DEA/G. CAPPELLI/DE AGOSTINI/GETTY IMAGES

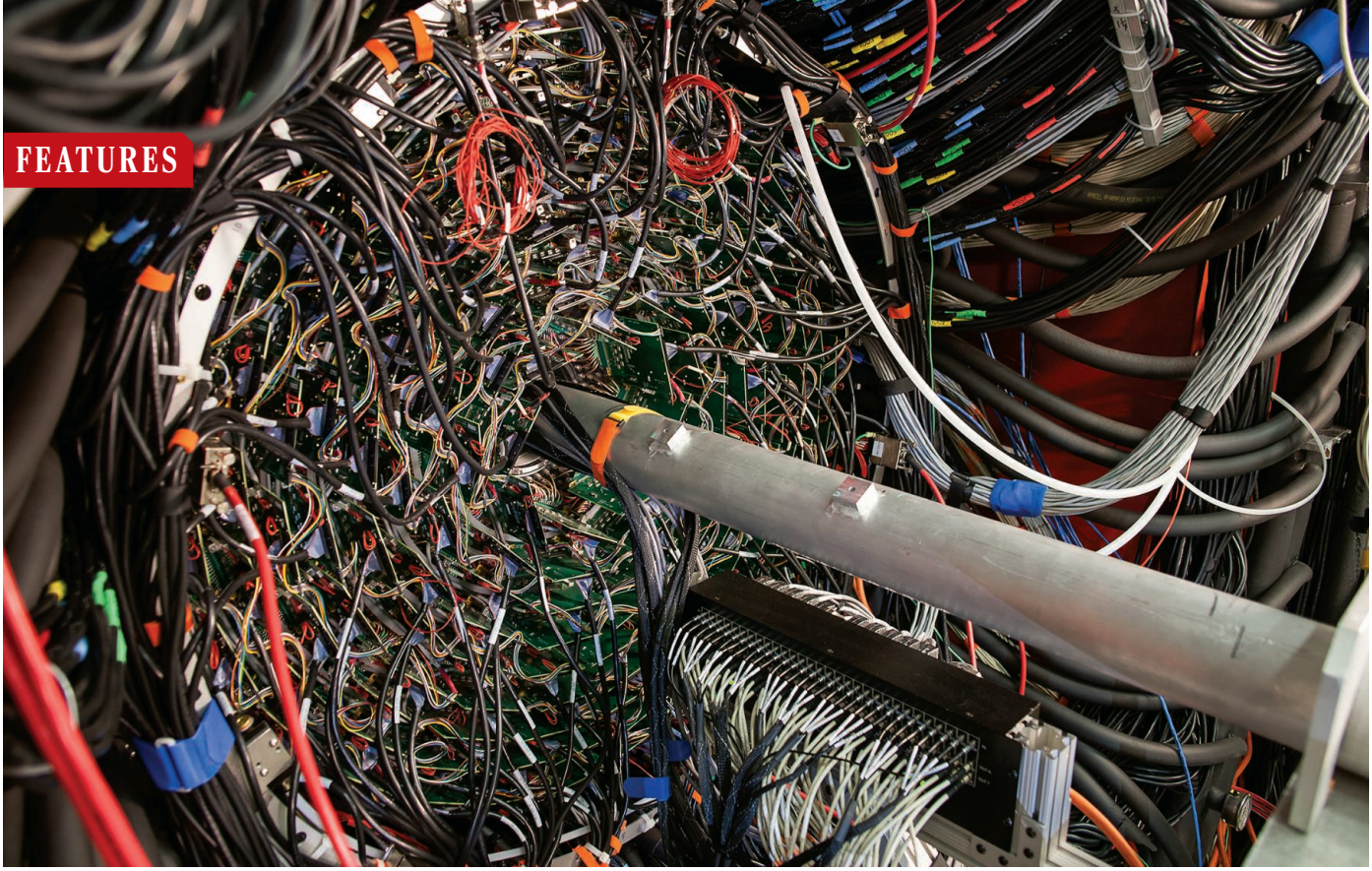


PHOTO: BILL KELLEY III

Probing the proton

A newly upgraded accelerator explores the seething maelstrom at the heart of matter

By **Adrian Cho**, *in Newport News, Virginia*

The GlueX detector will search for weird new particles predicted by the theory of the strong force.

Hall A here at Thomas Jefferson National Accelerator Facility is everything a science fan could want from a physics lab. The cylindrical cavern, 24 meters from floor to dome and twice as wide, echoes like a cathedral. From high on the wall, a silvery pipe extends fingerlike to the center, from which fan out two vast machines, multitiered assemblages of steel, pipes, wires, and electronics that sweep up to the roof.

Think of it all as part of a giant electron microscope, designed to probe two of the most familiar yet mysterious constituents of the universe: the proton and the neutron. The pipe carries a beam of high-energy electrons from an accelerator, which smash into targets mounted at the room's center. The hall's two 1400-tonne spectrometers capture the subatomic particles blasted from the collisions, tracking their paths and energies. For 20 years, Jefferson Lab physicists have battered the nucleus without quite cracking its secrets. But starting this year, they will train a more powerful microscope on their quarry.

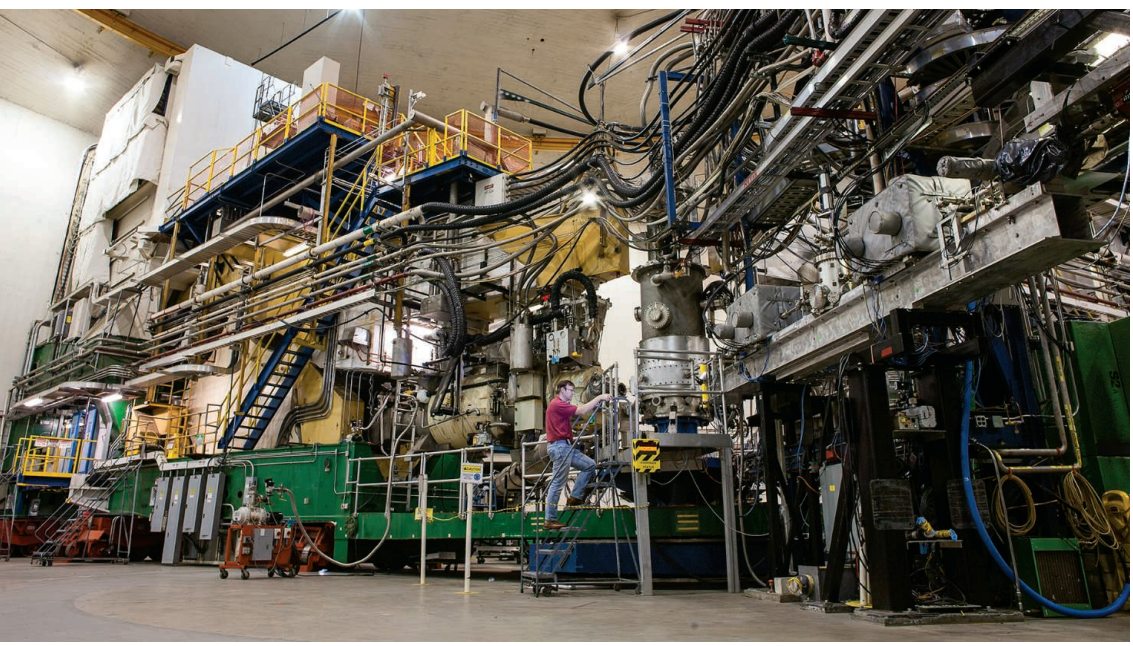
In the cartoon view, the positively charged proton and the uncharged neutron both consist of trios of particles called up quarks and down quarks. (Two ups and a down make a proton; two downs and an up make a neutron.) But earlier experiments have shown that those "valence" quarks are just a small part of the story. A nucleon—a proton or neutron—is really a pullulating mass of countless quarks, antiquarks, and gluons, particles that convey the strong nuclear force that holds quarks together. A proton or neutron is so messy that physicists can't say exactly how its most basic properties, such as its mass and spin, emerge from the tangle.

Jefferson Lab physicists are finishing a $338 million upgrade to their particle accelerator, the Continuous Electron Beam Accelerator Facility (CEBAF), to double its energy and probe the innards of protons and neutrons with unprecedented precision. Compared with meccas of particle physics such as CERN, Europe's particle physics lab near Geneva, Switzerland, Jefferson Lab is relatively obscure. But it will be the global focal point for this sort of nuclear physics. "It will be a very exciting program," says Fabienne Kunne, a physicist at the French Atomic Energy Commission in Saclay, who is not involved with the project. "It should make enormous progress."

Unlike particle physicists' quest for the famed Higgs boson, spotted in 2012 at CERN, Jefferson Lab's mission requires not one grand experiment but a mosaic of measure-

ments. The diversity shows in the lab's four experimental halls: Two, including Hall A, feature spectrometers that can be reconfigured to study specific interactions, and two contain detectors designed to capture every interaction so that they can be sorted out later. "Jefferson Lab is a little bit more Wild West," says Cynthia Keppel, a physicist who oversees the two spectrometer halls. "You've got settlers coming in from all over with all manner of ideas."

The mess inside a nucleon arises from the peculiar nature of the strong force. At first blush, that force works something like the electromagnetic force, which binds electrons to nuclei in atoms and produces light. Within an atom, the negatively charged electrons cling to the positively charged nucleus by exchanging massless quantum particles called photons. In the same way, the quarks within the nucleon cling to one another by exchanging massless particles called gluons, which carry the strong force.



The huge, reconfigurable spectrometers in Jefferson Lab's Hall A swivel around on massive steel wheels.

But the strong force is far more complex than the electromagnetic force. Unlike the passive photons, gluons themselves exchange gluons—lots of them. Moreover, a nucleon's three valence quarks aren't the only ones inside it. Untold quark-antiquark pairs also flit into and out of existence. Those fleeting "sea" quarks need not be up and down quarks, but can be heavier strange and charm quarks, too. Thus, each valence quark is shrouded by a cloud of quarks and gluons in which "what you see depends on the scale at which you look," says Michael Pennington, chief theorist at Jefferson Lab.

That complexity makes calculating anything involving the strong force nearly impossible. For example, most of a nucleon's mass comes not from the valence quarks, but from the energy of the quark-gluon cloud (thanks to Einstein's equivalence of mass and energy). But only recently have theorists accounted for it in numerical simulations (*Science*, 19 December 2008, p. 1772). As for nucleons' spin, physicists once thought it came from the spins of the valence quarks. But since 1987, they have known that most of it must originate in some other way—either in the swirling motion of the quarks or in the gluons.

Unfortunately, physicists can't decipher nucleons by simply taking them apart. The strong force is so strong that they can't isolate a quark or a gluon the way they might pluck an electron from an atom. Strike a nucleon hard enough to smash it, and it will spew particles containing either three quarks or a quark-antiquark pair, forcing researchers to tease information from those composite particles instead.

The "confinement" of quarks and gluons within other particles remains mysterious, says Colin Morningstar, a theorist at Carnegie Mellon University in Pittsburgh, Pennsylvania. Experiments and numerical simulations leave little doubt that it holds, but nobody has proved mathematically that it must or explained how the "field" of gluons arranges itself to make it happen. The Clay Mathematics Institute in Providence offers a $1 million prize for a proof. "I've demonstrated it numerically," Morningstar quips, "but they haven't sent me a check."

Jefferson Lab physicists are hardly the first to try to pierce the confusion. From 1967 until 1975, physicists at SLAC National Accelerator Laboratory in Menlo Park, California, used their linear accelerator to fire electrons into nucleons, discovering the quark in 1968. Since 1978, physicists at CERN have fired muons—heavier, unstable cousins of electrons—into nucleons, discovering the proton's "spin crisis." From 1992 to 2007, researchers at Germany's Electron Synchrotron lab in Hamburg collided beams of electrons and protons in their Hadron Electron Ring Accelerator.

But those efforts raised as many questions as they answered, Pennington says. "It's a little bit like the discovery of North America," he says. "These other experiments showed that there was something there, and we're going to map it out." Jefferson Lab physicists say the CEBAF accelerator is the ideal tool for the job.

Instead of a traditional circular accelerator or synchrotron, CEBAF consists of twin 235-meter linear accelerators, like the straights on a racetrack, connected by arcing beamlines. The electrons make five laps, passing through a different set of arcs each lap. The resulting beam has an extremely narrow range of energies, says Arne Freyberger, Jefferson Lab's director of accelerator operations. That narrow energy spread enables physicists to precisely measure the changes in a scattered electron's energy and momentum. And because the beam is continuous rather than pulsed like most accelerator beams, particle collisions don't bunch up in time, enabling researchers to use timing techniques to sift out rare decays.

Just to build CEBAF, which turned on in 1995 and cost $515 million, Jefferson Lab helped pioneer a whole new technology. Within an accelerator, charged particles gain energy by surfing radio waves sloshing in hollow "RF cavities" the size of a big salami. Until the 1990s, accelerators used copper cavities, but Jefferson Lab physicists opted for cavities of superconducting niobium. Although they must be cooled to near absolute zero with liquid helium, they consume just 5% as much power as copper cavities, says Robert Rimmer, an accelerator physicist at Jefferson Lab. "For a machine like CEBAF, you couldn't afford the power bill if you used copper cavities," he says. Superconducting cavities have become standard for accelerators.

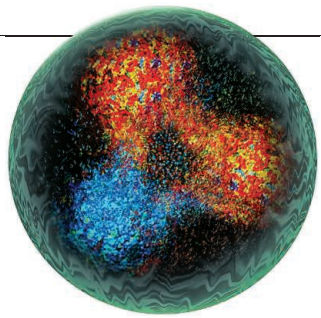
PHOTO: BILL KELLEY III

In upgrading CEBAF, physicists doubled the accelerator's energy from 6 to 12 gigaelectron volts simply by adding new cavities in empty space at the end of each linac. That extra energy will be a boon to the more than 70 experiments approved for the upgraded machine, most of which are devoted to fathoming the nucleon. In particular, the energy will literally add new dimensions to physicists' picture of nucleons.

For decades, physicists have made do with essentially one-dimensional views of nucleons. Ping electrons off a nucleon gently, and their deflections reveal the distribution of the quarks within—but only perpendicular to electrons' original path. Strike the nucleon hard enough to break it apart, and the electrons' deflections and energy losses will reveal the momenta of quarks and gluons within the nucleon—but only along the electrons' initial path.

In nuclear physics, higher energy means finer resolution. So the upgraded CEBAF will enable physicists to trace the distribution of the quarks' position and momentum in three dimensions. To do so, they must capture not only scattered electrons, but also other particles that shoot out of certain rare collisions. For example, Julie Roche of Ohio University, Athens, and colleagues will focus on a process in which an electron interacts with a quark by exchanging a photon that then rebounds out of the interaction. Working in the lab's Hall A, they aim to capture both the electron and the photon.

Tracking quarks in three dimensions would probe their swirling motion and could help explain the nucleon's spin. It would mark a "major advance" that could help unify physicists' understanding of the strong force, says Amanda Cooper-Sarkar, a physicist at the University of Oxford in the United Kingdom who does not work at Jefferson Lab.



The proton (artist's concept) is a bit like a troubled teenager: a mess inside and nearly incomprehensible.

For other clues to the strong force, Jefferson Lab researchers will look for new, exotic particles made of quarks and gluons. Physicists know of only two general kinds of particles cemented by the strong force: baryons, which contain three quarks or three antiquarks, and mesons, made of a quark-antiquark pair. But simulations suggest other combinations are possible. Particles called glueballs would consist entirely of gluons, and "hybrid" mesons and baryons would contain an extra "valence" gluon. They would have distinctive combinations of mass, spin, and symmetry properties that would show through in their decays.

A dedicated detector called GlueX housed in the new Hall D will search for those odd particles. GlueX researchers will first fire the CEBAF electron beam into a diamond crystal to produce polarized high-energy photons. These will then crash into a liquid hydrogen target to produce particles familiar and novel. If newcomers appear, the GlueX team aims to study the pattern of their properties, says Eugene Chudakov, a physicist at Jefferson Lab who oversees Hall D. "You cannot say anything [about the theory] if you see only one," he says. "You really need to be able to see several."

Jefferson Lab physicists will also explore a long-standing mystery: why an isolated proton or neutron behaves differently from one in a nucleus. In 1983, researchers with the European Muon Collaboration (EMC) at CERN fired muons both into deuterium nuclei, which contain one proton and one neutron, and into iron nuclei, which contain 26 protons and typically 30 neutrons. In theory, the momentum distribution of the quarks should be the same in the nucleons in each nucleus. Instead, researchers found a deficit of higher momentum quarks in the larger one. That "EMC effect" is unexplained.

Or Hen of Tel Aviv University in Israel and Douglas Higinbotham of Jefferson Lab have an idea of how it comes about. In 2008, researchers working in Hall A found that in the nucleus of carbon-12, the six protons and six neutrons tend to form fleeting proton-neutron pairs. Last October, they spotted similar pairs in heavier nuclei, too. The paired particles overlap spatially, and the commingling of their quark-gluon clouds could cause the EMC effect.

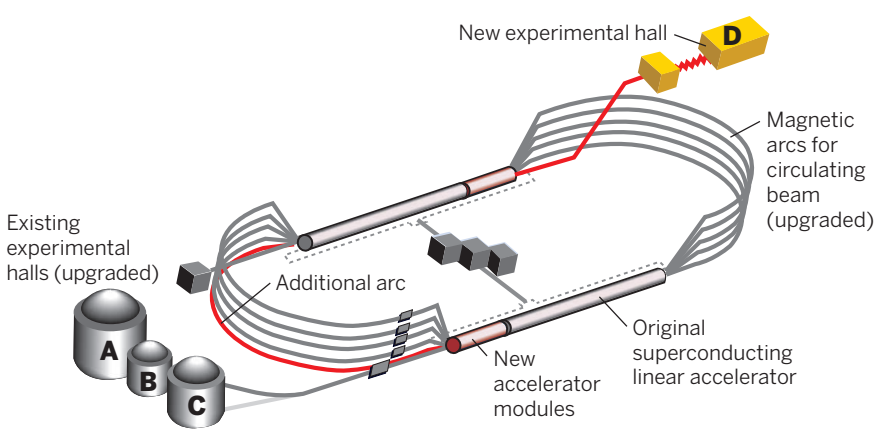
To test that idea, the physicists will fire electrons from CEBAF into deuterium nuclei to probe the quarks inside the neutron. The electrons will also break up the nuclei. In each event, measuring the momentum of the recoiling proton should tell physicists how much the proton and neutron in the nucleus overlapped when the electron struck. If the number of high-momentum quarks in the neutron decreases when the neutron and proton overlap more, then such pairing could explain the EMC effect, Hen says.

Will the upgraded CEBAF completely decipher the proton and the neutron? Not quite. The machine should be adept at mapping the quarks. But to fully probe the gluons will require higher energy collisions than CEBAF can muster. Jefferson Lab researchers now hope to build a second accelerator that would fire protons or nuclei into their electron beam, boosting the collision energies. But they have competition: Physicists at Brookhaven National Laboratory in Upton, New York, want to build an electron-ion collider by adding an electron accelerator to their $1.1 billion 4-kilometer-long atom smasher, the Relativistic Heavy Ion Collider (*Science*, 19 October 2012, p. 324).

Wherever it winds up, such a machine is at least a decade away. For now, the Department of Energy has asked both labs to work together on developing the science case for building one, says Jefferson Lab Director Hugh Montgomery: "Unless we can convince people that it's worth doing this, there's no point in discussing where it could be built." Meanwhile, Jefferson Lab physicists are ready to ramp up their new experiments. For the proton, clarity may be coming. ■

Electron racetrack

The CEBAF accelerator comprises two linear accelerators connected by arcing beam pipes. Electrons take up to five laps of the track. Thanks to a last-minute design change to the original accelerator, physicists had room to double CEBAF's energy just by extending each accelerator.



IMAGES: COURTESY OF JEFFERSON LAB (2)

THE INSURGENT

Animal rights activist Justin Goodman is using science's own tools to combat animal research

By **David Grimm**

From a small wooden desk in a row home a few miles north of the U.S. Capitol, Justin Goodman is waging war against animal research. Goodman, 34 and thin, with a scruffy beard, close-shaved head, and colorful tattoos covering most of his skin, directs laboratory investigations at People for the Ethical Treatment of Animals (PETA). Whereas other activists stick to protests and publicity stunts, he and his team have spent the past 5 years challenging scientists on their own turf. In talks and papers published in the peer-reviewed literature, they marshal data in an attempt to show researchers that animal experimentation is flawed, cruel, and just plain worthless. "We're trying to make inroads with a community that has historically despised us," he says.

Goodman never earned a Ph.D., but he's notched some significant successes as a PETA scientist. His papers have questioned the validity of the university committees that oversee animal research, encouraged U.S. allies to explore alternatives to animals in military medical training, and wounded the reputation of the world's largest accreditor of lab animal welfare.

Many researchers are unswayed, countering that although Goodman's studies have the veneer of science, they're anything but. They say the work is biased, methodologically flawed, and deeply misleading. "PETA is trying to make a point, and they've done a good job making that point," says Taylor Bennett, the senior scientific adviser for the National Association for Biomedical Research, who's based in Hinsdale, Illinois. "But I don't consider what they're doing science."

Science or not, Goodman's campaign could weaken the public's shaky support

PHOTO: PAUL FETTERS

Justin Goodman at his primary workspace—a desk in his Washington, D.C., home.

PHOTO: PAUL FETTERS

for animal studies, Bennett says. The effort is already hurting research, adds Christian Newcomer, who leads the accrediting organization Goodman has targeted. "The point is to attack everything about biomedical research at every level," he says. "They're putting up barriers to scientific progress."

But Goodman is just getting started.

"LET ME PLAY YOU SOMETHING," Goodman says, as he tilts open a laptop on his dining room table and pulls up Spotify. He clicks on a song from 1989 called "Cats and Dogs" by the hardcore punk band Gorilla Biscuits. Breakneck drumming and guitar thrash from the speakers as a nasal-voiced singer screams, "My true compassion is for all living things/And not just the ones who are cute so I do what I can/I wanna save lives and I've got a plan." Goodman nods along with the tempo. "My whole ethic as a human being can be traced back to records like this," he says.

Goodman began thinking about animal rights as a teenager. Although he grew up eating meat, something changed when, at 15, he started going to punk rock shows in dive bars and smoky nightclubs in New York City. His immediate family was wracked by drug addiction and mental illness, and the gigs and the sense of community they fostered offered escape. They were also a window into a new worldview.

"I would go see a show on a Saturday night, and there would be a table set up about animal rights issues," he says.

Pamphlets trumpeted undercover PETA investigations at chimpanzee labs; a TV hooked up to a VHS player looped videos with titles like "Meat is Murder." Bands talked about vivisection between songs and wrote about the evils of factory farming in their album liner notes. It was a world that nurtured many animal rights advocates.

"It connected with something in me," says Goodman, who soon went vegan. His voice is deep and raspy, as if his vocal cords are still strained from years of howling along with his favorite groups. "I was trying to figure out who I was and what I cared about, and it was all laid out for me on a silver platter."

He wanted to follow punk bands around the country, but forced himself to go to college instead. His grandfather had practically raised him, taking him to movies and museums when things got volatile at home. He taught Goodman how to oil paint and write letters, and he schooled his young grandson in civic engagement, bringing him along as he visited politicians in Queens to complain about everything from cracks in the sidewalk to the national debt. "After he retired, he spent 30 years auditing classes at a local university," Goodman says. "The whole reason I went to college is because I didn't want to disappoint my grandfather."

Still, Goodman struggled. He dropped out of three schools before entering San Diego City College in 2002. There, in a speech communications class, he had to give a presentation on how something is made. "I had never spoken publicly about animal issues before," he says, "but I decided to talk about where meat comes from." He delivered a 5-minute presentation about making a hamburger, from the cow to the meat patty, with no grisly details omitted. "A couple of years later, I ran into a guy who sat next to me in that class," Goodman says. "He told me he had gone vegetarian because of my talk. I realized the power an individual could have when talking about animal rights."

The issue became Goodman's calling. After college, he moved back to the East Coast with his wife, where they attended the University of Connecticut (UConn), Storrs. Both pursued Ph.D.s: his wife in experimental psychology, he in sociology. Though Goodman had begun thinking about animal rights, he had never participated in, much less organized, a protest. That changed when Herbert Terrace gave a talk on campus. The Columbia University psychologist was famous for trying to teach a chimpanzee, named Nim Chimpsky, to communicate with humans, and Goodman felt Terrace had mistreated the animal. He founded a group called the UConn Animal Rights Collective, a name that belied its modest composition: Goodman, his wife, and one other student.

Nevertheless, the group made an impact, handing out flyers at the door of Terrace's talk that blasted primate "imprisonment" in U.S. labs. "People thought we were passing out information about the talk," Goodman grins. "You should have seen their faces when they sat down."

Goodman with his dog Zoe. The "WOLF FIRE" tattoos on his knuckles refer to a book about radical environmentalism.

But what he considers his biggest coup came when he found out that a lab at UConn was doing invasive research on monkeys. Scientists had implanted devices into the skulls of a handful of rhesus macaques to study how the brain records eye movements. "The lab wasn't on anyone's radar," Goodman says, "so I started looking into it." He asked questions around campus, scoured the Internet, and reached out to a scientist at PETA, the world's largest animal rights organization, with headquarters in Norfolk, Virginia.

Although PETA was known for its campaigns against fur and factory farming, it had made its reputation fighting animal re-

search. In 1981, one of PETA's co-founders went undercover at a monkey lab in Silver Spring, Maryland, exposing animal welfare violations that resulted in the first conviction of a U.S. researcher for animal cruelty. (It was later overturned.)

The PETA scientist taught Goodman how to file Freedom of Information Act (FOIA) requests to get lab documents. "I was reading sloppily written vet records, experimental protocols, and grant applications," Goodman says. The lab files revealed welfare violations including injured animals and inadequate veterinary care, he says. Two monkeys died during the research.

Goodman's group hammered the lab in a 4-year campaign that included leafleting the campus, staging a mock funeral while wearing monkey masks, and disrupting the university's 125th anniversary festivities, shouting on bullhorns, "There's nothing to celebrate. UConn kills primates!" Goodman also sent a 40-page complaint to the U.S. Department of Agriculture. His efforts were covered by Connecticut's largest newspaper and *The New York Times*. The National Institutes of Health (NIH) eventually ordered the lab to return more than $65,000 in grant money, and the research shut down in 2006.

Empowered by the victory, Goodman decided to pursue animal rights full time. He got the words "WOLF FIRE" tattooed on his knuckles—a reference to a book about radical environmentalism—and dropped out of his Ph.D. program, leaving UConn after 5 years with just a master's degree. Then he applied for a job at PETA.

GOODMAN'S DESK SITS IN THE BACK of his house, snuggled into a corner underneath a single window on the first floor. Inherited from his grandfather, the workspace overlooks a narrow yard, where a black-and-white border collie and a chestnut-brown mutt sun themselves on the grass. On the desktop is a stack of PETA business cards, a laptop, and a framed postcard from 1928 that features a portrait of a white cat. "It says something about a mysterious and delicate creature ending up on the table of a vivisector," Goodman says, translating the French text at the bottom.

Goodman has spent the past 7 years at this desk, ever since he and his wife moved to Washington, D.C., after leaving Connecticut. (She, unlike her husband, finished her Ph.D. and now teaches experimental psychology at Marymount University in Arlington, Virginia.) PETA hired him in 2007, and he spent his first couple of years filing FOIA requests to corroborate tips about animal welfare violations at university labs. But he soon realized he could do much more. "PETA had amassed tons of data on animal research, but most people were just focused on individual cases," he says. "I wanted to find a better way to use it."

That meant compiling the data, looking for patterns, and publishing in the peer-reviewed scientific literature. "A lot of things about animal research are taken at face value," Goodman says, noting for example that researchers often claim they are trying to reduce the number of animals they use. "I knew there was actual science we could bring to the table that would challenge decisions about funding and policymaking."

His methodology tends to be simple. In 2010 and 2011, for example, he and colleagues sent e-mail surveys to all 28 NATO nations, asking whether they used animals in military medical training, like gunshot surgery or treating injuries from chemical warfare. Most said they didn't. Goodman published the findings in *Military Medicine*—a move that pressured some of the remaining countries to form a working group to explore alternatives.

Goodman (*left*) and other PETA members protest outside the National Air and Space Museum in Washington, D.C., in 2010.

At about the same time, Goodman's team requested federal and state records to glean the composition of Institutional Animal Care and Use Committees (IACUCs), which oversee animal welfare at U.S. universities. The resulting paper, published in *Animals* in 2012, showed that IACUCs at the top NIH-funded schools were heavily dominated by animal researchers. "These committees are supposed to be unbiased," Goodman says. "Our work showed that there's no meaningful oversight."

Goodman made his biggest impact last year when he published a study in the *Journal of Applied Animal Welfare Science* that questioned the validity of inspections by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International, a private organization that bestows its seal of approval on more than 900 institutions in 39 countries. The paper—based on records requests and data from government databases—claimed that AAALAC-accredited labs were *more* likely to violate animal welfare guidelines than nonaccredited labs (*Science*, 29 August 2014, p. 988). Though scientists slammed the study, it received prominent media coverage and tarnished AAALAC's reputation.

Today, Goodman commands a small army of 12 researchers, and he'll be adding more soon. The group—which also coordinates undercover investigations, runs ad campaigns, and lobbies policymakers—has published six papers and presented

PHOTO: CQ ROLL CALL

more than two dozen posters at scientific conferences. More manuscripts are in the pipeline, including studies documenting a rise in animal use in U.S. labs and research into the mental impacts of captivity on primates. Goodman hopes that by publishing in the peer-reviewed literature, he can reach an audience that has traditionally viewed PETA as the enemy. "The public can push all it wants," he says, "but ultimately we need scientists to change their minds."

That's likely to be a hard sell.

THIS PAST NOVEMBER, at a convention center just outside Washington, D.C., PETA supporters disrupted a special address by Francis Collins. As the NIH director began his presentation—celebrating the 20th anniversary of the Association for Molecular Pathology—two women began shouting, "Why do you cause the suffering of baby monkeys, Francis Collins?," a reference to an NIH lab that has been measuring the psychological impact of removing young rhesus macaques from their mothers. A security officer escorted the women from the crowded room, as they yelled "Shame on you!" while holding signs that read, "Collins: Stop Abusing Baby Monkeys."

The stunt—coordinated by Goodman's department—and similar PETA tactics are one reason Goodman will have a hard time reaching scientists, says Bennett, who oversaw animal research at the University of Illinois, Chicago, for nearly 3 decades. He adds that PETA has rebuffed attempts by him and other scientists to engage in a dialogue, leading him to believe that the group's only goal is to demonize biomedical research. "Given PETA's reputation, many scientists don't even bother to read these things," he says of Goodman's papers.

But Bennett and others say there are even more fundamental problems with the studies. Take the NATO and IACUC papers. Bennett says both draw overblown conclusions from simplistic data sets. Most NATO countries, he notes, don't have a large military presence—so of course they don't use animals in training. "It's comparing apples and oranges." And although he acknowledges that IACUCs are stacked with animal researchers, he says that the committees require this type of expertise to properly evaluate animal protocols. He adds that Goodman's study found no evidence that federal guidelines are being violated. "The paper misrepresents the process."

Similar problems plague the AAALAC study, says Bennett, who notes that a lab can be cited for reasons that may have nothing to do with how its animals are actually treated, such as mundane paperwork errors. He adds that the paper's last author, the only one unaffiliated with PETA, was paid for his work. "That raises red flags for me."

Newcomer, AAALAC's executive director, says he asked Goodman for a copy of his raw data, but Goodman refused. So AAALAC painstakingly compiled a comparable data set, which Newcomer says showed that—in some cases—PETA could not have determined which labs were and were not AAALAC-accredited. "The underlying data points are hopelessly flawed," says Newcomer, who sent a letter to the journal. "They're trying to hoodwink scientists."

Yet AAALAC has already begun to feel the impact of the study, Newcomer says. Labs pay thousands of dollars a year to be accredited, and he says his employees have started getting questions about AAALAC's utility at conferences and during lab visits. If labs opt not to become certified based on the paper, he says, it will hurt animal welfare in the long run because they will no longer have AAALAC oversight. "We'll see more stray activities."

For his part, Goodman says that scientists aren't obligated to share their data and that AAALAC has a clear bias, too. "They make millions of dollars a year on their accreditations," he says. "Of course they're going to claim that their process works." He also takes issue with the idea that his papers overreach. "Our core conclusion with the AAALAC paper is that there needs to be more transparency with the accreditation process. Coming from PETA," he laughs, "that's pretty tame."

Goodman admits that if his AAALAC data had painted the organization in a positive light, he probably wouldn't have written it up. That doesn't make him different from any other researcher, he claims. "People don't publish stuff that isn't interesting. That's just the reality of science." He says he hopes researchers can get past PETA's reputation. "We *are* an interest group," he says. "But that in and of itself does not discredit the science we do. We shouldn't be held to a higher standard than everyone else."

BEFORE I LEAVE, GOODMAN takes me upstairs to show me his two cats, Mister and Baby, brown Maine Coons rescued from a shelter in San Diego. "Right after we got back from our honeymoon, we started a family," he smiles as he cuddles Baby. He says he sees where the desire to do animal research comes from. "I grew up with close family members who suffered from addiction, mental illness, and Alzheimer's," he says. "I'm not so far removed from the issue of human suffering that I can't understand why someone would be desperate to seek treatments and cures. I just think there's a better way." ■

PETA takes on animal research

Since its founding in 1980, People for the Ethical Treatment of Animals (PETA) has sought to end all animal research. Here are some of its notable campaigns.

1980
Ingrid Newkirk and Alex Pacheco found PETA and organize a protest against animal research: World Day for Laboratory Animals.

1981
Pacheco goes undercover at a Silver Spring, Maryland, laboratory, exposing injured monkeys in filthy conditions. Lead scientist is convicted for animal cruelty, but conviction later overturned.

1986
Maryland laboratory run by SEMA Inc. stops putting chimpanzees in isolation after PETA protests.

1996
By holding stock shares and proposing shareholder resolutions at annual meetings, PETA convinces Gillette to adopt a moratorium on animal testing.

2008
PETA announces $1 million prize for lab-grown meat, to spur research. Abandons prize in 2014 after limited interest.

2010
Justin Goodman becomes an associate director of research at PETA, soon begins publishing in scientific journals.

2012
PETA donates simulators to Egypt (pictured) so country will stop using animals in medical trauma training.

2013
PETA launches an International Science Consortium, which promotes and funds animal alternatives in biomedical research.

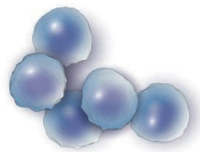
2014
PETA protests maternal deprivation experiments at NIH primate lab.

PHOTO: PETA

INSIGHTS

Human T cell plasticity *p. 371*

Debates over e-cigarettes *p. 375*



PERSPECTIVES

PLANETARY SCIENCE

Play it again, SAM

Lasers shine new light on the drying of Mars while reviving the mystery of methane

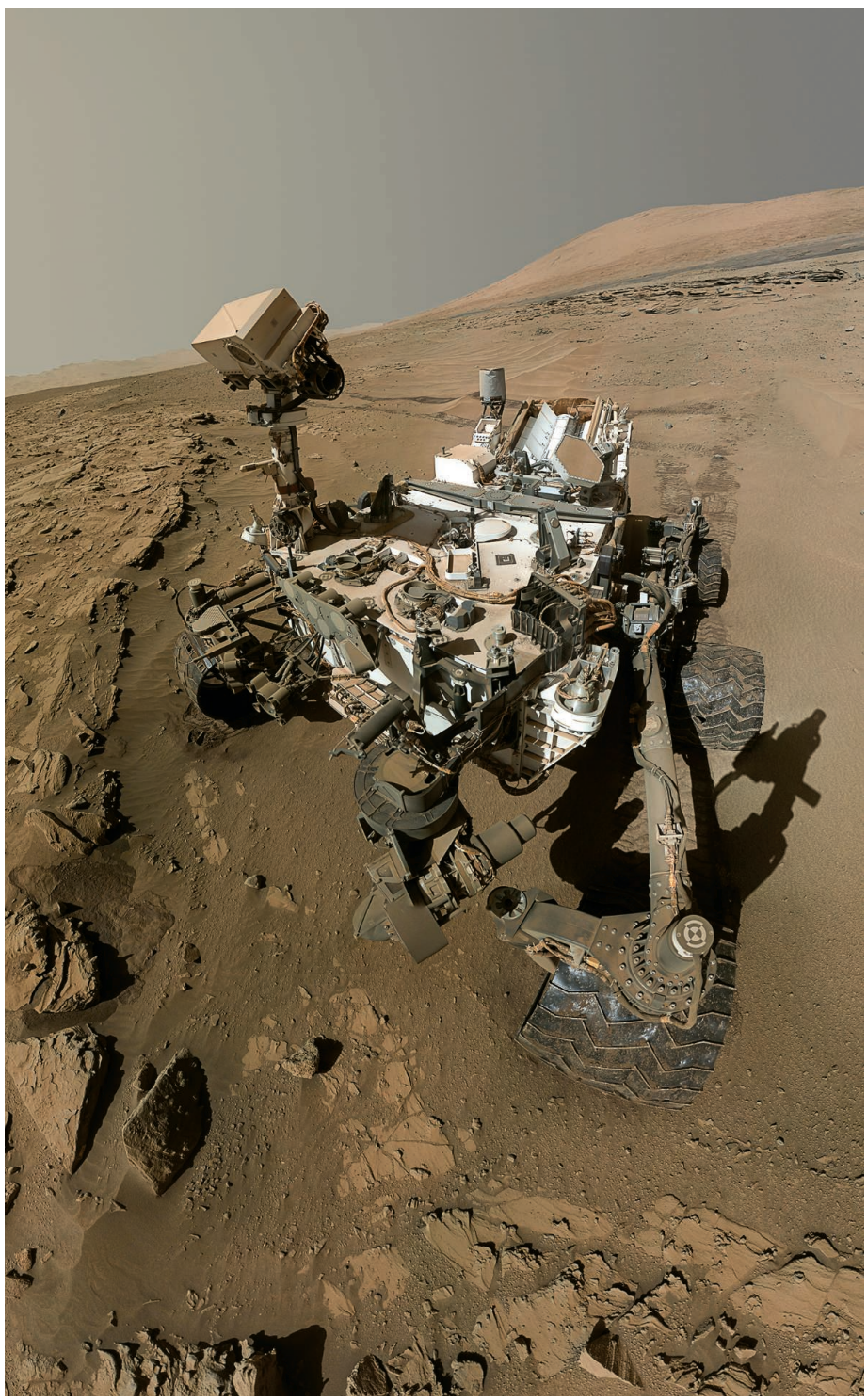
By **Kevin Zahnle**

Some discoveries are new, others old. Here, we consider one of each from NASA's Curiosity rover. The new discovery, reported by Mahaffy *et al.* (*1*) on page 412 of this issue, is a remarkable measurement of the deuterium-hydrogen (D/H) ratio in a Gale crater mudstone from 3 billion years ago. On page 415, Webster *et al.* (*2*) report on the latest chapter in the muddy matter of methane on Mars. What links them is that both were made using the tunable laser spectrometer (TLS), part of the SAM (Sample Analysis at Mars) package on the rover.

Martian water is enriched in deuterium (D), hydrogen's heavier isotope, compared with most water in the solar system. The cause of the enrichment is the preferential escape to space of H. The D/H ratio in Mars's air is now about 6 × SMOW (standard mean ocean water; that is, Earth) and was roughly the same 180 million years ago, as recorded by water trapped in martian meteorites (*3*). Curiosity has found in a single sample that D/H was only 3 × SMOW when a mudstone formed in a cold little pond at the bottom of Gale crater some 3 billion years ago (*1*).

Mars today holds the equivalent of 30 m of water in polar ice (*4*)—this is how deep the water would be if it were spread uniformly over the surface. Climate modelers think

Space Science and Astrobiology Division, NASA Ames Research Center, Moffett Field, CA, USA. E-mail: kevin.j.zahnle@nasa.gov



Looking back at you. A selfie taken by NASA's Curiosity rover. Whether the detected methane is from the rover itself or from the martian surface is not so clear.

PHOTO: NASA/JPL-CALTECH/MSSS

that all the polar water is exchangeable on 10-million-year time scales in response to Mars's Milankovich cycles (*5*). When D escape is taken into account, the implication is that Mars had at least ~90 m of exchangeable water at the time of the pond.

All this water will make many Mars scientists happy, but there are consequences. Today, H escape is equivalent to the loss of ~2 m of water per billion years, which is a lot less than ~60 m in 3 billion years. Faster H escape in the past, although not unexpected, needs to be balanced by an oxygen sink (*6*). Extreme oxidation is seen at the surface of Mars (for example, perchlorates), but there is no evidence that oxidation is deep enough and pervasive enough to accommodate all the oxygen from 60 m of water in the past 3 billion years. Apparently most of the oxygen escaped with the hydrogen (*6*).

"William of Ockham (12) would warn us to be wary of peekaboo methane when a known source—the rover—is so nearby."

The saga of methane on Mars begins with its first discovery in 1969. That announcement, based on a spectrum obtained 48 hours earlier by Mariner 7, was greeted by a front-page story in the *New York Times* (*7*). The team soon realized that they had actually seen a forbidden band in frozen CO_2. But the fascination with methane—the simplest, most stable, and most abundant organic molecule in the cosmos—has not gone away. Methane does not have a known chemical source in an atmosphere like Mars's, and its lifetime [standard photochemistry predicts 300 years (*8*)] is short enough that its presence in the atmosphere almost demands an exciting source. Pursuit has been vigorous, and there have been many subsequent discoveries of variable credibility and consistency (*2*). The reports describe an ephemeral gas with a lifetime of weeks or months rather than the expected 300 years, and the best of them describes phenomena seen only during the winter of 2003 (*9*). The reality of ephemeral methane has been contested nearly as vigorously (*10*, *11*).

The TLS/SAM experiment was intended to resolve the matter by looking for a distinctive pattern of spectral lines that uniquely identifies methane. The TLS looks at two slivers, one for the isotopes of H, C, and O, and the other for CH_4. The good news is that a lot of methane is seen. The bad news is that little of the methane is martian. Most is in the antechamber to the sample cell and comes from several sources, known and unknown, in the rover itself. Martian methane, when present, would be seen in the small difference between the signal obtained when there is martian air in the cell and the signal obtained when the cell is empty.

Figure 1 of Webster *et al.* chronicles the differences (*2*). At first, no martian methane was seen, both while the rover was awash in stowaway Florida air and then after the rover was evacuated. But later, as methane slowly built up again inside the rover [see column I, table S2, of the SM in (*2*)], methane appeared in five of six samples of Mars's air at levels on the order of 7 parts per billion by volume (ppbv). The statistics are marginal, but the measurements are self-consistent. After the fifth sighting, TLS/SAM performed the first of two higher-sensitivity enrichment experiments, only to find the methane nearly gone. A second enrichment experiment done 3 months later gives a low but nonzero CH_4 abundance of 0.9 ppbv with better statistics. Although 0.9 ppbv may not seem like much, it is probably more than can be supplied in steady state by the degradation of incoming exogenic matter.

It is an intriguing story. William of Ockham (*12*) would warn us to be wary of peekaboo methane when a known source—the rover—is so nearby. Because the concentration of methane inside the rover is approximately 1000 times as high as that in the martian air, it would not take much. But it would be a mistake to be too confident that the methane was never there. We know very little about Mars. The synchronous report that O_2 is seasonally variable [from Chemcam, another Curiosity instrument (*2*)] sends a clear warning that theory may be missing something fundamental about Mars's atmosphere. Methane could be caught up in this.

There is no denying that Mars has been something of a disappointment (*13*), at least if one had been hoping for a second Earth. The disappointment of today may turn to the wonder of tomorrow when we come to see Mars as the unique world that it is. ■

REFERENCES AND NOTES

1. P. R. Mahaffy *et al.*, *Science* **347**, 412 (2015).
2. C. R. Webster *et al.*, *Science* **347**, 415 (2015).
3. D. Bogard *et al.*, *Space Sci. Rev.* **96**, 425 (2001).
4. P. Christensen, *Elements* **2**, 151 (2006).
5. B. Levrard *et al.*, *J. Geophys. Res.* **112**, E06012 (2007).
6. H. Lammer *et al.*, *Int. J. Astrobiol.* **2**, 195 (2003).
7. W. Sullivan, *New York Times*, 8 August 1969, p. 1.
8. V. Krasnopolsky, J. Maillard, T. Owen, *Icarus* **172**, 537 (2004).
9. M. J. Mumma *et al.*, *Science* **323**, 1041 (2009).
10. K. Zahnle, R. Freedman, D. Catling, *Icarus* **212**, 493 (2011).
11. I have contested in print previous claims that methane was detected on Mars.
12. C. F. Chyba, *Nature* **348**, 113 (1990).
13. E. Anders, T. Owen, *Science* **198**, 453 (1977).

10.1126/science.aaa3687

IMMUNOLOGY

Flexibility for specificity

Human T cells show remarkable differentiation plasticity in responding to pathogens

By **Mark M. Davis**

Different types of T lymphocytes play a key role in many immune responses, such as killing virally infected or cancerous cells directly, inducing high-affinity antibody responses in B cells, and increasing or decreasing responses from other immune cells. This multiplicity of roles may relate to their recognition properties, which are very difficult to evade. Moreover, cells are very diverse—for $CD4^+$ T cells alone, there are at least six distinct subtypes. This raises the question of just how these different T cells are produced. Early evidence indicated that the type of T cell that dominates the response was dependent on the type of pathogen and route of entry (*1*, *2*). However, over the past several years, more and more flexibility has been observed in a T cell's phenotype (*3*, *4*). On page 400 of this issue, Becattini *et al.* (*5*) show that this flexibility is more the rule rather than the exception.

Most T cells use antibody-like T cell receptors (TCRs) on their cell surface to recognize degraded bits of proteins (peptides) or lipids that are bound to molecules of the major histocompatibility complex (MHC) expressed on the surface of various cell types. These MHC molecules are much like the people in Hollywood who go through the garbage cans of celebrities, looking for interesting bits of information to display to the tabloid press. But MHC molecules cannot tell whether the fragments they bind are from endogenous proteins of the cell; that is the job of T cells, with their diverse TCRs. The functional beauty of this system, and perhaps the reason for its centrality in so much of adaptive immunity, is that everything pertaining to a cell or a pathogen is eventually degraded, making it much harder for pathogens to evade detection. By contrast, antibodies must recognize intact antigens and thus are often diverted from the most important bits. In particular, MHC molecules bind small peptides—typically 8 to 10 amino acids—just enough information to be fairly unique, making it easier

that all the polar water is exchangeable on 10-million-year time scales in response to Mars's Milankovich cycles (*5*). When D escape is taken into account, the implication is that Mars had at least ~90 m of exchangeable water at the time of the pond.

All this water will make many Mars scientists happy, but there are consequences. Today, H escape is equivalent to the loss of ~2 m of water per billion years, which is a lot less than ~60 m in 3 billion years. Faster H escape in the past, although not unexpected, needs to be balanced by an oxygen sink (*6*). Extreme oxidation is seen at the surface of Mars (for example, perchlorates), but there is no evidence that oxidation is deep enough and pervasive enough to accommodate all the oxygen from 60 m of water in the past 3 billion years. Apparently most of the oxygen escaped with the hydrogen (*6*).

"William of Ockham (12) would warn us to be wary of peekaboo methane when a known source—the rover—is so nearby."

The saga of methane on Mars begins with its first discovery in 1969. That announcement, based on a spectrum obtained 48 hours earlier by Mariner 7, was greeted by a front-page story in the *New York Times* (*7*). The team soon realized that they had actually seen a forbidden band in frozen CO_2. But the fascination with methane—the simplest, most stable, and most abundant organic molecule in the cosmos—has not gone away. Methane does not have a known chemical source in an atmosphere like Mars's, and its lifetime [standard photochemistry predicts 300 years (*8*)] is short enough that its presence in the atmosphere almost demands an exciting source. Pursuit has been vigorous, and there have been many subsequent discoveries of variable credibility and consistency (*2*). The reports describe an ephemeral gas with a lifetime of weeks or months rather than the expected 300 years, and the best of them describes phenomena seen only during the winter of 2003 (*9*). The reality of ephemeral methane has been contested nearly as vigorously (*10*, *11*).

The TLS/SAM experiment was intended to resolve the matter by looking for a distinctive pattern of spectral lines that uniquely identifies methane. The TLS looks at two slivers, one for the isotopes of H, C, and O, and the other for CH_4. The good news is that a lot of methane is seen. The bad news is that little of the methane is martian. Most is in the antechamber to the sample cell and comes from several sources, known and unknown, in the rover itself. Martian methane, when present, would be seen in the small difference between the signal obtained when there is martian air in the cell and the signal obtained when the cell is empty.

Figure 1 of Webster *et al.* chronicles the differences (*2*). At first, no martian methane was seen, both while the rover was awash in stowaway Florida air and then after the rover was evacuated. But later, as methane slowly built up again inside the rover [see column I, table S2, of the SM in (*2*)], methane appeared in five of six samples of Mars's air at levels on the order of 7 parts per billion by volume (ppbv). The statistics are marginal, but the measurements are self-consistent. After the fifth sighting, TLS/SAM performed the first of two higher-sensitivity enrichment experiments, only to find the methane nearly gone. A second enrichment experiment done 3 months later gives a low but nonzero CH_4 abundance of 0.9 ppbv with better statistics. Although 0.9 ppbv may not seem like much, it is probably more than can be supplied in steady state by the degradation of incoming exogenic matter.

It is an intriguing story. William of Ockham (*12*) would warn us to be wary of peekaboo methane when a known source—the rover—is so nearby. Because the concentration of methane inside the rover is approximately 1000 times as high as that in the martian air, it would not take much. But it would be a mistake to be too confident that the methane was never there. We know very little about Mars. The synchronous report that O_2 is seasonally variable [from Chemcam, another Curiosity instrument (*2*)] sends a clear warning that theory may be missing something fundamental about Mars's atmosphere. Methane could be caught up in this.

There is no denying that Mars has been something of a disappointment (*13*), at least if one had been hoping for a second Earth. The disappointment of today may turn to the wonder of tomorrow when we come to see Mars as the unique world that it is. ■

REFERENCES AND NOTES

1. P. R. Mahaffy *et al.*, *Science* **347**, 412 (2015).
2. C. R. Webster *et al.*, *Science* **347**, 415 (2015).
3. D. Bogard *et al.*, *Space Sci. Rev.* **96**, 425 (2001).
4. P. Christensen, *Elements* **2**, 151 (2006).
5. B. Levrard *et al.*, *J. Geophys. Res.* **112**, E06012 (2007).
6. H. Lammer *et al.*, *Int. J. Astrobiol.* **2**, 195 (2003).
7. W. Sullivan, *New York Times*, 8 August 1969, p. 1.
8. V. Krasnopolsky, J. Maillard, T. Owen, *Icarus* **172**, 537 (2004).
9. M. J. Mumma *et al.*, *Science* **323**, 1041 (2009).
10. K. Zahnle, R. Freedman, D. Catling, *Icarus* **212**, 493 (2011).
11. I have contested in print previous claims that methane was detected on Mars.
12. C. F. Chyba, *Nature* **348**, 113 (1990).
13. E. Anders, T. Owen, *Science* **198**, 453 (1977).

10.1126/science.aaa3687

IMMUNOLOGY

Flexibility for specificity

Human T cells show remarkable differentiation plasticity in responding to pathogens

By **Mark M. Davis**

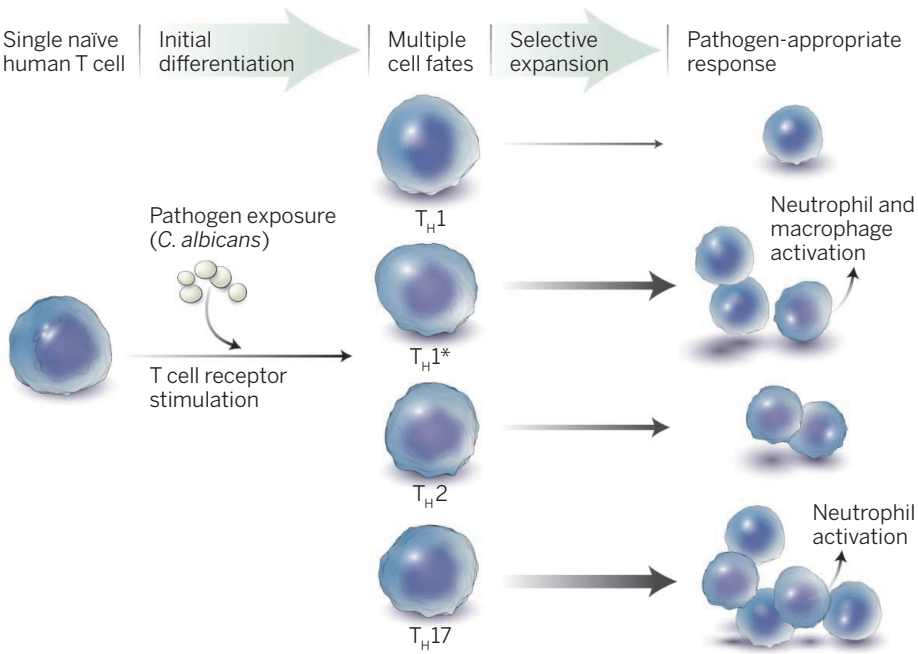
Different types of T lymphocytes play a key role in many immune responses, such as killing virally infected or cancerous cells directly, inducing high-affinity antibody responses in B cells, and increasing or decreasing responses from other immune cells. This multiplicity of roles may relate to their recognition properties, which are very difficult to evade. Moreover, cells are very diverse—for $CD4^+$ T cells alone, there are at least six distinct subtypes. This raises the question of just how these different T cells are produced. Early evidence indicated that the type of T cell that dominates the response was dependent on the type of pathogen and route of entry (*1*, *2*). However, over the past several years, more and more flexibility has been observed in a T cell's phenotype (*3*, *4*). On page 400 of this issue, Becattini *et al.* (*5*) show that this flexibility is more the rule rather than the exception.

Most T cells use antibody-like T cell receptors (TCRs) on their cell surface to recognize degraded bits of proteins (peptides) or lipids that are bound to molecules of the major histocompatibility complex (MHC) expressed on the surface of various cell types. These MHC molecules are much like the people in Hollywood who go through the garbage cans of celebrities, looking for interesting bits of information to display to the tabloid press. But MHC molecules cannot tell whether the fragments they bind are from endogenous proteins of the cell; that is the job of T cells, with their diverse TCRs. The functional beauty of this system, and perhaps the reason for its centrality in so much of adaptive immunity, is that everything pertaining to a cell or a pathogen is eventually degraded, making it much harder for pathogens to evade detection. By contrast, antibodies must recognize intact antigens and thus are often diverted from the most important bits. In particular, MHC molecules bind small peptides—typically 8 to 10 amino acids—just enough information to be fairly unique, making it easier

for T cells to discern what belongs in the body ("self") and what doesn't. Although this rationale is somewhat speculative, the diversity of T cell types is very clear.

Becattini *et al.* used two pathogens (*Candida albicans* and *Mycobacterium tuberculosis*) and a protein (tetanus toxin with the adjuvant alum) to show that a variety of different T cell types can be derived from a single initial T cell clone. Because TCRs do not mutate, unlike antibody genes, it is relatively easy to follow a particular T cell clone through its various divisions, and (in this case) the different functional properties of the daughter clones, just by sequencing the TCRs expressed by the clonal population and correlating sequences with phenotypic information.

Becattini *et al.* focused on four CD4$^+$ T cell types: T helper 1 (T_H1) cells, which produce the cytokine interferon IFN-γ to combat intracellular bacteria and viruses and activate macrophages; T_H2 cells, which make interleukins IL-4, IL-5, and IL-13 and help control parasites and activate eosinophils; a recently discovered T_H1^* type (also known as $T_H1/17$) that makes some IL-17 in addition to IFN-γ; and T_H17 cells, which produce IL-17 and IL-22, mobilize neutrophils, and combat fungi and extracellular bacteria. The authors took peripheral blood cells from up to five human donors, loaded them with carboxyfluorescein succinimidyl ester (CFSE, a fluorescent dye that allows one to track the number of cell divisions), and stimulated the cells with the fungus *C. albicans*. Memory T cells were recovered in bulk or isolated from each of the four different T cell populations surveyed, with the criterion of low CFSE staining to collect only those T cells that had divided in response to the pathogen. The TCRβ chains were then sequenced in bulk for each population, using high-throughput techniques. As expected for this pathogen, T_H17 cells were the most common, but T_H1^* cells also were abundant. Surprisingly, all four cell types were represented and were relatively equally diverse (595 to 976 unique TCRβ sequences in each population), indicating that at least initially, T cells of all predispositions were drawn upon. Most important, there was widespread sharing of TCRβ sequences in the different populations; in some cases, the same sequence occurred in all four T cell types. Subsequently, T cell cloning (deriving a population of T cells from a single member of each T cell population) confirmed these results (each cloned population had a T_H1, T_H2, T_H1^*, or T_H17 phenotype), as did determining that the same TCR sequences were present in the TCRs across these different cloned T cell populations. The same bulk approach was also used to analyze responses to *M. tuberculosis* and tetanus toxin, the former an intracellular mycobacterium and the latter a single protein. In the case of *M. tuberculosis*, the data were similar with respect to the initial diversity of the different TCR repertoires, but there was much less sharing of sequences between T cell types. For tetanus toxin, there was widespread sharing of TCR sequences, even more than observed with *C. albicans*.



One cell, multiple fates. A naïve human T cell primed with a pathogen in vitro gives rise to multiple T cell subtypes through a proposed two-phase process. In the first phase, specific T cells are stimulated to proliferate regardless of their phenotype. In the second phase, there is a selective expansion of T cells with the most appropriate functional properties, and multiple types can arise from the same original T cell clone. This model may explain how T cells acquire their specialized responses to infection and vaccines.

The extensive analysis of three different antigens makes a strong case that the recruitment of at least these types of CD4$^+$ T cells is relatively random at first, but that specific functional cells can be expanded preferentially by some process of selection and interaction with the pathogen or antigenic challenge (see the figure). That this principle is likely to extend to other CD4$^+$ T cell types and CD8$^+$ T cells is suggested by an earlier study that showed CD8$^+$ T cells specific for a particular viral epitope could express more than 100 different combinations of cytokines or chemokines (*6*). Although the study did not involve TCR sequencing, it does show that T cell responses can be much more diverse functionally than previously thought.

A question yet to be explored concerns how particular types of T cells are favored by pathogens (or by adjuvants, as is likely the case with the tetanus toxin exposure). Knowing the answer could help in the design of vaccines that boost the types of T cells that can do the most good. Another question is how flexible these T cell responses are: How likely are they to change with time, and under what influences? Data show that there is a delicate balance that can tip back and forth between T_H1- and T_H2-type cells (*7*). Other results show changes in the cytokine secretion of single CD4$^+$ T cells with time (*8*). And there is improved technology for deriving both TCR sequences and phenotypic information from single T cells (*9*), which will make studies like this much easier, and perhaps unravel even further the degree of heterogeneity in T cell responses to infection and vaccines. ■

Howard Hughes Medical Institute, Institute for Immunity, Transplantation and Infection, Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA. E-mail: mmdavis@stanford.edu

REFERENCES AND NOTES

1. S. L. Reiner, F. Sallusto, A. Lanzavecchia, *Science* **317**, 622 (2007).
2. J. J. O'Shea, W. E. Paul, *Science* **327**, 1098 (2010).
3. D. Bending *et al.*, *J. Clin. Invest.* **119**, 565 (2009).
4. N. J. Tubo *et al.*, *Cell* **153**, 785 (2013).
5. S. Becattini *et al.*, *Science* **347**, 400 (2015).
6. E. W. Newell, N. Sigal, S. C. Bendall, G. P. Nolan, M. M. Davis, *Immunity* **36**, 142 (2012).
7. Y. E. Antebi *et al.*, *PLOS Biol.* **11**, e1001616 (2013).
8. Q. Han *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1607 (2012).
9. A. Han, J. Glanville, L. Hansmann, M. M. Davis, *Nat. Biotechnol.* **32**, 684 (2014).

ACKNOWLEDGMENTS

Supported by the Howard Hughes Medical Institute, NIH, and the Bill and Melinda Gates Foundation.

10.1126/science.aaa5082

ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

Understanding artemisinin resistance

Genetic studies provide clues to how malaria parasites become drug resistant

By **Carol Hopkins Sibley**

The drug resistance specter looms over most infectious diseases. Malaria provides a particularly urgent example of increasing resistance and treatment failure. In the past decade, artemisinin combination therapies (ACTs) have contributed to impressive reductions in malaria morbidity and mortality (*1*). However, in 2009, Dondorp *et al.* found that when patients in western Cambodia infected with *Plasmodium falciparum* (the deadliest form of malaria) were treated with ACTs, they took longer than normal to clear their parasites. It is the artemisinin component that normally clears parasites quickly, and the authors therefore concluded that this component of the ACT was compromised (*2*). Slow-clearing parasites were also found in western Thailand and other parts of Cambodia (see the first figure) (*3*, *4*). Two reports in this issue, by Mok *et al.* (page 431) (*5*) and Straimer *et al.* (page 428) (*6*), apply specialized techniques to better understand the mechanisms that underlie this resistance to artemisinins.

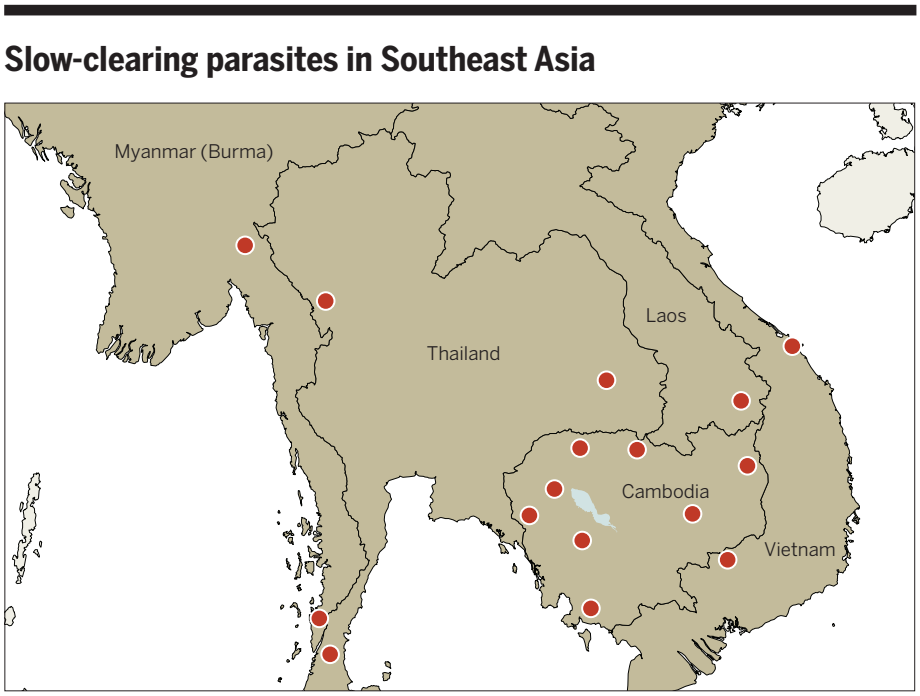
It is difficult and time-consuming to measure the speed of parasite clearance in patients. Collaborative efforts were therefore launched to define genetic changes that might quickly identify the resistant parasites. Studies focused on parasites from Cambodia found a number of broad genomic regions that seemed like good candidates for such molecular markers of resistance. Building on the work of an international consortium (*7*), Ariey *et al.* identified Kelch 13 (K13) as a gene strongly correlated with the slow-clearance phenotype in parasites from western Cambodia (*8*).

A molecular marker of resistance is usually defined by a simple, consistent pattern of genetic changes called single nucleotide polymorphisms (SNPs), but K13 is different. Studies in western Cambodia found that each slow-clearing parasite contained one of 15 different mutations that change a single amino acid in the "propeller region" of the K13 protein. As the geographic range of studies widened, more alleles have been discovered, further increasing the inventory of new mutations in the Mekong region (*9*, *10*). Even more worrisome, new populations of resistant parasites have emerged independently in Myanmar (*10*). It appears that the slow-clearing phenotype is spreading in the Mekong region, not because all parasites carry a single, highly resistant K13 mutant but mainly because a wide variety of K13 mutants are emerging independently.

Resistance to earlier antimalarial drugs spread from the Mekong region to Africa (*11*). To investigate whether artemisinin resistance had spread already, two groups assessed ~2000 parasite isolates from 25 sites across Africa (see the second figure) (*12*, *13*). They identified 20 additional K13 alleles. Most were rare (1 to 3% of the population); a few were shared among distant African sites, but most were observed in only one location. This observation raised an important question: Does artemisinin resistance already have a toehold in Africa, or are other changes in parasites also required?

The two reports in this issue help to clarify some aspects of this muddy situation. Straimer *et al.* used zinc-finger nucleases to change single SNPs and thereby create or revert K13 propeller mutants in cultured malaria parasites. They adapted a recently developed laboratory method (*7*) to test the artemisinin sensitivity of the modified parasites and found that when a resistant K13 mutation was repaired to wild type in Cambodian parasites, artemisinin resistance was lost. Conversely, conversion of a wild-type K13 gene to a propeller mutant rendered drug-sensitive lines more resistant to artemisinin treatment. This approach provides a laboratory test to determine whether K13 mutants like those seen in Africa can alone confer artemisinin resistance in a parasite.

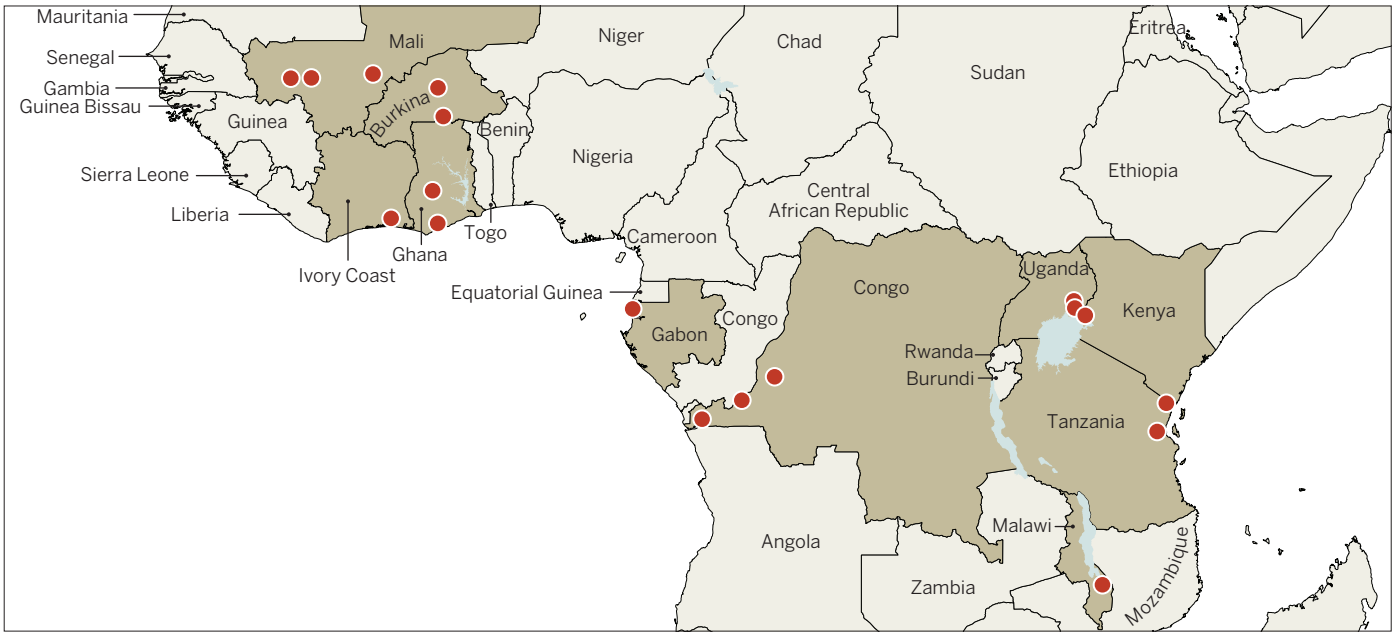
In other organisms, the Kelch proteins often act as interaction hubs that bind other proteins involved in a network of stress responses (*8*). Different K13 propeller mutants confer slow clearance on Mekong parasites (*8*–*10*), but supporting changes in stress responses might also be needed to render the parasites fully resistant to artemisinin. If so, the slow-clearing parasites should share other parts of their genome, even when they carry different K13 alleles. In fact, earlier genome-wide analyses found multiple genetic regions common to the slow-clearing parasites, which did not all carry the same K13



Spread of resistance in Southeast Asia. In the past 5 years, many different drug-resistant malaria parasites have emerged in Southeast Asia. Two articles in this issue (*5*, *6*) provide specialized approaches to define more fully the molecular basis of artemisinin resistance in malaria parasites. Data from (*8*, *9*, *10*).

Worldwide Antimalarial Resistance Network, Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. E-mail: carol.sibley@wwarn.org

K13 propeller mutations in Africa



On the path to resistance? African parasites carry some of the same mutations found in Southeast Asia, as well as novel ones, although artemisinin-resistant parasites have not yet been found there. Data from (*12*, *13*).

allele [see references in (*8*)].

The results of Straimer *et al.* offer direct support for this "permissive genome" idea. The authors used the new method to test the same Mekong-derived K13 alleles in cultured parasites from various geographic regions. They found that Mekong-derived parasites survived artemisinin treatment better than those from Africa, even when both were engineered to carry the same allele of K13.

Mok *et al.* address the question of whether other genes are also required to allow K13 mutant parasites to survive artemisinin exposure. Malaria parasites are unusual because their genes are transcribed in a set temporal pattern during the 72-hour cell cycle of *P. falciparum*. Using that metric, the authors identified three different patterns of transcription; these patterns correlated with progression through the cell cycle. One transcription pattern correlated with slow clearance and with delayed development early in the growth cycle. Given that artemisinins are metabolized in a few hours, this early delay could protect the resistant parasites during that brief, critical period (*14*).

The authors then measured the mRNA expression of genes during this early stage, comparing individual parasites from different geographic regions, and again related expression levels to their clearance phenotype. Genes involved in the oxidative stress response were prominent among those differentially expressed in slow- and fast-clearing parasites, providing strong evidence that genes in these networks are involved in the artemisinin response.

These findings suggest that the diverse mutations that have been observed in K13 may each diminish the capacity of the protein to function as a binding partner for proteins in the stress response network (*8*). Changes in expression of those stress proteins could compensate for the reduced K13 function and protect the parasites against artemisinin challenge.

Generally, drug resistance is tracked by identifying a few SNPs that uniquely identify resistant pathogens. The diversity of K13 mutant alleles found in slow-clearing parasites does not conform to that simple model. Even more puzzling, the slow-clearing phenotype is spreading in the Mekong region, not because a highly resistant parasite is spreading under selection but mainly due to the independent emergence of a wide variety of K13 mutants. Even more alleles of K13 have emerged in African parasites, but it is not clear whether a K13 mutation alone is enough to render an African parasite artemisinin-resistant or whether other genetic changes are also required.

The knowledge that—as Straimer *et al.* show—a parasite carrying a K13 propeller mutant is protected against artemisinin in the laboratory allows novel K13 alleles identified only by molecular methods to be tested in that assay. If African parasites lack "permissive genomes," the presence of a K13 mutant allele alone may not (yet) signal artemisinin resistance. African and Southeast Asian parasite genomes differ markedly, and identifying differences in the families of genes that respond to oxidative stress could thus help to find the full complement of genes required to confer artemisinin resistance.

Artemisinin resistance in the Mekong is now well established (*9*, *10*). Eliminating those resistant parasites before they spread more widely is crucial (*3*). The insights in the two reports in this issue (*4*, *5*) provide a basis for a clearer definition of a molecular signature. These markers can then be used to detect new foci of artemisinin-resistant parasites at the earliest possible time, and allow appropriate responses to be mobilized. ■

REFERENCES

1. World Health Organization, World Malaria Report, www.who.int/malaria/publications/world_malaria_report_2014/en; accessed 11 Dec. 2014.
2. A. M. Dondorp *et al.*, *N. Engl. J. Med.* **361**, 455 (2009).
3. A. M. Dondorp, P. Ringwald, *Trends Parasitol.* **29**, 359 (2013).
4. A. P. Phyo *et al.*, *Lancet* **379**, 1960 (2012).
5. S. Mok *et al.*, *Science* **347**, 431 (2015).
6. J. Straimer *et al.*, *Science* **347**, 428 (2015).
7. B. Witkowski *et al.*, *Lancet Infect. Dis.* **13**, 1043 (2013).
8. F. Ariey *et al.*, *Nature* **505**, 50 (2014).
9. E. A. Ashley *et al.*, *N. Engl. J. Med.* **371**, 411 (2014).
10. S. Takala-Harrison *et al.*, *J. Infect. Dis.* 10.1093/infdis/jiu491 (2014).
11. C. Roper *et al.*, *Science* **305**, 1124 (2004).
12. S. M. Taylor *et al.*, *J. Infect. Dis.* 10.1093/infdis/jiu467 (2014).
13. E. Kamau *et al.*, *J. Infect. Dis.* 10.1093/infdis/jiu608 (2014).
14. N. Klonis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5157 (2013).

10.1126/science.aaa4102

PUBLIC HEALTH

Smoke and fire over e-cigarettes

As nations adopt regulatory measures for e-cigarettes, it is imperative to understand how approaches to risk, cost-benefit, and trade-offs have shaped interpretations of evidence

By **Amy L. Fairchild****and* **Ronald Bayer**

In advance of a critical Framework Convention on Tobacco Control (FCTC) held in October 2014, two groups of scientists and public health experts launched a global battle royal over electronic cigarettes—devices that heat liquid nicotine but involve no tobacco.

Leaked documents appearing in the late spring indicated that the World Health Organization (WHO) was preparing to take an uncompromising stance against e-cigarettes at the FCTC meeting, treating them as a threat equal to tobacco cigarettes. This set the stage for 53 researchers from 18 nations, dominated by the United Kingdom, United States, and Australia, to send an urgent appeal to Dr. Margaret Chan, Director-General of WHO, on 26 May 2014. It was critical, they argued, to remain open to evidence regarding "low-risk noncombustible nicotine or tobacco products that may become viable alternatives to smoking in the future" as a potential harm-reduction strategy (*1*). There was "no evidence at present of material risk to health from vapour emitted from e-cigarettes" nor "credible evidence" that e-cigarettes would serve as a gateway to tobacco smoking. "We hope," concluded the letter that "WHO will be in the vanguard of science-based, effective and ethical tobacco policy, embracing harm reduction."

POLICY

Three weeks later, on 16 June 2014, a response was sent to Dr. Chan, organized by American researcher Dr. Stanton Glantz (*1*) and cosigned by 129 experts from more than 24 countries. E-cigarettes, the letter stated, were little more than a Trojan horse promoted by an industry bent only on "increasing profits" through "predatory" practices. The first letter to Chan, it charged, had made assertions about marketing, emissions, and harms that were "either contradicted by available evidence or for which no evidence is currently available." Indeed, the letter warned, the harm-reduction advocates had not cited "a single scientific study" (*1*). Charges and countercharges continued.

How can two groups, both of which seek to reduce the terrible burden of morbidity and mortality attributable to smoking, both of whom embrace the centrality of evidence-based policy, come to such different conclusions? To be sure, the strength of the current evidence and the basic soundness and logic of extrapolation based on that evidence are central. It may be years before these disagreements over the evidence—which involve profound contention over whether e-cigarettes will serve as "gateway" drugs, particularly for youth; whether "dual" use with tobacco cigarettes will undercut any morbidity and mortality benefits; and whether this potentially disruptive technology will prove to be an effective cessation tool—will be resolved.

Risk versus benefit.

But it is not evidence alone that accounts for this pitched battle. The opposing letters reflect very different understandings of what the protection of public health requires. Those who called upon WHO to remain open to the possibility that e-cigarettes could reduce the toll of smoking explicitly embraced harm reduction as an organizing principle. In contrast, the judgments of the second letter, while focused primarily on the science, were shaped implicitly by a precautionary impulse.

HARM REDUCTION AND PRECAUTION. arm reduction—an approach that embraces a posture of pragmatism and accepts that people will use drugs—has explicitly informed the position of those open to the use of e-cigarettes as a means to limit the toll of smoking-related morbidity and mortality (1). Some advocates argue that elimination of a habit like smoking should always be the goal; others maintain that risk minimization is sufficient (*2*). But regardless of how advocates position themselves on the question of cessation, the bar that must be met is not whether an alternative carries any risk, but whether there is enough evidence to suggest that the risks are less consequential than those of the behaviors in question. From this perspective, even uncertain evidence justifies action when the status quo—in the case of smoking, a projected one-billion deaths this century if left unchecked—is sufficiently threatening. Harm reduction requires that every piece of evidence be viewed against this deadly backdrop, including the harms of limiting or denying access to alternatives to tobacco cigarettes.

A very different approach informs those who want to ban or impose severe restrictions on e-cigarettes. For tobacco-control advocates taking a precautionary stance, those who would accept lesser harms are being duped by the industry, serving as little more than "naïve" pawns in a grand scheme to take back lost ground in the long battle over smoking (*3*). Given the long history of tobacco industry deception, such advocates assert that there can be no room for compromise when it comes to a product in which Big Tobacco has any interest.

This position echoes the Wingspread Statement of 1998, one of the foundational documents in the history of precautionary thinking: "When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically" (*4*). The Wingspread authors were particularly concerned that assessment of risks be conducted in advance of allowing the introduction of new products or practices. This, of course, is the logic that supports the system of drug regulation in Europe and the United States, which requires that products be

*Department of Sociomedical Sciences, Rosenfield Building, Columbia University, New York, NY 10032, USA. *Corresponding author. E-mail: alf4@columbia.edu*

PHOTO: THINKSTOCK

proven "safe and effective"—not in absolute terms, but, as in the case of harm reduction, in a kind of risk-risk calculus—before being allowed onto the market. Nicotine replacement therapies—such as the patch and nicotine gum, for which the field has accepted the potential risks of long-term use—are tolerable given the consequences of smoking and comfortably fit within this model.

There are multiple understandings of the Precautionary Principle, some of which acknowledge the necessity of trade-offs and risk-benefit analysis. Other versions find the very notion of trade-offs morally offensive (*5*, *6*). The question is, then, which version of the Precautionary Principle is at issue here?

In its strictest interpretation—sometimes called "deep green" precaution, reflecting the concept's roots in environmental protection—the principle holds that any suspicion of harm should be sufficient to trigger prohibition, even "in the absence of any scientific evidence" and "without regard to cost" (*6*, *7*). Deep green precaution has the virtue of consistency, demanding and prohibiting certain courses of action when evidence is contested or unavailable. In the recent clash, the fiercest opposition to e-cigarettes has reflected a logic resembling a "deep green" version of the Precautionary Principle. Opponents of this perspective, most notably Cass Sunstein, have referred to the Precautionary Principle as the "paralysis principle" (*8*), arguing that it substitutes intuitive fear for scientific proof and that its hostility to cost-benefit analysis would impose regulatory standards that, in the end, would be both socially costly and harmful. Indeed, in 2000, the European Commission offered guidelines that clearly sought to address the "mixed, and sometimes contradictory views" on precaution. What was essential in confronting risk, the guidelines stated, was "finding the correct balance" as a way to "avoid unwarranted recourse to the precautionary principle, as a disguised form of protectionism" (*9*).

Not all the supporters of a precautionary approach embrace an uncompromising conception of precaution. For example, the economist Frank Chalupka [who signed a letter from the opponents of harm reduction] maintains that his views are best reflected in the "middle ground" American Heart Association guidelines on e-cigarettes that he coauthored. This policy statement offers a nuanced view of the evidence (*10*) and suggests a balanced regulatory approach, much like that being vetted by the U.S. Food and Drug Administration (FDA), which acknowledges scientific uncertainty and tries to draft a flexible framework with the capacity to add regulation over time, in response to emerging evidence.

The opposing letters reflect very different understandings of what the protection of public health requires.

Like their harm reduction counterparts, those committed to precaution have been open about the role that values play in their judgments. Stanton Glantz, one of the chief antagonists of e-cigarettes, brings to the debate an intimate, unparalleled knowledge of tobacco industry documents that leaves no room for illusions about deception on the part of the tobacco companies. For Glantz and other signatories of the second letter to Dr. Chan, the "fundamental and irreconcilable conflict of interest" between tobacco and public health demands an unyielding stance (*1*). What such a strong position does not acknowledge is that this perspective also entails a cost: It only recognizes the potential benefits of erecting barriers to e-cigarettes without considering the potential toll measured by lives lost to combustible products. In a world of multiple risks, argues Jonathan Weiner, "precaution against one risk may induce other countervailing risks" and their associated burdens (*5*). Dr. Kenneth Warner, a signatory of the harm-reduction letter, argued in the 1990s, that it is critical to "recognize that what divides us are those judgments about trade-offs" (*11*).

CONCLUSION. Harm-reduction advocates, although not surprised, were bitterly disappointed, after the October 2014 FCTC meeting in Moscow, when WHO called on nation states to adopt very restrictive precautionary measures, urging that countries consider prohibiting e-cigarettes and banning advertising (*12*). Very different has been the response to the first Cochrane Review on e-cigarettes published on 17 December 2014. Although noting the need for more research, the review concluded that current evidence underscores the potential of e-cigarettes as cessation tools. Said one of the study's authors, "None of the studies in this review found that smokers who used electronic cigarettes short-term (2 years or less) had an increased health risk compared to smokers who did not use electronic cigarettes. We did not find any evidence from observational studies that people who used electronic cigarettes at the same time as using regular cigarettes were less likely to quit smoking" (*13*).

The clash between harm reduction and precaution is not limited to e-cigarettes. The comment period on Swedish Match's application to the FDA for Snus to be regarded as a Modified-Risk Tobacco Product (MRTP) remains open until 23 February 2015 (*14*). The current battle lines for this smokeless tobacco product are very much marked by the two letters to Dr. Chan over the summer. Warning labels on Snus and other smokeless tobacco products currently hew to precautionary syntax, stressing, "This product is not a safe alternative to cigarettes." Variants warn of gum disease, tooth loss, and oral cancer. If this first MRTP application is successful, the language of harm reduction with its emphasis on making clear both risks and benefits would prevail, with revised labels reading "No tobacco product is safe, but this product presents substantially lower risks to health than cigarettes" (*15*).

Ultimately, decisions about how to proceed will be made in the face of evolving evidence and the undeniable burden imposed by tobacco cigarettes. Decision-making may draw on elements of both precautionary thinking and harm reduction, but weighing the risks and benefits is unavoidable. It is imperative to recognize that deep precaution precludes that possibility. It has served as a kind of trump argument, hostile to the notion of trade-offs, seeing in them perilous compromise. Such a posture does not serve either science or policy well.

REFERENCES AND NOTES

1. Letters to WHO on nicotine science and public policy; http://nicotinepolicy.net/n-s-p/2003-glantz-letter-to-who-the-importance-of-dispassionate-presentation-and-interpretation-of-evidence.
2. A. Fairchild, J. Colgrove, *Am. J. Public Health* **94**, 192 (2004).
3. S. Chapman, M. A. Wakefield, *Tobacco Control* **22** (suppl. 1), i33 (2013).
4. Science and Environmental Health Network, www.sehn.org/state.html.
5. J. B. Weiner, in *Human and Ecological Risk Assessment: Theory and Practice*, D. J. Paustenbach, Ed. (Wiley, New York, 2002), pp. 1509–1531.
6. N. A. Ashford, in *Protecting Public Health and the Environment: Implementing the Precautionary Principle*, C. Raffensperger and J. Tickner, Eds. (Island Press, Washington, DC, 1999), pp. 198–206.
7. A. Jordan, T. O'Riordan, in *Protecting Public Health and the Environment: Implementing the Precautionary Principle*, C. Raffensperger and J. Tickner, Eds. (Island Press, Washington, DC, 1999), pp. 20, 25.
8. C. R. Sunstein, *Laws of Fear: Beyond the Precautionary Principle* (Cambridge Univ. Press, New York, 1995).
9. Commission of European Communities, "Communication from the Commission on the Precautionary Principle" (CEC, Brussels, 2000).
10. A. Bhatnagar *et al.*, *Circulation* **2014**, 1 (2014).
11. Cited in M. Pertschunk, *Smoke in Their Eyes: Lessons in Movement Leadership from the Tobacco Wars* (Vanderbilt Univ. Press, Nashville, TN, 2001), p. 119.
12. WHO, FCTC/COP/6/A/R/2, Item 4.4.2 (WHO, Geneva, 2014), pp. 8–9.
13. New Cochrane evidence shows that electronic cigarettes facilitate smoking cessation, 17 December 2014, [press release]; https://www.cochrane.org/features/new-cochrane-evidence-shows-electronic-cigarettes-facilitate-smoking-cessation.
14. Modified Risk Tobacco Product Applications, www.regulations.gov/#!documentDetail;D=FDA-2014-N-1051-0001.
15. M. Richtel, D. Jolly, *New York Times*, 30 November 2014, p. BU1.

10.1126/science.1260761

Brazil's Soy Moratorium

Supply-chain governance is needed to avoid deforestation

By **H. K. Gibbs[1*], L. Rausch[1], J. Munger[1], I. Schelly[1], D. C. Morton[2], P. Noojipady[3,4], B. Soares-Filho[5], P. Barreto[6], L. Micol[7], N. F. Walker[4]**

POLICY

Brazil's Soy Moratorium (SoyM) was the first voluntary zero-deforestation agreement implemented in the tropics and set the stage for supply-chain governance of other commodities, such as beef and palm oil [supplementary material (SM)]. In response to pressure from retailers and nongovernmental organizations (NGOs), major soybean traders signed the SoyM, agreeing not to purchase soy grown on lands deforested after July 2006 in the Brazilian Amazon. The soy industry recently extended the SoyM to May 2016, by which time they assert that Brazil's environmental governance, such as the increased enforcement and national implementation of the Rural Environmental Registry of private properties (Portuguese acronym CAR) mandated by the Forest Code (FC) (*1*), will be robust enough to justify ending the agreement (*2*). We argue that a longer-term commitment is needed to help maintain deforestation-free soy supply chains, as full compliance and enforcement of these regulations is likely years away. Ending the SoyM prematurely would risk a return to deforestation for soy expansion at a time when companies are committing to zero-deforestation supply chains (*3*).

Between 2001 and 2006, soybean fields expanded by one million hectares (Mha) in the Amazon biome, and direct conversion of forests for soy production contributed to record deforestation rates (*4*–*6*). Farms violating the SoyM were identified using a satellite and airborne monitoring system—developed by industry, NGOs, and government partners—and were blocked from selling to SoyM signatories. Monitoring data confirm high compliance with the SoyM (*6*).

ESTIMATING IMPACTS. In the 2 years preceding the agreement, nearly 30% of soy expansion occurred through deforestation rather than by replacement of pasture or other previously cleared lands. After the SoyM, deforestation for soy dramatically decreased, falling to only ~1% of expansion in the Amazon biome by 2014 (see the chart) (SM, table S1) (*6*). Soy increased by 1.3 Mha in the Amazon biome during this period (*5*).

In the Cerrado biome, where the SoyM does not apply, the annual rate of soy expansion into native vegetation remained sizable, ranging from 11 to 23% during 2007–2013 (SM, table S2). In Brazil's newest agricultural hotspot—the eastern Cerrado region in the states of Maranhão, Piauí, Tocantins, and Bahia (Mapitoba)—nearly 40% of total soy expansion (2007–2013) occurred at the expense of native vegetation (table S3). About half of the Cerrado biome has been converted for agricultural production in recent decades, and these woodlands and savannas have less protection than Amazon forests under environmental laws (*7*). Further study is needed to assess potential leakage into the Cerrado and other countries and to quantify the avoided deforestation from the SoyM.

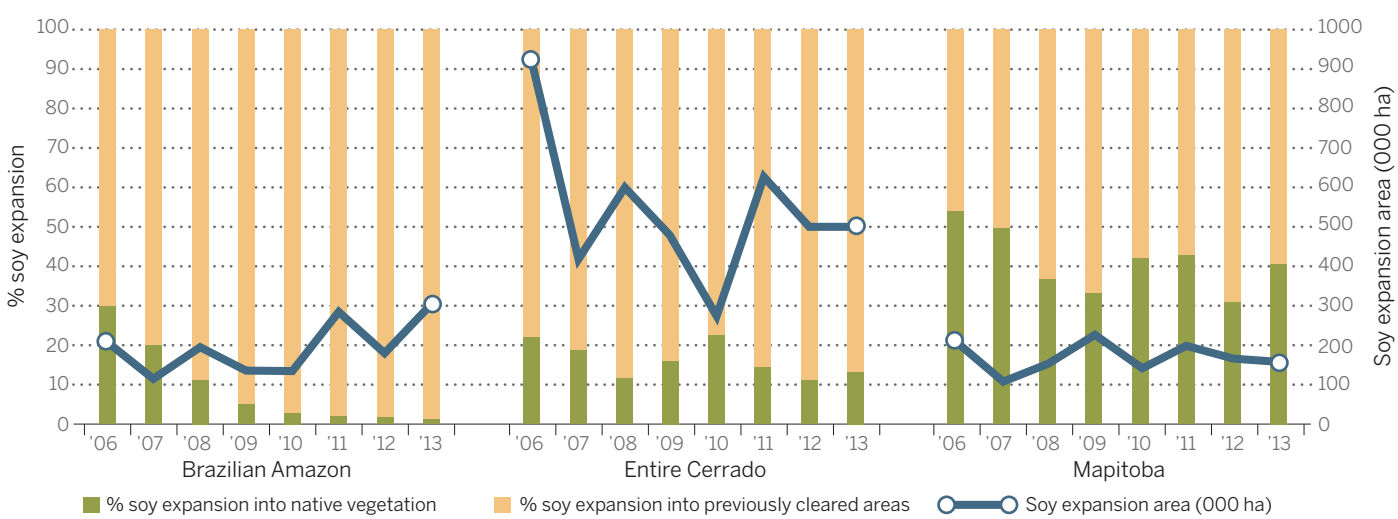
PROPERTY REGISTRATION. The CAR provides the first transparent mechanism to evaluate compliance with the FC and other regulations by linking a responsible landholder to land use on a particular property. All rural properties across Brazil are required to obtain the CAR by May 2016, although delays are expected, given the formidable task of demarcating more than 5 million properties. In Pará and Mato Grosso, the two states with the highest CAR participation, more than 65% and 48% of the agricultural land, respectively, is registered (SM).

Property registration alone, however, does not safeguard forests (*8*, *9*). In 2014, for example, nearly 25% of Amazon deforestation in Mato Grosso and 32% in Pará occurred within registered properties (*10*) (SM). In both states, nearly half of this clearing occurred in the Legal Reserve (LR) areas designated as set-asides required by the FC. Most of this clearing was illegal; few registered properties with deforestation in Mato Grosso (9%) or Pará (4%) had the ≥80% forest cover mandated by the FC (SM).

Comparing property-level compliance with the SoyM and the FC illustrates the relative response by soy farmers. In Mato Grosso, which accounts for 85% of the soy produced

Annual soy cultivation expansion by region

Annual soy expansion and land sources after the 2006 SoyM. Note that the Mapitoba area is included in the Cerrado (see SM).



in the Amazon biome, mapped farms with ≥25 ha of soy violated the FC, even while complying with the SoyM (table S4). Only 2% of mapped soy farms in Mato Grosso had sufficient LRs, making almost all deforestation illegal (table S5). At least 627 soy properties in Mato Grosso violated the FC and cleared forest illegally during the SoyM. Yet only 115 properties were excluded by soy traders for SoyM violations (*2*). This discrepancy can occur because the SoyM regulates only the portion of the property where soy is grown—not the entire property. The larger number of FC violations suggests that producers are more likely to comply with the SoyM.

LIMITED FEDERAL ENFORCEMENT. Without the SoyM, federal enforcement mechanisms would be the primary intervention against deforestation in the soy supply chain. Brazil's environmental protection agency, IBAMA, uses satellite data and field visits to issue fines and embargo economic activities on rural properties with illegal deforestation. The number of properties listed as embargoed more than tripled in the last 5 years (SM). However, thousands of deforestation events occur in the Brazilian Amazon each year across an area spanning 550 Mha, which makes it difficult to achieve enforcement (*11*). As of May 2014, roughly half of the registered properties with deforestation ≥25 ha, 2009–2013, were not embargoed (tables S6 and S7). Most of this deforestation was illegal. Government monitoring of embargoed properties is limited; in the embargoed list, for more than half of registered properties with embargoes, producer identification was inconsistent with the CAR system. Production could continue in embargoed areas and be transferred to another nonembargoed property or farmer for sale ("laundering"). Producer information was inconsistent between the embargoed list and the CAR system for more than half of the registered properties with embargoes. Soy traders and others use the CAR to check for embargoes; inconsistent information makes it difficult to avoid transactions with embargoed properties (SM).

Federal enforcement mechanisms are unlikely to be an effective substitute for the SoyM in the near term, because there is no simple way to identify properties that are in compliance with the FC. Recent changes to the FC have created the forest certificate (Portuguese acronym, CRA)–trading schemes, which allow landholders to purchase CRA from other properties and compensate for LR deficits accrued from illegal deforestation before 2008 (*1*). A system is not yet in place to monitor this off-property LR compensation. Enforcement is more straightforward under the SoyM, because all clearing for soy is prohibited. Of the existing policy and enforcement regimes, only the SoyM allows buyers to ensure deforestation-free supply chains over the next several years. Over the long term, elements of the SoyM and FC monitoring systems could be combined to satisfy market demands for information. However, even with eventual full compliance under the FC, legal deforestation could enter the soy supply chain without the SoyM (*1*).

> ***"Federal enforcement mechanisms are unlikely to be an effective substitute for the Soy Moratorium ..."***

VULNERABLE CERRADO. In the Amazon biome, there are an estimated 14.2 Mha of unprotected tropical forest considered suitable for soy production, and up to 2 Mha of this forest could be cleared legally under the FC (SM and fig S3). These forests would be vulnerable to soy expansion without the SoyM. However, the bank of eligible, previously cleared land suitable for soy production is more than six times the area planted in 2014 indicating the expansion is possible under the SoyM (table S8) (*12*).

More than 20 Mha of natural vegetation in the Cerrado are considered suitable for soy expansion, and up to 11 Mha of these lands could be legally converted under the FC. Large areas of cleared lands suitable for soy (42.5 Mha) also exist in the Cerrado, enough to triple current soy production, but these lands are not located in the regions with the most rapid recent expansion of soy into native vegetation. In the Mapitoba region, for example, there are fewer than 2 Mha of cleared lands considered suitable for soy production (fig. S3). If large-scale soy expansion continues in Mapitoba, remaining natural vegetation could be highly susceptible to soy conversion without additional safeguards. Expanding the SoyM could reduce the ongoing direct conversion of cerrado vegetation to soy.

By prohibiting new deforestation, the SoyM incentivizes soy expansion into already-cleared areas, which may displace pastures and could indirectly lead to more deforestation. Zero-deforestation agreements in the cattle sector, together with national and municipal policies, may partially mitigate the risk of this indirect deforestation (*11*). Ongoing efforts to increase production on existing pasturelands could free additional areas for production (*13*).

CONCLUSIONS. Since the SoyM's inception in 2006, only a small area of soy expansion in the Brazilian Amazon occurred in newly deforested areas. Soy farmers are about five times as likely to have violated the FC as the SoyM (115 versus 627 violations) (SM). The success of the SoyM is due to an array of factors, including (i) a limited number of soy buyers that exert considerable control over soy purchase and finance; (ii) simple requirements for compliance; (iii) streamlined and transparent monitoring and enforcement systems; (iv) simultaneous efforts by the Brazilian government to reduce deforestation; and (v) active participation by NGOs and government agencies (*14*). Monitoring and compliance mechanisms established by the SoyM offer a model for expanding supply-chain governance to other soy-producing regions and commodities.

We argue that the CAR and FC are not yet sufficient replacements and are unlikely to be fully implemented when the SoyM expires in 2016. Instead, the SoyM should be further extended and strengthened in the Amazon biome through expanded monitoring and exclusion of all deforestation on soy-producing properties, including small clearings and those located in indigenous lands and rural settlements, where soy production is expanding (SM). The SoyM should also be expanded to include the Cerrado biome to reduce conversion of remaining native vegetation.

[1]University of Wisconsin, Madison, WI 53726, USA. [2]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. [3]University of Maryland, College Park, MD 20742, USA. [4]National Wildlife Federation, Washington, DC 20006, USA. [5]Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil. [6]IMAZON Amazon Institute of People and the Environment, 66.060-162 Belém, Pará, Brazil. [7]Instituto Centro de Vida, 78045-055 Cuiabá, Mato Grosso, Brazil. *E-mail: hkgibbs@wisc.edu

REFERENCES AND NOTES

1. B. Soares-Filho *et al.*, *Science* **344**, 363 (2014).
2. Brazilian Vegetable Oil Industries Association, "New agenda for soybeans in the Amazon biome" (ABIOVE, São Paulo, 2014); www.abiove.com.br.
3. United Nations, New York Declaration on Forests (UN, New York, 2014); http://bit.ly/1KmeuRW.
4. D. C. Morton *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14637 (2006).
5. M. N. Macedo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1341 (2012).
6. B. F. T. Rudorff *et al.*, *Remote Sens.* **3**, 185 (2011).
7. G. F. Rocha *et al.*, *Rev. Bras. Cartogr.* **63**, 341 (2011).
8. A. A. Azevedo *et al.*, *Boletim Amazônia em Pauta* **3**, 1 (IPAM, Brasília, 2014); http://bit.ly/17j4REk.
9. R. Rajão *et al.*, *Public Adm. Dev.* **32**, 229 (2012).
10. In Pará, one-third of this was in INCRA settlements.
11. J. Börner *et al.*, *Glob. Environ. Change* **29**, 294 (2014).
12. E. F. Lambin *et al.*, *Glob. Environ. Change* **23**, 892 (2013).
13. B. B. Strassburg *et al.*, *Glob. Environ. Change* **28**, 84 (2014).
14. D. Nepstad *et al.*, *Science* **344**, 1118 (2014).

ACKNOWLEDGMENTS

See SM for thanks and funding.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/377/suppl/DC1

10.1126/science.aaa0181

PHOTO: THINKSTOCK

NEUROSCIENCE

Hard-wired for good?

Prosocial behavior has a long history in the brain

By **Frans B. M. de Waal**

Although the Bible depicts us as original sinners, and biology has for the longest time maintained that genuine altruism is an illusion, we are now seeing a sea change in ideas about what is called human "goodness." Science is increasingly looking at the prosocial side of human behavior, realizing that it must be more than a product of education, culture, and religion. Instead of a thin veneer that we risk losing at any moment, it seems that the seeds of morality may be contained in human nature.

As moral tendencies have been documented in human babies as well as our primate relatives, they don't seem to require language or religious dogma. Now, Donald W. Pfaff, an eminent American neuroscientist, has come along to place them squarely inside of our brains. Written in a highly readable style, *The Altruistic Brain* summarizes the current state of knowledge in the field and offers an excellent starting point for readers who are interested in learning more about the biology and neuroscience of human prosociality.

Pfaff reviews the dramatic increase in knowledge in this area over the past few decades, citing a range of biological and psychological studies on topics ranging from mirror neurons and perception-action mechanisms to oxytocin, empathy, and the internal rewards associated with doing good. Collectively, these data suggest that there are biological mechanisms that draw us into the sphere of others (i.e., promote empathy) by blurring the lines of identity. Pfaff argues that our ability to see others as ourselves explains altruistic behavior from maternal care to the heroism of firefighters during 9/11.

It was not so long ago that prominent neuroscientists, such as John Eccles, felt the need to invoke God to explain how our brains can produce altruistic tendencies (*1*). Others, including V. S. Ramachandran, have used the existence of human altruism as evidence that we are somehow special, set apart from other creatures and not "mere ape[s]" (*2*). Pfaff rejects these attempts to disconnect the human brain from the long evolution that preceded it. Instead, he recognizes that our brains are not that different from the brains of other mammals and that social attachment—which is at the root of empathy and prosociality—functions essentially the same in species from rodents to the higher primates.

The book also rightly emphasizes maternal care and sexual behavior as the first forms of social connection. These, too, we share with many other species. But Pfaff does not address how attachment and emotional processes—which do not directly translate into judgments of right and wrong—relate to human morality. He operates mostly at this emotional level; however, cognitive and societal processes are also essential factors in human morality. Biologically oriented philosophers, such as Patricia Churchland and Philip Kitcher, have offered more insightful accounts of how our species may have moved from prosocial tendencies to a system of prescribed standards that address the issue of right and wrong (*3*, *4*). Nevertheless, Pfaff's aversion to separating morality from emotion is refreshing and is consistent with the naturalistic approach to morality advanced by philosopher David Hume (*5*).

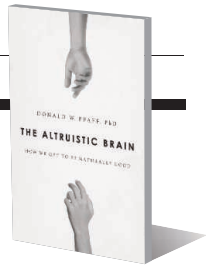
Pfaff does oversell his own theory, especially in the beginning of his book. He gives it an abbreviation: ABT, for the "Altruistic Brain Theory," while calling it a "perfect storm of audacity, awareness, and insight." He believes that ABT will change the world. But although the theory is compelling, there are earlier books and articles about how morality comes naturally to our species (*6*) and about the role of the brain in empathy and altruism (*7*).

We descend from a long line of primates that emphasize social cohesiveness, at least to an equal degree as selfish advantage. Group life requires compromise and a merging of interests, and the human brain reflects this long evolution.

The last part of the book offers wide-ranging reflections on the value of prosocial tendencies in daily life. Pfaff speaks of the role of these behaviors in building trust, their place in business, how empathy can be trained, and even how society can deal with bad behavior. It is obviously a rich topic and one that has been neglected for too long.

But in all of this, we should not forget that even if the brain is made for goodness, as the author argues, it is equally involved in our less-noble tendencies. It is the interaction and balance between these two that remains the greatest puzzle.

The reviewer is at the Department of Psychology, Emory University, Atlanta, GA 30322, USA. E-mail: dewaal@emory.edu



The Altruistic Brain
How We Are Naturally Good
Donald W. Pfaff
Oxford University Press, 2014. 305 pp.

REFERENCES

1. J. Eccles, *The Neurophysiological Basis of Mind* (Oxford Univ. Press, Oxford, 1953).
2. V. Ramachandran, *The Tell-Tale Brain: A Neuroscientist's Quest for What Makes Us Human* (Norton, New York, 2012).
3. P. Churchland, *Braintrust: What Neuroscience Tells Us about Morality* (Princeton Univ. Press, Princeton, NJ, 2011).
4. P. Kitcher. *Vaulting Ambition: Sociobiology and the Quest for Human Nature* (MIT Press, Cambridge, MA, 1985).
5. D. Hume. *A Treatise of Human Nature* (Oxford Univ. Press, Oxford, 1738).
6. F. B. M. de Waal, *Good Natured* (Harvard Univ. Press, Cambridge, MA, 1996).
7. S. D. Preston, F. B. M. de Waal, *Behav. Brain Sci.* **25**, 1 (2002).

10.1126/science.1261392

EDUCATION

Curriculum and the Cold War

A closer look at the "new math" movement of the 1960s

By **Jeremy Kilpatrick**

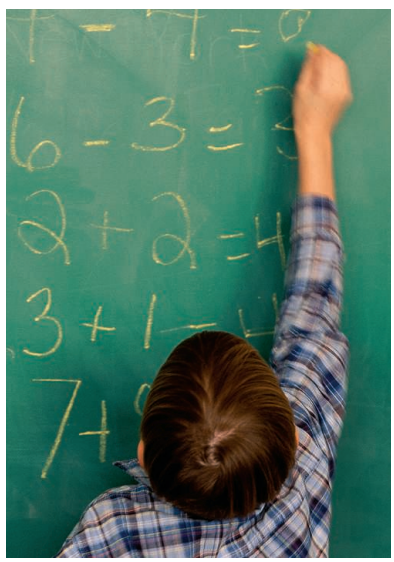
Spurred by concerns that the American education system was failing to keep pace with the Soviets after the "Sputnik crisis" of the late 1950s, a dramatic change occurred in the mathematics that was taught in American schools in the 1960s. This "new math," as it was known, was controversial from the start and ultimately short-lived. Now, more than a half century later, Christopher Phillips has looked back at that contentious era in *The New Math: A Political History*. The book sheds light on a time when changing political commitments were affecting what it meant to be prepared—mathematically—for citizenship in a modern, technology-laden society.

Phillips, a historian of science, views the new math phenomenon as indicative of the American belief that "curriculum could order and shape the mind, the family, the society, and the state." The book's subtitle, which establishes a political framework for the discussion, could have as easily referred to our social, cultural, or intellectual history, because he sees the curriculum debate as a broader argument about the "proper relationship between the content and purpose of education."

He focuses most of his attention on the largest and arguably most influential mathematics curriculum project, the School Mathematics Study Group (SMSG), which began at Yale in 1958 and then moved to Stanford in 1961 under the leadership of Edward G. Begle. The SMSG sought to improve the mathematics curriculum by producing new textbooks that emphasized abstract reasoning.

Phillips's analysis of SMSG's efforts represents a departure from the 1960s writer and cultural critic Benjamin DeMott. Writing from the battlefield of what he christened "the math wars," DeMott considered the textbooks produced by SMSG, although demanding, to "invite the student to understand mathematics as at once an area and a mode of thought—not as a clutch of senseless rules and problems that either do or don't 'come out'" (*1*). He also applauded the SMSG reformers as "among the exemplary intellectuals of this age" for being willing to invest so much time and effort in the writing of school textbooks. Although Phillips cites DeMott's commendation approvingly, he is less taken than DeMott with either the textbooks or the reformers' sacrifices. The new math was, in his view, an effort to instill mental discipline during "the Cold War of the classrooms."

Phillips goes on to analyze the origins of the new math, the textbooks the SMSG produced, how the new math was sold to educators and the public, and how the views of its detractors finally triumphed. He ends by bringing the story up to date, arguing that the mathematics classroom "will remain a political venue as long as learning math counts as learning to think."



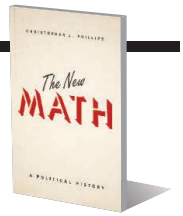
The book is based on extensive research and is incredibly well documented. Phillips has made considerable use of various archives, particularly the SMSG Records and the Moore Collection in the Archives of American Mathematics at the University of Texas.

By focusing on the SMSG project, Phillips has simplified his task of describing and analyzing the new math reforms, allowing him to go into greater detail about the social and political context in which they took place. Unfortunately, this narrow focus gives short shrift to the varied activities and contributions of other projects. For example, he mentions, but goes into no detail regarding, the University of Illinois Committee on School Mathematics, which emphasized precision of language and a form of discovery learning, and the Madison Project, which worked on developing the teacher's professional competence rather than creating textbook material. The Secondary School Mathematics Curriculum Improvement Study, which attempted to unify the secondary curriculum by doing away with separate courses in arithmetic, algebra, and geometry, is left out of the discussion altogether. He also fails to discuss the second round of SMSG, begun in 1966, which set out to produce instructional units that would demonstrate the relevance of mathematics to everyday life by having students in grades 7 to 12 construct mathematical models of realistic situations (*2*). In doing so, Phillips underplays the naiveté of the SMSG reformers' goal of dramatically changing school mathematics simply by producing new textbooks.

Phillips sees the SMSG textbook authors as attempting to promote mathematics as instilling mental discipline and mental habits, despite the fact that these terms do not appear in their statement of "philosophies and procedures" (*3*). It is fairer to say that those authors were principally concerned with attracting students to the study of mathematics by getting rid of jargon and making school mathematics more like the subject as then taught in the university. They believed that "the language of sets, relations, and functions would provide not only a more coherent discourse in the mathematics classroom but also a more meaningful structure for learning" (*4*).

Despite its limited scope and presumptions about the intentions of the individuals behind the new math movement, this book is likely to stand for a long time as the most thorough, authoritative account of this mid–20th-century phenomenon.

The New Math
A Political History
Christopher J. Phillips
University of Chicago Press, 2014. 241 pp.



The reviewer is in the Mathematics and Science Education Department, College of Education, 105 Aderhold Hall, University of Georgia, Athens, GA 30602-7124, USA. E-mail: jkilpatrick@uga.edu

REFERENCES

1. B. DeMott, *American Scholar*, **31** (2), 296 (1962).
2. School Mathematics Study Group, *Report of a Conference on Secondary School Mathematics, New Orleans, March 14–18, 1966* (SMSG Working Paper, SMSG, Stanford, CA, 1966).
3. School Mathematics Study Group, *Philosophies and Procedures of SMSG Writing Teams* (SMSG, Stanford, CA, 1965).
4. J. Kilpatrick, *NY State Math. Teachers' J.* **58**, 87 (2009).

10.1126/science.aaa1471

PHOTO: THINKSTOCK

LETTERS

Edited by **Jennifer Sills**

Carnivore coexistence: Value the wilderness

IN THEIR REPORT "Recovery of large carnivores in Europe's modern human-dominated landscapes" (19 December 2014, p. 1517), G. Chapron *et al.* recount encouraging news that brown bear, wolf, European lynx, and wolverine populations are stable or increasing in Europe. Still more encouraging, these carnivores now persist in human-dominated landscapes, a change they attribute to "land-sharing" practices.

Europe is undoubtedly becoming more hospitable for large carnivores, owing to effective law enforcement, shifting social values, and favorable habitat change (*1*). However, it is unclear whether land-sharing will routinely support self-sustaining large carnivore populations in isolation. Chapron *et al.* show that carnivores are permanently present in areas of high human density, but this does not imply intrinsic population growth. Rather, their presence might hinge on immigration from remaining wilderness areas.

The benefits of land-sharing can be contingent on connectivity between landscapes shared with humans and intact wilderness (*2*), particularly for long-lived, slow-reproducing, and strongly dispersive species like large carnivores (*3*), whose populations can be interconnected over huge distances (*4*). The recovery of the brown bear in Norway, for example, depends on females traveling from wilderness areas in Sweden, where human densities are extremely low (*5*). Similarly, the European lynx remains critically dependent on protected areas in Central Europe (*6*). Without detailed analyses linking land use and socioeconomic change to carnivore demography, it is premature to conclude that land-sharing offers an effective general model for carnivore conservation.

Protected areas are under increasing pressure worldwide (*7*); it is therefore essential that policy-makers continue to recognize the value of wilderness for biodiversity. Successful land-sharing should be applauded, but it does not represent a green light to roll back protected-area designation. Without the ongoing sparing of European wilderness zones, it is unlikely that favorable trends will continue for large carnivores.

James J. Gilroy,[1] Andrés Ordiz,[2] Richard Bischof[3]

[1]School of Environmental Science, University of East Anglia, NR4 7TJ, UK. [2]Grimsö Wildlife Research Station, Swedish University of Agricultural Sciences, SE–730 91, Sweden. [3]Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, NO-1432, Norway.

*Corresponding author. E-mail: james.gilroy1@gmail.com

REFERENCES

1. S. Deinet *et al.*, "Wildlife comeback in Europe: The recovery of selected mammal and bird species: Final report to rewilding Europe" (Zoological Society of London, London, 2013).
2. J. J. Gilroy *et al.*, *J. Appl. Ecol.* **51**, 1337 (2014).
3. H. S. Robinson, R. B Wielgus, H. S. Cooley, S. W. Cooley, *Ecol. Appl.* **18**, 1028 (2008).
4. M. K. Schwartz *et al.*, *Nature* **415**, 520 (2002).
5. R. Bischof, J. E. Swenson, *Ecol. Appl.* **22**, 361 (2012).
6. J. Müller *et al.*, *Biol. Conserv.* **177**, 210 (2014).
7. J. E. Watson, N. Dudley, D. B. Segan, M. Hockings, *Nature* **515**, 67 (2014).

The stability of brown bear populations in populated areas may depend on the preservation of nearby wilderness.

PHOTO: ATHANASIOS GIOUMPASIS/GETTY IMAGES

Carnivore coexistence: America's recovery

IN THEIR REPORT "Recovery of large carnivores in Europe's modern human-dominated landscapes" (19 December 2014, p. 1517), G. Chapron *et al.* attribute the return of predators in Europe to a "coexistence" model for conservation, in which carnivores and humans inhabit shared landscapes. They contrast this model with the "separation" model, in which carnivores and humans do not intermingle. Chapron *et al.* suggest that the separation model originated in North America and serves as the basis for management approaches in Asia, Africa, and Neotropical countries. We consider this dichotomization artificial and posit that there is no such North American model.

In North America, black and grizzly bear, cougar, and gray wolf populations are all expanding. Large protected areas in North America have long supported carnivores living separately from people (*1*), but many of the recent recoveries have occurred in landscapes shared with people. For example, New Jersey, the most densely populated state in the United States, supports densities of 12.4 black bears/100 km^2 (*2*). The broader North American recovery has benefited from cultural acceptance and favorable legislation (*3*), and thus is similar to coexistence-based recoveries reported for Europe.

Chapron *et al.* argue that coexistence is more effective than separation, but true separation of carnivores and people, such as fencing to isolate African lions from humans, has been shown to be a successful conservation model in some regions of the globe (*4*), particularly where large carnivores regularly kill people. For example, in Tanzania alone, lions attacked more than 1000 people between 1990 and 2007 (*5*).

The approaches that Chapron *et al.* identified to distinguish coexistence from separation do not derive from North American management policies. Rather, recovery of large carnivores in North America reflects the legacy of large protected areas, low risk of carnivores killing humans, and an improvement of the public's views toward carnivores.

Matthew E. Gompper,[1]* Jerrold L. Belant,[2] Roland Kays[3]

[1]Department of Fisheries and Wildlife Sciences, University of Missouri, Columbia, MO 65211, USA. [2]Carnivore Ecology Laboratory, Mississippi State University, Mississippi State, MS 39762, USA. [3]North Carolina State University and North Carolina Museum of Natural Sciences, Raleigh, NC 27601, USA.

*Corresponding author. E-mail: gompperm@missouri.edu

REFERENCES

1. W. D. Newmark, *Conserv. Biol.* **9**, 512 (1995).
2. New Jersey Division of Fish and Wildlife, Black bear management and the comprehensive black bear management policy FAQs (www.njfishandwildlife.com/bearpolicy10_faq.htm).
3. J. D. C. Linnell, J. E. Swenson, R. Anderson, *Anim. Conserv.* **4**, 345 (2001).
4. C. Packer *et al.*, *Ecol. Lett.* **16**, 635 (2013).
5. H. Kushnir, H. Leitner, D. Ikanda, C. Packer, *Human Dim. Wildl.* **15**, 315 (2010).

Carnivore coexistence: Trophic cascades

G. CHAPRON *ET AL.* present a compelling case that large carnivores and people can successfully share the same landscape ("Recovery of large carnivores in Europe's modern human-dominated landscapes," Report, 19 December 2014, p. 1517). However, the conclusions of Chapron *et al.* are a beginning, not an end, to an inquiry into the possibilities and implications of coexistence. In remote areas with low human densities, large carnivores play vital roles through trophic cascades in regulating prey, smaller carnivores, and ecosystems (*1*). A key question in light of the findings of Chapron *et al.* is whether large carnivores in human-dominated landscapes perform a similar ecological role. The research to date suggests that they may not, or that common models of coexistence inhibit large carnivores fulfilling their roles as apex predators.

Humans can interfere with wolves' predator-prey relationships.

PHOTO: MATT CARDY/GETTY IMAGES

For example, where high human densities are accompanied by relaxed environmental policies, large carnivores frequently supplement their diets with garbage, crops, livestock, and carcasses managed by humans (*2*). This can influence their abundance, life histories, and space use, in turn altering how they interact with co-occurring predators and prey (*2*). Human presence and human infrastructure can also alter predator-prey relationships by providing an opportune "shield" for one species from the other (*3*). This has given rise to a three-way interaction involving people, predators, and prey. Complicating things further is that humans often facilitate hybridization between large carnivores (especially canids) and nonindigenous or domesticated species, either through purposeful introductions or habitat modifications (*4*).

If we seek transformative change about the way in which large carnivores are managed, we need to better understand how to prevent ecological outcomes that undermine restoration goals. A successful model of coexistence will need to achieve not merely the fact of coexistence, but one that preserves to the greatest extent possible the critical role played by large carnivores on ecosystem processes.

T. M. Newsome[1,2]* and W. J. Ripple[1]

[1]Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA. [2]Desert Ecology Research Group, School of Biological Sciences, The University of Sydney, NSW 2006, Australia.

*Corresponding author. E-mail: tnew5216@uni.sydney.edu.au

REFERENCES

1. W. J. Ripple *et al.*, *Science* **343**, 151 (2014).
2. T. M. Newsome *et al.*, *Glob. Ecol. Biogeogr.* **24**, 1 (2014).
3. J. Berger, *Biol. Lett.* **3**, 620 (2007).
4. J. M. Rhymer, D. Simberloff, *Annu. Rev. Ecol. Syst.* **27**, 83 (1996).

Ferns to fulfillment

AT FIRST, I was saddened when I read the Working Life article "For the love of ferns" (19 December 2014, p. 1586), in which K. Perkins describes getting into science by accident, pursuing her graduate degree in a top lab at Albert Einstein College of Medicine, accepting a tenure-track position at Johns Hopkins, and subsequently withdrawing from biomedical sciences. The pressures of funding and establishing one's own lab, along with life circumstances including an unexpected divorce, ended what by all appearances was a promising career.

However, after a little reflection, I saw the story through a different lens. Perkins' relatively convoluted path led to a career as a high school physiology teacher, which she describes as "a life of work but also a life of people and play." It seems she ultimately found the elusive work-life balance, and that is indeed something to celebrate.

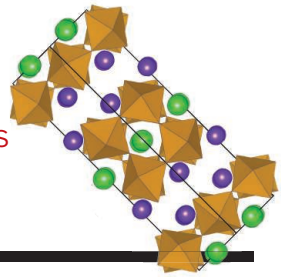
Lawrence Reynolds

Department of Animal Sciences, North Dakota State University, Fargo, ND 58108, USA. E-mail: larry.reynolds@ndsu.edu

RESEARCH

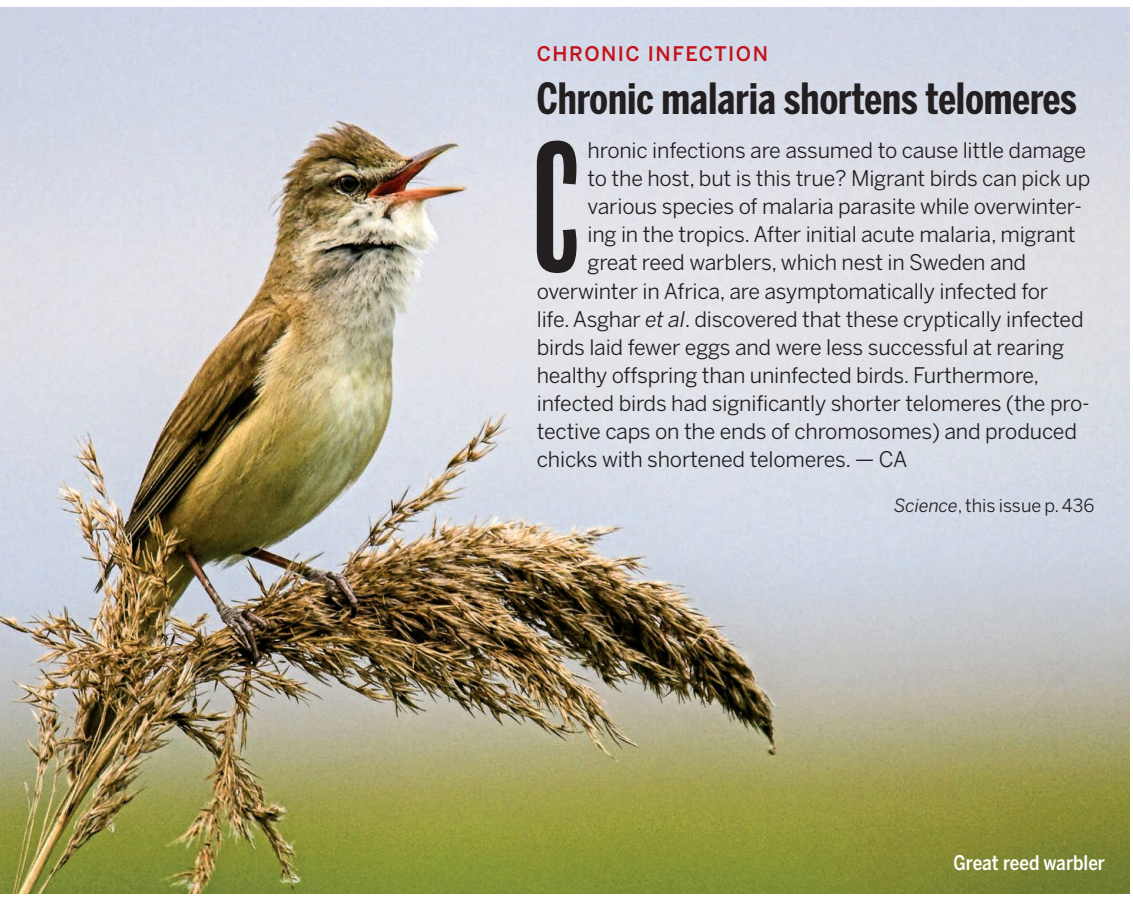
Chemically engineered layered perovskite crystals

Pitcher et al., *p. 420*



IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**



Great reed warbler

CHRONIC INFECTION

Chronic malaria shortens telomeres

Chronic infections are assumed to cause little damage to the host, but is this true? Migrant birds can pick up various species of malaria parasite while overwintering in the tropics. After initial acute malaria, migrant great reed warblers, which nest in Sweden and overwinter in Africa, are asymptomatically infected for life. Asghar *et al*. discovered that these cryptically infected birds laid fewer eggs and were less successful at rearing healthy offspring than uninfected birds. Furthermore, infected birds had significantly shorter telomeres (the protective caps on the ends of chromosomes) and produced chicks with shortened telomeres. — CA

Science, this issue p. 436

HUMAN EVOLUTION

Getting a grip

The evolution of the hand—particularly the opposable thumb—was key to the success of early humans. Without a precise grip, involving forceful opposition of thumb with fingers, tool technology could not have emerged. Skinner *et al*. analyzed the internal bone structure of Pliocene *Australopithecus* hands, dated at 3.2 million years old. Internal bone structure reveals the patterns and directions of forces operating on the hand, providing clues to the kinds of activities performed. Modern human-like hand postures consistent with the habitual use of tools appeared about half a million years earlier than the first archaeological evidence of stone tools. — AMS

Science, this issue p. 395

T CELL IMMUNITY

For T cells, variety is the spice of life

$CD4^+$ helper T cells come in a variety of flavors. This allows them to respond in a manner that is tailored to the pathogen they encounter. Becattini *et al*. wondered whether multiple "flavors" of human $CD4^+$ T cells respond to specific stimuli or if just one flavor dominates. To find out, they stimulated human memory $CD4^+$ T cells with a fungus, a bacteria, or a vaccine antigen. Multiple helper cell subsets participated in each response. T cell receptor sequencing revealed that in some cases, T cells with the same specificity acquired different helper cell fates. Thus, there is more heterogeneity in human T cell responses than previously appreciated. — KLM

Science, this issue p. 400

ASTROPHYSICS

A light on the origin of cosmic rays

There's a new lab for studying the origins of cosmic rays: our neighbor galaxy, the Large Magellanic Cloud. Astronomers are now making progress on this topic by examining the gamma rays that are produced when cosmic rays interact with gas or lower-energy photons. The H.E.S.S. Collaboration has detected three sources of gamma rays in a variety of forms in the galactic satellite to the Milky Way. The sources include the pulsar wind nebula of N 157B, the supernova remnant N 132D, and the superbubble 30 Dor C. Oddly, supernova 1987A was not detected. — MMM

Science, this issue p. 406

PLANETARY SCIENCE

Melting silica in massive planets

To simulate the extreme conditions inside large planets requires extreme experiments. Millot *et al*. used high-pressure shock waves almost twice that of the center of Earth to melt silica, one of the primary components of planetary interiors. This was possible only by shocking a very dense form of silica called stishovite. Determining the melting point of silica is vital for developing better computational models of the interior of planets several times the mass of Earth. The high-pressure liquid was electrically conductive, a property that may contribute to magnetic dynamos in very large terrestrial exoplanets. — BG

Science, this issue p. 418

Published by AAAS

MATERIALS CHEMISTRY

Soldering semiconductor nanoparticles

The optical and electronic properties of semiconductor nanoparticles can be tuned through changes in their size and composition. However, poor contact between interfaces can degrade nanoparticle performance in devices. Dolzhnikov *et al*. report the synthesis of a gel-like "solder" for metal chalcogenide nanoparticles, such as cadmium selenide and lead telluride, by cross-linking molecular wires of these materials. — PDS

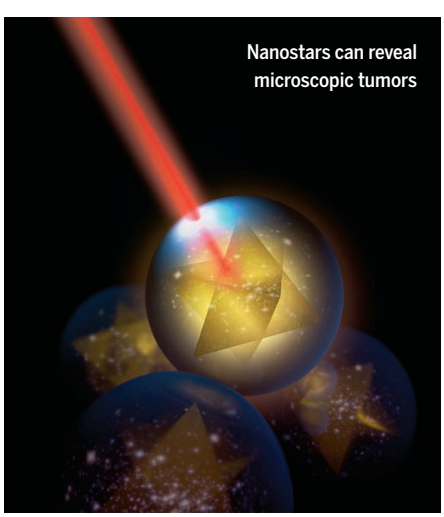
Science, this issue p. 425

CANCER IMAGING

Seeing nanostars

Microscopic tumors may be invisible to the naked eye, but they are no match for nanosized imaging agents that penetrate cancer tissue and signal the presence of disease. Harmsen *et al.* created a new generation of cancer-imaging agents called "nanostars"—star-shaped gold cores and Raman reporter molecules wrapped in silica—that can be visualized by Raman scattering. A new feature that puts the nanostars in resonance with the near-infrared window lets them outshine any previous nanoparticles. With these surface-enhanced resonance Raman scattering nanostars, the authors visualized microscopic lesions in animal models of pancreatic, breast, and prostate cancer, as well as sarcomas. — MLF

Sci. Transl. Med. **7**, 271ra7 (2015).



Nanostars can reveal microscopic tumors

CREDITS: (LEFT TO RIGHT) V. ALTOUNIAN/*SCIENCE*; © DIMITAR GORGEV/ALAMY

PROTEASOMES

A detailed look at proteasomes in situ

The 26*S* proteasome is a protein machine that degrades intracellular proteins in the cytosol. The proteasome is critical for protein quality control and for the regulation of numerous cellular processes in eukaryotic cells. The structure of isolated proteasomes is well established, but how intact proteasomes look within the cell is less clear. Asano *et al*. used an improved approach to electron cryotomography to look at proteasomes in intact hippocampal neurons. Their analysis suggests that these cells only use about 20% of their proteasomes in an unstressed state, which leaves significant spare capacity to deal with proteotoxic stress. — SMH

Science, this issue p. 439

CELL BIOLOGY

A binding partner for an orphan receptor

Anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase that is important during the development of the nervous system. ALK can also be aberrantly activated in certain types of cancers, including neuroblastoma and lung adenocarcinoma. Until now, ALK was an "orphan" receptor because its ligand was not known. Now, Murray *et al.* show that heparin binds directly to ALK and stimulates its activity (see the Focus by Lemke). An antibody that blocked the binding prevented heparin from activating this receptor tyrosine kinase, providing a potential avenue for therapeutic intervention. — JDB

Sci. Signal. **8**, ra6; see also fs2 (2015).

SIGNAL TRANSDUCTION

How a mouse's nose feels the cold

Guanylyl cyclases (GC) are a family of transmembrane proteins that regulate physiological processes as diverse as blood pressure and our sense of smell. Neurons in the Grueneberg ganglion, a sensory organ in the mouse nose involved in detecting cold temperatures, express a subtype of GCs called GC-Gs, but scientists do not fully understand what activates these proteins. Chao *et al*. now show that cool temperatures can directly activate GC-Gs, probably by driving GC-G proteins to oligomerize (group together). Mouse pups exposed to cool temperatures normally make ultrasonic cries that encourage maternal care, but pups engineered to lack GC-G proteins could not. — VV

EMBO J. 10.15252/embj.201489652 (2014).

Guanylyl cyclase proteins in the mouse nose sense cold

OPTICS

A one-two approach to air-lasing

Standoff spectroscopy involves the excitation and detection of chemical species at a distance. From security screening to environmental monitoring, there is a need to obtain accurate readings, which usually entails getting as much of a signal back as possible. Powerful laser pulses can excite pockets of air at great distances, allowing the remote sensing of the surrounding atmosphere, but their utility is limited by the small amount of light that gets back to the detector. Laurain *et al.* provide a solution to this problem, in the form of a two-color sequence of intense laser pulses. The first pulse, at infrared wavelengths, efficiently dissociates the air molecules; it is followed by a second ultraviolet pulse that excites the dissociated molecules and induces lasing. The induced laser light is emitted directionally, providing the possibility of greatly enhanced detection efficiency. — ISO

Phys. Rev. Lett. **113**, 253901 (2014).

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

PROTEOMICS

Protein expression across human tissues

Sequencing the human genome gave new insights into human biology and disease. However, the ultimate goal is to understand the dynamic expression of each of the approximately 2000 protein-coding genes and the function of each protein. Uhlén *et al.* now present a map of protein expression across 32 human tissues. They not only measured expression at an RNA level, but also used antibody profiling to precisely localize the corresponding proteins. An interactive website allows exploration of expression patterns across the human body. — VV

Science, this issue p. 394

SEA LEVEL CHANGE

Volume and shape combine to find a level

How can we understand the geological record of sea level change? Sea level varies on time scales from decades to millions of years. These changes have local, regional, and global components and are caused by a wide variety of earth processes. Cloetingh and Haq review how the views of stratigraphers (who interpret the record of marine sediments) and geodynamicists (who consider changes in the shape of Earth caused by lithospheric and mantle processes) have begun to complement each other and are moving toward a more coherent interpretation of the history of sea level. They focus on cyclic sea-level changes 0.5 to 3.0 million years in duration that occurred in the Cretaceous period, approximately 145 to 65 million years ago. — HJS

Science, this issue p. 393

DRUG RESISTANCE

Mechanisms propelling drug resistance

If it were to spread, resistance to the drug artemisinin would seriously derail the recent gains of global malaria control programs (see the Perspective by Sibley). Mutations in a region called the K13-propeller are predictive for artemisinin resistance in Southeast Asia. Mok *et al*. looked at the patterns of gene expression in parasites isolated from more than 1000 patients sampled in Africa, Bangladesh, and the Mekong region. A range of mutations that alter protein repair pathways and the timing of the parasite's developmental cycle were only found in parasites from the Mekong region. Straimer *et al*. genetically engineered the K13 region of parasites obtained from recent clinical isolates. Mutations in this region were indeed responsible for the resistance phenotypes. — CA

Science, this issue p. 431, p. 428; see also p. 373

MARS ATMOSPHERE

Of water and methane on Mars

The Curiosity rover has been collecting data for the past 2 years, since its delivery to Mars (see the Perspective by Zahnle). Many studies now suggest that many millions of years ago, Mars was warmer and wetter than it is today. But those conditions required an atmosphere that seems to have vanished. Using the Curiosity rover, Mahaffy *et al*. measured the ratio of deuterium to hydrogen in clays that were formed 3.0 to 3.7 billion years ago. Hydrogen escapes more readily than deuterium, so this ratio offers a snapshot measure of the ancient atmosphere that can help constrain when and how it disappeared. Most methane on Earth has a biological origin, so planetary scientists have keenly pursued its detection in the martian atmosphere as well. Now, Webster *et al*. have precisely confirmed the presence of methane in the martian atmosphere with the instruments aboard the Curiosity rover at Gale crater. — MMM

Science, this issue p. 412, p. 415; see also p. 370

MAGNETIC MATERIALS

Tilting toward two properties

Opposing electronic and symmetry constraints can make it difficult to combine some pairs of material properties in a single crystalline material. Magnetization and electrical polarization are such a pair, but their combination could be useful for applications such as magnetoelectric information storage. Pitcher *et al*. now show that careful design of chemical substitutions in a layered perovskite are both electrically polar and weakly ferromagnetic at temperatures up to 330 K. — PDS

Science, this issue p. 420

Livestock, such as cattle, cause strife among the Nyangatom people of East Africa

ANTHROPOLOGY

The evolutionary benefits of warfare

Many human societies engage in warfare, but given the mortal risks involved, many evolutionary anthropologists have wondered why. Is there an evolutionary benefit to warfare? Glowacki and Wrangham tackled this question by studying the Nyangatom, a nomadic society in East Africa. Nyangatom men carry out livestock raids to pay for the right to marry. Men who were active cattle raiders had more wives and children than men who were not. But they had to wait for this benefit. Young raiders give stolen livestock as gifts to paternal relatives. They only benefitted later in life by inheriting the larger herds they helped to build. — GR

Proc. Natl. Acad. Sci. U.S.A. 10.1073/pnas.1412287112 (2014).

MINERAL PHYSICS

A lower-mantle water cycle component

Water has a dramatic impact on mantle dynamics and chemistry. Despite speculation about water in Earth's lower mantle, a mechanism for supplying water to this remote region has been elusive. Pamato *et al.* suggest that a hydrous mineral called "phase D" may fill that role. Aluminum-rich hydrous phase D was found to be stable at the right temperature-pressure conditions and should be present in rocks with compositions that make up subducted ocean crust. Subducted slabs could therefore introduce water-rich patches into an otherwise dry lower mantle, providing an alternative explanation for some seismic and geochemical anomalies. — BG

Nat. Geosci. 10.1038/ngeo2306 (2015).

ENVIRONMENTAL SCIENCE

Droughts and dead zones on the rise

Nutrient-rich agricultural runoff or wastewater discharge can lead to the formation of harmful algal blooms or oxygen-depleted (hypoxic) "dead zones" in surface waters. But now, Zhou *et al.* show that low tributary river discharge was the largest contributor to a record-breaking hypoxic event in Lake Erie in 2012. Drought conditions across much of North America that year decreased water flow into the lake, exacerbating the effects of persistent nutrient runoff. Because drought and other extreme climate events are expected to increase with climate change, management strategies need to consider all factors that may degrade future water quality. — NW

Environ. Sci. Technol. 10.1021/es503981n (2014).

PLANT GENOMICS

Probing plant evolution by GC content

Scientists use GC content (that is, the percentage of guanine or cytosine residues in a genome) as a proxy to measure many elements relating to gene evolution. Within the major group of flowering plants called monocots, which includes many agriculturally important species, the GC content of the genomes of grasses decreases from the 5' to 3' end of the gene. In order to better understand how the distribution of GC content evolved in monocots, Clément *et al*. examined orthologous genes across 10 monocot species. They found that the specific pattern of GC distribution seen in grasses is in fact not grass-specific—it is ancestral to the monocots. — LMZ

Genome Biol. Evol. 10.1093/gbe/evu278 (2014).

Lake Erie as seen from space

HIV ERADICATION

A cure from which there is no escape

One way to cure HIV-1 would be to pharmacologically drive virus production in latently infected cells and then to get $CD8^+$ T cells to kill these virus-spewing cells. But HIV-1 mutates to escape $CD8^+$ T cell responses, so is this strategy even possible? To find out, Deng *et al*. tested whether people chronically infected with HIV-1 harbor $CD8^+$ T cells that can recognize unmutated portions of latent HIV-1. They found that indeed, they do, and that stimulating these cells led them to kill cells infected with HIV-1 derived from latently infected cells, both in culture and in mice engineered to have a human immune system. — KLM

Nature 10.1038/nature14053 (2015).

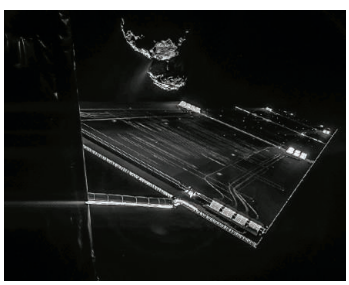
PHOTOS: (TOP TO BOTTOM) RANDY OLSON/NATIONAL GEOGRAPHIC CREATIVE; NASA/JEFF SCHMALTZ, LANCE/EOSDIS MODIS RAPID RESPONSE

Rosetta begins its COMET TALE

By **M. G. G. T. Taylor,[1]* C. Alexander,[2] N. Altobelli,[3] M. Fulle,[4] M. Fulchignoni,[5] E. Grün,[6] P. Weissman[2]**

Comets are the best sample of primitive solar nebula material presently available to us, dating back 4.57 billion years to the origin of our planetary system. Past missions to comets have all been "fast flybys": They provided only a snapshot view of the dust and ice nucleus, the nebulous coma surrounding it, and how the solar wind interacts with both of these components. Such space-based investigations of comets began in the 1980s with a flotilla of spacecraft: the European Space Agency's (ESA's) first deep space mission, Giotto, which pursued comet 1P/Halley; Deep Space 1 at 19P/Borrelly; Stardust at 81P/Wild 2; Deep Impact and Stardust NeXT at 9P/Tempel; and EPOXI at 103P/Hartley 2.

Rosetta is now taking a more prolonged look. The spacecraft is an ESA mission, with contributions from member states and from NASA, and it currently orbits the Jupiter family comet 67P/Churyumov-Gerasimenko (67P). Rosetta met the comet nucleus on 6 August 2014, at 3.7 astronomical units (AU) from the Sun, and delivered the Philae lander to the nucleus surface on 12 November 2014, when the comet was 3.0 AU from the Sun.



Rosetta is uniquely positioned to further the understanding of these primitive bodies, having revealed an unusual and fascinating object. After rendezvous, the Rosetta spacecraft moved from 100 km above the comet to a bound orbit only ~10 km away. This early period of the mission has revealed previously unseen details of a comet nucleus, as Rosetta's instruments recorded measurements that were once impossible. This issue of *Science* contains the first published scientific results from Rosetta at comet 67P.

The surface of the comet shows evidence of many active processes and is highly complex. The solid nucleus is an object for which neither horizontal nor vertical variations are modest (Thomas *et al.*, this issue). The current comet shape model suggests that the mass is 10^{13} kg (about 100 million times the mass of the international space station), with a bulk density of ~470 kg/m^3 (similar to cork, wood, or aerogel). The low mass and density values strongly constrain the composition and internal structure of the nucleus, implying a relatively fluffy nature, with a porosity of 70 to 80% (Sierks *et al.*, this issue). The nucleus surface itself appears rich in organic materials, with little sign of water ice (Capaccioni *et al.*, this issue).

The coma produced by ices sublimating from the nucleus is highly variable, displaying large diurnal and possibly seasonal changes. For example, both atomic H and O have been detected close to the nucleus and vary with time, probably stemming from electron impact dissociation of venting H_2O vapor. The total H_2O gas production rates varied from 1×10^{25} molecules per second in early June 2014 to 4×10^{25} molecules per second in early August, broadly consistent with predictions. In August, water outflow from the surface varied by a factor of at least 5, owing to the effects of terrain, comet shape, and daily illumination changes and possibly other factors (Gulkis *et al.*, this issue). The science team reports the detection of several molecules, including $H_2{}^{17}O$, $H_2{}^{18}O$, CO, and CO_2, and assessed their time variability and heterogeneous distribution (Hässig *et al.*, this issue). A high D/H ratio in water, 5.3×10^{-4}, was measured, which precludes the idea that Jupiter family comets contain solely Earth ocean–like water (Altwegg *et al.*, this issue). As observed at 3.6 AU from the Sun, a cloud of about 10^5 grains (larger than 5 cm) surrounds the nucleus in bound orbits, likely from the previous perihelion passage. The nucleus currently emits dust grains up to 2 cm in size, giving a dust/gas mass ratio of 4 ± 2 averaged over the sunlit nucleus surface (Rotundi *et al.*, this issue). This is higher than generally accepted for comets. In a progressive series of observations, Rosetta observed the emergence of an energetic ion environment from a low-activity comet nucleus under the influence of the solar wind (Nilsson *et al.*, this issue).

The data presented here allow us to build a detailed portrait of comet 67P. These initial observations provide a reference description of the global shape, the surface morphology and composition, and the bulk physical properties of the nucleus. Subsequent measurements with the orbiter and with the Philae lander will further describe the comet over time. Rosetta will follow the comet at close range through its closest approach to the Sun, perihelion, in August 2015, and then as the comet moves away from the Sun. The spacecraft will perform many flybys that will allow the onboard instruments to measure the evolution of the nucleus and coma with respect to the comet's initial state, defined by the data presented here.

The Rosetta mission has begun to explore our origins, thanks to the efforts of thousands of people at ESA, NASA, industrial partners, and space agencies and to engineers and scientists from around the world. For more than 25 years, they dreamed of these moments when they designed, developed, and launched the Rosetta spacecraft and then followed its interplanetary journey, watched over its long sleep, and woke it from hibernation. These first papers are dedicated to all of them.

10.1126/science.aaa4542

PHOTO: ESA/ROSETTA/PHILAE/CIVA

[1]*European Space Research and Technology Centre, Science and Robotic Exploration Directorate, European Space Agency, Noordwijk, Netherlands.* [2]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.* [3]*European Space Astronomy Centre, Science and Robotic Exploration Directorate, European Space Agency, Madrid, Spain.* [4]*INAF–Osservatorio Astronomico, Trieste, Italy.* [5]*Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique, Observatoire de Paris, France.* [6]*Max-Planck-Institut für Kernphysik, Heidelberg, Germany.* ***Corresponding author. E-mail: mtaylor@cosmos.esa.int***

On the nucleus structure and activity of comet 67P/Churyumov-Gerasimenko

H. Sierks,* C. Barbieri, P. L. Lamy, R. Rodrigo, D. Koschny, H. Rickman, H. U. Keller, J. Agarwal, M. F. A'Hearn, F. Angrilli, A.-T. Auger, M. A. Barucci, J.-L. Bertaux, I. Bertini, S. Besse, D. Bodewits, C. Capanna, G. Cremonese, V. Da Deppo, B. Davidsson, S. Debei, M. De Cecco, F. Ferri, S. Fornasier, M. Fulle, R. Gaskell, L. Giacomini, O. Groussin, P. Gutierrez-Marques, P. J. Gutiérrez, C. Güttler, N. Hoekzema, S. F. Hviid, W.-H. Ip, L. Jorda, J. Knollenberg, G. Kovacs, J.-R. Kramm, E. Kührt, M. Küppers, F. La Forgia, L. M. Lara, M. Lazzarin, C. Leyrat, J. J. Lopez Moreno, S. Magrin, S. Marchi, F. Marzari, M. Massironi, H. Michalik, R. Moissl, S. Mottola, G. Naletto, N. Oklay, M. Pajola, M. Pertile, F. Preusker, L. Sabau, F. Scholten, C. Snodgrass, N. Thomas, C. Tubiana, J.-B. Vincent, K.-P. Wenzel, M. Zaccariotto, M. Pätzold

Images from the OSIRIS scientific imaging system onboard Rosetta show that the nucleus of 67P/Churyumov-Gerasimenko consists of two lobes connected by a short neck. The nucleus has a bulk density less than half that of water. Activity at a distance from the Sun of >3 astronomical units is predominantly from the neck, where jets have been seen consistently. The nucleus rotates about the principal axis of momentum. The surface morphology suggests that the removal of larger volumes of material, possibly via explosive release of subsurface pressure or via creation of overhangs by sublimation, may be a major mass loss process. The shape raises the question of whether the two lobes represent a contact binary formed 4.5 billion years ago, or a single body where a gap has evolved via mass loss.

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: sierks@mps.mpg.de Cite as H. Sierks *et al.*, *Science* **347**, aaa1044 (2015).
Read the full article at http://dx.doi.org/10.1126/science.aaa1044

Dust measurements in the coma of comet 67P/Churyumov-Gerasimenko inbound to the Sun

A. Rotundi,* H. Sierks, V. Della Corte, M. Fulle, P. J. Gutierrez, L. Lara, C. Barbieri, P. L. Lamy, R. Rodrigo, D. Koschny, H. Rickman, H. U. Keller, J. J. López-Moreno, M. Accolla, J. Agarwal, M. F. A'Hearn, N. Altobelli, F. Angrilli, M. A. Barucci, J.-L. Bertaux, I. Bertini, D. Bodewits, E. Bussoletti, L. Colangeli, M. Cosi, G. Cremonese, J.-F. Crifo, V. Da Deppo, B. Davidsson, S. Debei, M. De Cecco, F. Esposito, M. Ferrari, S. Fornasier, F. Giovane, B. Gustafson, S. F. Green, O. Groussin, E. Grün, C. Güttler, M. L. Herranz, S. F. Hviid, W. Ip, S. Ivanovski, J. M. Jerónimo, L. Jorda, J. Knollenberg, R. Kramm, E. Kührt, M. Küppers, M. Lazzarin, M. R. Leese, A. C. López-Jiménez, F. Lucarelli, S. C. Lowry, F. Marzari, E. Mazzotta Epifani, J. A. M. McDonnell, V. Mennella, H. Michalik, A. Molina, R. Morales, F. Moreno, S. Mottola, G. Naletto, N. Oklay, J. L. Ortiz, E. Palomba, P. Palumbo, J.-M. Perrin, J. Rodríguez, L. Sabau, C. Snodgrass, R. Sordini, N. Thomas, C. Tubiana, J.-B. Vincent, P. Weissman, K.-P. Wenzel, V. Zakharov, J. C. Zarnecki

Critical measurements for understanding accretion and the dust/gas ratio in the solar nebula, where planets were forming 4.5 billion years ago, are being obtained by the GIADA (Grain Impact Analyser and Dust Accumulator) experiment on the European Space Agency's Rosetta spacecraft orbiting comet 67P/Churyumov-Gerasimenko. Between 3.6 and 3.4 astronomical units inbound, GIADA and OSIRIS (Optical, Spectroscopic, and Infrared Remote Imaging System) detected 35 outflowing grains of mass 10^{-10} to 10^{-7} kilograms, and 48 grains of mass 10^{-5} to 10^{-2} kilograms, respectively. Combined with gas data from the MIRO (Microwave Instrument for the Rosetta Orbiter) and ROSINA

Published by AAAS

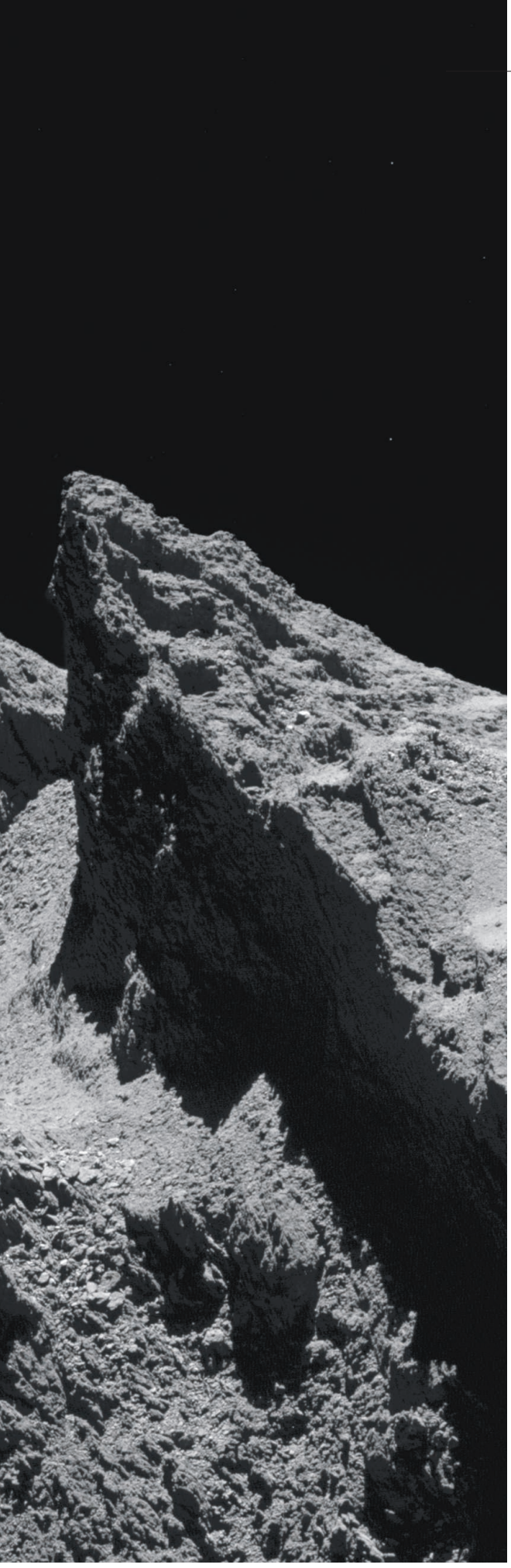


PHOTO: ESA/ROSETTA/MPS FOR OSIRIS TEAM MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

(Rosetta Orbiter Spectrometer for Ion and Neutral Analysis) instruments, we find a dust/gas mass ratio of 4 ± 2 averaged over the sunlit nucleus surface. A cloud of larger grains also encircles the nucleus in bound orbits from the previous perihelion. The largest orbiting clumps are meter-sized, confirming the dust/gas ratio of 3 inferred at perihelion from models of dust comae and trails.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: rotundi@uniparthenope.it** Cite as A. Rotundi *et al.*, *Science* **347**, aaa3905 (2015). Read the full article at http://dx.doi.org/10.1126/science.aaa3905

The organic-rich surface of comet 67P/Churyumov-Gerasimenko as seen by VIRTIS/Rosetta

F. Capaccioni,* A. Coradini, G. Filacchione, S. Erard, G. Arnold, P. Drossart, M. C. De Sanctis, D. Bockelee-Morvan, M. T. Capria, F. Tosi, C. Leyrat, B. Schmitt, E. Quirico, P. Cerroni, V. Mennella, A. Raponi, M. Ciarniello, T. McCord, L. Moroz, E. Palomba, E. Ammannito, M. A. Barucci, G. Bellucci, J. Benkhoff, J. P. Bibring, A. Blanco, M. Blecka, R. Carlson, U. Carsenty, L. Colangeli, M. Combes, M. Combi, J. Crovisier, T. Encrenaz, C. Federico, U. Fink, S. Fonti, W. H. Ip, P. Irwin, R. Jaumann, E. Kuehrt, Y. Langevin, G. Magni, S. Mottola, V. Orofino, P. Palumbo, G. Piccioni, U. Schade, F. Taylor, D. Tiphene, G. P. Tozzi, P. Beck, N. Biver, L. Bonal, J.-Ph. Combe, D. Despan, E. Flamini, S. Fornasier, A. Frigeri, D. Grassi, M. Gudipati, A. Longobardo, K. Markus, F. Merlin, R. Orosei, G. Rinaldi, K. Stephan, M. Cartacci, A. Cicchetti, S. Giuppi, Y. Hello, F. Henry, S. Jacquinod, R. Noschese, G. Peter, R. Politi, J. M. Reess, A. Semery

The VIRTIS (Visible, Infrared and Thermal Imaging Spectrometer) instrument on board the Rosetta spacecraft has provided evidence of carbon-bearing compounds on the nucleus of the comet 67P/Churyumov-Gerasimenko. The very low reflectance of the nucleus (normal albedo of 0.060 ± 0.003 at 0.55 micrometers), the spectral slopes in visible and infrared ranges (5 to 25 and 1.5 to 5% $kÅ^{-1}$), and the broad absorption feature in the 2.9-to-3.6–micrometer range present across the entire illuminated surface are compatible with opaque minerals associated with nonvolatile organic macromolecular materials: a complex mixture of various types of carbon-hydrogen and/or oxygen-hydrogen chemical groups, with little contribution of nitrogen-hydrogen groups. In active areas, the changes in spectral slope and absorption feature width may suggest small amounts of water ice. However, no ice-rich patches are observed, indicating a generally dehydrated nature for the surface currently illuminated by the Sun.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: fabrizio.capaccioni@iaps.inaf.it** Cite as F. Capaccioni *et al.*, *Science* **347**, aaa0628 (2015). Read the full article at http://dx.doi.org/10.1126/science.aaa0628

Birth of a comet magnetosphere: A spring of water ions

H. Nilsson,* G. Stenberg-Wieser, E. Behar, C. S. Wedlund, H. Gunell, M. Yamauchi, R. Lundin, S. Barabash, M. Wieser, C. Carr, E. Cupido, J. Burch, A. Fedorov, J.-A. Sauvaud, H. Koskinen, E. Kallio, J.-P. Lebreton, A. Eriksson, N. Edberg, R. Goldstein, P. Henri, C. Koenders, P. Mokashi, Z. Nemeth, I. Richter, K. Szego, M. Volwerk, C. Vallat, M. Rubin

The Rosetta mission shall accompany comet 67P/Churyumov-Gerasimenko from a heliocentric distance of >3.6 astronomical units through perihelion passage at 1.25 astronomical units, spanning low and maximum activity levels. Initially, the solar wind permeated the thin comet atmosphere formed from sublimation, until the size and plasma pressure of the ionized atmosphere

defined its boundaries: A magnetosphere is born. Using the Rosetta Plasma Consortium Ion Composition Analyzer, we traced the evolution from the first detection of water ions to when the atmosphere begins repelling the solar wind (~3.3 astronomical units), and we report the spatial structure of this early interaction. The near-comet water population comprises accelerated ions (<800 electron volts), produced upstream of Rosetta, and lower-energy locally produced ions; we estimate the fluxes of both ion species and energetic neutral atoms.

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: hans.nilsson@irf.se Cite as H. Nilsson *et al.*, *Science* **347**, aaa0571 (2015).
Read the full article at http://dx.doi.org/10.1126/science.aaa0571

The morphological diversity of comet 67P/Churyumov-Gerasimenko

N. Thomas,* H. Sierks, C. Barbieri, P. L. Lamy, R. Rodrigo, H. Rickman, D. Koschny, H. U. Keller, J. Agarwal, M. F. A'Hearn, F. Angrilli, A.-T. Auger, M. A. Barucci, J.-L. Bertaux, I. Bertini, S. Besse, D. Bodewits, G. Cremonese, V. Da Deppo, B. Davidsson, M. De Cecco, S. Debei, M. R. El-Maarry, F. Ferri, S. Fornasier, M. Fulle, L. Giacomini, O. Groussin, P. J. Gutierrez, C. Güttler, S. F. Hviid, W.-H. Ip, L. Jorda, J. Knollenberg, J.-R. Kramm, E. Kührt, M. Küppers, F. La Forgia, L. M. Lara, M. Lazzarin, J. J. Lopez Moreno, S. Magrin, S. Marchi, F. Marzari, M. Massironi, H. Michalik, R. Moissl, S. Mottola, G. Naletto, N. Oklay, M. Pajola, A. Pommerol, F. Preusker, L. Sabau, F. Scholten, C. Snodgrass, C. Tubiana, J.-B. Vincent, K.-P. Wenzel

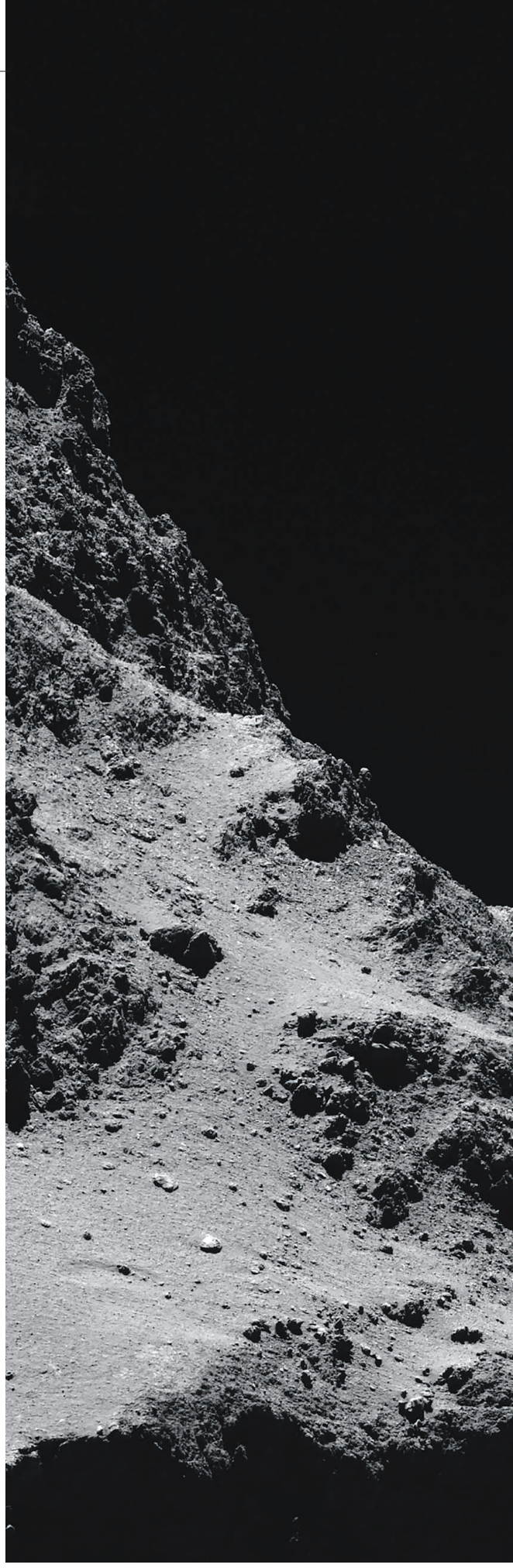
Images of comet 67P/Churyumov-Gerasimenko acquired by the OSIRIS (Optical, Spectroscopic and Infrared Remote Imaging System) imaging system onboard the European Space Agency's Rosetta spacecraft at scales of better than 0.8 meter per pixel show a wide variety of different structures and textures. The data show the importance of airfall, surface dust transport, mass wasting, and insolation weathering for cometary surface evolution, and they offer some support for subsurface fluidization models and mass loss through the ejection of large chunks of material.

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: nicolas.thomas@space.unibe.ch Cite as N. Thomas *et al.*, *Science* **347**, aaa0440 (2015). Read the full article at http://dx.doi.org/10.1126/science.aaa0440

67P/Churyumov-Gerasimenko, a Jupiter family comet with a high D/H ratio

K. Altwegg,* H. Balsiger, A. Bar-Nun, J. J. Berthelier, A. Bieler, P. Bochsler, C. Briois, U. Calmonte, M. Combi, J. De Keyser, P. Eberhardt, B. Fiethe, S. Fuselier, S. Gasc, T. I. Gombosi, K.C. Hansen, M. Hässig, A. Jäckel, E. Kopp, A. Korth, L. LeRoy, U. Mall, B. Marty, O. Mousis, E. Neefs, T. Owen, H. Rème, M. Rubin, T. Sémon, C.-Y. Tzou, H. Waite, P. Wurz

The provenance of water and organic compounds on Earth and other terrestrial planets has been discussed for a long time without reaching a consensus. One of the best means to distinguish between different scenarios is by determining the deuterium-to-hydrogen (D/H) ratios in the reservoirs for comets and Earth's oceans. Here, we report the direct in situ measurement of the D/H ratio in the Jupiter family comet 67P/Churyumov-Gerasimenko by the ROSINA mass spectrometer aboard the European Space Agency's Rosetta spacecraft, which is found to be $(5.3 \pm 0.7) \times 10^{-4}$—that is, approximately three times the terrestrial value. Previous cometary measurements and our new finding suggest a wide range of D/H ratios in the



Published by AAAS

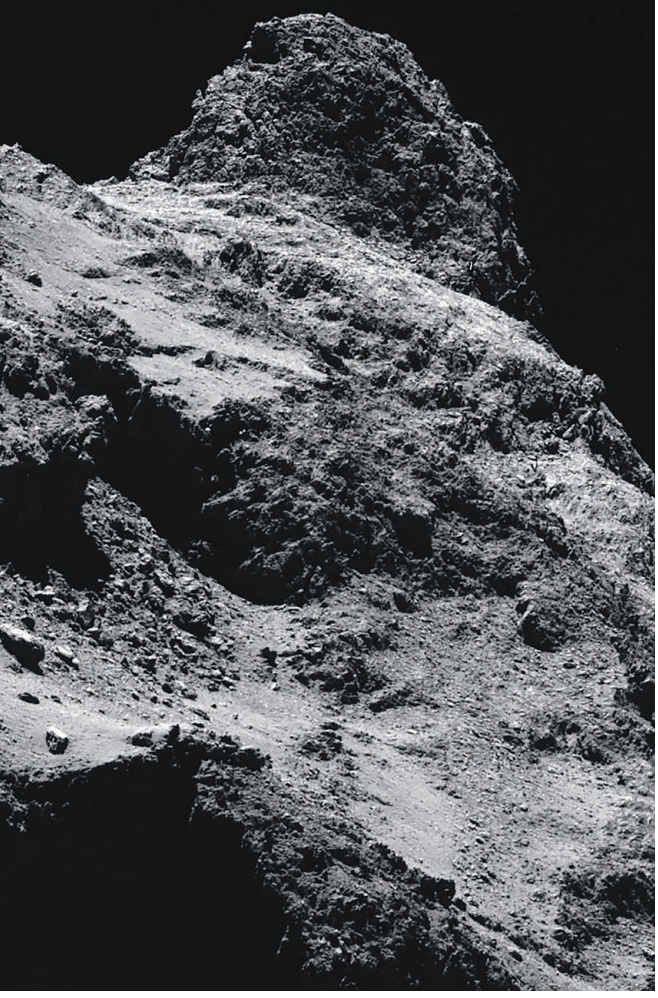


PHOTO: ESA/ROSETTA/MPS FOR OSIRIS TEAM MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA

water within Jupiter family objects and preclude the idea that this reservoir is solely composed of Earth ocean–like water.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: altwegg@space.unibe.ch** Cite as K. Altwegg *et al.*, *Science* **347**, 1261952 (2015).
Read the full article at http://dx.doi.org/10.1126/science. 1261952

Time variability and heterogeneity in the coma of 67P/ Churyumov-Gerasimenko

M. Hässig,* K. Altwegg, H. Balsiger, A. Bar-Nun, J. J. Berthelier, A. Bieler, P. Bochsler, C. Briois, U. Calmonte, M. Combi, J. De Keyser, P. Eberhardt, B. Fiethe, S. A. Fuselier, M. Galand, S. Gasc, T. I. Gombosi, K. C. Hansen, A. Jäckel, H. U. Keller, E. Kopp, A. Korth, E. Kührt, L. Le Roy, U. Mall, B. Marty, O. Mousis, E. Neefs, T. Owen, H. Rème, M. Rubin, T. Sémon, C. Tornow, C.-Y. Tzou, J. H. Waite, P. Wurz

Comets contain the best-preserved material from the beginning of our planetary system. Their nuclei and comae composition reveal clues about physical and chemical conditions during the early solar system when comets formed. ROSINA (Rosetta Orbiter Spectrometer for Ion and Neutral Analysis) onboard the Rosetta spacecraft has measured the coma composition of comet 67P/ Churyumov-Gerasimenko with well-sampled time resolution per rotation. Measurements were made over many comet rotation periods and a wide range of latitudes. These measurements show large fluctuations in composition in a heterogeneous coma that has diurnal and possibly seasonal variations in the major outgassing species: water, carbon monoxide, and carbon dioxide. These results indicate a complex coma-nucleus relationship where seasonal variations may be driven by temperature differences just below the comet surface.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: myrtha.haessig@swri.org** Cite as M. Hässig *et al.*, *Science* **347**, aaa0276 (2015).
Read the full article at http://dx.doi.org/10.1126/science.aaa0276

Subsurface properties and early activity of comet 67P/Churyumov-Gerasimenko

S. Gulkis,* M. Allen, P. von Allmen, G. Beaudin, N. Biver, D. Bockelée-Morvan, M. Choukroun, J. Crovisier, B. J. R. Davidsson, P. Encrenaz, T. Encrenaz, M. Frerking, P. Hartogh, M. Hofstadter, W.-H. Ip, M. Janssen, C. Jarchow, S. Keihm, S. Lee, E. Lellouch, C. Leyrat, L. Rezac, F. P. Schloerb, T. Spilker

Heat transport and ice sublimation in comets are interrelated processes reflecting properties acquired at the time of formation and during subsequent evolution. The Microwave Instrument on the Rosetta Orbiter (MIRO) acquired maps of the subsurface temperature of comet 67P/Churyumov-Gerasimenko, at 1.6 mm and 0.5 mm wavelengths, and spectra of water vapor. The total H_2O production rate varied from 0.3 kg s^{-1} in early June 2014 to 1.2 kg s^{-1} in late August and showed periodic variations related to nucleus rotation and shape. Water outgassing was localized to the "neck" region of the comet. Subsurface temperatures showed seasonal and diurnal variations, which indicated that the submillimeter radiation originated at depths comparable to the diurnal thermal skin depth. A low thermal inertia (~10 to 50 J K^{-1} m^{-2} $s^{-0.5}$), consistent with a thermally insulating powdered surface, is inferred.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: samuel.gulkis@jpl.nasa.gov** Cite as S. Gulkis *et al.*, *Science* **347**, aaa0709 (2015).
Read the full article at http://dx.doi.org/10.1126/science.aaa0709

COMETARY SCIENCE

Time variability and heterogeneity in the coma of 67P/Churyumov-Gerasimenko

M. Hässig,[1,2*] K. Altwegg,[1,3] H. Balsiger,[1] A. Bar-Nun,[4] J. J. Berthelier,[5] A. Bieler,[1,6] P. Bochsler,[1] C. Briois,[7] U. Calmonte,[1] M. Combi,[6] J. De Keyser,[8,9] P. Eberhardt,[1†] B. Fiethe,[10] S. A. Fuselier,[2] M. Galand,[11] S. Gasc,[1] T. I. Gombosi,[6] K. C. Hansen,[6] A. Jäckel,[1] H. U. Keller,[12,14] E. Kopp,[1] A. Korth,[13] E. Kührt,[14] L. Le Roy,[3] U. Mall,[13] B. Marty,[15] O. Mousis,[16] E. Neefs,[17] T. Owen,[18] H. Rème,[19,20] M. Rubin,[1] T. Sémon,[1] C. Tornow,[14] C.-Y. Tzou,[1] J. H. Waite,[2] P. Wurz[1]

Comets contain the best-preserved material from the beginning of our planetary system. Their nuclei and comae composition reveal clues about physical and chemical conditions during the early solar system when comets formed. ROSINA (Rosetta Orbiter Spectrometer for Ion and Neutral Analysis) onboard the Rosetta spacecraft has measured the coma composition of comet 67P/Churyumov-Gerasimenko with well-sampled time resolution per rotation. Measurements were made over many comet rotation periods and a wide range of latitudes. These measurements show large fluctuations in composition in a heterogeneous coma that has diurnal and possibly seasonal variations in the major outgassing species: water, carbon monoxide, and carbon dioxide. These results indicate a complex coma-nucleus relationship where seasonal variations may be driven by temperature differences just below the comet surface.

Initially, comets were classified depending on the location where they formed in the protoplanetary disc (*1*, *2*). This classification assumed a similar composition of the nucleus within a given formation region. No cometary nucleus composition has been sampled in situ. Rather, it is implicitly assumed that measurements of the outgassing of comets reveal the composition of the volatile components of the nucleus. However, compositional homogeneity of at least one comet was confirmed by studying outgassing from the fragments of the broken-up comet Schwassmann-Wachmann 3 (*3*). Detailed observations of other cometary comae indicated that there is evidence of heterogeneity. Missions to comet Halley detected release of volatiles in multiple jet-like features that were dominantly seen on the sunlit side of the nucleus (*4*, *5*). The Deep Impact mission detected asymmetries in composition in the coma of Tempel 1 (*6*). In particular, these remote sensing observations at Tempel 1 indicated an absence of correlation between H_2O and CO_2 in the coma.

Detailed, close-up cometary images have also showed visible differences between different areas of cometary nuclei. These images suggested that heterogeneity in the coma of a comet may be related to heterogeneity of the nucleus. Observations by EPOXI (Extrasolar Planet Observation and Deep Impact Extended Investigation) at Hartley 2 in 2010 near perihelion indicated that the nucleus is complex, with two different sized lobes separated by a middle waist region that is smoother and lighter in color (*7*). Outgassing from sunlit surfaces of the nucleus revealed that the waist and one of the lobes were very active. A CO_2 source was detected at the small lobe of the comet, whereas the waist was more active in H_2O and had a considerably lower CO_2 content. Based on these coma observations, it has been tentatively suggested that the heterogeneity in the comet's nucleus was primordial (*7*). Seasonal effects could not be ruled out because the observations also showed a complex rotational state for the comet (*7*). The smaller of the two lobes may have been illuminated differently because of this complex rotation (*7*). In support of the findings at Hartley 2, there are indications of a heterogeneous nucleus for comet Tuttle and a heterogeneous coma (*7*, *8*).

The Stardust mission to comet P81/Wild 2, on the other hand, showed a large mixing of materials on the scale of grains and therefore a homogenized mix of the refractory material in the comet (*9*). The results at Hartley 2 and at

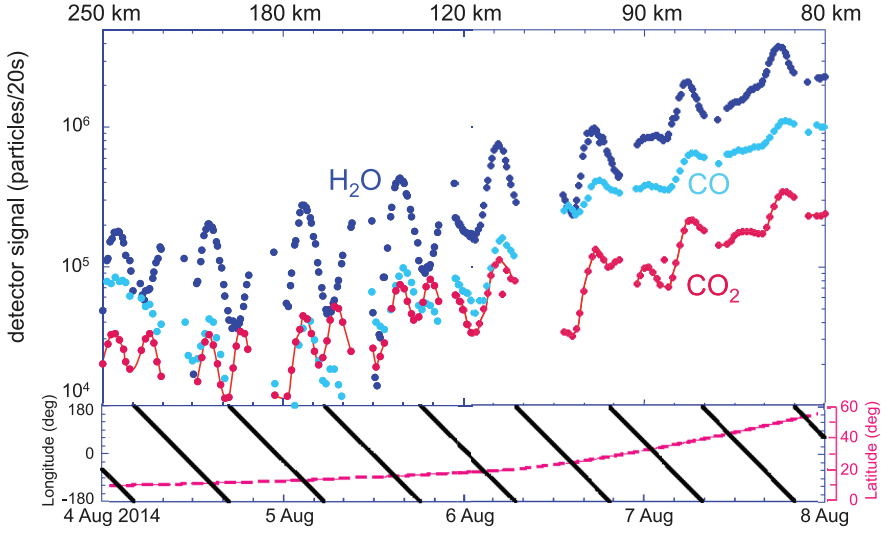


Fig. 1. H_2O, CO, and CO_2 measurements for 4 to 8 August 2014. The upper panel shows the signal on the DFMS detector for H_2O, CO, and CO_2, and the lower panel shows the latitude and longitude of the nadir view of the spacecraft. At the top is the distance from the spacecraft to the comet. The signal increases with decreasing distance to the comet, and diurnal variations are also visible. CO_2 has a different periodicity than H_2O, as seen around 4 to 6 August.

[1]Physikalisches Institut, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland. [2]Southwest Research Institute, 6220 Culebra Road, San Antonio, TX 78238, USA. [3]Center for Space and Habitability (CSH), University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland. [4]Department of Geosciences, Tel-Aviv University, Ramat-Aviv, Tel-Aviv, Israel. [5]Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS), Institute Pierre Simon Laplace (IPSL), Centre national de recherche scientifique (CNRS), Université Pierre et Marie Curie (UPMC), Université de Versailles Saint-Quentin-en-Yvelines (UVSQ), BP 102, UPMC, 4 Place Jussieu, F-75252 Paris Cedex 05, France. [6]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, 2455 Hayward Street, Ann Arbor, MI 48109, USA. [7]Laboratoire de Physique et Chimie de l'Environnement et de l'Espace (LPC2E), UMR 7328 CNRS – Université d'Orléans, France. [8]Belgian Institute for Space Aeronomy (BIRA-IASB), Ringlaan 3, B-1180 Brussels, Belgium. [9]Center for Plasma Astrophysics, KULeuven, Celestijnenlaan 200D, 3001 Heverlee, Belgium. [10]Institute of Computer and Network Engineering (IDA), TU Braunschweig, Hans-Sommer-Straße 66, D-38106 Braunschweig, Germany. [11]Department of Physics, Imperial College London, Prince Consort Road, London SW7 2AZ, UK. [12]Institute for Geophysics and Extraterrestrial Physics, Technische Universität (TU) Braunschweig, 38106 Braunschweig, Germany. [13]Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany. [14]German Aerospace Center, Institute of Planetary Research, Asteroids and Comets, Rutherfordstraße 2, 12489 Berlin, Germany. [15]Centre de Recherches Pétrographiques et Géochimiques (CRPG), 15 Rue Notre Dame des Pauvres, BP 20, 54501 Vandoeuvre lès Nancy, France. [16]Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille), UMR 7326, 13388, Marseille, France. [17]Engineering Division, BIRA-IASB, Ringlaan 3, B-1180 Brussels, Belgium. [18]Institute for Astronomy, University of Hawaii, Honolulu, HI 96822, USA. [19]Université de Toulouse, Université Paul Sabathier (UPS), Observatoire de Midi-Pyrénées (OMP), Institut de Recherche en Astrophysique et Planétologie (IRAP), Toulouse, France. [20]CNRS, IRAP, 9 Avenue du Colonel Roche, BP 44346, F-31028 Toulouse Cedex 4, France.
***Corresponding author. E-Mail: myrtha.haessig@swri.org**
†Deceased.

P81/Wild 2 raise the larger question of whether heterogeneity in the coma is a common feature in comets and whether this reveals an underlying heterogeneity in the composition of the nucleus, which would point to general transport of cometesimals in the early solar system.

In August, the European Space Agency's mission Rosetta arrived at its target comet 67P/Churyumov-Gerasimenko (67P) after a 10-year journey (*10*). Rosetta provides an excellent opportunity for long-term study during the comet's sunward approach to perihelion. The observations presented here are from a 2-month period beginning near the initial encounter at about 3.5 astronomical units from the Sun.

Like Hartley 2, the nucleus of 67P appears complex in shape. The comet 67P consists of two lobes of different sizes, connected by a neck region. The lobes are much larger, more rugged, and darker than the neck region and the overall shape has been compared to a rubber duck (*11*). The structural similarities of 67P and Hartley 2 suggest the possibility of another heterogeneous comet and, by virtue of the extended observations at 67P, a chance to determine whether heterogeneity in the coma and nucleus are related.

Here, we show compositional variations in H_2O, CO, and CO_2 at comet 67P observed with ROSINA/DFMS (Rosetta Orbiter Spectrometer for Ion and Neutral Analysis/Double Focusing Mass Spectrometer) (*12*). ROSINA/DFMS is a mass spectrometer that measures the in situ neutral and plasma coma composition at the position of the spacecraft (see the supplementary materials). During Rosetta's approach to 67P, ROSINA/DFMS measured the neutral coma composition with a time resolution (>10 measurements per rotation) much finer than the rotation period of the comet of ~12.4 hours (*13*). In August, the spacecraft scanned the comet at northern summer hemisphere (positive latitudes) from about 10° up to almost 90° (Coordinates: Cheops System). In September, the spacecraft made a similar scan at southern winter hemisphere (negative latitudes) down to about −50°. Two data sets are shown in Figs. 1 and 2 to illustrate the diurnal and latitudinal variations and heterogeneity of the cometary coma.

During this approach and latitude scan, the H_2O, CO, and CO_2 signals from the comet increased by more than an order of magnitude, roughly in agreement with a $1/R^2$ dependence on the coma density, where R is the distance from the comet's center. Overall, the H_2O signal is the strongest; however, there are clearly periods when the CO or CO_2 signals rival that of H_2O.

Superposed on this general increase in signal are large, diurnal variations for all three neutral species. For H_2O, these variations are periodic, initially with half the rotation rate of the comet (~6.2 hours) and then, after 6 August, at the rotation rate (~12.4 hours). This change in periodicity in the signal is interpreted as a latitudinal effect of the sampling position. Peaks occur at ±90° longitude. For the most part, the CO signal follows the H_2O signal, but the variations are smaller. CO_2 shows a different periodicity. Initially, a CO_2 peak is observed in association with an H_2O peak, and a second CO_2 peak occurs approximately 3 hours later. After August 6, a single CO_2 peak is observed; however, this peak is not exactly coincident with the H_2O peak. The two CO_2 peaks merge, resulting in a shoulder on the main peak and a slight shift of the main CO_2 peak relative to that of H_2O (~45 min or one measurement point). Statistical uncertainties ($\sqrt{\#\text{particles}}$) in the signal detected by ROSINA/DFMS are smaller than the dots in Figs. 1 to 3, and contributions to the signal due to spacecraft outgassing (*14*) are subtracted.

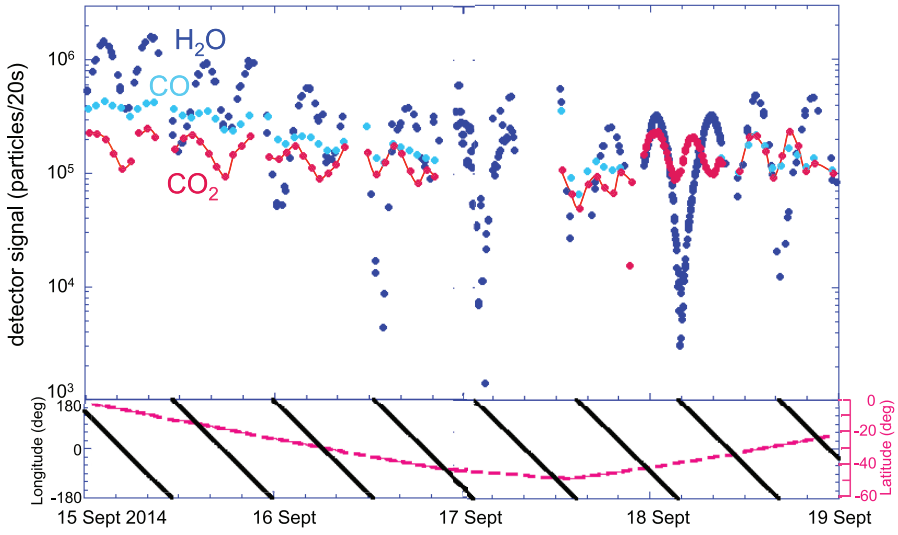


Fig. 2. H_2O, CO, and CO_2 measurements for 15 to 19 September 2014. Over this 4-day period, the spacecraft remained at a nearly fixed distance from the comet and executed a southern latitude scan from about 0° to −45° latitude. H_2O and CO_2 have different periodicities, and there are deep minima in the H_2O signal. CO follows the CO_2 profile with less variation.

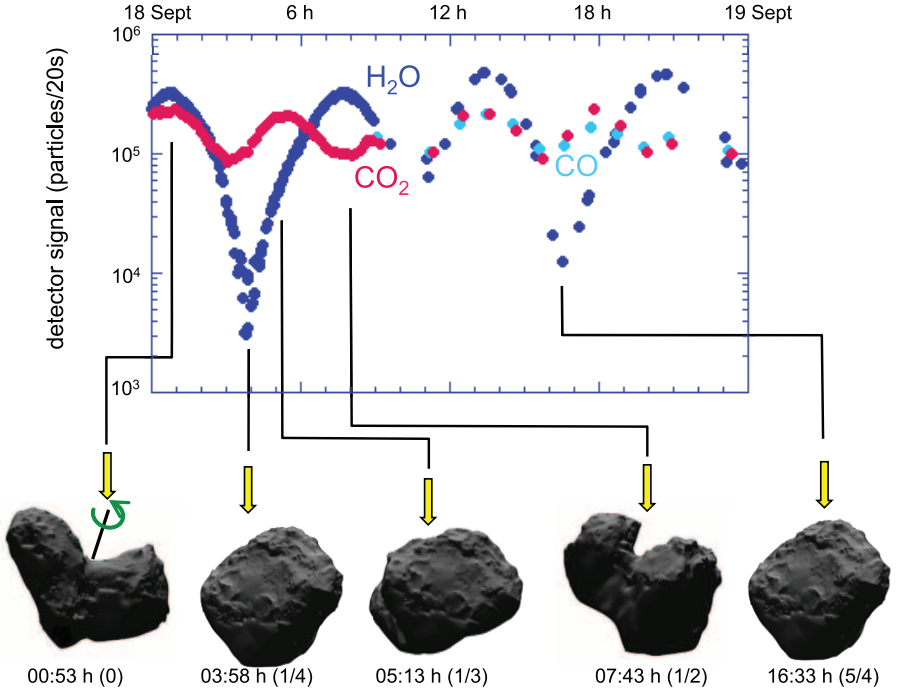


Fig. 3. H_2O, CO, and CO_2 profiles for 18 September 2014. The Sun is shining on the comet from the top middle of the pictures. The snapshots of the spacecraft view of the comet show that H_2O peaks are observed when the neck region is in view. The separate CO_2 peak and the deep minimum in H_2O occur when the spacecraft views the larger of the two lobes and the neck region is blocked. [Shape model credit: ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA].

The diurnal variations at half the rotation rate of the comet that are seen in August are also observed at southern latitudes in the September time frame (Fig. 2). The H_2O peaks in Fig. 3 are nearly equal, and there is a deep minimum between the two peaks. As in the first data set, CO follows H_2O. However, there is much less variation in CO than in H_2O, resulting in times when the CO signal is greater than that for H_2O. The best example of the differences between H_2O and CO_2 are seen just after 18 September (Figs. 2 and 3). The nearly equal H_2O peaks and the deep minimum in the H_2O signal are evident, as is the clear offset between the second CO_2 and H_2O peaks.

We have combined the signal and the spacecraft perspective over the 18 to 19 September 2014 window to illustrate which side of the comet is in view when the peaks occur (Fig. 3). The peaks in H_2O signal are observed when the neck of the comet is in view of the spacecraft. The deep minimum in H_2O signal is observed when the spacecraft views the southern hemisphere of the larger of the two lobes. This large lobe blocks a direct view of the neck of the comet. The separate, second CO_2 enhancement is observed when the spacecraft views the underside of the body of the larger of the two lobes of the comet. The CO signal in the second rotation of the comet follows the CO_2 profile, and CO and CO_2 have very similar intensities.

We see from this data (Figs. 1 to 3) that the coma composition of 67P is highly heterogeneous. H_2O, CO, and CO_2 variations are strongly tied to the rotation period of the comet and to the observing latitude. At large negative latitudes, the H_2O signal varies by at least two orders of magnitude (Fig. 3). Also, the H_2O minima are not as deep when the spacecraft is at mid and high positive latitudes because there is a view of the neck region over the edge of the larger lobe (see Fig. 1 and the observations on 15 September in Fig. 2).

The separate CO_2 peak also occurs when the spacecraft views the bottom of the larger of the two lobes of the comet (see Fig. 3 at 5 hours). CO follows H_2O at positive latitudes and follows both H_2O and CO_2 at negative latitudes.

The separate CO_2 peak, the large variations in the H_2O signal, and the weaker variations in CO result in large changes in the relative concentration of H_2O, CO, and CO_2 in the heterogeneous coma of 67P (see the supplementary materials). For example, the CO/H_2O number density ratio is 0.13 ± 0.07, and the CO_2/H_2O ratio is 0.08 ± 0.05 in the last H_2O peak on 7 August at 18 hours in Fig. 1 (measured high in the northern summer hemisphere). However, The CO/H_2O ratio changes from 0.56 ± 0.15 to 4 ± 1 and back to 0.38 ± 0.15 within the second cometary rotation (Fig. 3), between 12 and 24 hours on 18 September, measured low in the southern winter hemisphere. Similarly, the CO_2/H_2O ratio changes from 0.67 ± 0.15 to 8 ± 2 and back to 0.39 ± 0.15 over the same rotation. These are large changes within a short amount of time, which indicate a strongly heterogeneous and time-variable coma.

The similarities in the structure of the nuclei and the heterogeneous comae of 67P and Hartley 2 are striking. The behavior in terms of the H_2O dominant outgassing at the neck versus CO_2 outgassing at one of the lobes described here was also found for Hartley 2 (*7*).

The compositional differences in the Hartley 2 coma were interpreted as evidence for a heterogeneous cometary nucleus (*7*). However, seasonal effects could not be ruled out. With observations over a wide range of latitudes at 67P, we can distinguish between compositional differences and seasonal effects; to do so, we have mapped the CO_2/H_2O density ratio from 17 August through 22 September onto the shape model (Fig. 4).

Although a direct mapping of the signal observed in the coma onto the comet surface is oversimplified, a generalized interpretation reveals features of the outgassing of the comet. Seasonal effects on the CO_2/H_2O ratio are clearly evident (Fig. 4). On the upper half of the comet, the CO_2/H_2O ratio is less than 1, indicating a higher sublimation of H_2O from positive latitude regions that receive more illumination during northern hemisphere summer on the comet. A broad region of high CO_2/H_2O ratio occurs at negative latitudes in the winter hemisphere, likely the result of deep minima in the H_2O signal (such as the one shown in Fig. 3 on 18 September at 4 hours). This winter hemisphere of the comet is poorly illuminated by the Sun. With limited illumination, this region of the comet nucleus may be considerably colder than other regions, including the neck and smaller lobe. The temperature at and below the surface of the nucleus may be sufficient to sublimate CO and CO_2 but not sufficient to sublimate water. The weak, periodic illumination of this region may be sufficient to drive CO and CO_2 sublimation, producing the separate CO and CO_2 peak (Fig. 3 at 18 hours). However, the compositional asymmetry in the two H_2O peaks cannot be explained in a similar way and might be the strongest indication for heterogeneity in the comet nucleus. The strong heterogeneity in the coma of comet 67P is likely driven by seasonal effects on the comet nucleus. However, the smaller variation of CO and CO_2 compared with H_2O might indicate that CO and CO_2 ices sublimate from a greater depth, whereas H_2O ice sublimates closer to the surface and experiences more direct temperature differences due to sunlight. Furthermore, that lack of overall correlation between H_2O, CO, and CO_2 implies that the outgassing from the nucleus is not correlated or that CO and CO_2 are not strictly embedded in H_2O. For Tempel 1, material was found in layers and supports the above idea (*15*).

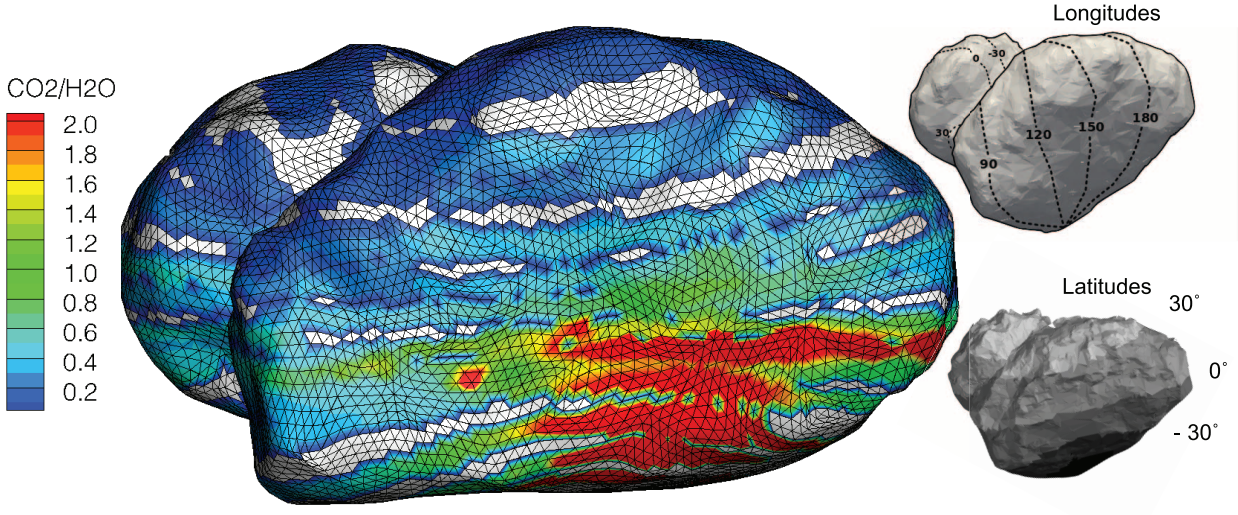


Fig. 4. The nadir point for each pair of CO_2/H_2O measurements over the time period from 17 August through 22 September was mapped to the model surface. The mapping is shown for the bottom side of the larger of the two lobes of the comet, and cometary latitudes run approximately vertically in this. The layering is due to spacecraft rastering above the comet nucleus. A high ratio is measured for the lower part that is poorly sunlit in southern hemisphere winter. [Shape model credit: ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA].

In summary, the coma composition has been measured over many rotational periods of the comet and a wide range of latitudes with high time resolution and compositional detail. Concentrations of the three molecules change over the rotational period of the comet and indicate a strongly heterogeneous coma. For the most part, H_2O dominates, but CO and CO_2 can at times dominate in the coma. These observations also indicate that there are substantial diurnal and latitudinal variations in the coma. Peaks in the H_2O signal are observed, along with deep minima at high negative latitudes when the neck region of the nucleus is blocked from view of the spacecraft. A separate peak in CO_2 signal occurs when the winter hemisphere of the larger lobe of the comet faces the spacecraft. The diurnal and latitudinal variations suggest that compositional differences in the coma may be seasonal and may indicate different subsurface temperatures in the nucleus.

Further observations may distinguish seasonal effect from nucleus heterogeneity. As the comet approaches the Sun, the overall temperature of the nucleus will increase, and as the seasons change, there may be considerable changes in the H_2O, CO, and CO_2 outgassing, with the current high CO_2/H_2O ratio region shown in Fig. 4. In addition, differences in the sublimation of species similar in sublimation temperatures could demonstrate the extend of heterogeneity in the nucleus independent of seasonal changes.

REFERENCES AND NOTES

1. M. F. A'Hearn, R. C. Millis, D. O. Schleicher, D. J. Osip, R. V. Birch, The ensemble properties of comets: Results from narrowband photometry of 85 comets, 1976-1992. *Icarus* **118**, 223–270 (1995). doi: 10.1006/icar.1995.1190
2. M. J. Mumma, S. B. Charnley, The chemical composition of comets: Emerging taxonomies and natal heritage. *Annu. Rev. Astron. Astrophys.* **49**, 471–524 (2011). doi: 10.1146/annurev-astro-081309-130811
3. N. Dello Russo *et al.*, Compositional homogeneity in the fragmented comet 73P/Schwassmann-Wachmann 3. *Nature* **448**, 172–175 (2007). doi: 10.1038/nature05908; pmid: 17625560
4. H. U. Keller *et al.*, First Halley Multicolour Camera imaging results from Giotto. *Nature* **321** (6067s), 320–326 (1986). doi: 10.1038/321320a0
5. H. A. Weaver, M. J. Mumma, H. P. Larson, D. S. Davis, Postperihelion observations of water in comet Halley. *Nature* **324**, 441–444 (1986). doi: 10.1038/324441a0
6. L. M. Feaga, M. F. A'Hearn, J. M. Sunshine, O. Groussin, T. L. Farnham, Asymmetreis in the distribution of H_2O and CO_2 in the inner coma of Comet 9P/Tempel 1 as observed by Deep Impact. *Icarus* **190**, 345–356 (2007). doi: 10.1016/j.icarus.2007.04.009
7. M. F. A'Hearn *et al.*, EPOXI at comet Hartley 2. *Science* **332**, 1396–1400 (2011). doi: 10.1126/science.1204054; pmid: 21680835
8. B. P. Bonev *et al.*, The peculiar volatile composition of comet 8P/Tuttle: A contact binary of chemically distinct cometesimals? *Astrophys. J.* **680**, L61–L64 (2008). doi: 10.1086/589649
9. M. F. A'Hearn, Whence comets? *Science* **314**, 1708–1709 (2006). doi: 10.1126/science.1137083; pmid: 17170287
10. K.-H. Glassmeier, H. Boehnhardt, D. Koschny, E. Kührt, I. Richter, The Rosetta Mission: Flying towards the origin of the solar system. *Space Sci. Rev.* **128**, 745–801 (2007). doi: 10.1007/s11214-006-9140-8
11. N. Thomas *et al.*, The morphological diversity of comet 67P/Churyumov-Gerasimenko. *Science* **347**, aaa0440 (2014). doi: 10.1126/science.aaa0440
12. H. Balsiger *et al.*, ROSINA-ROSETTA-orbiter-spectrometer-for-ion-and-neutral-analysis. *Space Sci. Rev.* **128**, 1–21 (2007). doi: 10.1007/s11214-006-8335-3
13. S. Mottola *et al.*, The rotation state of 67P/Churyumov-Gerasimenko from approach observations with the OSIRIS cameras on Rosetta. *Astron. Astrophys.* **569**, L2 (2014). doi: 10.1051/0004-6361/201424590
14. B. Schläppi *et al.*, Influence of spacecraft outgassing on the exploration of tenuous atmospheres with in situ mass spectrometry. *J. Geophys. Res.* **115**, A12313 (2010). doi: 10.1029/2010JA015734
15. M. F. A'Hearn *et al.*, Deep Impact: Excavating comet Tempel 1. *Science* **310**, 258–264 (2005). doi: 10.1126/science.1118923; pmid: 16150978

ACKNOWLEDGMENTS

The authors thank the following institutions and agencies, which supported this work: Work at the University of Bern was funded by the State of Bern, the Swiss National Science Foundation, and the European Space Agency (ESA) PRODEX Program. Work at the Max-Planck Institut für Sonnensystemforschung (MPS) was funded by the Max-Planck-Gesellschaft (MPG) and Bundesministerium für Wirtschaft und Energie (BMWI) under contract 50QP1302. Work at Southwest Research institute (SwRI) was supported by subcontract 1496541 from the Jet Propulsion Laboratory (JPL) and under NASA prime contract NNX148F71G. Work at BIRA-IASB was supported by the Belgian Science Policy Office via PRODEX/ROSINA PEA 90020. Work at Imperial College London has been partially funded by the Science and Technology Facilities Council (STFC). This work has been carried out thanks to the support of the A*MIDEX project (ANR-11-IDEX-0001-02) funded by the Investissements d'Avenir French Government program, managed by the French National Research Agency (ANR). This work was supported by CNES grants at IRAP, LATMOS, LPC2E, Laboratoire d'Astrophysique de Marseille (LAM) and CRPG and by the European Research Council (grant 267255). A.B.-N. thanks the Ministry of Science and the Israel Space agency. Work at the University of Michigan was funded by NASA under contract JPL-1266313. Work by J.H.W. at SwRI was supported by NASA JPL subcontract NAS703001TONMO710889. ROSINA would not give such outstanding results without the work of the many engineers, technicians, and scientists involved in the mission, in the Rosetta spacecraft, and in the ROSINA instrument over the past 20 years whose contributions are gratefully acknowledged. Rosetta is an ESA mission with contributions from its member states and NASA. We thank the OSIRIS team for giving permission to use the shape model. We acknowledge herewith the work of the ESA Rosetta team. All ROSINA data are available on request until it is released to the PSA archive of ESA and to the PDS archive of NASA.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/aaa0276/suppl/DC1
Materials and Methods
Supplementary Text
References

10 October 2014; accepted 31 December 2014
10.1126/science.aaa0276

COMETARY SCIENCE

Birth of a comet magnetosphere: A spring of water ions

Hans Nilsson,[1,2]* Gabriella Stenberg Wieser,[1] Etienne Behar,[1,2] Cyril Simon Wedlund,[3] Herbert Gunell,[4] Masatoshi Yamauchi,[1] Rickard Lundin,[1] Stas Barabash,[1] Martin Wieser,[1] Chris Carr,[5] Emanuele Cupido,[5] James L. Burch,[6] Andrei Fedorov,[7] Jean-André Sauvaud,[7] Hannu Koskinen,[8,9] Esa Kallio,[3,9] Jean-Pierre Lebreton,[10] Anders Eriksson,[11] Niklas Edberg,[11] Raymond Goldstein,[4] Pierre Henri,[10] Christoph Koenders,[12] Prachet Mokashi,[6] Zoltan Nemeth,[13] Ingo Richter,[12] Karoly Szego,[13] Martin Volwerk,[14] Claire Vallat,[15] Martin Rubin[16]

The Rosetta mission shall accompany comet 67P/Churyumov-Gerasimenko from a heliocentric distance of >3.6 astronomical units through perihelion passage at 1.25 astronomical units, spanning low and maximum activity levels. Initially, the solar wind permeates the thin comet atmosphere formed from sublimation, until the size and plasma pressure of the ionized atmosphere define its boundaries: A magnetosphere is born. Using the Rosetta Plasma Consortium ion composition analyzer, we trace the evolution from the first detection of water ions to when the atmosphere begins repelling the solar wind (~3.3 astronomical units), and we report the spatial structure of this early interaction. The near-comet water population comprises accelerated ions (<800 electron volts), produced upstream of Rosetta, and lower energy locally produced ions; we estimate the fluxes of both ion species and energetic neutral atoms.

An active comet emits gas and dust that expand into the surrounding space and form the comet atmosphere or coma. Its neutral component—which is mainly composed of H_2O, CO, and CO_2—is thought to result from the sublimation of near surface ice or icy grains when the comet draws closer to the Sun (*1*). Interplanetary space is permeated by the solar wind: a stream of charged particles (plasma) emitted from the Sun. The electromagnetic forces of the solar wind act on the charged particles in the environment of the comet. Charged particles are created as the atmosphere is ionized by extreme ultraviolet light as well as through collisions between the solar wind and the comet atmosphere. The partially ionized and collisional atmosphere of a well-developed comet is electrically conductive. The conductive part of the atmosphere acts as an obstacle to the solar wind. When the solar wind meets this obstacle, plasma boundaries such as a bow shock and an ionopause (contact surface) form, separating the solar wind domain from the atmospheric plasma (*2*–*4*). This resembles the interaction between the solar wind and the unmagnetized planets Mars and Venus. The Sun distance at which the comet–solar wind interaction changes to this more "active" stage depends on the comet. For comet 67P/Churyumov-Gerasimenko at the time of the Rosetta rendezvous, the nucleus outgassing rate is low (on the order of 10^{26} molecules per second). The resulting comet atmosphere is thin enough that the ion–neutral collision frequency is negligible. The solar wind should be largely unaffected by the comet atmosphere, and the plasma boundaries will not yet have formed (*5*, *6*).

Because of weak gravity, the neutral atmosphere of a comet extends well into the solar wind, beyond the plasma boundaries once formed. If a constituent of an atmosphere is ionized within the solar wind domain, the newborn ion is accelerated by the solar wind electric field: The ion is "picked up" by the solar wind. Newly picked-up ions may form ring distributions in velocity space, if the source region is large compared with the ion gyro radius. This is not the situation we expect at a comet with a low outgassing rate (*6*). The accelerated ion gains energy provided by the solar wind. If the energy taken from the solar wind becomes substantial, the solar wind is slowed down noticeably and deflected in a process known as mass loading (*7*). Mass loading is believed to be important for planets around young stars, where the solar extreme ultraviolet radiation heating the upper part of an atmosphere is generally more intense than for an older star such as the Sun (*8*), causing a higher temperature and an extended scale height of the upper atmosphere. More of the upper atmosphere may thus reach into the solar wind domain. Another possible situation is that the dynamic pressure of a stellar wind is substantially higher than the ionospheric plasma pressure of a planet (*9*), leading to a situation in which the stellar wind magnetic field permeates the entire atmosphere of the planet. A weak comet represents an example of the situation when the local plasma pressure cannot balance the solar wind dynamic pressure. The solar wind itself, as well as ion and neutral fluxes resulting from the interaction of the atmosphere and solar wind, may affect atmospheric and surface chemistry at a comet or a planet with a very thin atmosphere.

The Rosetta mission reached comet 67P/Churyumov-Gerasimenko at a heliocentric distance of ~3.6 astronomical units (AU) when the comet was still in a low-activity state. For all previous in situ spacecraft comet encounters, the targeted comet had high activity, with an enormous atmosphere extending millions of kilometers into space, and well-formed plasma boundaries with a bow shock and an ionopause. On 13 March 1986, the Giotto spacecraft flew by comet Halley at a sun distance of 0.89 AU. The relative velocity was a few tens of kilometers per second, with a closest approach distance of ~600 km. The outgassing rate of Halley at that time was estimated to be on the order of 10^{30} molecules per second. The bow shock was observed at ~1 million km from the nucleus, and the ionopause at ~5000 km (*10*, *11*). Rosetta thus enters a wholly different domain, approaching a low-activity comet as close as 10 km from the nucleus and remaining in its close vicinity at a walking pace. We study this environment using the ion composition analyzer, which is part of the Rosetta Plasma Consortium (RPC-ICA) (*12*, *13*). The RPC-ICA can distinguish between different mass groups of ions with a mass-per-charge ratio of ~1, 2, 4, 8, 16, and 32 atomic mass units (amu)/*e* and covers an energy range per charge of 10 eV/*e* to 40 keV/*e* in 12 s. The basic field of view is near 2.8π sr, with an electrostatic entrance elevation sweep providing a ±45° view away from a central 360° viewing plane divided into 16 sectors of 22.5°. Because of spacecraft blocking, the actual field of view is closer to 2π sr or half the sky. A complete angular scan is performed in 16 steps, taking 192 s to complete. Because of instrument limitations, the elevation field of view is restricted to only small elevation angles in the lowest-energy bins. Water ions are expected to be one of the most common ions in the near vicinity of the comet (*1*–*5*, *14*). The RPC-ICA cannot distinguish water ions from O^+ or H_3O^+ ions, but for convenience, we will call ions in the mass group in the vicinity of mass 16 amu/*e* "water ions" throughout this paper (the mass of water molecules is 18 amu). Similarly, for convenience we will express energy as electron volts, which is really eV/*e*, energy per charge.

[1]Swedish Institute of Space Physics, Post Office Box 812, 981 28 Kiruna, Sweden. [2]Luleå University of Technology, Department of Computer Science, Electrical and Space Engineering, Rymdcampus 1, 981 28 Kiruna, Sweden. [3]Aalto University, School of Electrical Engineering, Department of Radio Science and Engineering, Post Office Box 13000, FI-00076 Aalto, Finland. [4]Belgian Institute for Space Aeronomy, Avenue Circulaire 3, 1180 Brussels, Belgium. [5]Imperial College London, Exhibition Road, London SW7 2AZ, UK. [6]Southwest Research Institute, 6220 Culebra Road, San Antonio, TX 78238, USA. [7]Institut de Recherche en Astrophysique et Planétologie, Toulouse, France. [8]Department of Physics, University of Helsinki, Post Office Box 64, FI-00014 Helsinki, Finland. [9]Finnish Meteorological Institute, Post Office Box 503, FI-00101 Helsinki, Finland. [10]Laboratoire de Physique et Chimie de l'Environnement et de l'Espace (LPC2E), UMR 7328 CNRS–Université d'Orléans, France. [11]Swedish Institute of Space Physics, Ångström Laboratory, Lägerhyddsvägen 1, Uppsala, Sweden. [12]Technicsche Universität–Braunschweig, Institute for Geophysics and Extraterrestrial Physics, Mendelssohnstraße 3, D-38106 Braunschweig, Germany. [13]Wigner Research Centre for Physics, 1121 Konkoly Thege Street 29-33, Budapest, Hungary. [14]Space Research Institute, Austrian Academy of Sciences, Schmiedlstraße 6, 8042 Graz, Austria. [15]Rosetta Science Ground Segment, Science and Robotic Exploration (SRE-OOR), Office A006, European Space Astronomy Centre, Post Office Box 78, 28691 Villanueva de la Cañada, Madrid, Spain. [16]Physikalisches Institut, University of Bern, Sidlerstraße 5, CH-3012 Bern, Switzerland.
*Corresponding author. E-mail: hans.nilsson@irf.se

The first detectable traces of water ions were observed on 7 August 2014, at a distance of 100 km from the center of the nucleus of comet 67P/Churyumov-Gerasimenko (Fig. 1A). The Sun–comet distance was 3.6 AU. The cyclical nature of the data is due to the entrance angle sweep. The solar wind ions are only seen when the instrument probes the solar wind direction. Inspection of the full data set reveals that the ions at ~650 eV are protons (H^+), those at ~1300 eV are alpha particles (He^{2+}), and the ions in the lower-energy bins are water ions. The water ions are coming from a direction 90° from that of the solar wind direction. This is consistent with the ions moving in the direction of the solar wind electric field. The water ions and the solar wind ions are for this event seen for the same elevation angle (same time in Fig. 1A). The lowest-energy ions can only be seen for small elevation angles because of instrument limitations. The Sun is always located at low elevation angles owing to constraints on spacecraft orientation imposed by the solar panels. The water ion flux for this first event was up to 3×10^8 m^{-2} s^{-1}. With a drift velocity in the range of 10 to 20 km s^{-1} (the latter corresponding to a drift energy of ~40 eV), the water ion density is ~10^4 to 10^5 m^{-3}. The data from 7 August are typical for the early observations, with an undisturbed solar wind. The weak water ion flux of a few 10^8 ions m^{-2} s^{-1} does not yet influence the solar wind.

By 11 August 2014, the water ions were accelerated to nearly 100 eV (Fig. 1B). The fact that no or few ions were seen at the lower energies indicates that there was some distance between the spacecraft and the strongest water ion source. The energy of the ion directly corresponds to the travel distance along the solar wind electric field, which is expected to be of the order of 1 V km^{-1} or less. The ions thus appear to have traveled several tens of kilometers from a main source to the observation points.

We observed enhanced activity on 21 September 2014 (Fig. 2A). The distance to the comet center was just below 28 km, and Rosetta was over the northern hemisphere of the comet. The sun distance was 3.3 AU. The repeating pattern of ions in the 700 to 800 eV range is solar wind H^+, at 1400 to 1600 eV it is He^{2+}, whereas at 3000 eV, a weak signal from He^+ can occasionally be seen. On this day, the RPC-ICA detected solar wind fluxes about an order of magnitude higher than average, up to 10^{12} m^{-2} s^{-1} for H^+ and 10^{11} m^{-2} s^{-1} for He^{2+}.

At around 18 UT, we saw a significant flux of water ions at the lowest energy we can measure. The observed water ion flux was 10^{10} m^{-2} s^{-1}. The ions were accelerated into the instrument by a negative spacecraft potential, determined by using the Langmuir probe (RPC-LAP) instrument (*15*) to be –5 to –10 V. The geometry of the observations and flow directions of all observed populations in the detector central plane is shown in Fig. 3. The solar wind is now slightly deflected because of interaction with the comet atmosphere and has a component of flow away from the comet. H^+ is somewhat more deflected than He^{2+}. The coarse sector resolution makes a precise

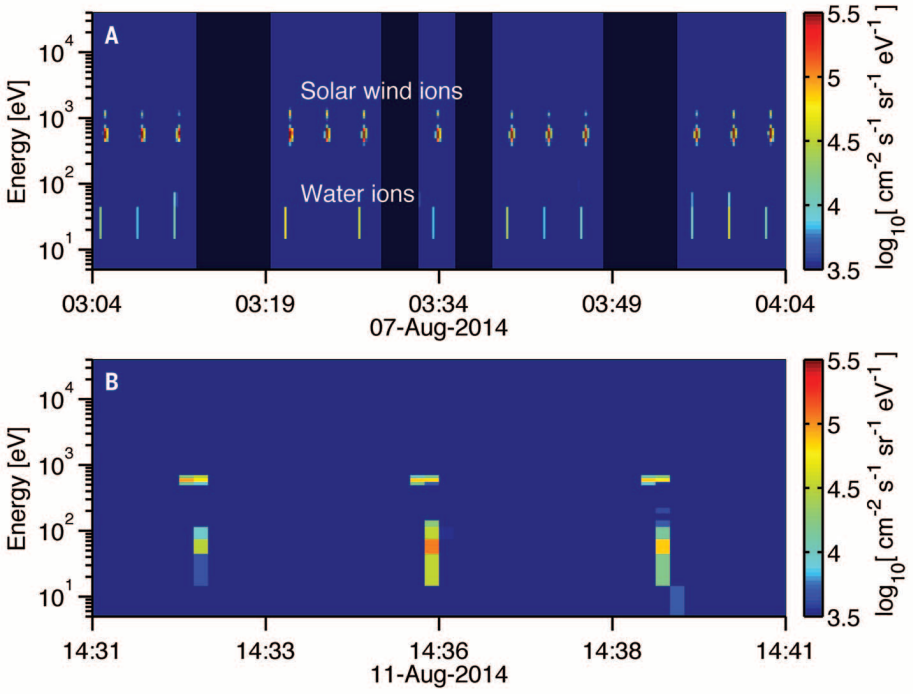


Fig. 1. Energy-time ion spectrograms. (**A**) Data obtained on 7 August 2014 between 03:04 and 04:04 UT, showing the first detection by RPC-ICA of water ions from the comet. The color scale shows differential flux summed over all azimuthal sectors and all masses. The water ions are the signal at just above 10 eV. Solar wind proton and alpha particles are visible at higher energies (650 and 1300 eV). The black patches indicate times of decoder errors. (**B**) Energy-time spectrogram from 11 August 2014 between 14:31 and 14:41 UT. The water ions are accelerated to nearly 100 eV.

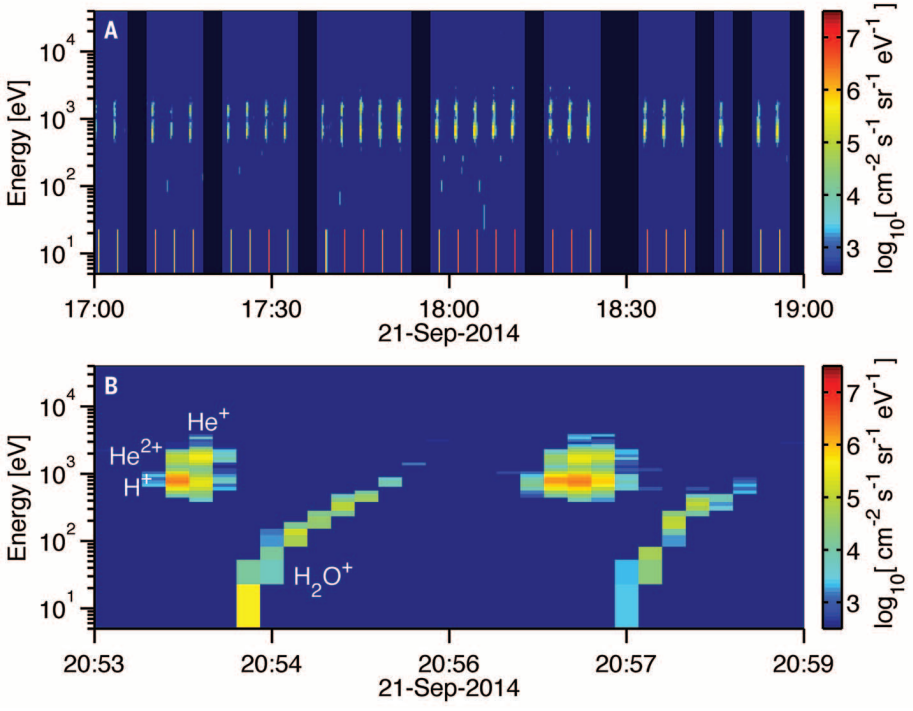


Fig. 2. Energy-time ion spectrograms. (**A**) Data from 21 September 2014 between 17:00 and 19:00 UT. The color scale is differential flux. The solar wind ions are visible in the range 700 to 800 eV (H^+) and 1400 to 1600 eV (He^{2+}). At even higher energies (3000 eV), a weak signal from He^+ is sometimes seen. In the lowest-energy bin, we see water ions. (**B**) Details of the time period 20:53 to 20:59 on 21 September. The solar wind components are clearly seen at different energies with a faint He^+ signal at higher energy. The ions seen in the energy range 100 to 1000 eV, forming curves, are comet-origin water ions with an elevation angle of 20° to 30°.

angle difficult to determine. The deflection is also seen in the higher-angular-resolution elevation scans; protons are typically deflected by ±25° in elevation angle from the sun line and He^{2+} by about half that angle. The elevation angle is the angle out of the plane depicted in Fig. 3.

We observed the low-energy water ions in the lowest energy bin only. An upper limit to the drift energy is thus 10 eV or a velocity of ~10 km s^{-1}, corresponding to a water ion density of 10^6 m^{-3}. The RPC-LAP Langmuir probes (*15*) indicate a total plasma density of 10^7 m^{-3}, which is consistent with our observations if the drift velocity is 1 km s^{-1}. The neutral density determined by using Rosetta Orbiter Spectrometer for Ion and Neutral Analysis Cometary Pressure Sensor (ROSINA COPS) (*16*) was a few times 10^{13} m^{-3}. The ionization time for water molecules at 3.3 AU was ~10^7 s. Thus, the observed ion density would be produced in 1 to 10 s over a distance of 10 km, which is the order of magnitude of the distances involved. A water ion density of 10^7 m^{-3} would correspond to about the same particle density as that of the solar wind H^+, and the mass density would be an order of magnitude higher than that of the solar wind. The kinetic energy density of water ions is still tiny compared with that of the solar wind, with a kinetic energy ratio of 10^{-4} to 10^{-3} for 1 and 10 km s^{-1} water ions, respectively.

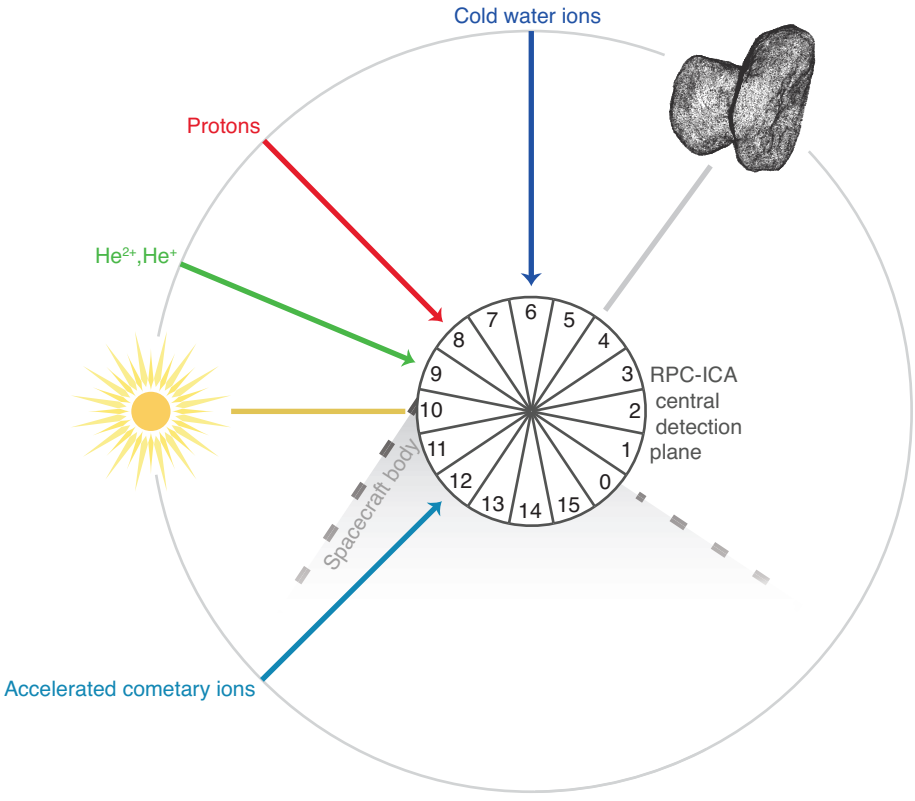


Fig. 3. The geometry of the RPC-ICA observations made on 21 September 2014. The cold ions were mainly observed between 17 and 19 UT. The accelerated ions were observed between 21:00 and 21:30 UT. Sector numbers identify the field of view of the corresponding instrument sector. [Comet shape model credit: ESA/Rosetta/NAVCAM.]

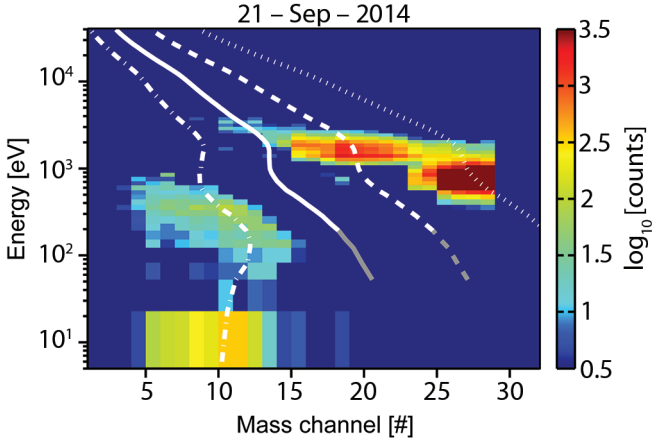


Fig. 4. Energy-mass matrix from 21 September 2014. The observed counts are integrated over a time period including the data shown in Fig. 2B. The white lines show the position of the peak of the signal for ions with mass 1 amu/*e* (dotted), 2 amu/*e* (dashed), 4 amu/*e* (solid), and 16 amu/*e* (dash-dotted), respectively. For a given energy, lighter ions are seen at higher mass channel numbers. H^+, He^{2+}, and He^+ are seen at ~750, 1500, and 3000 eV, respectively. Cometary ions are seen at 10 eV. Gray lines show extrapolation beyond 300 eV, the lowest energy available from laboratory calibration. The line at mass ~16 amu/*e* has been determined through in-flight calibration. The signal at 16 amu/*e* has an asymmetric shape, with the peak to the right and a rather sharp cut-off toward higher-mass channel numbers. This is observed both in the laboratory and in flight.

We also examined the narrower time period 20:53 to 20:59 UT on 21 September 2014 (Fig. 2B). Because of instrument angular scans, we see repetitive structures of accelerated ions forming curves from the lowest energy up to that of solar wind H^+ at 800 eV (the ions are seen in sector 12; a schematic overview is provided in Fig. 3). The ions appear to follow an angular dispersion, with higher-energy ions seen for larger elevation deflection angles (seen as a time series in the energy spectrogram). Careful analysis reveals that this is an artifact caused by an offset in our energy table. All of these ions are seen just at the limit of observable angles for that energy. The ions come from a small range of angles of 20° to 30°. The ions reach energies of 800 eV, which implies acceleration over a distance of a few 100 km, assuming that the solar wind electric field is the underlying acceleration mechanism. The integrated flux of the accelerated cometary ions is 10^9 m^{-2} s^{-1}, which is an order of magnitude less than the cold water ions.

By constructing an energy mass matrix from the time period when accelerated cometary ions were observed on 21 September 2014, we see that the different ion populations are seen where they are expected: H^+ at mass per charge 1, alpha particles at 2 amu/*e*, He^+ at 4 amu/*e*, and the accelerated ions around 16 amu/*e* (Fig. 4). The highest- and lowest-energy bins with significant fluxes of cometary ions appear also to include ions with mass of well above 16 amu/*e*. This indicates the presence of some heavier ions (CO^+ and CO_2^+) for this occasion.

The observed He^+ are created by charge exchange between He^{2+} and H_2O and allow us to make an estimate of the total H_2O atmosphere the ions have passed through of 2.5 × 10^{17} m^{-2} (*17*). This corresponds to an in situ neutral H_2O density of ~1 × 10^{13} m^{-3}, which is in good agreement with ROSINA COPS measurements (*16*). The charge exchange of H^+ into H is more important for the near-comet environment because this constitutes a much larger flux. It has commonly been observed at other solar system objects such as Mars (*18*). RPC-ICA cannot measure these energetic neutral atoms, but we can infer the approximate flux of H by comparing with the effect of the charge exchange product that we can observe, He^+. The flux ratio of He^+ to He^{2+}

for 30 September 2014, the day when we observed the clearest He^+ signal, was 2.1% (*17*). Using models of charge exchange, we then estimate that at the same point in space, 20 km from the center of the comet nucleus, we should have a H-to-H^+ ratio of 4.4%. Scaling this value with a model of the comet atmosphere density, we find that the H-to-H^+ ratio close to the nucleus reaches up to 33% at the terminator—a large fraction of the impinging solar wind will have charge-exchanged with water molecules close to the nucleus, leading to the formation of a relatively dense cloud of energetic neutral hydrogen atoms.

At 100 km from a low-activity comet nucleus, where we first detected the traces of cometary water ions, the source of the observed ions is already spatially structured. This is revealed through the absence of ion populations at the lowest energies. Thus, at 100 km from the nucleus the number of locally produced ions was below the measurement threshold of the instrument. This view is supported by time variations of the local neutral density as measured with the ROSINA instrument, which measures substantial variations related to the comet nucleus rotation period (*19*).

At a distance of 28 km from the center of the nucleus, we observed a comet plasma of similar number density to that of the solar wind. The neutral atmosphere had a density of 10^{13} m^{-3}, which is similar to the F region in Earth's ionosphere (*20*). This atmosphere is still essentially collision-less, but charge-exchange collisions are already noticeable. We are witnessing the initial stages of an ionosphere being born.

In this nascent stage, before the plasma boundaries are formed, the comet nucleus and near-comet atmosphere are affected by a relatively direct flux of solar wind. However, at this stage the solar wind is already somewhat deflected by the comet atmosphere. The deflection is ~20° from the Sun-comet line for H^+ and about half of that for He^{2+}. Simple calculations indicate that a large fraction of the solar wind impinging on the nucleus and lowest part of the atmosphere will undergo charge exchange: 30% for our example day of 30 September 2014. Energetic cometary ions were observed at 28 km from the nucleus. The continuous presence of ions throughout the energy interval from near 0 to 800 eV indicates a continuous source extending over a few 100 km. Thus, similar fluxes can be expected to hit the nucleus. According to our observations, the nucleus is consequently hit by the solar wind, by charge-exchanged H, and a flux of 10^9 m^{-2} s^{-1} of accelerated ions with a cometary origin, with an energy of up to 800 eV per charge. The neutral hydrogen flux will not be affected by the enhanced magnetic field caused by the solar wind slowing down and by associated magnetic field line draping. The important role of a heterogeneous and dynamic comet atmosphere, the deflection of the solar wind at an early stage, and the presence of accelerated ions present challenges for comet–solar wind interaction models (*6*, *21*). This energetic ion environment represents the first stage of the birth of the comet magnetosphere.

REFERENCES AND NOTES

1. K. Altwegg, H. Balsiger, J. Geiss, Composition of the volatile material in Halley's coma from in situ measurements. *Space Sci. Rev.* **90**, 3–18 (1999). doi: 10.1023/A:1005256607402
2. T. Mukai, W. Miyake, T. Terasawa, M. Kitayama, K. Hirao, Plasma observation by Suisei of solar-wind interaction with comet Halley. *Nature* **321** (6067s), 299–303 (1986). doi: 10.1038/321299a0
3. F. M. Neubauer, Giotto magnetic field results on the boundaries of the pile-up region and the magnetic cavity. *Astron. Astrophys.* **187**, 73–79 (1987).
4. K. I. Gringauz *et al.*, First in situ plasma and neutral gas measurements at comet Halley. *Nature* **321** (6067s), 282–285 (1986). doi: 10.1038/321282a0
5. A. J. Coates, Ionospheres and magnetospheres of comets. *Adv. Space Res.* **20**, 255–266 (1997). doi: 10.1016/S0273-1177(97)00543-7
6. M. Rubin *et al.*, Plasma environment of a weak comet—Predictions for Comet 67P/Churyumov–Gerasimenko from multifluid-MHD and hybrid models. *Icarus* **242**, 38–49 (2014). doi: 10.1016/j.icarus.2014.07.021
7. K. Szegö *et al.*, Physics of mass loaded plasmas. *Space Sci. Rev.* **94**, 429–671 (2000). doi: 10.1023/A:1026568530975
8. I. Ribas, E. F. Guinan, M. Gudel, M. Audard, Evolution of the solar activity over time and effects on planetary atmospheres. I. High-energy irradiances (1-1700 Å). *Astrophys. J.* **622**, 680–694 (2005). doi: 10.1086/427977
9. J. G. Luhmann, The solar wind interaction with unmagnetized planets: A tutorial. *Geophys. Monogr.* **58**, 401–411 (1991).
10. R. Reinhard, The Giotto encounter with comet Halley. *Nature* **321** (6067s), 313–318 (1986). doi: 10.1038/321313a0
11. A. Johnstone *et al.*, Ion flow at comet Halley. *Nature* **321** (6067s), 344–347 (1986). doi: 10.1038/321344a0
12. C. Carr *et al.*, RPC: The Rosetta plasma consortium. *Space Sci. Rev.* **128**, 629–647 (2007). doi: 10.1007/s11214-006-9136-4
13. H. Nilsson *et al.*, RPC-ICA: The ion composition analyzer of the Rosetta plasma consortium. *Space Sci. Rev.* **128**, 671–695 (2007). doi: 10.1007/s11214-006-9031-z
14. A. J. Coates *et al.*, Pickup water group ions at comet Grigg-Skjellerup. *Geophys. Res. Lett.* **20**, 483–486 (1993). doi: 10.1029/93GL00174
15. A. Eriksson *et al.*, RPC-LAP: The Rosetta Langmuir probe instrument. *Space Sci. Rev.* **128**, 729–744 (2007). doi: 10.1007/s11214-006-9003-3
16. H. Balsiger *et al.*, Rosina – Rosetta orbiter spectrometer for ion and neutral analysis. *Space Sci. Rev.* **128**, 745–801 (2007). doi: 10.1007/s11214-006-8335-3
17. Materials and methods, and calculation of energetic neutral atom (ENA) fluxes, are provided as supplementary materials on *Science* Online.
18. E. Kallio *et al.*, Energetic Neutral Atoms (ENA) at Mars: Properties of the hydrogen atoms produced upstream of the martian bow shock and implications for ENA sounding technique around non-magnetized planets. *Icarus* **182**, 448–463 (2006). doi: 10.1016/j.icarus.2005.12.019
19. M. Hässig *et al.*, Time variability and heterogeneity in the coma of 67P/Churyumov-Gerasimenko. *Science*.
20. M. H. Rees, *Physics and Chemistry of the Upper Atmosphere* (Cambridge Univ. Press, 1989).
21. K. C. Hansen *et al.*, The plasma environment of comet 67P/Churyumov-Gerasimenko throughout the Rosetta main mission. *Space Sci. Rev.* **128**, 133–166 (2007). doi: 10.1007/s11214-006-9142-6

ACKNOWLEDGMENTS

Rosetta is a European Space Agency (ESA) mission with contributions from its member states and the National Aeronautics and Space Administration (NASA). Rosetta's Philae lander is provided by a consortium led by Deutsches Zentrum für Luft- und Raumfahrt, Max-Planck Society, Centre National d'Études Spatiales, and Agenzia Spaziale Italiana. The work on RPC-ICA was funded by the Swedish National Space Board under contracts 108/12, 112/13, 149/12, and 135/13. Work at BIRA-IASB was supported by the Belgian Science Policy Office through the Solar-Terrestrial Centre of Excellence and by PRODEX/ROSETTA/ROSINA PEA 4000107705. The work was also supported by the UK Space Agency, the Academy of Finland through contract 251573, the Swedish Research Council through contract 621-2013-4191, NASA through the Jet Propulsion Laboratory, California Institute of Technology, the Bundesministerium für Wirtschaft und Energie and DLR under contract 50QP1001 for ROSETTA, and the Hungarian ESA PECS agreement 4200098080. We are indebted to the whole Rosetta mission team, Science Ground Segment, and Rosetta Mission Operation Control for their hard work making this mission possible. The RPC data will be made available through the Planetary Science Archive of ESA and the Planetary Data System archive of NASA and will be available upon request until that point.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/aaa0571/suppl/DC1
Materials and Methods
Supplementary Text
Fig. S1
Table S1
References (*22*–*24*)

10 October 2014; accepted 19 December 2014
10.1126/science.aaa0571

COMETARY SCIENCE

The organic-rich surface of comet 67P/Churyumov-Gerasimenko as seen by VIRTIS/Rosetta

F. Capaccioni,[1*] A. Coradini,[1†] G. Filacchione,[1] S. Erard,[2] G. Arnold,[3] P. Drossart,[2] M. C. De Sanctis,[1] D. Bockelee-Morvan,[2] M. T. Capria,[1] F. Tosi,[1] C. Leyrat,[2] B. Schmitt,[4] E. Quirico,[4] P. Cerroni,[1] V. Mennella,[5] A. Raponi,[1] M. Ciarniello,[1] T. McCord,[6] L. Moroz,[3] E. Palomba,[1] E. Ammannito,[7] M. A. Barucci,[2] G. Bellucci,[1] J. Benkhoff,[8] J. P. Bibring,[9] A. Blanco,[10] M. Blecka,[11] R. Carlson,[12] U. Carsenty,[3] L. Colangeli,[8] M. Combes,[2] M. Combi,[13] J. Crovisier,[2] T. Encrenaz,[2] C. Federico,[14] U. Fink,[15] S. Fonti,[10] W. H. Ip,[16] P. Irwin,[17] R. Jaumann,[3,18] E. Kuehrt,[3] Y. Langevin,[9] G. Magni,[1] S. Mottola,[3] V. Orofino,[10] P. Palumbo,[19] G. Piccioni,[1] U. Schade,[20] F. Taylor,[17] D. Tiphene,[2] G. P. Tozzi,[21] P. Beck,[4] N. Biver,[2] L. Bonal,[4] J.-Ph. Combe,[6] D. Despan,[2] E. Flamini,[22] S. Fornasier,[2] A. Frigeri,[1] D. Grassi,[1] M. Gudipati,[12,23] A. Longobardo,[1] K. Markus,[3] F. Merlin,[2] R. Orosei,[24] G. Rinaldi,[1] K. Stephan,[3] M. Cartacci,[1] A. Cicchetti,[1] S. Giuppi,[1] Y. Hello,[2] F. Henry,[2] S. Jacquinod,[2] R. Noschese,[1] G. Peter,[25] R. Politi,[1] J. M. Reess,[2] A. Semery[2]

The VIRTIS (Visible, Infrared and Thermal Imaging Spectrometer) instrument on board the Rosetta spacecraft has provided evidence of carbon-bearing compounds on the nucleus of the comet 67P/Churyumov-Gerasimenko. The very low reflectance of the nucleus (normal albedo of 0.060 ± 0.003 at 0.55 micrometers), the spectral slopes in visible and infrared ranges (5 to 25 and 1.5 to 5% kÅ^{-1}), and the broad absorption feature in the 2.9-to-3.6–micrometer range present across the entire illuminated surface are compatible with opaque minerals associated with nonvolatile organic macromolecular materials: a complex mixture of various types of carbon-hydrogen and/or oxygen-hydrogen chemical groups, with little contribution of nitrogen-hydrogen groups. In active areas, the changes in spectral slope and absorption feature width may suggest small amounts of water-ice. However, no ice-rich patches are observed, indicating a generally dehydrated nature for the surface currently illuminated by the Sun.

Comets are probably the least-altered objects that survive from the origin of the solar system, and they therefore carry the record of the physical processes that have led to their formation. The solid components we observe on the nuclei may have existed before the solar system formed, as interstellar grains, or they could be materials that condensed in the early protosolar nebula (*1*). Previous information on the nucleus composition was mainly derived from measurements of the coma's gas and dust components, either from ground-based observations, from in situ measurements (*2*–*4*) or from analysis of grains of cometary origin (*5*) or presumed cometary origin, as stratospheric interplanetary dust particles and Antarctic micrometeorites (*6*–*8*). Unfortunately, only a few direct observations of comet nuclei have been carried out, limiting the amount of information known about their surface composition (*9*, *10*).

With the arrival, in early August 2014, of the Rosetta spacecraft at 67P/Churyumov-Gerasimenko (hereafter 67P), the Visible, Infrared and Thermal Imaging Spectrometer (VIRTIS) (*10*) began a mapping campaign to provide direct measurements of the surface composition of the nucleus. VIRTIS is a dual channel spectrometer. VIRTIS-M (M for mapper) is a hyper-spectral imager covering a wide spectral range with two detectors: a charge-coupled device detector [visible (VIS)] ranging from 0.25 through 1.0 μm and an HgCdTe detector [infrared (IR)] covering the 1.0-to-5–μm region. VIRTIS-M uses a slit and a scan mirror to generate images with spatial resolution of 250 μrad over a field of view of 3.7°. The second channel is VIRTIS-H (H for high resolution), a point spectrometer with high spectral resolution ($\lambda/\Delta\lambda$ = 3000 at 3 μm) in the range 2 to 5 μm (*11*).

The VIRTIS observations described in this paper were obtained from August to September 2014 (3.6 to 3.3 astronomical units from the Sun), with a surface spatial resolution varying between 15 and 30 m per pixel, over the instrument spectral range 0.5 to 5.0 μm. The three major regions of the nucleus—the "head," "neck," and "body"—are presented in Fig. 1, with spectra representative of the composition of the three areas. The spectra show several common features: a very low albedo, clear absorption in the range 2.9 to 3.6 μm, and a slightly reddish spectral slope with a slope change around 1.0 μm. The reflectance spectra also show definite evidence of thermal emission at 3.5 to 5.0 μm; once the temperature is calculated from the radiance data (*12*), we obtain surface temperatures in the range 180 to 230 K during daytime. A notable feature of all spectra is the absence of water-ice absorption bands (at 1.5, 2.0, and 3.0 μm). This indicates that no water-ice–rich patches are present, across the nucleus surface, at a scale larger than ~10 m, with an upper limit on the water-ice abundance of ~1%. The lack of ice absorption features in our spectra, along with the relatively high surface temperature mentioned above, indicate that the top layers of the surface (estimated up to few hundred micrometers), which are probed by the reflected light, are composed mainly of dark dehydrated refractory materials.

The normal albedo derived from VIRTIS data is 0.060 ± 0.003 at 0.55 μm, in excellent agreement with the value obtained by the Optical, Spectroscopic, and Infrared Remote Imaging System (OSIRIS) team (*13*). The reflectance spectra in the VIS region display a positive slope ("red" slope) over the range 0.5 to 0.8 μm with a coefficient of 5 to 25% kÅ^{-1}, which is also in good agreement with the OSIRIS results. The spectrum has a knee at ~1.0 μm and displays a more neutral (1.5 to 5% kÅ^{-1}) spectral slope in the range 1.0 to 2.0 μm. If compared to the reflectance of trans-Neptunian objects, 67P would be associated, according to the present taxonomy described in (*14*), with the two slightly red classes (denoted BR and IR) being substantially different from the extreme cases of the very red objects (RR) or the neutral slopes objects (BB).

To take full advantage of the mapping capability of VIRTIS, we mapped the spectral slopes over the ranges 0.5 to 0.8 μm and 1.0 to 2.0 μm, onto the shape model (Fig. 2). The region located in the neck, which is also the region associated with the first sign of activity in August to September 2014 (*13*), displays a less steep slope than the rest of the surface in both VIS and IR ranges.

Moving into the IR region, the most prominent feature observed is a wide absorption band extending from 2.9 to 3.6 μm. The band is fairly asymmetric with a steeper slope toward the shorter-wavelength region, a band depth (relative

[1]Istituto di Astrofisica e Planetologia Spaziali, Istituto Nazionale di Astrofisica (INAF), Rome, Italy. [2]Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique, Observatoire de Paris/CNRS/Université Pierre et Marie Curie/Université Paris-Diderot, Meudon, France. [3]Institute for Planetary Research, Deutsches Zentrum für Luft- und Raumfahrt (DLR), Berlin, Germany. [4]Université Grenoble Alpes, CNRS, Institut de Planétologie et d'Astrophysique de Grenoble, Grenoble, France. [5]Osservatorio di Capodimonte, INAF, Napoli, Italy. [6]Bear Fight Institute, Winthrop, WA 98862, USA. [7]University of California, Los Angeles, Los Angeles, CA 90095, USA. [8]European Space Agency (ESA), European Space Research and Technology Centre (ESTEC), Noordwijk, Netherlands. [9]Institut d'Astrophysique Spatial, CNRS, Orsay, France. [10]Dipartimento di Matematica e Fisica "Ennio De Giorgi," Università del Salento, Italy. [11]Space Research Centre, Polish Academy of Sciences, Warsaw, Poland. [12]NASA Jet Propulsion Laboratory, Pasadena, CA 91109, USA. [13]Space Physics Research Laboratory, The University of Michigan, Ann Arbor, MI 48109, USA. [14]Università di Perugia, Perugia, Italy. [15]Lunar Planetary Laboratory, University of Arizona, Tucson, AZ 85721, USA. [16]National Central University, Taipei, Taiwan. [17]Departement of Physics, Oxford University, Oxford, UK. [18]Free University of Berlin, Institute of Geosciences, Malteserstraße 74-100, Building Haus A, 12249 Berlin, Germany. [19]Università "Parthenope," Napoli, Italy. [20]Helmholtz-Zentrum Berlin für Materialien und Energie, Berlin, Germany. [21]Osservatorio Astrofisico di Arcetri, INAF, Firenze, Italy. [22]Agenzia Spaziale Italiana, Rome, Italy. [23]Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA. [24]Istituto di Radioastronomia, INAF, Bologna, Italy. [25]Institut für Optische Sensorsysteme, DLR, Berlin, Germany.
***Corresponding author. E-mail: fabrizio.capaccioni@iaps.inaf.it**
†Deceased.

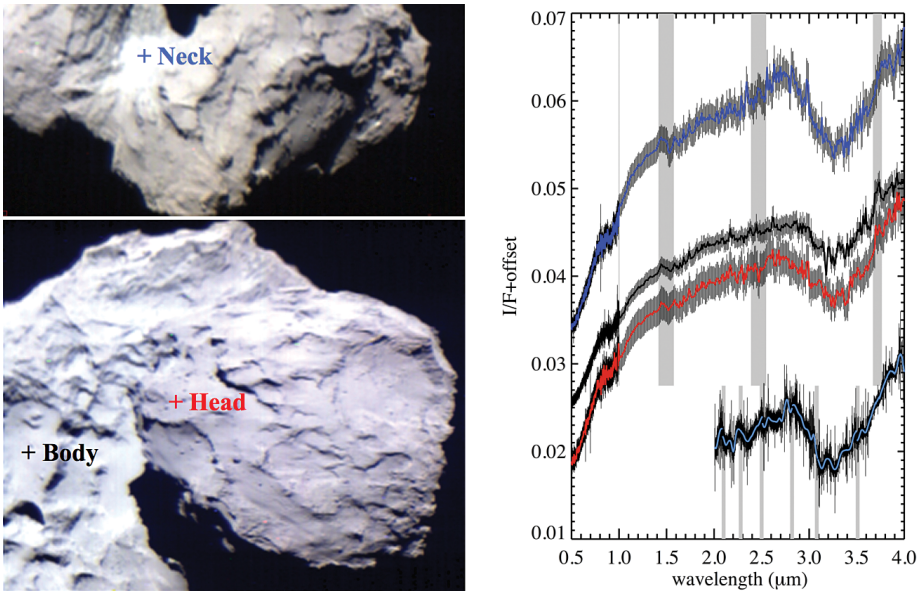


Fig. 1. Examples of images acquired by VIRTIS-M and representative spectra. (**Left**) VIRTIS-M VIS RGB images of the nucleus acquired on the neck (top panel, acquisition V1_00366693519 taken on 15 August 2014 from 03:20 to 03:54, spatial resolution 25 m per pixel) and on the head and body regions (bottom panel, acquisition V1_00368219081 taken on 1 September 2014 from 19:05 to 20:30, spatial resolution 12.5 m per pixel). (**Right**) Representative reflectance spectra (thermal emission removed) of the three regions in the 0.5-to-4.0–μm range by VIRTIS-M (red, black, and blue curves) and in the 2-to-4–μm range by VIRTIS-H (light blue curve). The color code corresponds to the location of the crosses reported in the images. The vertical scale is given in radiance factor (I/F), and offsets are added for clarity: black and blue spectra shifted by 0.01, light blue spectrum shifted by −0.02. Vertical light gray bars indicate instrumental bridging zones and order sorting filters gaps (*11*). The spectra extend to 5.0 μm but their analysis is limited to 4.0 μm, where the spectral interpretation is unaffected by the emissivity of the surface. Error bars associated with the measurements are also reported.

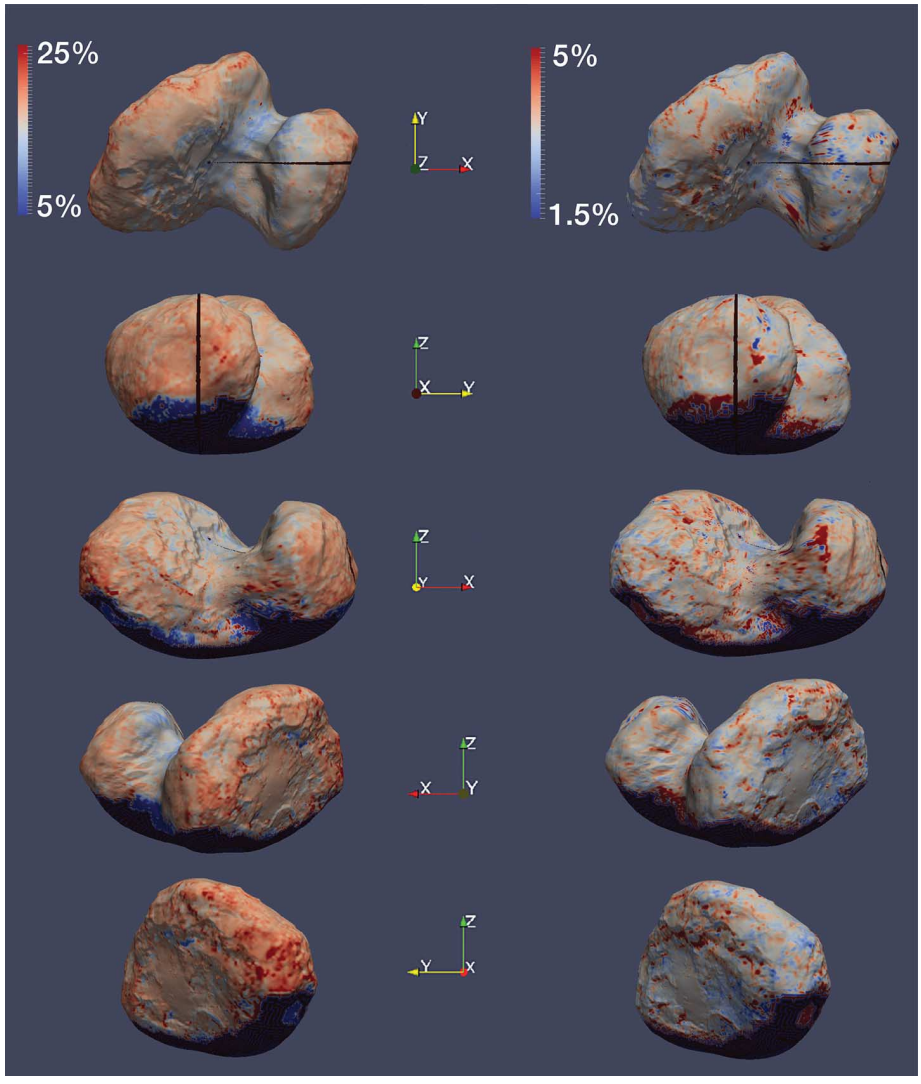


Fig. 2. Spectral slopes, sampled at 1° by 1° spatial resolution, are shown projected on the nucleus shape model for different viewing orientations. The spectral slopes in the VIS (0.5 to 0.8 μm, left column) and IR (1 to 2.0 μm, right column) are reported in units of percent per kiloangstrom (%/kÅ^{-1}). Slopes have been calculated from 160 observations acquired from 7 August 2014 to 2 September 2014, limiting the samples to the pixels having incidence and emission angles less than 80°. [Credit for nucleus shape model: ESA/Rosetta/MPS for OSIRIS Team MPS/UPD/LAM/IAA/SSO/INTA/UPM/DASP/IDA]

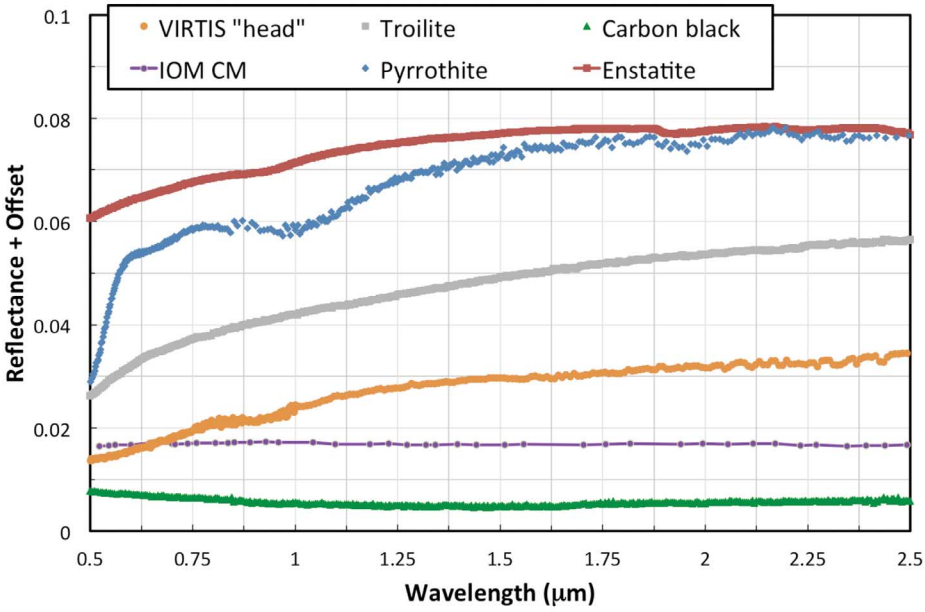


Fig. 3. The spectrum of the head shown in Fig. 1 is compared (in the spectral range 0.5 to 2.5 μm) to the spectra of several other compounds described in the text. Enstatite, pyrrothite, and troilite spectra are scaled down by 100, 75, and 50%, respectively. The Murchison IOM is from (*16*), enstatite spectrum from (*29*), troilite and carbon black spectra from (*30*), and pyrrothite spectrum from (*31*).

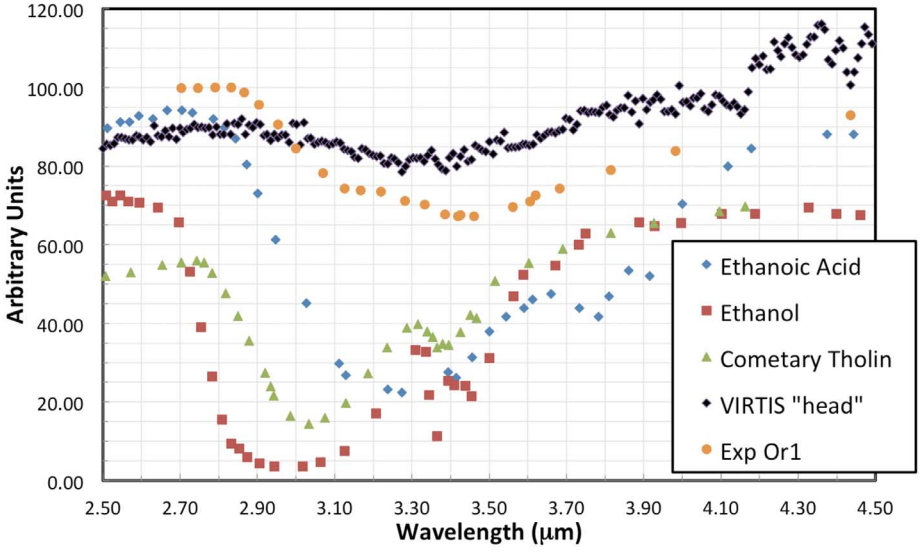


Fig. 4. The spectrum of the head in the spectral range 2.5 to 4.5 μm is compared to several other organic compounds described in the text. The VIRTIS spectrum is rescaled in arbitrary units to compare the X-H stretch region with ethanol and ethanoic (acetic) acid spectra (*32*), a cometary tholins (obtained after ion irradiation of a mixture of 80% H_2O, 16% CH_3OH, 3.2% CO_2, and 0.8% C_2H_6) (*33*), and a refractory residue (labeled "Exp Or1") obtained after UV irradiation of a mixture of H_2O:CH_3OH:NH_3:CO:CO_2 in the ratio 2:1:1:1:1 (*34*).

to the local continuum) on the order of 20%, and a band center located at 3.2 to 3.3 μm. The band is observed in both VIRTIS-M and VIRTIS-H spectra, independently calibrated, with similar band depth and shape.

When this band is mapped over the surface of 67P, the data taken so far indicate a limited variability across the observable nucleus, mainly localized in the neck region. The absorption band of spectra taken in the neck region is deeper, broader, and shifted toward shorter wavelengths (2.8 to 3.6 μm) with respect to the other regions. The short-wavelength shoulder of the neck band is most likely due to the addition, in this region, of a small percentage of water-ice mixed with the dark material responsible for the overall appearance of the spectrum. This is in agreement with the slight increase in the reflectance (on the order of a few percent), the increase in band depth (Fig. 1), and the previously mentioned decrease of the VIS spectral slope for this same region (Fig. 2, left column top projection). Furthermore, it is consistent with localized activity driven by volatiles (including water) in the uppermost few millimeters of the nucleus surface. This process is quite distinct from the diffuse activity observed by OSIRIS (*13*) in regions where VIRTIS observations indicate a dehydrated crust. We note that an increase of the amount of organic matter may also contribute to part of the band-depth increase in the neck. In contrast to the neck region, the spectral features observed on the head and body (Fig. 1) are markedly similar and indicate a compositionally homogeneous surface at the present spatial resolution.

Comparison with laboratory spectra of carbonaceous chondrites of the CI, CM, and CR types reveals that none of the typical features of their spectra are compatible with the spectra of 67P [for a thorough analysis of carbonaceous chondrites spectra, see (*15*–*17*)]. Spectra of bulk carbonaceous chondrites display a change of slope at ~0.6 μm and are basically flat in the near-IR range and thus do not match the red slope seen on 67P. Furthermore, carbonaceous chondrites' broad 3-μm band due to structural hydroxyl or water is shifted substantially toward shorter wavelengths. Thus, the material that composes the surface of 67P is not identical to the main groups of primitive carbonaceous chondrites. The reflectance spectra of the insoluble organic matter (IOM) extracted from the Murchison chondrite are indeed dark and flat in both the visible and IR ranges (*15*), but the absolute abundance of IOM in primitive chondrites is limited to a few percent and cannot account for the overall appearance of the 67P spectra.

Numerous asteroidal bodies have an absorption feature around 2.7 to 2.8 μm (*18*) associated with the presence of hydrated minerals. In the case of small bodies such as 24 Themis, 65 Cybele, and a few other objects (*18*–*20*), the absorption-band maxima are shifted toward higher wavelength (~3.1 μm). For these observations, the absorption feature is explained by the presence of water-ice, with a possible contribution of organic compounds (~3.4 μm). Still, the band shape and position of the maximum of absorption measured by VIRTIS for 67P is substantially different from these objects and, at present, distinct across the observed small bodies.

The very low reflectance of 67P surface material throughout VIRTIS spectral ranges requires the presence of a darkening agent in the refractory component. Fe-bearing opaque minerals are known to absorb visible and near-IR light, and their presence in cometary grains is established from the analysis of grains from the comet 81P/Wild2 collected by the STARDUST mission, stratospheric dust particles, and Antarctic micrometeorites (*6*–*8*). IOM also contains some insoluble opaque minerals (e.g., sulfides, iron oxides) that are not dissolved by the chemical procedure of extraction from the bulk chondrites. Sulfides (troilite, pyrrhotite, pentlandite), Fe-Ni alloys, Fe-bearing crystalline, and amorphous silicates have been identified in those dust grains. In particular, the

analysis of the cometary dust samples indicates the presence of amorphous silicates associated with Fe-Ni inclusions and of crystalline silicates, mainly Mg-rich olivines and low-Ca pyroxenes (Fo_{51-100} and En_{58-100}). Similar compositions (Fo_{48-100} and En_{65-99}) have been measured for Antarctic micrometeorites of presumed cometary origin (*7*, *8*). Iron sulfides (e.g., troilite, pyrrhotite) are common in meteorites, interplanetary dust particles, and cometary dust; they are highly absorbing opaque phases and may play an important role in darkening and reducing spectral contrasts of diagnostic absorption bands if these minerals are fine-grained [(*21*) and references therein]. In addition to darkening the IR range, sulfides display a reflectance spectra slope over the 0.5-to-1.0–μm range that may approximate the slope seen in VIRTIS spectra (Fig. 3).

In many previous studies, the reddish surfaces of outer solar system bodies, such as Kuiper belt objects, have been modeled using tholins produced from N_2:CH_4 gas mixtures (*22*). Although the reflectance spectra of those tholins may qualitatively account for the reddish slope of VIRTIS data, they are definitely too bright in the IR to be considered as viable materials present at the surface of 67P.

The broad absorption band observed at ~3.3 μm in all VIRTIS spectra is a congested spectral region that results from the contribution of X-H vibrations in different chemical groups. CH and OH groups can be reasonably suspected (*23*). On the other hand, a major contribution of NH or NH_2 is less likely, as nitrogen-rich carbonaceous materials devoid of oxygen are not representative of the bulk macromolecular fraction of Jupiter family comets (JFCs) (*24*). The absorption bands of water-ice, structural water, and hydroxyl trapped in silicates and iron oxides or oxy-hydroxides occur at shorter wavelengths compared with VIRTIS spectra, and their contribution should be low, if any, except in the neck active area where slightly wider and stronger bands are observed. The substructure in the observed absorption band points to a complex mixture of both aromatic and aliphatic C-H bonds (*25*) (but their presence cannot yet be firmly established); in addition, the broad feature may also be consistent with the presence of OH groups, either as carboxylic groups or alcoholic groups, inserted in a macromolecular organic solid (*23*). These chemical groups are known to account for the faint and broad feature in the IR spectra of IOM extracted from primitive meteorites (*26*, *27*). The position and shape of the band may depend on the alcohol/carboxylic group ratio and on the groups onto which they are branched. They may account for some of the observed variability of the band shape (Fig. 4).

VIRTIS clearly observed a comet that is different from the other JFCs encountered so far. Previous spectroscopic observations of the surfaces of JFCs indicated the presence of water-ice (*10*) and tentatively identified a weak band near 2.39 μm that is possibly compatible with hydrocarbons (*9*). Mass spectrometry of solid grains ejected from comet nuclei showed various species of organic materials ranging from aliphatic and aromatic compounds to carboxylic acids and alcohols (*28*), but these compounds have not been observed in absorption on comet nuclei. From this point of view, 67P represents a different species in the cometary zoo.

The compositional homogeneity of the surface observed by VIRTIS, despite the rejuvenating processes acting on the nucleus (at least on the active portion of it) at every passage close to the Sun, indicates that space weathering plays only a minor role in determining the observed composition. The refractory compounds so widespread on the surface of 67P are then representative of the bulk pristine material of the nucleus. Organic compounds are formed upon irradiation [either by ultraviolet (UV) photons or energetic particles] of ices or by the polymerization of mixtures of ices kept at low temperatures (*21*). The VIRTIS observations indicate that high-volatility ices (CH_4, CO, CO_2, CH_3OH, etc.), along with water, must have been readily available at the time of the formation of the solid materials that ended up forming the nucleus of 67P in the early phases of the protosolar nebula. This suggests that the larger abundance of organics on the surface of 67P, with respect to other JFCs, could be correlated with a formation scenario in which most of the material that constitutes the nucleus of 67P is produced in a low-temperature environment consistent with large distances from the Sun, such as the Kuiper belt region.

REFERENCES AND NOTES

1. M. J. Mumma, S. J. Charnley, The chemical composition of comets—Emerging taxonomies and natal heritage. *Annu. Rev. Astron. Astrophys.* **49**, 471–524 (2011). doi: 10.1146/annurev-astro-081309-130811
2. M. C. Festou, H. U. Keller, H. A. Weaver, in *Comets II*, M. C. Festou, H. U. Keller, H. A. Weaver, Eds. (Univ. of Arizona Press, Tucson, AZ, 2004), chap. 1.
3. S. A. Sandford *et al*., Organics captured from comet 81P/Wild 2 by the Stardust spacecraft. *Science* **314**, 1720–1724 (2006). doi: 10.1126/science.1135841; pmid: 17170291
4. M. Combes *et al*., Infrared souding of comet Halley from Vega 1. *Nature* **321**, 266–268 (1986). doi: 10.1038/321266a0
5. D. Brownlee *et al*., Comet 81P/Wild 2 under a microscope. *Science* **314**, 1711–1716 (2006). doi: 10.1126/science.1135840; pmid: 17170289
6. F. J. M. Rietmeijer, in *Planetary Materials*, J. J. Papike Ed., (Mineralogical Society of America, Chantilly, VA, 1998), chap. 2.
7. E. Dobrică *et al*., Connection between micrometeorites and Wild 2 particles from Antarctic snow to cometary ices. *Meteorit. Planet. Sci.* **44**, 1643–1661 (2009). doi: 10.1111/j.1945-5100.2009.tb01196.x
8. M. E. Zolensky *et al*., Mineralogy and petrology of comet 81P/Wild 2 nucleus samples. *Science* **314**, 1735–1739 (2006). doi: 10.1126/science.1135842; pmid: 17170295
9. L. A. Soderblom *et al*., Short-wavelength infrared (1.3–2.6 μm) observations of the nucleus of comet 19P/Borrelly. *Icarus* **167**, 100–112 (2004). doi: 10.1016/j.icarus.2003.08.019
10. J. M. Sunshine *et al*., Exposed water ice deposits on the surface of comet 9P/Tempel 1. *Science* **311**, 1453–1455 (2006). doi: 10.1126/science.1123632; pmid: 16456037
11. A. Coradini *et al*., VIRTIS: An imaging spectrometer for the Rosetta mission. *Space Sci. Rev.* **128**, 529–559 (2007). doi: 10.1007/s11214-006-9127-5
12. F. Tosi *et al*., Thermal measurements of dark and bright surface features on Vesta as derived from Dawn/VIR. *Icarus* **240**, 36–57 (2014). doi: 10.1016/j.icarus.2014.03.017
13. H. Sierks *et al*., On the nucleus structure and activity of comet 67P/Churiumov-Gerasimenko. *Science* **347**, aaa1044 (2015). doi: 10.1126/science.aaa1044
14. M. A. Barucci, I. N. Belskaya, M. Fulchignoni, M. Birlan, Taxonomy of Centaurs and Trans-Neptunian objects. *Astron. J.* **130**, 1291–1298 (2005). doi: 10.1086/431957
15. E. A. Cloutis, T. Hiroi, M. J. Gaffey, C. M. O'D. Alexander, P. Mann, Spectral reflectance properties of carbonaceous chondrites: 1. CI chondrites. *Icarus* **212**, 180–209 (2011). doi: 10.1016/j.icarus.2010.12.009
16. E. A. Cloutis, P. Hudon, T. Hiroi, M. J. Gaffey, P. Mann, Spectral reflectance properties of carbonaceous chondrites: 2. CM chondrites. *Icarus* **216**, 309–346 (2011). doi: 10.1016/j.icarus.2011.09.009
17. E. A. Cloutis, P. Hudon, T. Hiroi, M. J. Gaffey, Spectral reflectance properties of carbonaceous chondrites: 3. CR chondrites. *Icarus* **217**, 389–407 (2012). doi: 10.1016/j.icarus.2011.11.004
18. D. Takir, J. P. Emery, Outer main belt asteroids: Identification and distribution of four 3-μm spectral groups. *Icarus* **219**, 641–654 (2012). doi: 10.1016/j.icarus.2012.02.022
19. A. S. Rivkin, J. P. Emery, Detection of ice and organics on an asteroidal surface. *Nature* **464**, 1322–1323 (2010). doi: 10.1038/nature09028; pmid: 20428165
20. H. Campins *et al*., Water ice and organics on the surface of the asteroid 24 Themis. *Nature* **464**, 1320–1321 (2010). doi: 10.1038/nature09029; pmid: 20428164
21. C. De Bergh, B. Schmitt, L. V. Moroz, E. Quirico, D. P. Cruikshank, "Laboratory data on ices, refractory carbonaceous materials, and minerals relevant to transneptunian objects and Centaurs," in *The Solar System Beyond Neptune*, A. Barucci, H. Boehnhardt, D. Cruikshank, A. Morbidelli, Eds. (Univ. of Arizona Press, Tucson, AZ, 2008), pp. 483–506.
22. C. Dalle Ore *et al*., Composition of KBO (50000) Quaoar. *Astron. Astrophys.* **501**, 349–357 (2009). doi: 10.1051/0004-6361/200911752
23. Y. J. Pendleton, L. J. Allamandola, The organic refractory material in the diffuse interstellar medium: Mid-infrared spectroscopic constraints. *Astrophys. J. Suppl. Ser.* **138**, 75–98 (2002). doi: 10.1086/322999
24. E. Dartois *et al*., Ultracarbonaceous Antarctic micrometeorites, probing the solar system beyond the nitrogen snow line. *Icarus* **224**, 243–252 (2013). doi: 10.1016/j.icarus.2013.03.002
25. L. V. Moroz, G. Arnold, A. V. Korochantsev, R. Wäsch, Natural solid bitumens as possible analogs for cometary and asteroid organics: Reflectance spectra of pure bitumens. *Icarus* **134**, 253–268 (1998). doi: 10.1006/icar.1998.5955
26. Y. Kebukawa, S. Nakashima, M. E. Zolensky, Kinetics of organic matter degradation in the Murchison meteorite for the evaluation of parent-body temperature history. *Meteorit. Planet. Sci.* **45**, 99–113 (2010).
27. F. R. Orthous-Daunay *et al*., Mid-infrared study of the molecular structure variability of insoluble organic matter from primitive chondrites. *Icarus* **223**, 534–543 (2013). doi: 10.1016/j.icarus.2013.01.003
28. J. Kissel, F. R. Krueger, J. Silén, B. C. Clark, The Cometary and Interstellar Dust Analyzer at comet 81P/Wild 2. *Science* **304**, 1774–1776 (2004). doi: 10.1126/science.1098836;pmid: 15205526
29. R. L. Klima, C. M. Pieters, M. D. Dyar, Spectroscopy of synthetic Mg-Fe pyroxenes I: Spin-allowed and spin-forbidden crystal field bands in the visible and near-infrared. *Meteorit. Planet. Sci.* **42**, 235–253 (2007). doi: 10.1111/j.1945-5100.2007.tb00230.x
30. The HoserLab Database is available at http://psf.uwinnipeg.ca/Home.html.
31. The U.S. Geological Survey SpecLab Database is available at http://speclab.cr.usgs.gov.
32. The Spectral Database System for Organic Compounds at the National Institute of Materials and Chemical Research in Japan is available at http://sdbs.db.aist.go.jp/sdbs/cgi-bin/cre_index.cgi.
33. G. D. McDonald *et al*., Production and chemical analysis of cometary ice tholins. *Icarus* **122**, 107–117 (1996). doi: 10.1006/icar.1996.0112
34. G. M. Muñoz Caro, W. A. Schutte, UV-photoprocessing of interstellar ice analogs: New infrared spectroscopic results. *Astron. Astrophys.* **412**, 121–132 (2003). doi: 10.1051/0004-6361:20031408

ACKNOWLEDGMENTS

We thank the following institutions and agencies for support of this work: Italian Space Agency (ASI, Italy), Centre National d'Etudes Spatiales (CNES, France), DLR (Germany), NASA (USA) Rosetta Program, and Science and Technology Facilities Council (UK). VIRTIS was built by a consortium, which includes Italy, France, and Germany, under the scientific responsibility of the Istituto di Astrofisica e Planetologia Spaziali of INAF, Italy, which also guides the scientific operations. The VIRTIS instrument development has been funded and managed by ASI, with contributions from Observatoire de Meudon financed by CNES, and from DLR. We thank the Rosetta Science Ground Segment and the Rosetta Mission Operations Centre for their support throughout the early phases of the mission. The VIRTIS calibrated data will be available through the ESA's Planetary Science Archive Website (www.rssd.esa.int/index.php?project=PSA&page=index) and is available upon request until posted to the archive. The VIRTIS Team wishes to dedicate this paper to the memory of Angioletta Coradini, conceiver of the instrument, our leader, and friend.

10 October 2014; accepted 19 December 2014
10.1126/science.aaa0628

COMETARY SCIENCE

Subsurface properties and early activity of comet 67P/Churyumov-Gerasimenko

Samuel Gulkis,[1]‡ Mark Allen,[1] Paul von Allmen,[1] Gerard Beaudin,[2] Nicolas Biver,[3] Dominique Bockelée-Morvan,[3] Mathieu Choukroun,[1] Jacques Crovisier,[3] Björn J. R. Davidsson,[4] Pierre Encrenaz,[2]* Therese Encrenaz,[3]* Margaret Frerking,[1] Paul Hartogh,[5] Mark Hofstadter,[1] Wing-Huen Ip,[6] Michael Janssen,[1] Christopher Jarchow,[5] Stephen Keihm,[1] Seungwon Lee,[1] Emmanuel Lellouch,[3] Cedric Leyrat,[3] Ladislav Rezac,[5] F. Peter Schloerb,[7,1] Thomas Spilker[1]†

Heat transport and ice sublimation in comets are interrelated processes reflecting properties acquired at the time of formation and during subsequent evolution. The Microwave Instrument on the Rosetta Orbiter (MIRO) acquired maps of the subsurface temperature of comet 67P/Churyumov-Gerasimenko, at 1.6 mm and 0.5 mm wavelengths, and spectra of water vapor. The total H_2O production rate varied from 0.3 kg s^{-1} in early June 2014 to 1.2 kg s^{-1} in late August and showed periodic variations related to nucleus rotation and shape. Water outgassing was localized to the "neck" region of the comet. Subsurface temperatures showed seasonal and diurnal variations, which indicated that the submillimeter radiation originated at depths comparable to the diurnal thermal skin depth. A low thermal inertia (~10 to 50 J K^{-1} m^{-2} $s^{-0.5}$), consistent with a thermally insulating powdered surface, is inferred.

The Rosetta Orbiter spacecraft carries 12 science instruments to study comet 67P/Churyumov-Gerasimenko with a long-range goal of better understanding the origin of our solar system. Comet 67P/Churyumov-Gerasimenko is a typical Jupiter family comet (JFC). JFCs are believed to have been injected from the Kuiper belt outside the orbit of Neptune because of gravitational interactions with the giant planets or collisions. Currently, comet 67P is in a low-inclination orbit with a semimajor axis of 3.46 AU and an orbital period of 6.45 years. Its present orbital configuration was achieved in February 1959 during a close encounter with Jupiter that converted its perihelion distance from 2.7 AU to 1.28 AU (*1*). Its prior history is uncertain. According to some statistical studies, comet 67P might have been subject to mass erosion driven by thermal sublimation with depletion of the surface materials up to a depth of a few kilometers (*2*). As a consequence, a dust mantle would gradually build up and cover the nucleus surface, where ice mixtures of H_2O, CO_2, and CO would be found (*3*).

The instrument suite on Rosetta was chosen to study physical and chemical properties of the nucleus, the evolution of the outgassing from the nucleus, and development of the coma as a function of the heliocentric distance, as well as the interaction of the solar wind with the target comet. The development of the gas and dust comas along the cometary orbit depends critically on the energy transport through the dust mantle and the differential outgassing of different species according to their volatility. The Microwave Instrument on the Rosetta Orbiter (MIRO) (*4*) has been designed to study heat transport in the nucleus, outgassing from the nucleus, and development of the coma as interrelated processes. The measurements and monitoring of the interplay between gas emissions and surface temperature evolution by MIRO therefore bring key information to the temporal evolution of the subsurface ice-dust complex. This work reports on early measurements of comet 67P with the MIRO instrument from June to September 2014.

The MIRO instrument consists of a 30-cm diameter, offset parabolic reflector telescope followed by two heterodyne receivers. Center-band operating frequencies of the receivers are near 190 GHz (1.6 mm) and 562 GHz (0.5 mm). Throughout this Report, we refer to the two bands as the mm (millimeter) and smm (submillimeter) bands, respectively. Broadband continuum channels are implemented in both the mm and smm bands for measurements of near-surface temperatures and temperature gradients in the nucleus of comet 67P. The smm band also includes a very high resolution spectrometer (44 kHz channel width) for measuring spectral lines of $H_2{}^{16}O$, $H_2{}^{18}O$, and other molecules in the coma. Here, we report the first measurements of water $H_2{}^{16}O$ and $H_2{}^{18}O$ in the coma and subsurface thermal measurements of the nucleus obtained during the time interval from June to September 2014. We refer to the June-July observations as pre-rendezvous observations, characterized here as the time when the nucleus was unresolved by the MIRO smm beam and the entire coma could be observed at once. The August-September period is identified as the postrendezvous time interval, during which both the nucleus and coma were observed with spatial resolutions that varied from 500 to 40 m.

Observations of comet 67P with MIRO began on 24 May 2014. The instrument mode was dual-continuum and spectroscopic. Initial observations

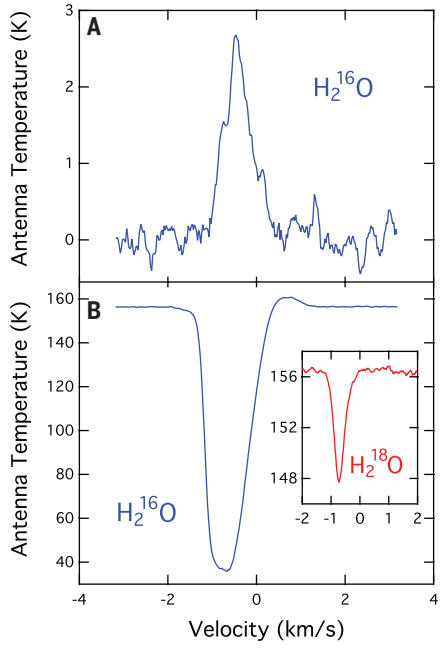


Fig. 1. Sample spectra of the 1(1,0) – 1(0,1) transitions of $H_2{}^{16}O$ and $H_2{}^{18}O$ lines at 556.939 GHz and 547.676 GHz, respectively, obtained with MIRO in nadir-viewing geometry. All spectra seen to date are biased toward negative gas velocities (toward the spacecraft), which is interpreted as preferential outgassing from the day side of the nucleus. (**A**) Line observed in emission on 23 June 2014 UT when the Rosetta distance to the comet is Δ = 128,000 km, the heliocentric distance is r_H = 3.84 AU, the spatial resolution is Ω = 260 km, and the phase (Sun-comet observer) angle is Φ = 32°. (**B**) (Inset plot) Lines observed in absorption against the nucleus subsurface thermal emission on 19 August 2014 UT (Δ = 81 km, r_H = 3.5 AU, Φ = 37°, Ω = 160 m). The line intensity is expressed in the main-beam TB scale. The velocity scale is given relative to the nucleus velocity. The spectral resolution is 24 m s^{-1}. The $H_2{}^{16}O$ spectrum of 23 June is consistent with an expansion velocity of 0.68 ± 0.02 km s^{-1}. The optical thickness of the $H_2{}^{16}O$ line observed on 23 June is 0.3. On 19 August, the $H_2{}^{16}O$ line is optically thick near the line center (opacity > 10), whereas the $H_2{}^{18}O$ line is optically thin (opacity < 0.1).

[1]Jet Propulsion Laboratory/California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA. [2]LERMA, Observatoire de Paris, PSL Research University, UPMC Université Paris 06, CNRS, UMR8112, F-75014 Paris, France. [3]LESIA-Observatoire de Paris, CNRS, UPMC, Université Paris–Diderot, 5 place Jules Janssen, 92195 Meudon, France. [4]Department of Physics and Astronomy, Uppsala University, Box 516, SE-75120 Uppsala, Sweden. [5]Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany. [6]National Central University, Jhongli, Taoyuan City 32001, Taiwan. [7]University of Massachusetts, 619 Lederle Graduate Research Tower, Amherst, MA 01003, USA.
*Distinguished Visiting Scientist at the Jet Propulsion Laboratory
†Retired. ‡Corresponding author. E-mail: samuel.gulkis@jpl.nasa.gov

were carried out with the instrument pointed in the nadir direction of the comet and off of the comet to provide background reference.

In the June-July 2014 time interval, the Rosetta spacecraft approached the comet from 550,000 km to 973 km. The phase angle of the observations during this period ranged from 35° to 1°, and the heliocentric distance changed from 3.95 to 3.62 AU. The MIRO smm single-pixel, one-dimensional (1D) field of view ranged from 2400 to 4.2 km, larger than the nucleus radius of 3.8 km. It was possible to measure and monitor the total water outgassing rate of comet 67P without spatial mapping during this time interval.

Water vapor was first detected with MIRO on 6 June 2014, when the comet was at 3.92 AU from the Sun. A weak emission line from the 1(1,0) – 1(0,1) rotational transition of $H_2{}^{16}O$ at the frequency of 556.936 GHz was detected, which then rapidly increased as the distance of Rosetta to the comet decreased (Fig. 1A). A large blue shift (frequency increase) of the spectral line in the comet rest velocity frame (about –0.4 km s^{-1} in the June period) was observed, indicative of water outgassing, essentially toward the hemisphere facing Rosetta and the Sun. This suggests large temperature contrasts between daytime and nighttime, with temperatures on the night side too low for rapid sublimation of water ice. The water production rate displayed periodic variations on short time scales, typically by 50% in magnitude with two maxima and minima per period (Fig. 2), which are correlated with the 12.4-hour rotation of the nucleus (*5*). This variation is related to the complex shape of the nucleus, for which the total area exposed to the Sun varies during the rotation (*5*). Both the line Doppler shift and time variations suggest that the nucleus regions exposed to solar illumination were the main source of water vapor in the June-July period. However, surface inhomogeneity in ice content and material properties may also play a significant role. Indeed, local jet activity was observed by the OSIRIS camera (*5*), specifically in the so-called "neck" region of 67P, where the spectral and reflectance properties indicate the presence of water ice (*5*, *6*). Simple models considering either a homogeneous distribution of water ice on the surface or one concentrated in the neck region (Fig. 2) can explain the observed time variation in a first approximation (*7*).

As mentioned earlier, in the June-July time period, the nucleus was unresolved and it was possible to sample the inner coma in a single-pixel field of view. The spectra observed during this time period show blue-shifted water lines (Fig. 1) in emission, which are indicative of water outgassing mainly into the hemisphere facing the sun. Therefore, for purposes of calculating the mean total water production rate over one nucleus rotation using this data set, we used the approximation of hemispheric outgassing. We inferred from radiative transfer calculations (*8*, *9*) that the mean total water production rate over one nucleus rotation was 1×10^{25} molecules/s (0.3 kg s^{-1}) in early June and reached 2×10^{25} molecules/s (0.6 kg s^{-1}) at the end of July 2014. For more localized production patterns, higher values are derived as a result of higher opacities. We estimate that the above values might underestimate the actual production rate by a factor of two, at most. Based on the rotational variations in Fig. 2, we estimate the outgassing variability to be on the order of ±20%. We expect to be able to refine this estimate, as we are able to better define the parameters that enter in the radiative transfer calculations, i.e., the temperature and velocity profiles in the coma and the global extent of the outgassing.

Rosetta was at a distance of 100 km from the comet on August 6 and continued to approach rapidly. The nucleus was well resolved at that time (and throughout August and September), with a spatial resolution of <200 m in the smm channel. When viewing the resolved nucleus through the coma, the continuum radiation (from the nucleus) was observed to be in absorption at the $H_2{}^{16}O$ and $H_2{}^{18}O$ spectral line frequencies present in the coma. Absorption of the nucleus thermal emission in the coma provides clear evidence that the excitation temperatures of the absorbing water molecules are lower than the nucleus brightness temperature (TB) (*9*). Starting in early August, the spectral line due to $H_2{}^{18}O$ provided a more sensitive measure of the water production rate than the $H_2{}^{16}O$ line, which became optically thick (Fig. 1B). Therefore, we focus here on the $H_2{}^{18}O$ observations to derive quantitative information on water production for this time period. The frequency of the peak intensity of the water lines is diagnostic of the gas expansion velocity, whereas the width contains information on the radial profiles of the gas velocity and temperature.

The water column density along the MIRO line of sight exhibited considerable variability, as it varied by a factor of 5 or more over a time scale of 10 hours during the August 7.4 to 9.5 UTC period (Fig. 3). Considering both nadir and limb observations acquired from August 6 to September 1, we estimate the mean water production for this period (heliocentric distance ranged from 3.6 AU to 3.46 AU) to be about 4×10^{25} molecules/s. In this time frame, the MIRO footprint unevenly covered nucleus latitudes from 0° to 80° N and longitudes from 0° to 360° (Fig. 3 and figs. S1 and S2). However, maximum column densities are generally observed above regions of cometocentric coordinates at latitudes of 65° to 70° and longitudes of 0° to 100°, i.e., close to the North pole and to the neck, specifically at the place where the nucleus presents its deepest concavity (the neck is the region connecting the small head and large body of the nucleus in reference to the ducklike shape of 67P). This is in line with water outgassing mostly from the neck region and responsible for the prominent dust jet seen in OSIRIS images. Water production from icy grains has been observed in a number of comets (*10*), and therefore, sublimating icy grains from this dust jet could also be significant contributors to the observed water production. To our surprise, the highest water column densities are often observed above shadowed neck regions where the nucleus thermal continuum—and, consequently, the surface and subsurface

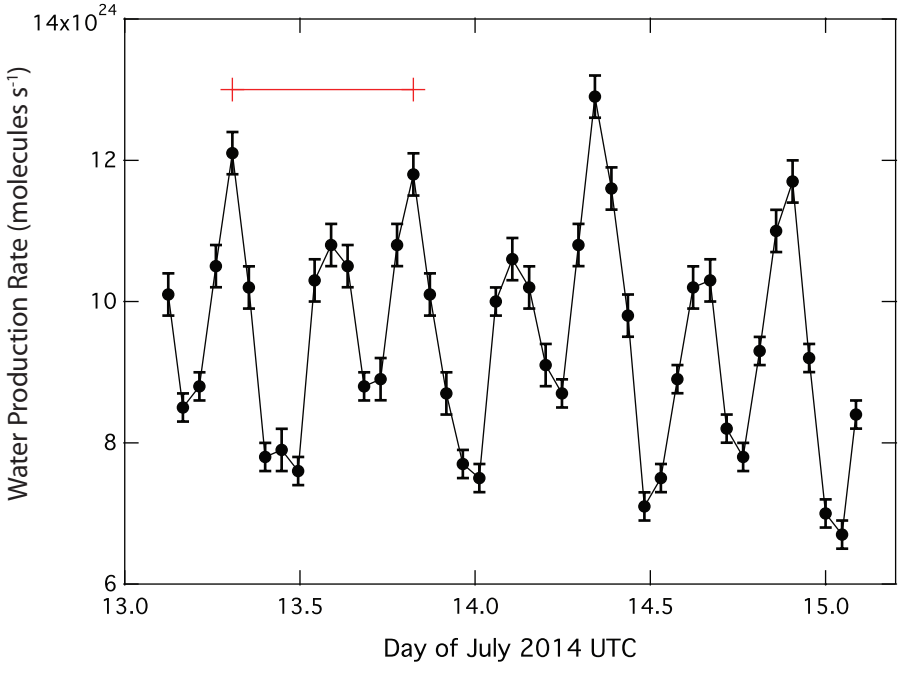


Fig. 2. Variations of the total water production rate as the nucleus rotates. The rates shown above are based on MIRO $H_2{}^{16}O$ observations between 13 and 15 July 2014. Based on coarse maps, the pointing direction of the MIRO field of view was 300 arc sec southwest from the nucleus center at the time of the observations. In one nucleus rotation period (12.4 hours) (*5*), the water production rate exhibits two minima and two maxima. The rotation period of 67P (12.4 hours) is shown with a red bar. Early models to explain the variations are found in (*7*).

temperatures—is low (Fig. 3 and fig. S2). It is possible that observed gases are originating from nearby insolated regions. Inversely, low column densities are measured above some illuminated regions (on the body and on the head), again supporting an activity driven by more than solar insolation and confirming the inhomogeneity in ice distribution observed by other Rosetta instruments (*5*, *6*).

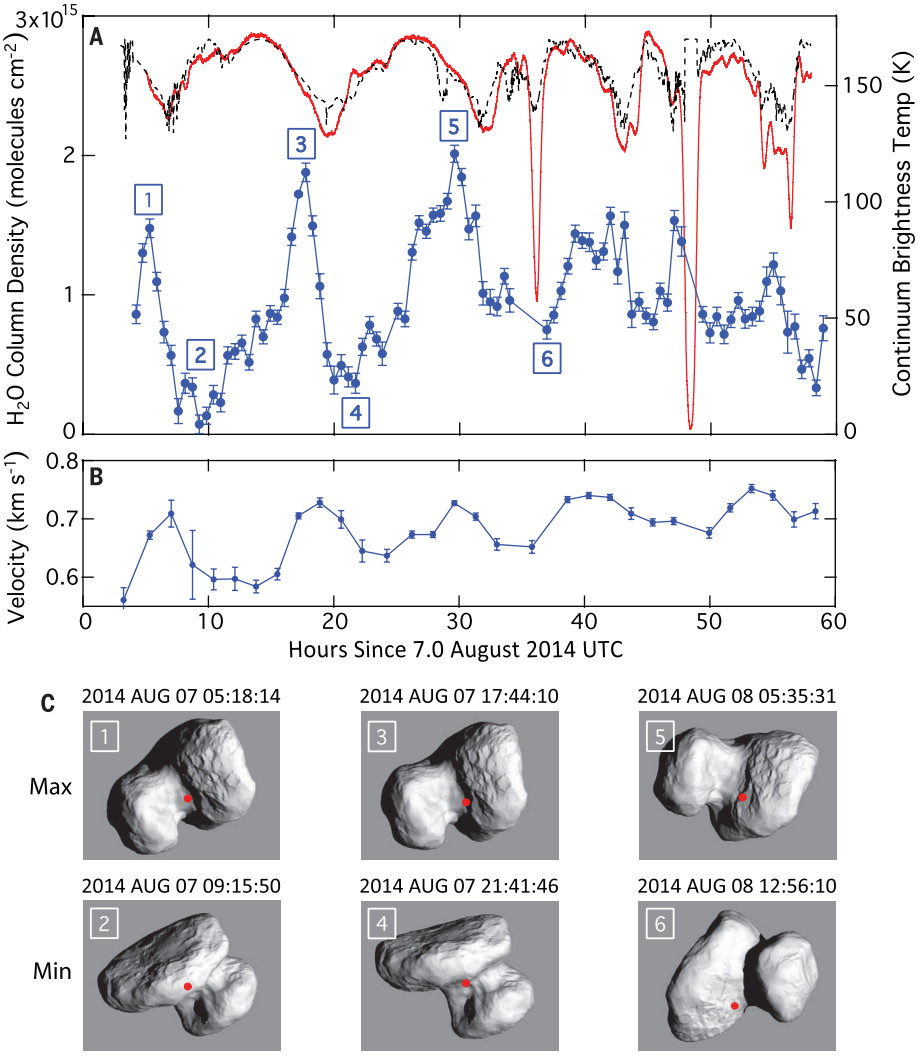


Fig. 3. Water column density and expansion velocity measured by MIRO from 7.4 to 9.5 August 2014 UTC by using nadir $H_2{}^{18}O$ spectra. Significant spatial variability is seen, with maxima correlated with lines of sight near the neck region of the nucleus. (**A**) Column density (blue dots with error bars), nucleus thermal continuum temperatures (red curve), and cosine of the incidence angle in the center of MIRO beam (dotted black curve). Large drops in the smm continuum (red) temperature are observed when the MIRO line of sight crosses the limb. (**B**) Mean water expansion velocity given by the velocity at maximum absorption. (**C**) (Bottom) Illumination maps of the nucleus at the specified times The SHAP2 digital shape model provided by the OSIRIS team was used. The submillimeter MIRO beam position is shown by the red dot on the illumination maps. The *x*,*y* plane is perpendicular to the MIRO line of sight and contains the comet-Sun line.

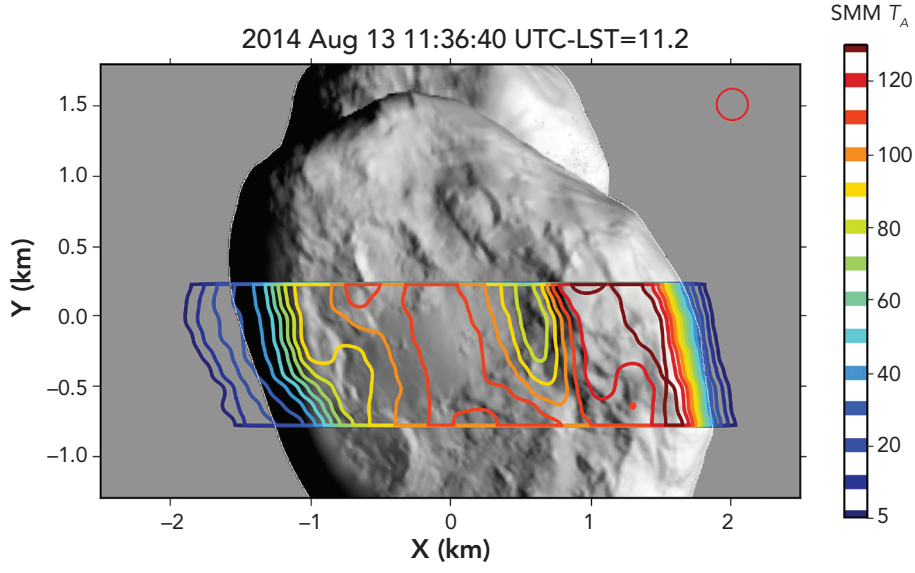


Fig. 4. Illumination map of the digital shape model SHAP2 with contour plot of the MIRO submillimeter antenna temperature. MIRO sees subsurface brightness features correlated with surface slopes, with regions facing the Sun generally brighter (warmer) than those tilted away from the Sun. This is an indication of rapid response to solar illumination and, hence, a low thermal inertia. Excellent overlap of the MIRO and shape model limbs is obtained on the Sun-lit (right) side of the nucleus. The left side of the nucleus is in polar night at this time, and the shape model—based on visible imaging—cannot determine its true shape here. The extension of the MIRO data beyond the limb on the night side indicates the true size of the nucleus in this region. The shape model is presented as an illumination map with the intensity given by the cosine of the solar illumination angle projected onto a plane perpendicular to the line joining the spacecraft to the barycenter of the comet.

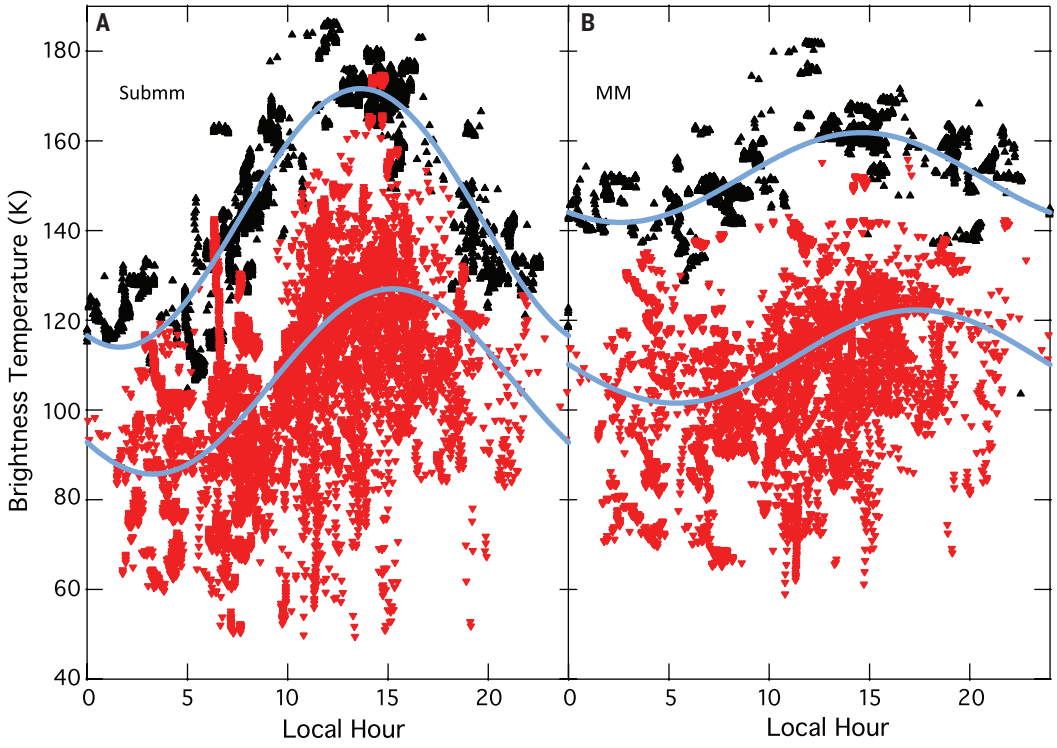


Fig. 5. Brightness temperatures as a function of local solar time, from MIRO continuum measurements of September 2014, are shown for effective latitude bins of 20° to 30° N (black data points) and 20° to 30° S (red data points). The solid curves are diurnal sinusoidal fits to the data. Both effective latitudes and the local solar time are computed from the shape model–derived surface orientation at the MIRO beam center. Only points for which the MIRO beams lie entirely within the nucleus are included. The data are restricted to the 100° to 200° longitude band in order to eliminate data in the neck region, where extreme shadowing conditions obscure the interpretation of the diurnal heating curve. (**A**) Submillimeter data. (**B**) Millimeter data.

Line profiles provide information on the kinematics of the gas flow. Note that the line-peak velocity (−0.6 to −0.75 km s^{-1}) is observed to be higher along lines of sight where higher column densities (Fig. 3) are recorded, whereas the line width decreases. Therefore, projection effects of the gas flow are minimized for these lines of sight, and the line-peak velocity should be close to the expansion velocity. These findings also demonstrate that the observed column-density enhancements are related to larger sublimation fluxes from the surface and not to local density enhancements caused by shocks between flows that would decelerate the gas.

Continuum observations from the MIRO 0.5- and 1.6-mm bands were geo-referenced to the SHAP2 digital shape model provided by the OSIRIS team. The map of the antenna temperature (Fig. 4) was derived by scanning the MIRO beam across the nucleus and then combining the scans into a map using the "on-the-fly" technique developed for mapping with single-dish radio telescopes (*11*). The antenna temperature contours correlate well with illumination patterns on the nucleus surface, which are, in turn, determined by topography and orientation with respect to the Sun.

Subsurface temperatures, which determine the MIRO smm and mm thermal fluxes, may have three components. (i) The diurnal component—variation with local solar time—is confined to a shallow layer defined by the thermal 1/*e* penetration depth—on the order of 1 to 2 cm for 67P's 12.4-hour rotation period and plausible thermal properties of a particulate regolith. (ii) An annual signal, dependent on obliquity, orbital period, and perihelion location, with latitude-dependent amplitude and a 1/*e* attenuation depth of the order 1 m. The seasonal variations in temperatures determine the MIRO emissions from penetration depths below the diurnal layer. (iii) The "static" or "subseasonal" component is the approximately isothermal layer beneath the seasonally varying layer, and its temperature depends on the latitude-dependent orbit variation of the solar flux.

Plotting the MIRO mm and smm brightness temperatures as a function of the local solar time and latitude (Fig. 5) reveals three key features of the nucleus thermal environment. (i) Both the smm and mm channels reveal sensitivity to the diurnal heating cycle, peaking in early to midafternoon, although with amplitudes diminished by a factor of approximately 2 to 4 from the smm to mm channel. (ii) The observed 40- to 50K smm day-night peak-to-peak variations represent 35% or more of the predicted surface temperature variations for a completely insulating surface (zero thermal inertia) at the September Sun distance. For realistic nonzero thermal inertia values, the ratio of the smm TB/surface temperature peak-to-peak values will be higher. This result implies that the smm 1/*e* penetration depth lies at or above the 1/*e* thermal attenuation depth for diurnal variations. (iii) Seasonal effects are apparent in the reduction of the diurnally averaged TB values from the 20°N to 30°N latitude bin to the 20°S to 30°S latitude bin. We conclude that, on a global scale, the MIRO smm emissions originate primarily at depths comparable to the diurnal layer thickness, whereas the mm emissions originate at depths roughly three times the thermal attenuation depth, consistent with electrical absorption properties, which scale inversely with wavelength. The latitude dependence of the measured brightness temperatures is consistent with expectations from seasonal variations of the solar flux given 67P's orbital parameters.

Temperature depth profiles were computed by solving the 1D heat equation for a range of values of the thermal inertia, and the resulting antenna temperatures in the mm and smm bands were calculated from evaluating the transfer of the thermal radiation emitted in the subsurface region. Comparing these theoretical antenna temperatures with MIRO observations leads to an estimated thermal inertia in the range of ~10 to 50 J K^{-1} m^{-2} $s^{-0.5}$ (*12*).

With this range of values, large (up to ~50 K) differences between the MIRO measured dayside subsurface temperatures and the temperature at the surface can be inferred, highlighting the key insulating role played by the dusty surface: It largely restricts the amount of heat transferred to the interior, including ices within. This certainly contributes greatly to the longevity of 67P and, probably, of comets in general. The importance of measuring the temperatures below the surface of a comet—and particularly below its diurnal layer—is illustrated by these data.

MIRO's spectroscopic and continuum observations provide interrelated data on heat transport, sublimation of ices, and outgassing by providing data on the nucleus subsurface temperature, location, time, and fluxes of water being emitted from the nucleus surface. Less than 1% of the 67P nucleus surface is needed to explain the water gas production rates discussed earlier if water ice were located on the surface (*13*). A substantial portion of the measured outgassing originates from the neck of the nucleus.

REFERENCES AND NOTES

1. A. Carusi, L. Kresak, E. Perozzi, G. B. Valsecchi, *Long-Term Evolution of Short-Period Comets* (Adam Hilger, Ltd., Accord, MA, 1985).
2. O. Groussin, G. Hahn, P. L. Lamy, R. Gonczi, G. B. Valsecchi, The long-term evolution and initial size of comets 46P/ Wirtanen and 67P/Churyumov–Gerasimenko. *Mon. Not. R. Astron. Soc.* **376**, 1399–1406 (2007). doi: 10.1111/ j.1365-2966.2007.11553.x
3. H. L. F. Houpis, D. A. Mendis, W.-H. Ip, The chemical differentiation of the cometary nucleus: The process and its consequences. *Astrophys. J.* **295**, 654–667 (1985). doi: 10.1086/163409
4. S. Gulkis *et al*., MIRO: Microwave instrument for the Rosetta orbiter. *Space Sci. Rev.* **128**, 561–597 (2007). doi: 10.1007/ s11214-006-9032-y
5. H. Sierks *et al*., On the nucleus structure and activity of comet 67P/Churyumov-Gerasimenko. *Science* **347**, aaa1044 (2015).
6. F. Capaccioni *et al*., The organic-rich surface of comet 67P/ Churyumov-Gerasimenko as seen by VIRTIS/Rosetta. *Science* **347**, aaa0628 (2015).
7. See supplementary materials on *Science Online*, Models used for the water production observed on 13 to 15 July 2014.
8. V. Zakharov, D. Bockelée-Morvan, N. Biver, J. Crovisier, A. Lecacheux, Radiative transfer simulation of water rotational excitation in comets: Comparison of the Monte Carlo and escape probability methods. *Astron. Astrophys.* **473**, 303–310 (2007). doi: 10.1051/0004-6361:20066715
9. S. Lee, P. von Allmen, L. Kamp, S. Gulkis, B. Davidsson, Non-LTE radiative transfer for submillimeter water lines in Comet 67P/Churyumov-Gerasimenko. *Icarus* **215**, 721–731 (2011). doi: 10.1016/j.icarus.2011.07.007
10. K. J. Meech *et al*., EPOXI: Comet 103P/Hartley 2 observations from a worldwide campaign. *Astrophys. J.* **734**, L1–L9 (2011). doi: 10.1088/2041-8205/734/1/L1
11. J. G. Mangum, D. T. Emerson, E. W. Greisen, The On The Fly imaging technique. *Astron. Astrophys.* **474**, 679–687 (2007). doi: 10.1051/0004-6361:20077811
12. See supplementary materials on *Science Online*, Method for determining the thermal inertia.
13. See supplementary materials on *Science Online*, Ice coverage and content on 67P.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance and support of the entire MIRO project teams at the Jet Propulsion Laboratory, the Max-Planck-Institut für Sonnensystemforschung, LESIA-Observatoire de Paris, and LERMA-Observatoire de Paris–Meudon. We thank Y. Anderson, T. Koch, R. Nowicki, and L. Pan for their efforts in scheduling, operations, and support of the MIRO instrument. The authors recognize here the enormous contributions that our friend and colleague, Lucas Kamp, made to the MIRO science and engineering efforts before his recent death. The authors acknowledge support from their institutions and funding sources. A part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. A part of the research was carried out at the Max-Planck-Institut für Sonnensystemforschung with financial support from Deutsches Zentrum für Luft- und Raumfahrt and Max-Planck-Gesellschaft. Parts of the research were carried out by LESIA and LERMA, Observatoire de Paris, with financial support from Centre National d'Études Spatiales (CNES) and CNRS/Institut National des Sciences de l'Univers (INSU); at the National Central University with funding from the Taiwanese National Science Counsel grant 101-2111-M-008-016; at the University of Massachusetts, Amherst, USA; and at Uppsala University with financial support from the Swedish National Space Board. We acknowledge personnel at ESA's European Space Operations Center (ESOC) in Darmstadt, Germany, and at ESA and NASA/JPL tracking stations for their professional work in communication with and directing the Rosetta spacecraft, thereby making this mission possible. We acknowledge the excellent support provided by the Rosetta teams at the ESOC in Germany and the European Space Astronomy Center (ESAC) in Spain. Special mention is made of C. Vallat, whose superb efforts have made many difficult MIRO observing sequences possible; B. Grieger, whose calculations and visualizations have enabled and optimized our science return; and N. Altobelli and M. Küppers, whose early liaison support and later oversight responsibilities have greatly benefited the mission as a whole. The authors thank Holger Sierks and the OSIRIS team for permission to use the SHAP2 shape model for analysis purposes and for permission to publish several images here. All MIRO data are released through the PSA archive of ESA and the PDS archive of NASA. The data reported here are scheduled for delivery to the archive on 19 May 2015 and are available on request until then.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/aaa0709/suppl/DC1
Supplementary Text
Figs. S1 to S4
References (*14*–*17*)

12 October 2014; accepted 27 December 2014
10.1126/science.aaa0709

COMETARY SCIENCE

67P/Churyumov-Gerasimenko, a Jupiter family comet with a high D/H ratio

K. Altwegg,[1]* H. Balsiger,[1] A. Bar-Nun,[2] J. J. Berthelier,[3] A. Bieler,[1,4] P. Bochsler,[1] C. Briois,[5] U. Calmonte,[1] M. Combi,[4] J. De Keyser,[6] P. Eberhardt,[1]† B. Fiethe,[7] S. Fuselier,[8] S. Gasc,[1] T. I. Gombosi,[4] K.C. Hansen,[4] M. Hässig,[1,8] A. Jäckel,[1] E. Kopp,[1] A. Korth,[9] L. LeRoy,[1] U. Mall,[9] B. Marty,[10] O. Mousis,[11] E. Neefs,[6] T. Owen,[12] H. Rème,[13,14] M. Rubin,[1] T. Sémon,[1] C.-Y. Tzou,[1] H. Waite,[8] P. Wurz[1]

The provenance of water and organic compounds on Earth and other terrestrial planets has been discussed for a long time without reaching a consensus. One of the best means to distinguish between different scenarios is by determining the deuterium-to-hydrogen (D/H) ratios in the reservoirs for comets and Earth's oceans. Here, we report the direct in situ measurement of the D/H ratio in the Jupiter family comet 67P/Churyumov-Gerasimenko by the ROSINA mass spectrometer aboard the European Space Agency's Rosetta spacecraft, which is found to be (5.3 ± 0.7) × 10^{-4}—that is, approximately three times the terrestrial value. Previous cometary measurements and our new finding suggest a wide range of D/H ratios in the water within Jupiter family objects and preclude the idea that this reservoir is solely composed of Earth ocean–like water.

The delivery of water and organic compounds to Earth and other terrestrial planets is still under debate (*1*–*4*). Existing scenarios range from negligible (*1*, *2*) to substantial (*3*, *4*) cometary contributions to terrestrial water. Hence, the comparison of the deuterium-to-hydrogen ratio (D/H) in water between the different populations of comets and Earth's oceans is crucial if one wants to distinguish among these scenarios. Previous D/H measurements have been made for a dozen comets from the Oort cloud and the Jupiter family [for example, (*5*) and references therein]. So far, only one measurement has been made in situ: in the coma of the Oort-cloud comet 1P/Halley, via the mass spectrometers present aboard the European Space Agency (ESA) Giotto spacecraft, and based on an assumption made on the oxygen isotopic composition (*6*, *7*). Here, we report the direct in situ measurement of the D/H ratio in the Jupiter family comet (JFC) 67P/Churyumov-Gerasimenko.

The mass spectrometer ROSINA-DFMS (Rosetta Orbiter Sensor for Ion and Neutral Analysis, Double Focusing Mass Spectrometer) on the European cometary space mission Rosetta is designed to measure isotopic ratios (*8*). Its mass resolution and high dynamic range enable it to detect very rare species such as $HD^{18}O$ relative to the most abundant isotope $H_2{}^{16}O$ (*9*). ROSINA has the capability to measure all isotopic ratios in water independently (D/H, $^{17}O/^{16}O$, and $^{18}O/^{16}O$), and the D/H ratio can be deduced from two different species, namely $HD^{16}O/H_2{}^{16}O$ and $HD^{18}O/H_2{}^{18}O$.

Rosetta has a neutral gaseous background due to spacecraft outgassing. The permanent particle density in the close vicinity of the spacecraft far away from the comet is ~10^6 cm^{-3}, consisting mostly of water but also of organic material, fragments of hydrazine, and vacuum grease (fluorine). Even after 10 years in space and after hibernation, the background from Rosetta can be measured and characterized with ROSINA (*10*). The D/H ratio in water outgassed from the Rosetta spacecraft is compatible with the terrestrial value of 1.5 × 10^{-4}, as expected (*9*), and did not vary with time of degassing, indicating negligible isotope fractionation. These observations demonstrate the capability of DFMS and of the analysis method. On 1 August 2014, Rosetta was within 1000 km from comet 67P/Churyumov-Gerasimenko at 3.6 astronomical units (AU) from the Sun, and the coma was still hidden beneath the spacecraft background for in situ measurements. However, a few days later it approached the comet to within 100 km. The factor of 100 larger densities at this distance surpassed the spacecraft background by more than a factor of 2; the difference between background (26 May) and coma (22 August) at mass 19 dalton is evident (Fig. 1). There are four peaks on mass 19 dalton: fluorine, which is due to background (vacuum grease) from the spacecraft (*10*); $H^{18}O$, a fragment from $H_2{}^{18}O$ due to electron impact ionization in the instrument and a minor contribution from photodissociation of water in the coma; $H_2{}^{17}O$; and $HD^{16}O$. All mass peaks have the same shape. $H_2{}^{17}O$ is hidden in the shoulder of

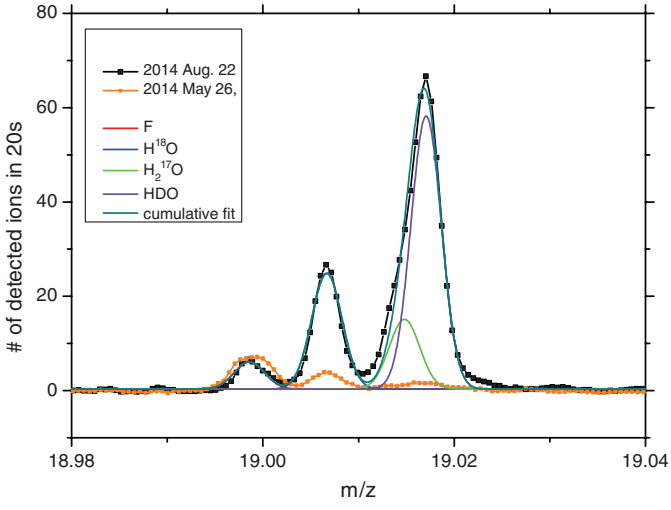


Fig. 1. Typical mass/charge 19 daltons/*e* spectra from DFMS data are shown for 26 May 2014, when the spacecraft was at 800,000 km from the nucleus, and for 22 August, with the spacecraft at 60 km from the nucleus. The integration time was 20 s. The number of ions detected are plotted as a function of *m*/*z*. The peak fits for 26 August are also shown for fluorine (F), $H^{18}O$, $H_2{}^{17}O$, and HDO, using the same peak width for all species.

[1]Physikalisches Institut, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland. [2]Department of Geosciences, Tel-Aviv University, Ramat-Aviv, Tel-Aviv, Israel. [3]Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS), 4 Avenue de Neptune, F-94100 Saint-Maur, France. [4]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, 2455 Hayward, Ann Arbor, MI 48109, USA. [5]Laboratoire de Physique et Chimie de l'Environnement et de l'Espace (LPC2E), UMR 6115 CNRS–Université d'Orléans, France. [6]Space Physics Division, Belgisch Instituut voor Ruimte-Aëronomie (BIRA)–Institut d'Aéronomie Spatiale de Belgique (IASB), Ringlaan 3, B-1180 Brussels, Belgium. [7]Institute of Computer and Network Engineering (IDA), Technicsche Universität Braunschweig, Hans-Sommer-Straße 66, D-38106 Braunschweig, Germany. [8]Department of Space Science, Southwest Research Institute, 6220 Culebra Road, San Antonio, TX 78228, USA. [9]Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany. [10]Centre de Recherches Pétrographiques et Géochimiques, Centre de Recherches Pétrographiques et Géochimiques (CRPG)–CNRS, Université de Lorraine, 15 rue Notre Dame des Pauvres, BP 20, 54501 Vandoeuvre lès Nancy, France. [11]Université de Franche-Comté, Institut Univers, Transport, Interfaces, Nanostructures, Atmosphère et Environnement, Molécules (UTINAM), CNRS/Institut National des Sciences de l'Univers (INSU), UMR 6213 Besançon Cedex, France. [12]Institute for Astronomy, University of Hawaii, Honolulu, HI 96822, USA. [13]Université de Toulouse, Université Paul Sabatier (UPS)–Observatoire Midi-Pyrénées (OMP), Institut de Recherche en Astrophysique et Planétologie (IRAP), Toulouse, France. [14]CNRS, IRAP, 9 Avenue du Colonel Roche, BP 44346, F-31028 Toulouse cedex 4, France.
*Corresponding author. E-mail: altwegg@space.unibe.ch
†Deceased.

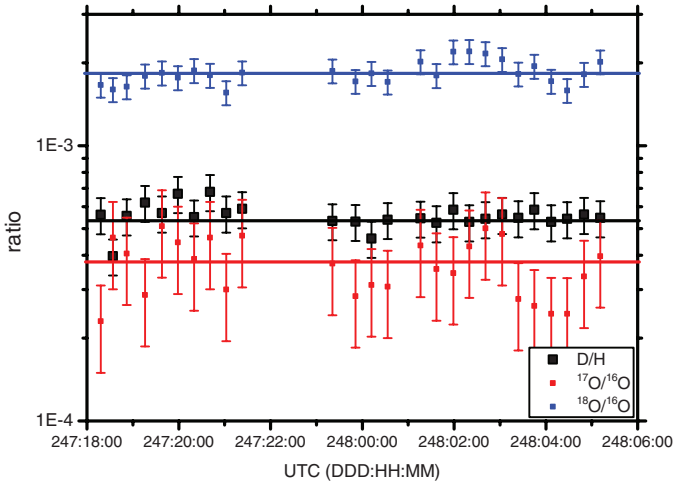


Fig. 2. Ratios of D/H, $^{17}O/^{16}O$, and $^{18}O/^{16}O$ for 4 and 5 September. The period of 12 hours corresponds almost to a comet rotation (12.4-hour period). The times given are those for the measurement of *m/z* 18 ($H_2{}^{16}O$). *m/z* 19 (HDO and $H_2{}^{17}O$) and 20 ($H_2{}^{18}O$) are measured 30 s and 1 min later, respectively. The error bars represent statistical errors from the low count rates and errors from the fit. Errors due to background and due to uncertainties in the detector gain, which are of a systematic nature, are only considered for the mean ratios given in the text.

the much larger peak of HDO. Because the positions of the masses are known, the only fit parameters remaining are the amplitudes. In this way, $H_2{}^{17}O$ can clearly be separated from $HD^{16}O$.

Before 5 August, all four species had a very similar intensity. However, on 8 August the intensity of $HD^{16}O$ was more than double the height of the ^{18}OH peak, whereas fluorine stayed constant. This can only be attributed to a much higher D/H ratio in water. By 22 August, the background was almost negligible because the spacecraft was now within 50 km from the nucleus. However, the signal on mass 21 ($HD^{18}O$) was still very low. Water peaks on mass 18 and 19 dalton were therefore analyzed by using more than 50 spectra taken between 8 August and 5 September, leading to a derivation of D/H from $HD^{16}O/H_2{}^{16}O$ [analysis is described in (*9*)]. Uncertainties of the measurements are carefully estimated by error propagation, taking into account statistical uncertainties in the measured signal and uncertainties originating from calibration, background subtraction, and fitting methods. The biggest contribution is probably the uncertainty for the background because Rosetta is now permanently in the cometary coma, and background corrections have to be done with data from before August 2014. Analysis of the spacecraft background over the 10 years of the cruise phase, however, has shown that for a stable spacecraft attitude, the background remains stable over very long times outside of reaction wheel offloadings (*10*). The value derived from our analysis for D/H is $(5.3 \pm 0.7) \times 10^{-4}$ (2σ error, where σ is the SD as described above).

By deriving HDO relative to $H_2{}^{16}O$, we also found the $^{17}O/^{16}O$ ratio. Additionally, the ratio of $^{18}O/^{16}O$ follows from the analysis of mass-to-charge ratio (*m/z*) = 20 daltons, which contains two well-separated peaks: $H_2{}^{18}O$ (20.0148 daltons) and HF (20.0062 daltons) (Fig. 2). HF is almost entirely due to spacecraft background. The results for the oxygen isotopic ratios in cometary water are compatible with solar system values, with $^{17}O/^{16}O = (3.7 \pm 0.9) \times 10^{-4}$ and $^{18}O/^{16}O = (1.8 \pm 0.2) \times 10^{-3}$. Although from the figure one might see a small modulation for the ratios as a function of rotation of the comet, statistics are too poor to come to a conclusion. Once the comet activity increases, it should be feasible to narrow down the values for the heavy isotopes of water.

The D/H ratio shows dramatic variations among solar system reservoirs of water (Fig. 3). The protosolar nebula (PSN) D/H value is estimated to be $(2.1 \pm 0.5) \times 10^{-5}$ based on measurements of H_2 in the atmosphere of Jupiter (*11*) and (^{3}He+D)/H in the solar photosphere (*12*). This value is close to interstellar D/H ratios of H_2 around 2.0×10^{-5} to 2.3×10^{-5} (*13*). In contrast, most solar system objects are enriched in deuterium (Fig. 3), with an enrichment factor *f* (defined as the ratio $[D/H]_{object}/[D/H]_{PSN}$) averaging 6 for the inner solar system (including Earth, the Moon, and volatile-rich primitive meteorites such as carbonaceous chondrites). Comets analyzed so far, mostly long-period ones, display higher *f* values, typically in the 10 to 20 range. The cause of the deuterium enrichment in solar system bodies is usually attributed to water-ice rich in deuterium infalling from the presolar cloud onto the nebula disk (*5*). Thereafter, because comets may have accreted ice with various chemical histories (*14*), several mechanisms have been proposed that would induce a deuterium fractionation in the early PSN. Because a part of the ice accreted by comets could have vaporized and recondensed within the PSN, an isotopic exchange could have occurred between the initially deuterium-rich water and molecular hydrogen in the warm regions of the disk (*15*). At low temperatures, this reaction favors the concentration of deuterium in HDO, but the extremely slow kinetics tend to inhibit the reaction. In these models, isotopic exchange occurs as long as H_2O does not crystallize, implying that the observed D/H ratios should be representative of the local values where and when the building blocks of the host objects condensed (*16*). Alternatively, a part of the ice accreted by comets could have remained pristine (*14*). Under these circumstances, gas-grain reactions could have induced deuterium fractionation in the cold outer part of the PSN (*17*, *18*). Regardless of the fractionation mechanism, all of these models are consistent with a deuterium enrichment profile following a radial increase throughout the PSN from low values close to the Sun to high values in the outer part of the disk (*16*–*18*).

Most D/H ratio measurements in water in comets come from long-period comets, presumably originating from the Oort cloud [Oort Cloud Comets (OCCs)]. Population of this cometary reservoir is attributed to the scattering of icy bodies originally located in the Uranus-Neptune formation region between ~10 and 15 AU in the PSN (*19*), although a nonsolar, external origin for a large fraction of OCCs has recently been proposed (*20*). In contrast, JFCs are expected to have formed in the Kuiper Belt region beyond Neptune (*21*). D/H ratios of OCCs, analyzed either in situ by means of mass spectrometry in the case of comet Halley (*6*, *7*) or spectroscopy (*5*), show a range varying from ~1.3 to 2.9 times the terrestrial value (*f* ~ 9.8 to 21.9) (Fig. 3). In addition, the D/H ratio was found to be similar to the comet Halley in in situ measurements in the water plume of Saturn's satellite Enceladus (*22*). These values support the predicted D/H radial increase with distance from the Sun and the origin of OCCs from a common, localized region of the disk (*16*–*18*).

Recent D/H measurements in water in the two JFCs analyzed so far—namely, $(1.61 \pm 0.24) \times 10^{-4}$ (*f* ~ 7.8) for 103P/Hartley 2 (*23*) and an upper limit of 2.0×10^{-4} (*f* < 9.5) for 45P/Honda–Mrkos–Pajdušáková (*5*)—contradict this view. The hydrogen isotope composition of 103P/Hartley 2 is closer to the terrestrial value than the OCCs average, reviving the possibility of a cometary, rather than asteroidal, origin for the oceans. These data lead to two possible conclusions: Either JFCs originate from the Kuiper Belt and the chemical models developed so far (*16*–*18*) are not representative, or these comets formed over a wide range of heliocentric distances in the outer part of the PSN. With regard to the first possibility, a recent chemical model leading to a nonmonotonic *f* profile throughout the PSN (*24*) matches these observations if the JFCs were formed in the Kuiper Belt (*21*). In contrast with previous PSN models evolving as closed systems, this model assumes that the disk continues to be fed by material infalling from the presolar cloud. Alternatively, it has been proposed that JFCs and OCCs could originate from the same extended outer region of the PSN (*25*), so that 103P/Hartley 2 and 45P/Honda–Mrkos–Pajdušáková (45P/HMP) may simply have formed in the inner part of this common reservoir. In this case, the range of D/H ratios measured in JFCs should be similar to the one found in OCCs, as suggested by observations.

The new D/H value of $(5.3 \pm 0.7) \times 10^{-4}$ (*f* ~ 25.2) from comet 67P/Churyumov-Gerasimenko is not consistent with previous JFC data and is even higher than values characteristic of OCCs (~30

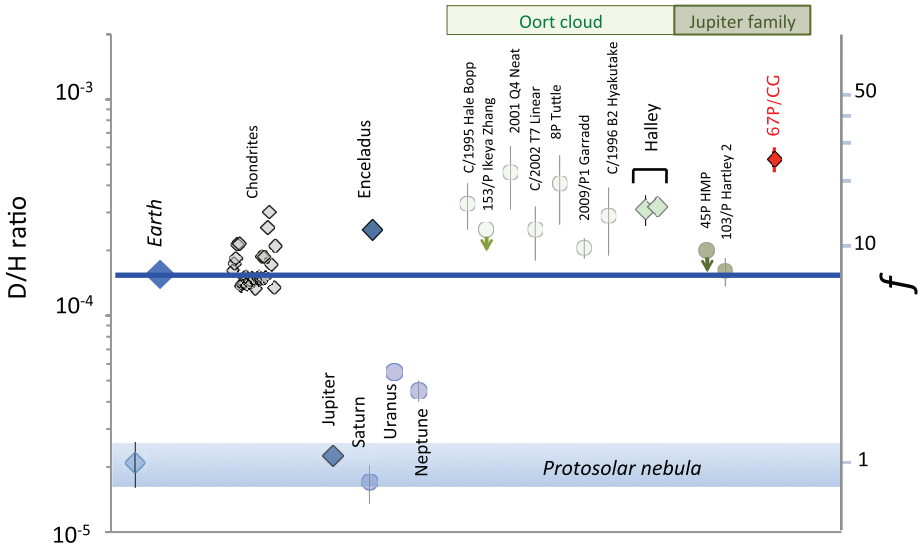


Fig. 3. D/H ratios in different objects of the solar system. Data are from (*1*, *2*, *5*–*7*, *26*–*28*) and references therein. Diamonds represent data obtained by means of in situ mass spectrometry measurements, and circles refer to data obtained with astronomical methods.

to 120% higher than that of comet Halley). In contrast to previous JFC measurements, this estimate matches models that predict a monotonic radial increase of the enrichment profile (*16*–*18*). From the ROSINA measurements on comet 67P/Churyumov-Gerasimenko, we conclude that the D/H values of JFCs may be highly heterogeneous, possibly reflecting the diverse origins of JFCs. If this is the case, then the new measurement supports models advocating an asteroidal (carbonaceous chondrite–like) rather than cometary origin for the oceans, and by extension for the terrestrial atmosphere (*1*, *2*).

REFERENCES AND NOTES

1. B. Marty, The origins and concentrations of water, carbon, nitrogen and noble gases on Earth. *Earth Planet. Sci. Lett.* **313**, 56–66 (2012). doi: 10.1016/j.epsl.2011.10.040
2. C. M. O. Alexander *et al*., The provenances of asteroids, and their contributions to the volatile inventories of the terrestrial planets. *Science* **337**, 721–723 (2012). doi: 10.1126/science.1223474; pmid: 22798405
3. N. Dauphas, The dual origin of the terrestrial atmosphere. *Icarus* **165**, 326–339 (2003). doi: 10.1016/S0019-1035(03)00198-2
4. T. Owen, A. Bar-Nun, I. Kleinfeld, Possible cometary origin of heavy noble gases in the atmospheres of Venus, Earth and Mars. *Nature* **358**, 43–46 (1992). doi: 10.1038/358043a0; pmid: 11536499
5. C. Ceccarelli, P. Caselli, D. Bockelee-Morvan, O. Mousis, S. Pizzarello, F. Robert, D. Semenov, Deuterium fractionation: the Ariadne's thread from the pre-collapse phase to meteorites and comets today. Protostars and Planets VI (2014); available at http://arxiv.org/pdf/1403.7143v1.pdf.
6. H. Balsiger, K. Altwegg, J. Geiss, D/H and O-18/O-16 ratio in the hydronium ion and in neutral water from in situ ion measurements in comet Halley. *J. Geophys. Res.* **100** (A4), 5827–5834 (1995). doi: 10.1029/94JA02936
7. P. Eberhardt, M. Reber, D. Krankowsky, R. R. Hodges, The D/H and $^{18}O/^{16}O$ ratios in water from comet P/Halley. *Astron. Astrophys.* **302**, 301 (1995).
8. H. Balsiger *et al*., ROSINA Rosetta Orbiter Spectrometer for Ion and Neutral Analysis. *Space Sci. Rev.* **128**, 745–801 (2002). doi: 10.1007/s11214-006-8335-3
9. M. Hässig *et al*., ROSINA/DFMS capabilities to measure isotopic ratios in water at comet 67P/Churyumov-Gerasimenko. *Planet. Space Sci.* **84**, 148–152 (2013). doi: 10.1016/j.pss.2013.05.014
10. B. Schläppi *et al*., Influence of spacecraft outgassing on the exploration of tenuous atmospheres with in situ mass spectrometry. *J. Geophys. Res.* **115** (A12), A12313 (2012). doi: 10.1029/2010JA015734
11. H. F. Levison, M. J. Duncan, R. Brasser, D. E. Kaufmann, Capture of the Sun's Oort cloud from stars in its birth cluster. *Science* **329**, 187–190 (2010). doi: 10.1126/science.1187535; pmid: 20538912
12. P. R. Mahaffy, T. M. Donahue, S. K. Atreya, T. Owen, H. B. Niemann, Galileo Probe measurements of D/H and 3He/4He in Jupiter's atmosphere. *Space Sci. Rev.* **84**, 251–263 (1998). doi: 10.1023/A:1005091806594
13. J. Geiss, G. Gloeckler, Abundances of deuterium and helium-3 in the protosolar cloud. *Space Sci. Rev.* **84**, 239–250 (1998). doi: 10.1023/A:1005039822524
14. T. Prodanovic, G. Steigman, B. D. Fields, The deuterium abundance in the local interstellar medium. *Mon. Not. R. Astron. Soc.* **406**, 1108–1115 (2010).
15. R. Visser, S. D. Doty, E. F. van Dishoeck, The chemical history of molecules in circumstellar disks. II. Gas-phase species. *Astron. Astrophys.* **534**, A132 (2011). doi: 10.1051/0004-6361/201117249
16. J. Geiss, H. Reeves, Deuterium in the solar system. *Astron. Astrophys.* **93**, 189–199 (1981).
17. J. J. Kavelaars, O. Mousis, J.-M. Petit, J. A. Weaver, On the formation location of Uranus and Neptune as constrained by dynamical and chemical models of comets. *Astrophys. J.* **734**, L30 (2011). doi: 10.1088/2041-8205/734/2/L30
18. K. Furuya, Y. Aikawa, H. Nomura, F. Hersant, V. Wakelam, Water in protoplanetary disks: Deuteration and turbulent mixing. *Astrophys. J.* **779**, 11 (2013). doi: 10.1088/0004-637X/779/1/11
19. T. Albertsson, D. Semenov, T. Henning, Chemodynamical deuterium fractionation in the early solar nebula: The origin of water on Earth and in asteroids and comets. *Astrophys. J.* **784**, 39 (2014). doi: 10.1088/0004-637X/784/1/39
20. L. Dones, P. R. Weissman, H. F. Levison, M. J. Duncan, Oort cloud formation and dynamics. *Comets* **II**, 153–174 (2004).
21. M. J. Duncan, H. F. Levison, A disk of scattered icy objects and the origin of Jupiter-family comets. *Science* **276**, 1670–1672 (1997). doi: 10.1126/science.276.5319.1670; pmid: 9180070
22. J. H. Waite Jr. *et al*., Liquid water on Enceladus from observations of ammonia and 40Ar in the plume. *Nature* **460**, 487–490 (2009). doi: 10.1038/nature08153
23. P. Hartogh *et al*., Ocean-like water in the Jupiter-family comet 103P/Hartley 2. *Nature* **478**, 218–220 (2011). doi: 10.1038/nature10519; pmid: 21976024
24. L. Yang, F. J. Ciesla, C. M. O. Alexander, The D/H ratio of water in the solar nebula during its formation and evolution. *Icarus* **226**, 256–267 (2013). doi: 10.1016/j.icarus.2013.05.027
25. R. Brasser, A. Morbidelli, Oort cloud and Scattered Disc formation during a late dynamical instability in the Solar System. *Icarus* **225**, 40–49 (2013). doi: 10.1016/j.icarus.2013.03.012
26. J. Aléon, J., Multiple origins of nitrogen isotopic anomalies in meteorites and comets. *Astrophys. J.* **722**, 1342–1351 (2010). doi: 10.1088/0004-637X/722/2/1342
27. D. Bockelée-Morvan *et al*., Herschel measurements of the D/H and $^{16}O/^{18}O$ ratios in water in the Oort-cloud comet C/2009 P1 (Garradd). *Astron. Astrophys.* **544**, L15 (2012). doi: 10.1051/0004-6361/201219744
28. H. Feutchgruber *et al*., The D/H ratio in the atmospheres of Uranus and Neptune from Herschel PACS observations. *Astron. Astrophys.* **551**, A26 (2013).

ACKNOWLEDGMENTS

A. Morbidelli is acknowledged for helpful discussion. The authors thank the following institutions and agencies, which supported this work: Work at the University of Bern was funded by the State of Bern, the Swiss National Science Foundation, and the ESA PRODEX Program. Work at Max-Planck-Institut für Sonnensystemforschung was funded by the Max Planck Society and BMWI under contract 50QP1302. Work at Southwest Research institute was supported by subcontract 1496541 from the Jet Propulsion Laboratory. Work at BIRA-IASB was supported by the Belgian Science Policy Office via PRODEX/ROSINA PEA 90020. This work was supported by Centre National d'Études Spatiales grants at IRAP, LATMOS, LPC2E, UTINAM, CRPG, and by the the European Research Council (grant 267255). A.B.-N. thanks the Ministry of Science and the Israel Space agency. Work at the University of Michigan was funded by NASA under contract JPL-1266313. ROSINA would not give such outstanding results without the work of the many engineers and technicians involved in this instrument and in the mission over the years whose contributions are gratefully acknowledged. Rosetta is an ESA mission with contributions from its member states and NASA. The work of the ESA Rosetta team is herewith acknowledged. All ROSINA data are available on request until they are released to the PSA archive of ESA and to the PDS archive of NASA.

1 October 2014; accepted 1 December 2014
Published online 11 December 2014;
10.1126/science.1261952

COMETARY SCIENCE

The morphological diversity of comet 67P/Churyumov-Gerasimenko

Nicolas Thomas,[1*] Holger Sierks,[2] Cesare Barbieri,[3] Philippe L. Lamy,[4] Rafael Rodrigo,[5] Hans Rickman,[6] Detlef Koschny,[7] Horst Uwe Keller,[8,22] Jessica Agarwal,[2] Michael F. A'Hearn,[9] Francesco Angrilli,[10] Anne-Therese Auger,[11] M. Antonella Barucci,[12] Jean-Loup Bertaux,[13] Ivano Bertini,[14] Sebastien Besse,[7] Dennis Bodewits,[9] Gabriele Cremonese,[15] Vania Da Deppo,[16] Björn Davidsson,[17] Mariolino De Cecco,[18] Stefano Debei,[10] Mohamed Ramy El-Maarry,[1] Francesca Ferri,[14] Sonia Fornasier,[12] Marco Fulle,[19] Lorenza Giacomini,[20] Olivier Groussin,[11] Pedro J. Gutierrez,[21] Carsten Güttler,[2] Stubbe F. Hviid,[22,2] Wing-Huen Ip,[23] Laurent Jorda,[24] Jörg Knollenberg,[22] J.-Rainer Kramm,[2] Ekkehard Kührt,[22] Michael Küppers,[25] Fiorangela La Forgia,[3] Luisa M. Lara,[21] Monica Lazzarin,[3] Josè J. Lopez Moreno,[21] Sara Magrin,[3] Simone Marchi,[26] Francesco Marzari,[3] Matteo Massironi,[20,14] Harald Michalik,[27] Richard Moissl,[25] Stefano Mottola,[22] Giampiero Naletto,[28,14,16] Nilda Oklay,[2] Maurizio Pajola,[14] Antoine Pommerol,[1] Frank Preusker,[22] Lola Sabau,[29] Frank Scholten,[22] Colin Snodgrass,[30,2] Cecilia Tubiana,[2] Jean-Baptiste Vincent,[2] Klaus-Peter Wenzel[7]

Images of comet 67P/Churyumov-Gerasimenko acquired by the OSIRIS (Optical, Spectroscopic and Infrared Remote Imaging System) imaging system onboard the European Space Agency's Rosetta spacecraft at scales of better than 0.8 meter per pixel show a wide variety of different structures and textures. The data show the importance of airfall, surface dust transport, mass wasting, and insolation weathering for cometary surface evolution, and they offer some support for subsurface fluidization models and mass loss through the ejection of large chunks of material.

The European Space Agency's Rosetta spacecraft entered a close orbit about the Jupiter family comet 67P/Churyumov-Gerasimenko (hereafter 67P) on 6 August 2014 at 3.60 astronomical units (AU) from the Sun. It carries OSIRIS (Optical, Spectroscopic and Infrared Remote Imaging System) (*1*), which has acquired images of the surface at scales of <0.8 meter per pixel in the subsequent 2 months, thereby providing large improvements in resolution, coverage, and data quality over previous flyby missions of cometary nuclei (*2*–*4*). The concept of cometary nuclei as rather uniform, pristine, protoplanetesimals that may have been subjected to collisional processing is persistent, despite evidence of regional differences on individual objects such as comet 9P/Tempel 1 (*4*) and comet 103P/Hartley 2 (*5*) and major differences between objects (*6*). The OSIRIS observations have revealed an irregular-shaped, processed nucleus surface with morphologically diverse units. Here we describe the basic morphology of the surface and some of the structural heterogeneity evident in the highest-resolution data acquired until September 2014, when the comet was 3.27 AU from the Sun.

Combining the orbit of Rosetta with OSIRIS imaging sequences, we derived shape models of almost 100% of the illuminated area (~70% of the surface) using different reconstruction techniques. We used a model obtained using a stereophotogrammetric (SPG) approach (*7*–*9*). Figure S1 displays different shaded relief views of a global SPG-based shape model of ~5 to 7 m spatial resolution and is discussed in detail in (*10*). The remaining 30% of the surface (the unilluminated southern polar region) will receive the highest solar fluxes around perihelion.

Regional-scale terrains

Nineteen regions have been defined on the nucleus at this time (Fig. 1) (*11*). It is expected that others will be mapped once the Sun crosses the comet's equator and illuminates the southern hemisphere completely. The regions can be grouped into five basic categories: dust-covered terrains, brittle materials with pits and circular structures, large-scale depressions, smooth terrains, and exposed consolidated surfaces (table S1).

Dust-covered terrains

The Ma'at and Ash regions (both blue in Fig. 1) appear to be coated with a smooth covering of dust (*12*). The thickness is uncertain and variable. The source of the dust is activity that produces low-velocity non-escaping dust particles in a form of airfall. Upward-facing surfaces in several other areas (such as Seth, Fig. 2, right, and Hathor, Fig. 4) also show smooth deposits, which are inferred to have the same source(s). The coating is not sufficiently thick to completely mask the large-scale structure of the substrate. The surface of 67P is almost devoid of recognizable impact craters. One example, which may be partially buried by airfall, leading to an estimated thickness for the airfall of 1 to 5 m using standard scaling laws, has been identified (fig. S2).We expect that any ice chunks or particles larger than a few millimeters that are emitted from the source will not sublimate before falling back and could therefore form visually bright, highly localized ice deposits and sources of gas on these surfaces. The importance of low-velocity airfall deposition of fluffy particles in reducing the thermal inertia should be considered in future thermal models. Given the likelihood of a high organic content for the airfall deposit (e.g., cometary tholin) and exposure to sunlight, the production of a form of organic, possibly polymerized, matrix from this deposit (after the loss of any volatiles) is also conceivable.

High-resolution data show evidence of dune-like structures (fig. S3) that may be the result of aeolian-driven surface transport of the dust (*13*). Images acquired in the Hapi region of aeolian

[1]Physikalisches Institut, Sidlerstrasse 5, University of Bern, 3012, Bern, Switzerland. [2]Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg, 3, 37077, Göttingen, Germany. [3]Department of Physics and Astronomy, University of Padova, Vicolo dell'Osservatorio 3, 35122 Padova, Italy. [4]Laboratoire d'Astrophysique de Marseille, UMR 7326, CNRS, and Aix Marseille Université, 38 Rue Frédéric Joliot-Curie, 13388 Marseille Cedex 13, France. [5]International Space Science Institute, Hallerstraße 6, 3012 Bern, Switzerland, and Centro de Astrobiología, Consejo Superior de Investigaciones Cientificas - Instituto Nacional de Tecnica Aeroespacial, 28850 Torrejón de Ardoz, Madrid, Spain. [6]Department of Physics and Astronomy, Uppsala University, Box 516, SE-75120 Uppsala, Sweden, and Polish Academy of Sciences Space Research Center, Bartycka 18A, PL-00716 Warszawa, Poland. [7]Scientific Support Office, European Space Agency, 2201, Noordwijk, Netherlands. [8]Institute for Geophysics and Extraterrestrial Physics, Technische Universität Braunschweig, 38106, Braunschweig, Germany. [9]Department of Astronomy, University of Maryland, College Park, MD 20742-2421, USA. [10]Department of Industrial Engineering, University of Padova, via Venezia 1, 35131 Padova, Italy. [11]Aix Marseille Université, CNRS, Laboratoire d'Astrophysique de Marseille, UMR 7326, 38 rue Frédéric Joliot-Curie, 13388 Marseille, France. [12]Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique, Observatoire de Paris, CNRS, Université Paris 06, Université Paris-Diderot, 5 Place J. Janssen, 92195, Meudon, France. [13]Laboratoire Atmosphères, Milieux, Observations Spatiales, CNRS/UVSQ/IPSL, 11 Boulevard d'Alembert, 78280, Guyancourt, France. [14]Centro di Ateneo di Studi ed Attività Spaziali, "Giuseppe Colombo" (CISAS), University of Padova, Via Venezia 15, 35131 Padova, Italy. [15]Istituto Nazionale di Astrofisica, Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, 35122 Padova, Italy. [16]CNR-IFN UOS Padova LUXOR, Via Trasea, 7, 35131 Padova, Italy. [17]Department of Physics and Astronomy, Uppsala University, 75120 Uppsala, Sweden. [18]UNITN, Universitá di Trento, Via Mesiano, 77, 38100 Trento, Italy. [19]INAF, Osservatorio Astronomico, Via Tiepolo 11, 34014 Trieste, Italy. [20]Dipartimento di Geoscienze, University of Padova, via G. Gradenigo 6, 35131 Padova, Italy. [21]Instituto de Astrofísica de Andalucía (CSIC), c/ Glorieta de la Astronomía s/n, 18008 Granada, Spain. [22]Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Planetenforschung, Asteroiden und Kometen, Rutherfordstraße 2, 12489 Berlin, Germany. [23]National Central University, Graduate Institute of Astronomy, 300 Chung-Da Road, Chung-Li 32054, Taiwan. [24]Laboratoire d'Astrophysique de Marseille, 38 Rue de Frédéric Joliot-Curie, 13388 Marseille Cedex 13, France. [25]Science Operations Department, European Space Astronomy Centre/European Space Agency, Post Office Box 78, 28691 Villanueva de la Canada, Madrid, Spain. [26]Solar System Exploration Research Virtual Institute, Southwest Research Institute, 1050 Walnut Street, Suite 300, Boulder, CO 80302, USA. [27]Institut für Datentechnik und Kommunikationsnetze der Technische Universität Braunschweig, Hans-Sommer-Strasse 66, 38106 Braunschweig, Germany. [28]University of Padova, Department of Information Engineering, Via Gradenigo 6/B, 35131 Padova, Italy. [29]Instituto Nacional de Tecnica Aeroespacial, Carretera de Ajalvir, 28850 Torrejon de Ardoz (Madrid), Spain. [30]Planetary and Space Sciences, Department of Physical Sciences, The Open University, Milton Keynes MK7 6AA, UK.
***Corresponding author. E-mail: nicolas.thomas@space.unibe.ch**

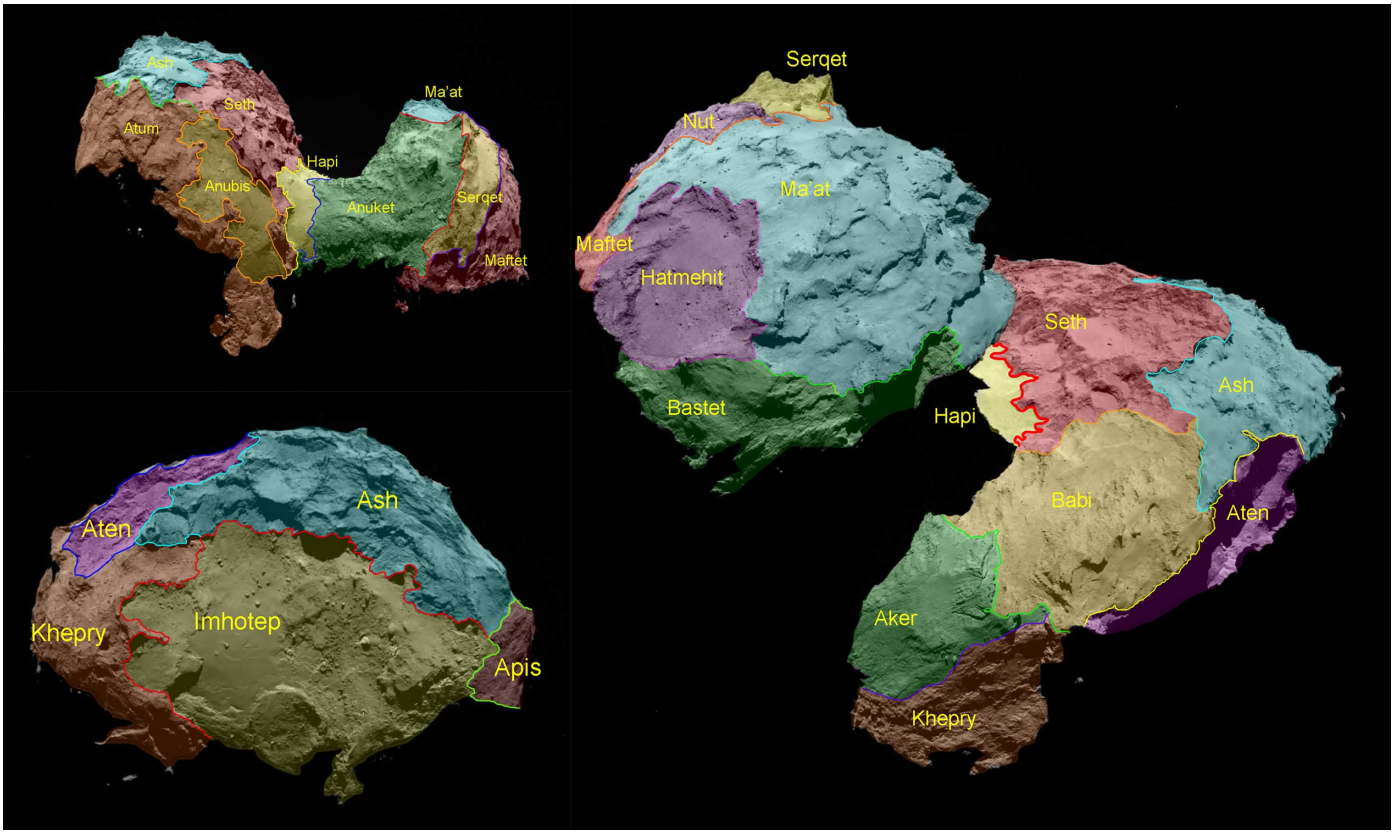


Fig. 1. Regional definitions based on large-scale unit boundaries. The nomenclature for the regions used here is also given. Images were acquired by OSIRIS (top left, image no. WAC_2014-09-05T02.29.12; bottom left, WAC_2014-09-05T06.29.13; right, NAC_2014-08-05T23.19.14).

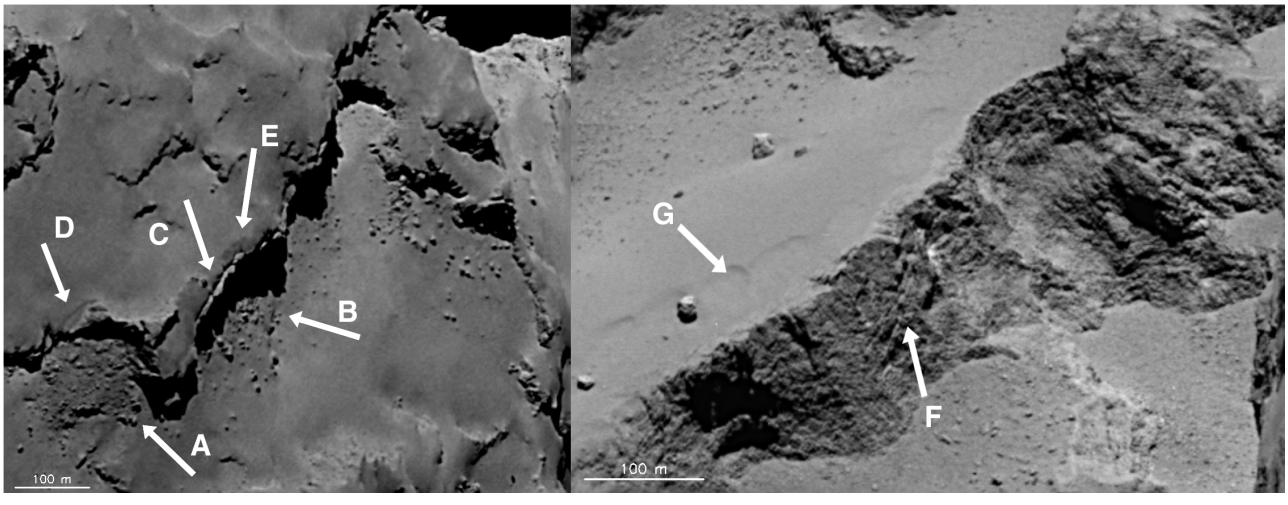


Fig. 2. (Left) The Ash-Seth boundary shows evidence of collapse with talus at the base (positions A and B). Fracturing to produce crevices (C) is also observed, with deflation of the surface (D and E) appearing as a possible precursor (NAC_2014-08-07T18.37.34.552Z_ID30_1397549700_F22). (**Right**) Oblique view of a circular depression in Seth, which appears to have been partially eroded. The face (F) appears rather uniform in texture, with some evidence of linear features. Note also the evidence of surface deflation (G) and that the horizontal slope appears dust-covered, whereas the vertical face appears "clean" (NAC_2014-09-02T21.44.22.575Z_ID10_1397549800_F22).

ripple structures, rocks with wind tails (fig. S4), and rocks with moats provide further support for localized gas-driven transport. To initiate saltation, the surface shear stress from the gas expansion of a vent must overcome the gravitational force and interparticle forces (*14*). Although the gas densities are low, following (*15*), velocities on the order of 300 m/s appear to be sufficient to move 100-μm particles and may be generated by localized sublimation. Electrostatic levitation in combination with horizontal electric fields across the terminator, as proposed for the Moon and 433 Eros, may be an alternative mechanism, although this may only be effective for smaller particles (*16*).

Brittle material

The Seth region and interfaces between Ma'at and Ash and other units (such as Imhotep) show

evidence of a morphologically distinct, consolidated brittle material (*17*). There is evidence in many places of fracturing and collapse of this material as if it is being undercut by mass wasting of a stratum below (Fig. 2, left). Where collapse or disruption of the material has occurred, debris is frequently observed.

We infer that this brittle material (*18*) is present everywhere underneath the airfall deposits in Ma'at and Ash. The evidence of collapse combined with the low gravity reinforces the concept of cometary material having low tensile strength. Simple mechanics has been used to estimate the strength necessary to maintain some overhangs, resulting in values locally of <20 Pa (*19*).

The Seth region is dominated by a series of circular, flat-floored, steep-walled depressions. In many cases, the Hapi facing wall is absent (Fig. 2, right). In some examples, the flat floors hold debris, presumably from collapse of the walls of the depressions. The flat floors, which are all almost orthogonal to the local gravity vector, all show evidence of the dust deposition seen in Ash and Ma'at, whereas the sides that are roughly parallel to the potential gradient are relatively clean. The largest example (Fig. 2, right) is 650 m in diameter and (if originally circular) has lost around one-third of its form, although there is little evidence of talus at its base, suggesting that the cut surface is composed of more-consolidated material. This surface also shows some faint horizontal and vertical lineation. There is some similarity to the circular depressions seen on 81P/Wild 2 (*3*), suggesting a common production mechanism such as sublimation or the collapse of internal voids. However, the quasicircular nature requires an organization that is not necessarily expected in either mechanism. The Seth region also contains a pit chain (*10*). Similar pits are seen in the Ma'at region.

Large-scale depression structures

Hatmehit is a circular depression, 1.0 km in diameter, on the head of the nucleus. The structure is rather shallow and hence very unlike the bowl-shaped impact structures commonly seen on other bodies. This may not rule out a collision origin, but the details of the creation of the rim and of the final surface distribution of fluidized ejecta require modeling in a parameter regime rarely explored numerically up to this point (*20*).

The head and the body of the nucleus have one large irregular-shaped depression structure each. The one on the body (Aten) is more distinctive than that on the head (Nut). Aten is surrounded by the brittle material of Ash, which is coated to a large extent by airfall. However, Aten shows no evidence of a similar dust coating. Furthermore, there is no evidence of the brittle material within Aten, with the interior being covered with boulders. At the edges of the depression, disrupted consolidated brittle material is seen. We suggest that Aten has been formed by one or more major mass loss events removing airfall and the mantling brittle material. It is important to recognize that the volume of the Aten depression as a whole is large (~0.12 km^3 from the shape model). We estimate that 67P can lose around 3 to 5 × 10^9 kg per orbit in its current orbit via sublimation. Assuming the density to be 440 kg/m^3 (*10*), the production of Aten via a sublimation mechanism alone would have required ~10 to 20 orbits at this production rate.

Smooth terrains

There are three areas characterized by extremely smooth material with no obvious impacts or circular flat-floored depressions and a paucity of boulders (Imhotep, Anubis, and Hapi). The Imhotep region is morphologically remarkable. It is dominated by a smooth surface that covers an area of >0.7 km^2 (Fig. 3). At its margins, the smooth material gives the impression that it is layered. Unlike the airfall deposits, the smooth material appears to be enclosed by more-consolidated material that surrounds it [consolidated cometary material (CCM) (*21*)]. The smooth material thins to one side and gives way to a terrain dominated by circular filled and unfilled structures. The filling material appears similar to the smooth material (Fig. 3, position D). On comet 9P/Tempel 1, the presence of pits expressed primarily as topographic and albedo rings with probable fill on the inside—a few having distinct flat floors—has been noted (*22*). The terrain on 67P has some similarities, and we note that a subsurface fluidization production mechanism has been proposed for 9P/Tempel 1 (*23*). However, the sharp-sided nature of the external surfaces of these circular structures may not be consistent with this mechanism in its simplest form, and other mechanisms [e.g., formations such as sun cups or dirt cones (*24*)] should be considered. Imhotep is the only region on the illuminated surface with these types of structures.

The smooth material of Imhotep is surrounded on three sides by fractured CCM, which has been subject to mass wasting with debris production. However, to the south, there is a pronounced raised circular structure, 650 m in diameter, which also contains smooth material that is partially eroded. The external surface of the circular structure is irregularly fractured. Smooth terrain is seen at several topographic levels in the region (*25*). The Anubis region appears similar

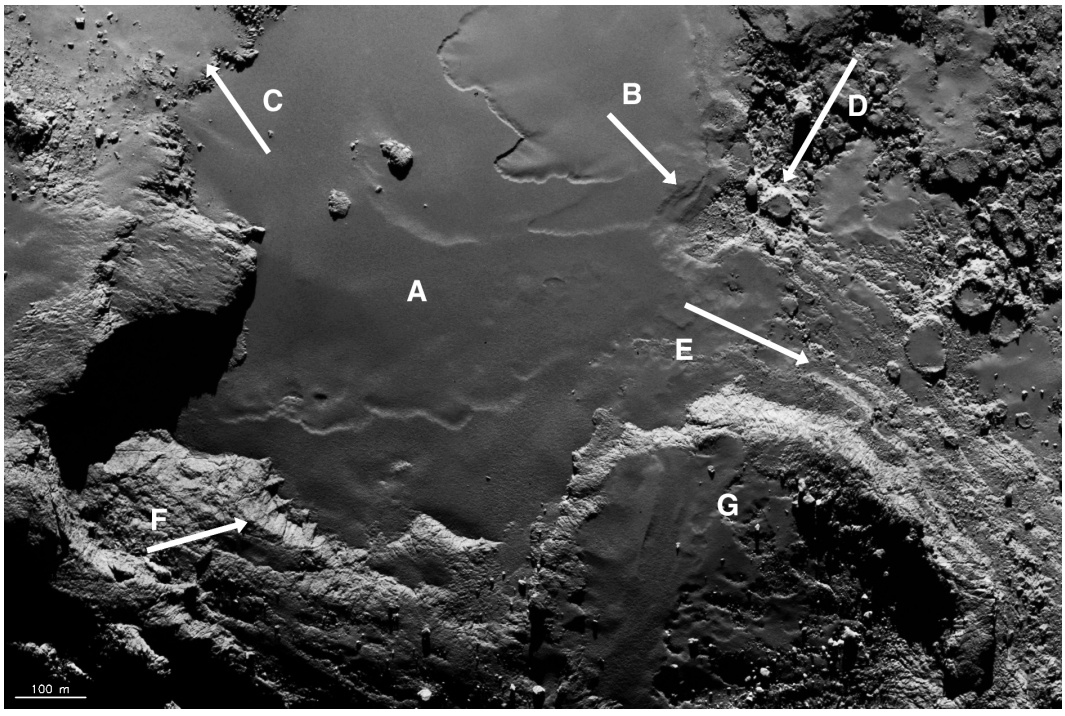


Fig. 3. Part of the Imhotep region. The area shows smooth terrain (A), layering of this material at its margin (B), smooth material on topographically higher surfaces (C), circular structures possibly related to those seen on 9P/Tempel 1 (D), and layered consolidated material (E) rising toward the 650–m-diameter raised semicircular structure (G). Fracturing of the consolidated material is evident throughout [e.g., position (F)] (NAC_2014-09-05T06.31.16.575Z_ID10_1397549600_F22).

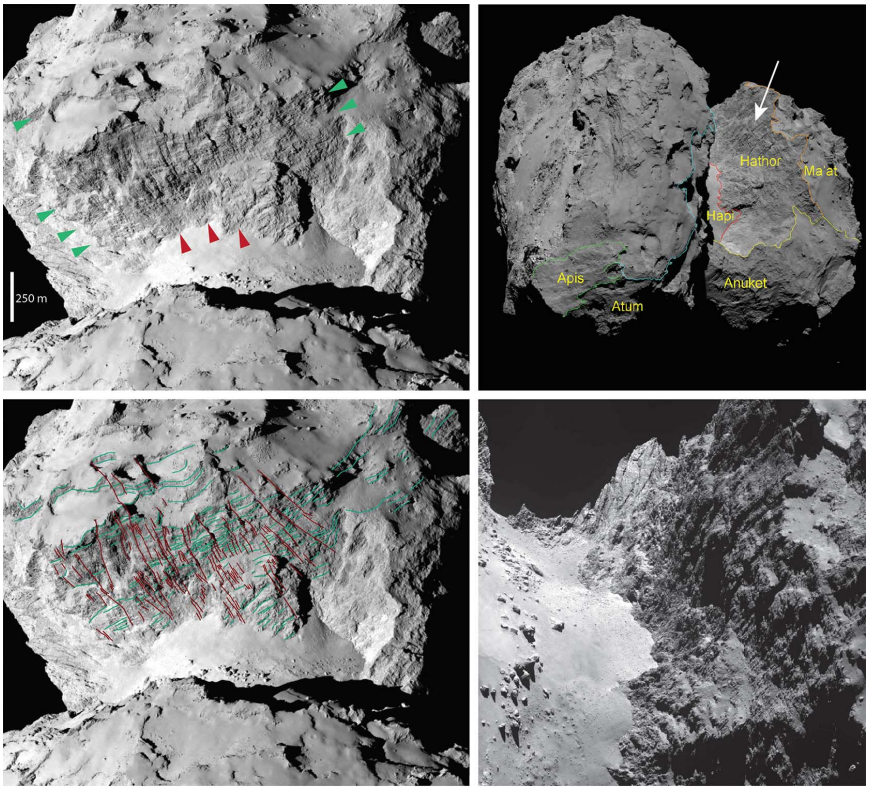


Fig. 4. The Hathor structure seen from different perspectives. (**Top right**) The position of Hathor in relation to other major regions. (**Lower right**) An oblique view of the Hathor face and the interface with the Hapi region. Notice the boulders strewn along the long axis of the Hapi region. (**Bottom left**) Mapping of the lineaments seen in the Hathor face. (**Top left**) Identification of the contacts of the heavily lineated terrain.

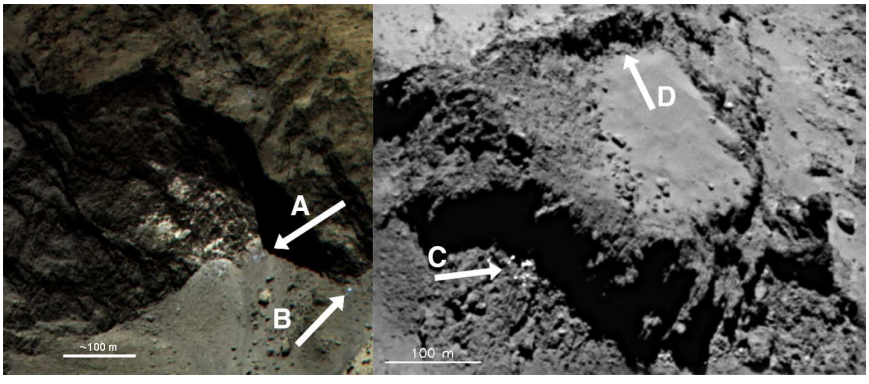


Fig. 5. (Left) Color composite of an alcove at the Hathor-Anuket boundary. The surface of the consolidated material in the alcove contains distinct spots of brighter material independent of viewing geometry. On the floor beneath the consolidated material, material 20% brighter and bluer than the surroundings can also be seen (e.g., positions A and B). Central wavelengths of the filters used to produce the composite were R (882 nm), G (700 nm), and B (481 nm). (Sequence acquired at 15:42:17 on 21 August 2014.) (**Right**) Bright, highly reflective meter-sized boulders at the Babi-Khepry boundary adjacent to mantle material (C). Note also the brittle fracture and collapse at position D (NAC_2014-09-03T06.44.22.578Z_ID10_1397549400_F22).

to Imhotep, but no circular structures are evident in the early data.

The Hapi region in the neck was the first unit to become visibly active (emitting dust) in OSIRIS data. The surface reflectance properties at visible wavelengths are mostly dominated by dust rather than the high (blue) reflectances expected of ices. (We discuss high-reflectivity particle "clusters" below.) The region is at least 2.2 km long, roughly 0.8 km wide, and forms a partial ring of ~140° around the nucleus, leading to its necklike appearance. A row of boulders is present along part of Hapi's length (Fig. 4, bottom right); strong dust emission from the vicinity of these boulders has been observed but not accurately localized.

Unlike Imhotep, the smooth material forming Hapi is piled up against the faces of Hathor and Seth in several places, suggesting that particles are falling or falling back from the faces. This may allow future studies of the coefficient of restitution. There is also what appears to be a small aeolian ripple field, and several boulders appear to have wind tails, indicating again that aeolian dust transport may be of significance (fig. S4). These observations suggest that Hapi is fundamentally different in character from the Imhotep and Anubis regions.

Exposed consolidated surfaces

Several regions on 67P give a rocky (CCM) appearance, although it is important to recognize that the bulk densities of the materials are probably factors of 5 to 10 lower than those of terrestrial silicates (*10*). The regions included in this category are listed in table S1. Hathor mostly consists of a 900-m-high "cliff" that rises up from the smooth Hapi region (Fig. 4). It is characterized by a set of aligned linear features, which run vertically upward for much of the height of the cliff, and by roughly perpendicular linear features aligned with small terraces, which might suggest inner layering. The cliff is roughly parallel to the local gravitational acceleration (fig. S1), indicating consolidated material of greater strength than the brittle material, with aligned weaknesses.

Of further interest within the Hathor region is the presence of an unlineated alcove, 250 to 300 m high, with unusual surface properties. The CCM appears to contain distinct locally bright spots <10 m in dimension (Fig. 5, left). These are seen at all observing geometries and not on adjacent terrain (fig. S6). At the base of the CCM (on the surface of Hapi), there is evidence of mass wasting. Within the talus, there are 2- to 5–m-sized structures that are ~20% brighter in the OSIRIS orange filter ($\lambda_{central}$ = 649 nm) than the surroundings and are appreciably bluer than the surrounding material. This may be evidence of more volatile-rich material being embedded in the CCM and would imply that not only is the surface of Hapi active (*10*) but also that the CCM has the potential to be active. Processes should be considered that involve the sublimation of freshly exposed material that has fallen onto the surface of Hapi from its surroundings.

Anuket borders Hathor and is also a consolidated surface but shows no evidence of the linear features seen in Hathor. Hathor borders the neck, whereas Anuket is part of the external surface of the head (fig. S1). Wide-angle camera views (fig. S5) show that the contact between Anuket and Ma'at is associated with a scarp (see also fig. S1). This implies that material similar to that of Anuket's surface extends underneath the dust-covered brittle material of Ma'at, although alternatives based on preferential erosion may also be viable. It further suggests that processes within the neck are leading to erosion of the Anuket material that forms the Hathor region. The brighter material in the Hathor region suggests that we are witnessing this erosion and consequently that Hathor is showing us the internal

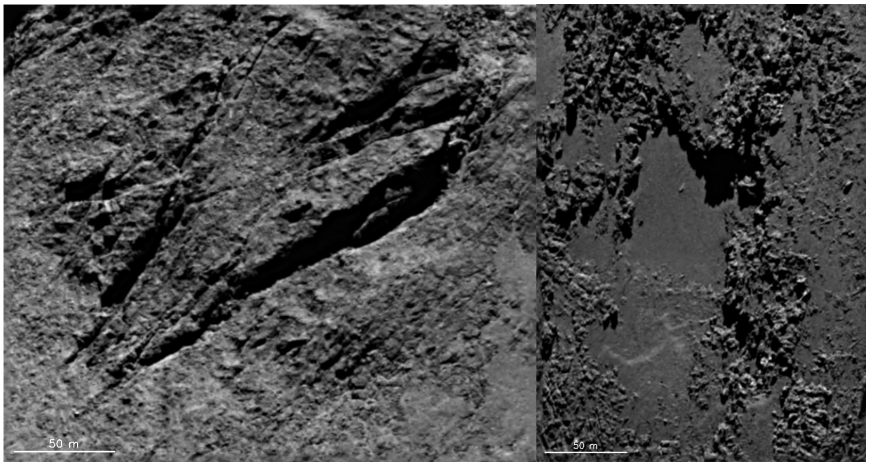


Fig. 6. (Left) Tectonic feature in the Aker region (NAC_2014-09-18T08.07.20.370Z_ID10_1397549000_F22). (**Right**) Ponded deposits in the Khepry region (NAC_2014-09-16T16.24.25.334Z_ID10_1397549500_F22).

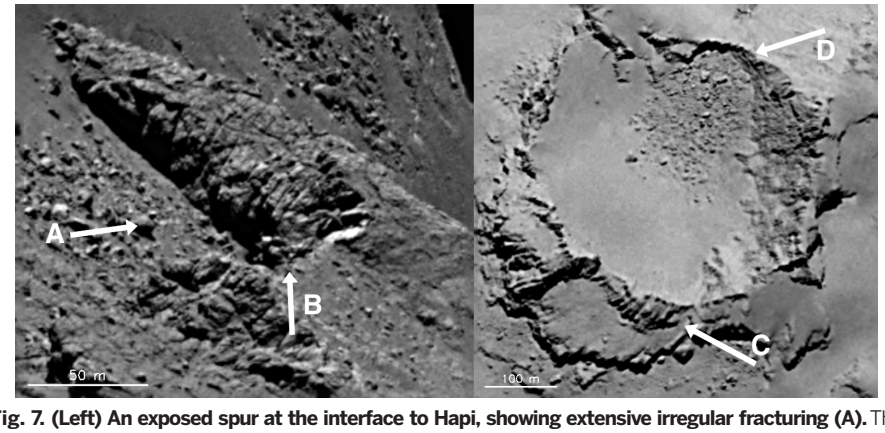


Fig. 7. (Left) An exposed spur at the interface to Hapi, showing extensive irregular fracturing (A). The cracking on the spur stops abruptly at a step (extending from position B) that seems to indicate a disruption of a mantling layer (NAC_2014-09-12T05.09.04.388Z_ID10_1397549600_F22). (**Right**) Circular depression that appears to have produced uplift with vertical fracture structures (position C). Note also the evidence of collapse and talus generation (position D) (NAC_2014-08-08T04.20.34.577Z_ID30_1397549900_F22).

structure of the head. This emphasizes the importance of trying to understand how the observed linear features in Hathor were produced. Anuket contains some notable cracks in its surface, which are roughly parallel to the neck (fig. S7). The location of these features within the neck region between the comet's two major lobes suggests that they may be the result of rotational- or activity-induced stresses in that particular part of the comet. Their importance for the future evolution of the nucleus is at present unclear (*10*).

On the body, the Aker region contains CCM and a set of tectonic fractures >200 m in length (Fig. 6, left). The entire Aker region and much of the neighboring Khepry region contain linear features. Brittle material appears to be completely absent. However, both Khepry and Aker show extremely smooth patches (50 to 100 m in dimension) of material that are in local topographic lows and surrounded by rougher material (Fig. 6, right). Boulders within their bounds are rare. These bear considerable similarity to ponded material observed on asteroid 433 Eros (*26*) in that they are smooth, flat, and of higher reflectance than the surroundings. For Eros, mechanisms such as electrostatic levitation of fine particles, seismic shaking, and disaggregation of local boulders have been proposed, with a current preference for electrostatic levitation (*27*).

Small-scale features

Local fracturing

Most of the consolidated materials on 67P are extensively fractured (e.g., Fig. 7 and Fig. 3, position F). Insolation weathering resulting from either thermal fatigue or thermal shock seems to be a viable mechanism, based on terrestrial studies (*28*), given the huge temperature ranges and temporal gradients likely to be experienced by the surface materials over diurnal and orbital time scales (*29*). However, fracturing is related to the thermal expansion coefficient of the material, which is itself a function of porosity. The more porous nature of cometary materials, as compared to terrestrial materials, will therefore reduce the magnitude of the stresses. Conversely, the inferred low tensile strength of cometary materials undoubtedly enhances their propensity to fracture.

With thermal fatigue, failure occurs along the preexisting lines of weakness and does not generate new failure planes (*28*). Thermal shock, however, can produce the seemingly random fracturing observed over most of the nucleus. A clear exception is the observed vertical jointing in material that appears to have been uplifted after the creation of a circular depression in the Ash region (Fig. 7, right).

Evidence for block/chunk movement and loss

Inspection of Fig. 7, left, also indicates that this particular fractured surface was once covered by a blanket of material, a few meters thick, which has been lost with little or no residual talus. This is reminiscent of the Aten depression. The Maftet region has several irregular-shaped pits, 10 to 20 m deep and 100 to 150 m in scale. On the surrounding terrain, there are blocks of similar size (fig. S8). A possible interpretation of these observations is that gas pressure, building in a subsurface pocket (*30*), overcomes the overburden and the cohesive strength of the material. A block or chunk is then ejected by the expanding gas. In the case of Fig. 7, left, and Aten, sufficient energy was apparently available to eject the material from the area completely. For Maftet, insufficient energy was available, and the block was therefore deposited on adjacent terrains. The overburden that the mechanism has to overcome is in all cases only a few tens of pascals at most, considering the observed thickness combined with the gravitational acceleration. Observed isolated boulders may be the result of a similar process. Further evidence of volatile activity in the Maftet region (a possible flow) can be seen in fig. S9, which offers some additional support for subsurface fluidization models.

Given the low thermal inertia inferred for cometary nuclei (*31*) and 67P in particular (*32*), it seems probable that the driving volatile is either CO or CO_2 but almost certainly not H_2O [based on thermodynamic properties (*33*)]. Clearly, models with a highly insulating airfall dust layer above a slightly more conductive consolidated layer should be developed.

High-reflectivity particle clusters

At more than 10 places on the surface, there are clusters of bright, mostly unresolved particles (e.g., Fig. 5, right). Particle sizes can reach several meters in diameter. The unresolved individual particles should be in the 10 cm to 1 m size range in order to have sufficient cross-sectional area (and therefore be sufficiently bright) to stand out against the background lower-reflectance material. If these clusters are bright because they are ice-rich, then they must also be new/fresh, as we expect most, if not all, surface ices to be removed during perihelion passage. This would indicate that mass wasting of the brittle material seen in Fig. 5, right, is active in the current epoch, and changes in surface appearance are expected at this position in the coming months.

Similarly, there are strong brightness differences (up to a factor of 10) between the smooth terrain of Imhotep and some of the surfaces of boulders nearby that are not readily explained by illumination. The spectral ratio between these bright surfaces and the surroundings indicates that they are relatively blue and therefore likely to be ice-rich.

Conclusions

Although we are still in a relatively early phase of the mission, the following conclusions appear to be robust. The surface shows a wide range of different surface textures, which point toward a variety of processes. The presence of airfall, dune/ripple-like structures, wind tails, and smooth depressions with ponded dust suggests that surface dust transport is of major importance in defining the uppermost surface layer in many regions. We suggest that this leads naturally to low surface thermal inertia and low but non-negligible gas production rates from most of the surface by the redistribution of icy chunks. Mass wasting of a consolidated brittle, possibly layered, material below this is pervasive and is enhanced by undercutting of this layer through erosion of its substrate in places. Fracturing of the surface is seen at all scales. In specific regions, the structural coherence of the surface material has led to extensive aligned lineament sets over several hundreds of meters. When viewed in relation to the observed mass wasting, it suggests major heterogeneity in tensile strength, which may also imply spatially variable thermal inertia (*34*). The observed, more random fracturing may result from insolation weathering and specifically from thermal shock. There is evidence to support surface erosion through the loss of large chunks of material (up to 10 to 100 m in scale), indicating that it may be a major mass loss process for the nucleus. The buildup of supervolatile (e.g., CO and/or CO_2) subsurface gas pockets to pressures exceeding the overburden and cohesive strength appears to be a plausible mechanism that could also explain evidence of flows. Small-scale high-reflectivity materials have been exposed in the surfaces of consolidated ("rocky") materials, at the bases of consolidated structures, and adjacent to disintegrating brittle material. There are spectral indications that these are ice-rich nodules, which we expect to sublimate as the comet moves toward perihelion in the coming months. A key outstanding question is whether all the observed diverse large-scale morphology can be explained by insolation alone or whether physical and/or chemical inhomogeneity (including large-scale layering) is required.

REFERENCES AND NOTES

1. H. U. Keller *et al*., OSIRIS - The scientific camera system onboard Rosetta. *Space Sci. Rev.* **128**, 433–506 (2007). doi: 10.1007/s11214-006-9128-4
2. H. U. Keller *et al*., Comet P/Halley's nucleus and its activity. *Astron. Astrophys.* **187**, 807–823 (1987).
3. D. E. Brownlee *et al*., Surface of young Jupiter family comet 81P/Wild 2: View from the Stardust Spacecraft. *Science* **304**, 1764–1769 (2004). doi: 10.1126/science.1097899; pmid: 15205524
4. P. C. Thomas *et al*., The shape, topography, and geology of Tempel 1 from Deep Impact observations. *Icarus* **187**, 4–15 (2007). doi: 10.1016/j.icarus.2006.12.013
5. P. C. Thomas *et al*., Shape, density, and geology of the nucleus of Comet 103P/Hartley 2. *Icarus* **222**, 550–558 (2013). doi: 10.1016/j.icarus.2012.05.034
6. A. Basilevsky, H. U. Keller, Comet nuclei: Morphology and implied processes of surface modification. *Planet. Space Sci.* **54**, 808–829 (2006). doi: 10.1016/j.pss.2006.05.001
7. F. Preusker *et al*., The northern hemisphere of asteroid (21) Lutetia - topography and orthoimages from Rosetta OSIRIS NAC image data. *Planet. Space Sci.* **66**, 54–63 (2012). doi: 10.1016/j.pss.2012.01.008
8. R. W. Gaskell *et al*., Characterizing and navigating small bodies with imaging data. *Meteorit. Planet. Sci.* **43**, 1049–1061 (2008). doi: 10.1111/j.1945-5100.2008.tb00692.x
9. A shape model using stereophotoclinometry (SPC) has also been generated. The SPC and SPG techniques are complementary, and detailed comparisons, outside the scope of this report, will be made at a later date. The exclusive use of SPG here is solely for internal consistency and simplicity and should not be interpreted as a choice of one approach over the other.
10. H. Sierks *et al*., *Science* **347**, aaa1044 (2015).
11. Although it is unusual for planetary geomorphology works, we do not use a latitude/longitude system for the nucleus at this time because of the ambiguity that would result when using a simple (latitude/longitude) system. An approach to solving the cartographic problem of the highly irregular shape is currently being formulated.
12. The sizes of the particles making up the surface layer are not known at this time and are formally constrained only by the resolution limit (~1 m). The use of the word "dust" should not be interpreted as implying a specific size range.
13. A. F. Cheng, C. M. Lisse, M. A'Hearn, Surface geomorphology of Jupiter Family Comets: A geologic process perspective. *Icarus* **222**, 808–817 (2013). doi: 10.1016/j.icarus.2012.10.004
14. R. Greeley, J. D. Iversen, *Wind as a Geological Process on Earth, Mars, Venus and Titan* (Cambridge Planetary Science Series, Vol. 4. Cambridge Univ. Press, Cambridge, 1985).
15. Y. Shao, H. Lu, A simple expression for wind erosion threshold friction velocity. *J. Geophys. Res.* **105**, 22,437–22,443 (2000). doi: 10.1029/2000JD900304
16. A. R. Poppe, M. Piquette, A. Likhanskii, M. Horányi, The effect of surface topography on the lunar photoelectron sheath and electrostatic dust transport. *Icarus* **221**, 135–146 (2012). doi: 10.1016/j.icarus.2012.07.018
17. Dust forms a thin veneer over the brittle material in the Ma'at and Ash regions. We here avoid the use of the word "mantle" because of the strong geological connotations and the implications of a compositionally different surface layer for which we have no evidence at this time. As noted in the text, layering is evident at several scales, but the origin (primordial or evolutionary) remains open.
18. We use the term "brittle material" to imply a rock that breaks/fractures with a high ratio of fine to coarse fragments. It is a weakly consolidated material that displays fracturing but more readily crumbles and creates debris deposits. Fractures contained in more-consolidated material (e.g., in the Aker region) show no evidence of fine fragments.
19. Assuming a rectangular cross-section, the tensile strength of an overhang can be estimated using $\sigma = 3\,\rho\,g\cos\alpha\,l^2/h$, where ρ is the density, g is the gravitational acceleration at the overhang, α is the angle between gravitational acceleration and the normal of the overhang, l is the length of the overhang, and h is its height. For larger, already collapsed, overhangs of ~200 m in length (e.g., at the Ash-Imhotep interface), we estimate values of $\sigma \leq 200$ Pa. Smaller (~10 m) overhangs are seen (with talus below in some cases where previous collapse has occurred), which suggest that collapse occurs at 10 to 20 Pa.
20. We are aware of some work in this field by (*35*) and unpublished calculations by M. Jutzi, which may be of significance for the future.
21. Although the morphology of many surfaces on 67P suggests a rock-like texture, we have avoided using that term here because of the expected very low density of the surfaces when compared to terrestrial rocks. We have adopted the term "consolidated cometary material" instead (CCM).
22. P. Thomas *et al*., The nucleus of Comet 9P/Tempel 1: Shape and geology from two flybys. *Icarus* **222**, 453–466 (2013). doi: 10.1016/j.icarus.2012.02.037
23. M. J. S. Belton, H. J. Melosh, Fluidization and multiphase transport of particulate cometary material as an explanation of the smooth terrains and repetitive outbursts on 9P/Tempel 1. *Icarus* **200**, 280–291 (2009). doi: 10.1016/j.icarus.2008.11.012
24. M. D. Betterton, Theory of structure formation in snowfields motivated by penitentes, suncups, and dirt cones. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **63**, 056129 (2001). doi: 10.1103/PhysRevE.63.056129; pmid: 11414983
25. Some smooth areas are also seen in the adjacent Khepry region (fig. S10), again at different gravitational heights. This argues against a single flow or a single source of fluidized material being responsible. More complex emplacement mechanisms seem to be required.
26. M. S. Robinson, P. C. Thomas, J. Veverka, S. Murchie, B. Carcich, The nature of ponded deposits on Eros. *Nature* **413**, 396–400 (2001). doi: 10.1038/35096518; pmid: 11574881
27. J. H. Roberts, L. M. Prockter, O. S. Barnouin, C. M. Ernst, E. Kahn, R. W. Gaskell, *Origin and Flatness of Ponds on Asteroid 433 Eros* [American Geophysical Union (AGU) Fall Meeting Abstracts, AGU, San Francisco, CA, 2013], abstr. P23E-1831.
28. K. Hall, C. E. Thorn, Thermal fatigue and thermal shock in bedrock: An attempt to unravel the geomorphic processes and products. *Geomorphology* **206**, 1–13 (2014). doi: 10.1016/j.geomorph.2013.09.022
29. Crack formation through a thermal contraction mechanism in frozen terrestrial soils is discussed in (*36*).
30. N. H. Samarasinha, A model for the breakup of Comet LINEAR (C/1999 S4). *Icarus* **154**, 540–544 (2001). doi: 10.1006/icar.2001.6685
31. O. Groussin *et al*., The temperature, thermal inertia, roughness and color of the nuclei of Comets 103P/Hartley 2 and 9P/Tempel 1. *Icarus* **222**, 580–594 (2013). doi: 10.1016/j.icarus.2012.10.003
32. P. L. Lamy *et al*., Spitzer Space Telescope observations of the nucleus of comet 67P/Churyumov-Gerasimenko. *Astron. Astrophys.* **489**, 777–785 (2008). doi: 10.1051/0004-6361:200809514
33. W. F. Huebner, J. Benkho, M.-T. Capria, A. Coradini, C. De Sanctis, R. Orosei, D. Prialnik, *Heat and Gas Diffusion in Comet Nuclei* [International Space Science Institute (ISSI) Scientific Report, ISSI, Bern Switzerland, 2006], report SR-004.
34. B. Davidsson *et al*., Thermal inertia and surface roughness of Comet 9P/Tempel 1. *Icarus* **224**, 154–171 (2013). doi: 10.1016/j.icarus.2013.02.008
35. D. Korycansky, E. Asphaug, Low-speed impacts between rubble piles modeled as collections of polyhedra, 2. *Icarus* **204**, 316–329 (2009). doi: 10.1016/j.icarus.2009.06.006
36. A. C. Maloof, J. B. Kellogg, A. M. Anders, Neoproterozoic sand wedges: Crack formation in frozen soils under diurnal forcing during a snowball Earth. *Earth Planet. Sci. Lett.* **204**, 1–15 (2002). doi: 10.1016/S0012-821X(02)00960-3

ACKNOWLEDGMENTS

OSIRIS was built by a consortium of the Max-Planck-Institut für Sonnensystemforschung, in Göttingen, Germany; Centro Interdipartimentale di Studi e Attività Spaziali–University of Padova, Italy; the Laboratoire d'Astrophysique de Marseille, France; the Instituto de Astrofísica de Andalucia, Consejo Superior de Investigaciones Cientificas, Granada, Spain; the Research and Scientific Support Department of the European Space Agency (ESA), Noordwijk, Netherlands; the Instituto Nacional de Técnica Aeroespacial, Madrid, Spain; the Universidad Politećhnica de Madrid, Spain; the Department of Physics and Astronomy of Uppsala University, Sweden; and the Institut für Datentechnik und Kommunikationsnetze der Technischen Universität Braunschweig, Germany. The support of the national funding agencies of Germany (Deutschen Zentrums für Luft- und Raumfahrt), France (Centre National d'Etudes Spatiales), Italy (Agenzia Spaziale Italiana), Spain (Ministerio de Educación, Cultura y Deporte), Sweden (Swedish National Space Board; grant no. 74/10:2), and the ESA Technical Directorate is gratefully acknowledged. H.R. was also supported bygrant no. 2011/01/B/ST9/05442 of the Polish National Science Center. W.-H.I acknowledges the Ministry of Science and Technology, Taiwan (grant no. NSC 101-2111-M-008-016). M.F.A. acknowledges NASA funding through Jet Propulsion Laboratory contract no. 1267923. We thank the ESA teams at European Space Astronomy Centre, European Space Operations Centre, and European Space Research and Technology Centre for their work in support of the Rosetta mission. The data will be placed in ESA's Planetary Sciences Archive after the proprietary period and are available on request until that time.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/aaa0440/suppl/DC1
Figs. S1 to S11
Table S1

10 October 2014; accepted 30 December 2014
10.1126/science.aaa0440

COMETARY SCIENCE

On the nucleus structure and activity of comet 67P/Churyumov-Gerasimenko

Holger Sierks,[1]* Cesare Barbieri,[2] Philippe L. Lamy,[3] Rafael Rodrigo,[4,5] Detlef Koschny,[6] Hans Rickman,[7,8] Horst Uwe Keller,[9,10] Jessica Agarwal,[1] Michael F. A'Hearn,[1,11,12] Francesco Angrilli,[13] Anne-Therese Auger,[14] M. Antonella Barucci,[15] Jean-Loup Bertaux,[16] Ivano Bertini,[17] Sebastien Besse,[6] Dennis Bodewits,[11] Claire Capanna,[14] Gabriele Cremonese,[18] Vania Da Deppo,[19] Björn Davidsson,[7] Stefano Debei,[13] Mariolino De Cecco,[20] Francesca Ferri,[17] Sonia Fornasier,[15] Marco Fulle,[21] Robert Gaskell,[22] Lorenza Giacomini,[17] Olivier Groussin,[14] Pablo Gutierrez-Marques,[1] Pedro J. Gutiérrez,[23] Carsten Güttler,[1] Nick Hoekzema,[1] Stubbe F. Hviid,[1,10] Wing-Huen Ip,[24,25] Laurent Jorda,[14] Jörg Knollenberg,[10] Gabor Kovacs,[1] J. Rainer Kramm,[1] Ekkehard Kührt,[10] Michael Küppers,[26] Fiorangela La Forgia,[2] Luisa M. Lara,[23] Monica Lazzarin,[2] Cédric Leyrat,[15] Josè J. Lopez Moreno,[23] Sara Magrin,[2] Simone Marchi,[27] Francesco Marzari,[2] Matteo Massironi,[17,28] Harald Michalik,[29] Richard Moissl,[26] Stefano Mottola,[10] Giampiero Naletto,[17,19,30] Nilda Oklay,[1] Maurizio Pajola,[17] Marco Pertile,[13,17] Frank Preusker,[10] Lola Sabau,[31] Frank Scholten,[10] Colin Snodgrass,[1]† Nicolas Thomas,[32,33] Cecilia Tubiana,[1] Jean-Baptiste Vincent,[1] Klaus-Peter Wenzel,[6] Mirco Zaccariotto,[13,17] Martin Pätzold[34]

Images from the OSIRIS scientific imaging system onboard Rosetta show that the nucleus of 67P/Churyumov-Gerasimenko consists of two lobes connected by a short neck. The nucleus has a bulk density less than half that of water. Activity at a distance from the Sun of >3 astronomical units is predominantly from the neck, where jets have been seen consistently. The nucleus rotates about the principal axis of momentum. The surface morphology suggests that the removal of larger volumes of material, possibly via explosive release of subsurface pressure or via creation of overhangs by sublimation, may be a major mass loss process. The shape raises the question of whether the two lobes represent a contact binary formed 4.5 billion years ago, or a single body where a gap has evolved via mass loss.

Since 23 March 2014 at a distance of 5 million km, the Optical, Spectroscopic, and Infrared Remote Imaging System [OSIRIS (*1*)] onboard the Rosetta spacecraft has been regularly imaging the nucleus of comet 67P/Churyumov-Gerasimenko (hereafter 67P) and its dust and gas environment in 25 broad- and narrow-band filters covering the wavelength range 240 to 1000 nm through its narrow- and wide-angle cameras, NAC and WAC [details in supplementary text 1 (SM1)]. The data thus far have provided a view of this comet at variance with our earlier knowledge [reviewed in (*2*)].

Bulk properties and rotation

The rotation period determined from early, unresolved observations was 12.4043 ± 0.0007 hours (*3*), implying a decrease of the spin period by 0.36 hours during (or since) the 2009 perihelion passage, consistent with the range predicted for the 2015 perihelion passage (*4*).

The nucleus was first resolved in the NAC (>1 pixel) on 16 June 2014, at a distance of 192,000 km, and the resolution has improved to 0.15 m/pixel at a distance of 10 km at the time of writing.

Starting in August 2014, when the comet was well enough resolved, a more accurate approach based on tracking stereo landmarks was used to determine the direction of the spin axis. Three different algorithms led to a prograde spin axis at RA = 69.3° ± 0.1°, Dec = 64.1° ± 0.1°, in good agreement with (*3*). We found no obvious evidence for complex rotation, and the current result constrains any motion of the spin axis to <0.3° over ~55 days. The lack of measurable change over such a short time scale is not yet a useful constraint on torques.

Early shape models of the nucleus were constructed from images acquired through September 2014, with a best resolution of 0.8 m/pixel. Stereophotogrammetry [(*5*) and references therein] and stereophotoclinometry (*6*) yielded high-resolution (5 to 10 m) shape models (Fig. 1). The shape has a very pronounced bilobe appearance, reminiscent of comets 8P/Tuttle (*7*) and 103P/Hartley 2 (*8*). The lobes in these two comets are aligned roughly along their longest axes, whereas in 67P the alignment appears more nearly perpendicular to the axes of the individual lobes, and the axis of rotation is closer to parallel to the long axes of the lobes. From the three-dimensional shape models, the larger lobe (the body) has a size of about 4.1 × 3.3 × 1.8 km, and the smaller lobe (the head) is 2.6 × 2.3 × 1.8 km; they are connected by a short "neck." The total volume, estimated by adding limb scans from the microwave instrument onboard Rosetta (MIRO) to fill in the unmapped portions around the southern (negative) rotational pole, is 21.4 ± 2.0 km^3.

The current shape model is not complete because the obliquity of the comet's rotational axis (52°) currently puts the southern (negative) pole in permanent shadow (it will have continuous sunlight at perihelion, 13 August 2015; equinox in May 2015). The volume of the model yields a mean density of 470 ± 45 kg/m^3 when combined with the mass, 1.0×10^{13} kg, determined by the Radio Science Investigation (RSI) instrument (*9*). Although the principal axes of the shape model are still uncertain because of the unmapped portion of the nucleus, they are consistent with simple rotation about the axis of maximum moment of inertia, assuming homogeneous density. The determined density implies high porosity, in the range of 70 to 80% depending on the adopted dust-to-ice mass ratio, or equivalently the bulk density for a solid mixture of ice and dust, which we assume to be 1500 to 2000 kg/m^3.

[1]Max-Planck-Institut für Sonnensystemforschung, 37077 Göttingen, Germany. [2]Department of Physics and Astronomy, University of Padova, 35122 Padova, Italy. [3]Laboratoire d'Astrophysique de Marseille, UMR 7326, CNRS & Aix Marseille Université, 13388 Marseille Cedex 13, France. [4]Centro de Astrobiologia, CSIC-INTA, 28850 Madrid, Spain. [5]International Space Science Institute, 3012 Bern, Switzerland. [6]Scientific Support Office, European Space Research and Technology Centre/ESA, 2201 AZ Noordwijk ZH, Netherlands. [7]Department of Physics and Astronomy, Uppsala University, 75120 Uppsala, Sweden. [8]PAS Space Research Center, 00716 Warszawa, Poland. [9]Institut für Geophysik und extraterrestrische Physik, Technische Universität Braunschweig, 38106 Braunschweig, Germany. [10]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Planetenforschung, 12489 Berlin, Germany. [11]Department of Astronomy, University of Maryland, College Park, MD 20742, USA. [12]Akademie der Wissenschaften zu Göttingen, 37077 Göttingen, Germany. [13]Department of Industrial Engineering, University of Padova, 35131 Padova, Italy. [14]Aix Marseille Université, CNRS, Laboratoire d'Astrophysique de Marseille, UMR 7326, 13388 Marseille, France. [15]LESIA–Observatoire de Paris, CNRS, Université Pierre et Marie Curie, Université Paris Diderot, 92195 Meudon, France. [16]LATMOS, CNRS/UVSQ/IPSL, 78280 Guyancourt, France. [17]Centro di Ateneo di Studi ed Attività Spaziali "Giuseppe Colombo" (CISAS), University of Padova, 35131 Padova, Italy. [18]INAF, Osservatorio Astronomico di Padova, 35122 Padova, Italy. [19]CNR-IFN UOS Padova LUXOR, 35131 Padova, Italy. [20]University of Trento, 38100 Trento, Italy. [21]INAF - Osservatorio Astronomico di Trieste, 34014 Trieste, Italy. [22]Planetary Science Institute, Tucson, AZ 85719, USA. [23]Instituto de Astrofisica de Andalucia (CSIC), c/ Glorieta de la Astronomìa s/n, 18008 Granada, Spain. [24]Graduate Institute of Astronomy, National Central University, Chung-Li 32054, Taiwan. [25]Space Science Institute, Macau University of Science and Technology, Macao, China. [26]Operations Department, European Space Astronomy Centre/ESA, 28691 Villanueva de la Canada, Madrid, Spain. [27]Solar System Exploration Research, Virtual Institute, Southwest Research Institute, Boulder, CO 80302, USA. [28]Dipartimento di Geoscienze, University of Padova, 35131 Padova, Italy. [29]Institut für Datentechnik und Kommunikationsnetze der TU Braunschweig, 38106 Braunschweig, Germany. [30]Department of Information Engineering, University of Padova, 35131 Padova, Italy. [31]Instituto Nacional de Tecnica Aeroespacial, 28850 Torrejon de Ardoz, Madrid, Spain. [32]Physikalisches Institut, University of Bern, 3012 Bern, Switzerland. [33]Center for Space and Habitability, University of Bern, 3012 Bern, Switzerland. [34]Rheinisches Institut für Umweltforschung, Abteilung Planetenforschung, Universität zu Köln, 50931 Köln, Germany.

***Corresponding author. E-mail: sierks@mps.mpg.de** †Present address: Planetary and Space Sciences, Department of Physical Sciences, Open University, Milton Keynes MK7 6AA, UK.

This low density confirms previous results for other comets based on indirect and less precise determinations, mostly from modeling nongravitational accelerations but also including the breakup of comet D/Shoemaker-Levy 9 (1993 F2),

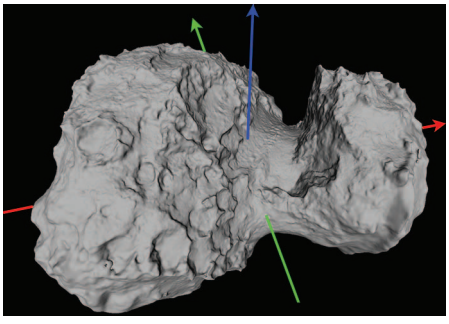


Fig. 1. Stereophotogrammetric shape models of comet 67P. The blue arrow indicates 67P's rotation axis *z*; the red and green arrows display its equatorial *x* and *y* axes (*x* according to the current zero-longitude definition). If the nucleus is homogeneous in density, the axis of minimum moment of inertia is consistent with being in the equatorial plane.

the fallback under gravity of the Deep Impact ejecta, and the lower limit on porosity (75%) of the upper 1 to 10 m from the faintness of the impact flash of Deep Impact [e.g., (*10*–*13*); review in (*14*)]. In fact, the mass is almost the same as deduced from the nongravitational acceleration but with larger uncertainty (*10*). Thus, we now constrain the porosity of a cometary nucleus even more tightly. The agreement between the axis of rotation and the axis of maximum moment of inertia reinforces the concept of weakly bonded icy dust aggregates, with porosity at small scales relative to the size of the nucleus (rather than large voids). This conclusion is further strengthened by the very low strength (10 Pa) deduced by (*15*).

We have calculated the gravitational and centrifugal forces on the same shape model. As shown (Fig. 2), the gravitational potential varies by less than a factor of 2. The centrifugal force varies from negligible to one-third of the gravitational force. However, the surface slope relative to local gravity varies markedly, with many areas exceeding 45° (Fig. 2). Escape velocity at the surface is poorly defined because nuclear rotation causes the gravitational acceleration to vary by a large factor in the time taken to move 1 nuclear radius. For a sphere with the same density and volume (radius = 1.72 km), the escape velocity would be 0.9 m/s, roughly three times the velocity deduced via the same approximation for comet 103P/Hartley 2 (*16*).

67P's two main lobes show considerable morphological diversity (*15*), and the neck is different from both. Geomorphological regions are identified in Fig. 3 [see (*15*) for more detail] and, following the ancient Egyptian theme of the Rosetta mission, they are named for Egyptian deities. The entire nucleus is dark, yielding a geometric albedo of the entire nucleus of only 5.9 ± 0.2% at 550 nm, similar to that of comet 9P/Tempel 1 (*17*). Despite the extreme morphological diversity, colorimetry of the nucleus is remarkably uniform, except in the Hapi region on the neck (SM2). Whether this is due to redeposited material on both lobes or implies that both lobes are intrinsically similar is still an open question.

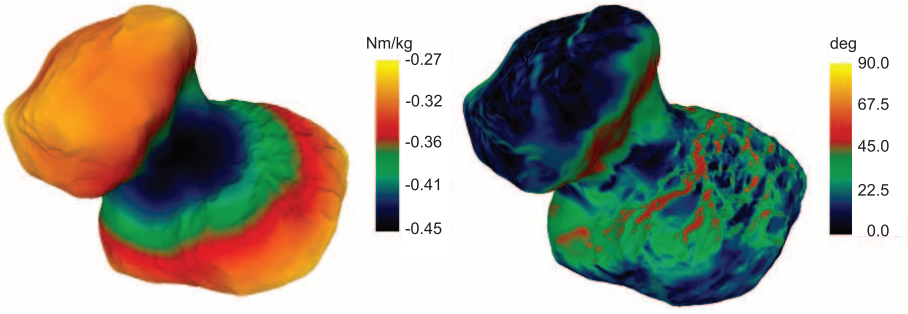


Fig. 2. Gravitational field. Left: The effective gravitational potential at the surface, including the centrifugal term due to rotation of the nucleus, is shown looking toward the neck from an intermediate latitude in the northern (positive) hemisphere (between the blue and red arrows in Fig. 1). Right: The slope of the terrain relative to local gravity.

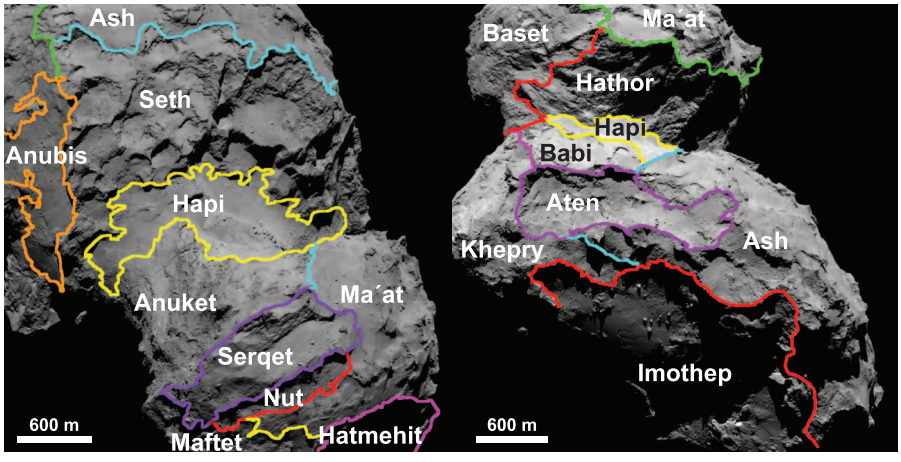


Fig. 3. Geomorphological map of comet 67P. Left: Region definition for one face of 67P, showing the Seth region on the body and the smooth Hapi region on the neck. Other regions are defined in more detail by (*15*). Right: Regional definition looking from the body (foreground) across to the head.

Activity and its source

OSIRIS has detected the presence of a dust coma since the pronounced outburst between 27 and 30 April 2014, or even earlier (*18*). We detected resolved features in the coma (SM2) in long WAC exposures with the 610-nm filter (SM1) at the end of July 2014, when the comet was 3.7 AU from the Sun and the spacecraft distance was 3000 km. Most of the activity was unambiguously coming out of the transition region between the small and large lobes of the nucleus, the Hapi region, or very close to it. This was consistent with ground-based observations of coma structures over the last two orbital periods, which implied an active region at high northern (positive) latitude (*19*). When the spacecraft distance dropped to 100 km (resulting in a resolution of 1.8 m/pixel in NAC, 10 m/pixel in WAC), it was possible to localize the active sources by inversion of the jets. Most of the jets arose from Hapi, at about +60° latitude (Fig. 4), although we found other minor active spots on both lobes of the nucleus (SM3). This major coma feature presents a diurnal variation of intensity due to changing insolation conditions and observational geometry: A planar fan-like jet appears brighter and more focused when viewed edge-on rather than face-on.

To investigate why the activity is dominant above Hapi, we calculated the energy incident on all parts of the nucleus over one rotation on 6 August 2014. According to the model as applied to the shape model of 67P (*20*), self-heating by thermal reradiation from the head and the body provides extra heating to the region of the neck and at the time of peak jet activity, and parts of Hapi and Hathor receive 10% higher flux than other regions. However, at this point in the orbit, the rotational axis is oriented so that the neck receives slightly less energy over a rotation than do other parts of the nucleus (Fig. 5). This suggests a large compositional (e.g., type of ice) or structural (e.g., depth of ice) difference in Hapi relative to other regions. More detailed models that take sublimation and thermal inertia into account are needed to clarify the situation. Because of the large obliquity of the rotational axis with respect to the orbital plane, it is likely that the currently

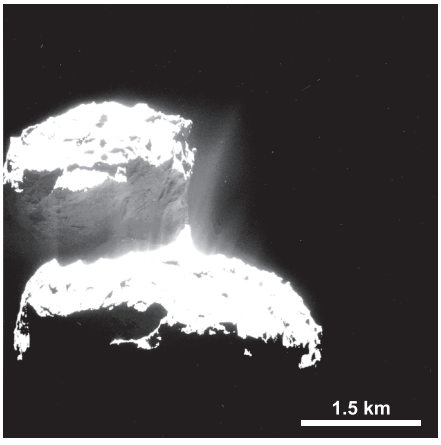


Fig. 4. Jets. Jets from the Hapi region as observed on 23 September 2014.

active northern (positive) hemisphere has evolved very differently than has the southern (negative, thus far only partially mapped) hemisphere.

Nuclear properties and geomorphology

In the Seth region, extending around the neck from one end of Hapi (Fig. 3), we observe a closely packed system of well-defined pits and depressions with remarkably flat floors. The walls of these features show linear structures that are parallel to their floors and extend laterally hundreds of meters, indicating a consolidated structure, but they also exhibit vertical striations. Some of these quasi-circular pits are minor sources of activity (e.g., Fig. 6). They are similarly seen in cliffs and are possibly erosion features related to activity. They range from 50 to 300 m in diameter and 10 to 200 m in depth. We find other pits everywhere on the nucleus. Such pits have been observed on comets 81P/Wild 2 and 9P/Tempel 1, and those were interpreted as activity-related features rather than impact craters because of their flat size distribution (*21*). The pits on the nucleus of 67P share a similar morphology, although the active ones tend to have a depth/diameter ratio close to 1, whereas the inactive pits are much shallower, seemingly filled with fine dust and multiple boulders. It is not clear whether these pits are inactive or whether they will "wake up" when they start receiving more illumination.

The head is characterized by impressive cliffs and sets of aligned linear structures >500 m in length exposed at the cliff facing the neck (Hathor) (Fig. 7). This cliff is opposite Seth, which has a different character. Fracturing is seen over the nucleus at all scales, except where the nucleus appears to be covered with smooth deposits. Small-scale fracturing appears more scattered than the larger-scale linear structures. Thermal shock may be a plausible mechanism given the very large variation in temperatures seen by cometary nuclei; the near-surface region of 67P undergoes fluctuations of up to 150 K each orbital period, much larger than those invoked to explain linear structures on Eros (*22*). The surface morphology strongly suggests that the loss of larger volumes of material, possibly via the explosive release of subsurface gas pressure or creation of overhangs by sublimation, may be a major mass loss process for the nucleus.

We observe irregular quasi-linear features at the Anuket-Hathor interface, which we infer to be cracks in the surface of the neck (Fig. 7). These cracks include ones that are roughly parallel to the neck. We note that within the Hapi region there is also an open crack, which is similarly aligned (Fig. 7) and extends well into Anuket. The fact that they are more or less parallel and across the neck suggests that they may be related to each other and to large-scale phenomena on the nucleus, such as flexure between the head and the body.

There are also clusters of small, bright spots (0.5 to 1 m), which might be ice-rich, and there are both groups and isolated examples of much larger clumps (up to 30 m).

Finally, we have found in several places on the nucleus, particularly on very steep slopes, a feature that we colloquially term "goosebumps" (Fig. 8). The bumps themselves exhibit a characteristic scale of about 3 m. Although we do not yet have an interpretation of these features, the fact that a single characteristic scale is present may be an important clue to formation processes, as many of the possible processes have characteristic scales.

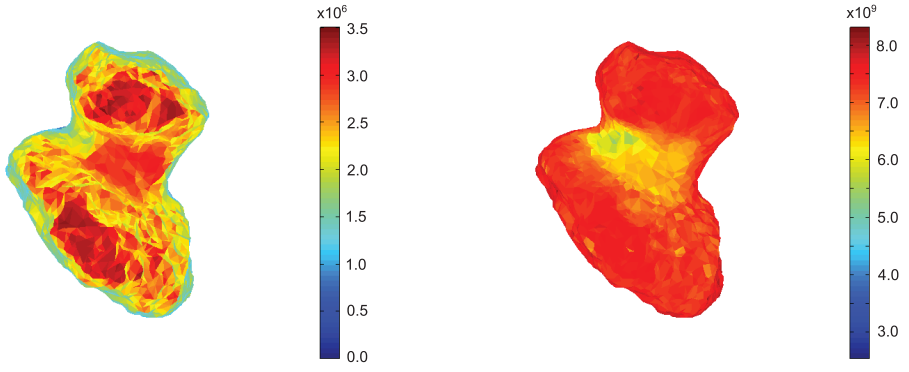


Fig. 5. Map of energy input. Left: A map looking at the northern (right-hand rule, positive) pole of 67P showing the total energy received from the Sun per rotation on 6 August 2014. The energy received includes thermal illumination by the surfaces of the comet itself. The base of the neck (Hapi) receives ~15% less energy than the most illuminated region, 3.5×10^6 J m^{-2} (per rotation). If self-heating were not included, the base of the neck would receive ~30% less total energy. Right: Similar to the left panel but showing total energy received over an entire orbital period in J m^{-2} (per orbit). This heating varies by only ~50% over the entire surface. Although not shown here, the opposite hemisphere receives essentially the same energy, but at a higher rate over a shorter time.

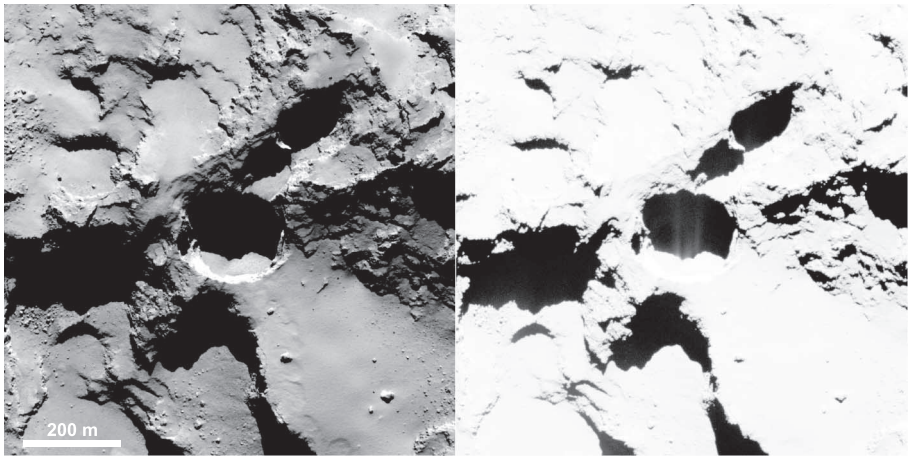


Fig. 6. Active pits detected in Seth region. NAC image acquired on 28 August 2014. Distance to the comet, 60 km; resolution, 1 m/pixel. Enhancing the contrast (right panel) shows fine structures in the shadow of the pit, interpreted as fine jets arising from the pit.

The origin of 67P/Churyumov-Gerasimenko

One key question about the origin of 67P is whether its nucleus is a primordial planetesimal or was formed from a much larger planetesimal. To address this, we must ask how the present 67P nucleus evolved since it formed eons ago.

Two modes of formation of original planetesimals in the outer parts of the solar system

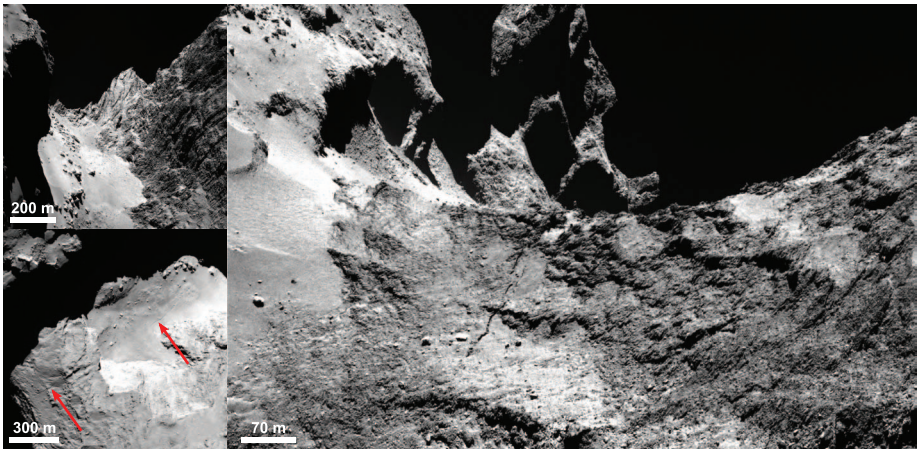


Fig. 7. Nucleus close-ups. Top left: The Hathor cliff face is to the right in this view. The aligned linear structures can be clearly seen. The smooth Hapi region is seen at the base of the Hathor cliff. Boulders are prevalent along the long axis of the Hapi region. Bottom left and right: Crack in the Hapi region. The left panel shows the crack (indicated by red arrows) extending across Hapi and beyond. The right panel shows the crack where it has left Hapi and is extending into Anuket, with Seth at the uppermost left and Hapi in the lower left.

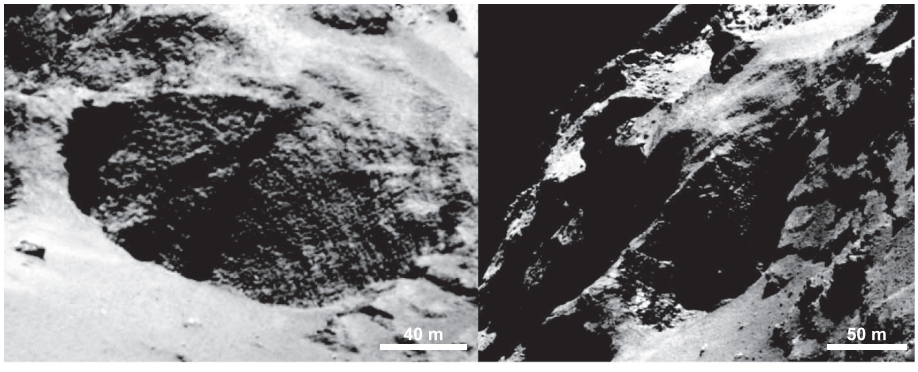


Fig. 8. Goosebumps. Characteristic scale of all the bumps is ~3 m, extending over >100 m. This example is in the active pit in the Seth region.

are well developed: hierarchical accretion (*23*) and clumping of "pebbles" due to gas-dust instabilities. In both cases, accretional velocities were likely very low. The pebble swarms caused by the dynamical instabilities may form objects as large as 10^2 to 10^3 km or as small as the 67P nucleus (*24*), although the goosebumps mentioned above suggest that a smaller scale could be important. Small objects could also form later as individual or reaccreted fragments produced by collisions involving large transneptunian objects (*25*), but these may be different from the small, original planetesimals, having experienced the radioactive heating of their large parent bodies and the impacts whereby they were ejected.

The amount of previous erosion of the 67P nucleus depends on the dynamical age of the comet—that is, the number of orbits it has performed as a member of the Jupiter family after its capture from the transneptunian reservoir (SM4). Unfortunately, this number is unknown and will remain so, because in the presence of frequent close encounters with Jupiter inducing strong orbital perturbations, the dynamical evolution of Jupiter family comets is strongly chaotic. Thus, the orbital history of 67P can be traced back past a close encounter in 1959 (minimum distance from Jupiter = 0.05 AU) [(*26*, *27*); see (*28*) and SM4 for sample calculations] but not much further.

A key question that Rosetta will address during the escort phase of the mission is whether the two lobes of the nucleus are separate cometesimals or whether the neck region has been carved out by erosion. If the head and the body turn out to be very different in composition or internal structure, this would argue for the idea that they are original cometesimals. If they are similar in composition and structure, the conclusion is less clear-cut. The two lobes, head and body, could still be separate cometesimals that formed at essentially the same heliocentric distance and time. Whether the neck could also be the result of erosion of a more convex body over many orbital revolutions remains to be clarified. Our study of energy input described above shows that this is not enough focused to the neck (SM5), but the question of whether fragmentation can still be sufficiently localized to cause the carving remains unresolved.

Some key features relevant to the formation are the high porosity implied by the bulk density (SM6), the pristine composition evidenced by the large abundance of CO (*29*), and the possible evidence of fracturing in the head and neck region (SM7). The primordial accretion model seems consistent with all observations. The scenario of colliding large bodies seems to require more special circumstances, tending to make it less attractive but still viable.

REFERENCES AND NOTES

1. H. U. Keller *et al*., OSIRIS—The scientific camera system onboard Rosetta. *Space Sci. Rev.* **128**, 433–506 (2007). doi: 10.1007/s11214-006-9128-4
2. P. L. Lamy *et al*., A portrait of the nucleus of comet 67P/Churyumov-Gerasimenko. *Space Sci. Rev.* **128**, 23–66 (2007). doi: 10.1007/s11214-007-9146-x
3. S. Mottola *et al*., The rotation state of 67P/Churyumov-Gerasimenko from approach observations with the OSIRIS cameras on Rosetta. *Astron. Astrophys*. **569**, id.L2 (2014). doi: 10.1051/0004-6361/201424590
4. P. J. Gutiérrez, L. Jorda, N. H. Samarasinha, P. Lamy, Outgassing-induced effects in the rotational state of comet 67P/Churyumov-Gerasimenko during the Rosetta mission. *Planet. Space Sci.* **53**, 1135–1145 (2005). doi: 10.1016/j.pss.2004.12.012
5. F. Preusker *et al*., The northern hemisphere of asteroid (21) Lutetia—topography and orthoimages from Rosetta OSIRIS NAC image data. *Planet. Space Sci.* **66**, 54–63 (2012). doi: 10.1016/j.pss.2012.01.008
6. R. W. Gaskell *et al*., Characterizing and navigating small bodies with imaging data. *Meteorit. Planet. Sci.* **43**, 1049–1061 (2008). doi: 10.1111/j.1945-5100.2008.tb00692.x
7. J. K. Harmon, M. C. Nolan, J. D. Giorgini, E. S. Howell, Radar observations of 8P/Tuttle: A contact-binary comet. *Icarus* **207**, 499–502 (2010). doi: 10.1016/j.icarus.2009.12.026
8. P. C. Thomas *et al*., Shape, density, and geology of the nucleus of Comet 103P/Hartley 2. *Icarus* **222**, 550–558 (2013). doi: 10.1016/j.icarus.2012.05.034
9. ESA, "Determining the mass of comet 67P/C-G" (blog entry); http://blogs.esa.int/rosetta/2014/08/21/determining-the-mass-of-comet-67pc-g/.
10. B. J. R. Davidsson, P. J. Gutiérrez, Nucleus properties of comet 67P/Churyumov-Gerasimenko estimated from non-gravitational force modeling. *Icarus* **176**, 453–477 (2005). doi: 10.1016/j.icarus.2005.02.006
11. E. Asphaug, W. Benz, Density of comet Shoemaker-Levy 9 deduced by modelling breakup of the parent 'rubble pile'. *Nature* **370**, 120–124 (2004). doi: 10.1038/370120a0
12. C. M. Ernst, P. H. Schultz, Evolution of the Deep Impact flash: Implications for the nucleus surface based on laboratory experiments. *Icarus* **190**, 334–344 (2007). doi: 10.1016/j.icarus.2007.03.030
13. J. E. Richardson, H. J. Melosh, C. M. Lisse, B. Carcich, A ballistics analysis of the deep impact ejecta plume: Determining comet Tempel 1's gravity, mass, and density. *Icarus* **190**, 357–390 (2007). doi: 10.1016/j.icarus.2007.08.001
14. M. F. A'Hearn, Comets as building blocks. *Annu. Rev. Astron. Astrophys.* **49**, 281–299 (2011). doi: 10.1146/annurev-astro-081710-102506
15. N. Thomas *et al*., The morphological diversity of comet 67P/Churyumov-Gerasimenko. *Science* **347**, aaa0440 (2015).
16. B. Hermalyn *et al*., The detection, localization, and dynamics of large icy particles surrounding comet 103P/Hartley 2. *Icarus* **222**, 625–633 (2013). doi: 10.1016/j.icarus.2012.09.030
17. J.-Y. Li *et al*., Deep Impact photometry of comet 9P/Tempel 1. *Icarus* **191**, 161–175 (2007). doi: 10.1016/j.icarus.2006.09.027
18. C. Tubiana *et al*., 67P/Churyumov-Gerasimenko: Activity between March and June 2014 as observed from Rosetta/OSIRIS. *Astron. Astrophys*. **573**, id.A62 (2015). doi: 10.1051/0004-6361/201424735
19. J.-B. Vincent, L. M. Lara, G. P. Tozzi, Z.-Y. Lin, H. Sierks, Spin and activity of comet 67P/Churyumov-Gerasimenko. *Astron. Astrophys.* **549**, id.A121 (2013). doi: 10.1051/0004-6361/201219350
20. B. J. R. Davidsson, H. Rickman, Surface roughness and three-dimensional heat conduction in thermophysical models. *Icarus* **243**, 58–77 (2014). doi: 10.1016/j.icarus.2014.08.039
21. M. J. S. Belton, P. D. Feldman, M. F. A'Hearn, B. Carcich, Cometary cryo-volcanism: Source regions and a model for the UT 2005 June 14 and other mini-outbursts on Comet 9P/Tempel 1. *Icarus* **198**, 189–207 (2008). doi: 10.1016/j.icarus.2008.07.009

22. D. A. J. Dombard, A. M. Freed, Thermally induced lineations on the asteroid Eros: Evidence of orbit transfer. *Geophys. Res. Lett.* **29**, 65-1–65-4 (2002). doi: 10.1029/2002GL015181
23. S. J. Weidenschilling, From icy grains to comets. In *Comets II*, M. C. Festou, H. U. Keller, H. A. Weaver, Eds. (University of Arizona Press, Tucson, 2004), pp. 97–104.
24. K. W. Janson, A. Johansen, Formation of pebble-pile planetesimals. *Astron. Astrophys.* **570**, A47 (2014). doi: 10.1051/0004-6361/201424369
25. D. R. Davis, P. Farinella, Collisional evolution of Edgeworth-Kuiper belt objects. *Icarus* **125**, 50–60 (1997). doi: 10.1006/icar.1996.5595
26. A. Carusi, L. Kresák, E. Perozzi, G. B. Valsecchi, *Long-Term Evolution of Short-Period Comets* (Adam-Hilger, Bristol, UK, 1985).
27. N. A. Belyaev, L. Kresak, E. M. Pittich, A. N. Pushkarev, *Catalogue of Short-Period Comets* (Astronomical Institute of the Slovak Academy of Sciences, Bratislava, 1986).
28. O. Groussin, G. Hahn, P. Lamy, R. Gonczi, G. B. Valsecchi, The long-term evolution and initial size of comets 46P/Wirtanen and 67P/Churyumov-Gerasimenko. *Mon. Not. R. Astron. Soc.* **376**, 1399–1406 (2007). doi: 10.1111/j.1365-2966.2007.11553.x
29. M. Haessig *et al.*, The variability and heterogeneity in the coma of 67P/Churyumov-Gerasimenko. *Science* **347**, aaa0276 (2015).

ACKNOWLEDGMENTS

In memory of our colleague and co-investigator, Michel Festou. We thank our retired co-investigators M. E. Bailey, P. Barthol, J. A. Burns, W. Curdt, F. Gliem, A. Llebaria, and J. P. Sivan for their contribution to the success of OSIRIS, as well as the team of optical, electrical, mechanical, thermal, and software engineers that turned the instrument into reality and supported it throughout in flight operations. OSIRIS was built by a consortium led by the Max-Planck-Institut für Sonnensystemforschung, Göttingen, Germany, in collaboration with CISAS, University of Padova, Italy, the Laboratoire d'Astrophysique de Marseille, France, the Instituto de Astrofísica de Andalucia, CSIC, Granada, Spain, the Scientific Support Office of the European Space Agency, Noordwijk, Netherlands, the Instituto Nacional de Técnica Aeroespacial, Madrid, Spain, the Universidad Politéchnica de Madrid, Spain, the Department of Physics and Astronomy of Uppsala University, Sweden, and the Institut für Datentechnik und Kommunikationsnetze der Technischen Universität Braunschweig, Germany. The support of the national funding agencies of Germany (DLR), France (CNES), Italy (ASI), Spain (MEC), Sweden (SNSB), and the ESA Technical Directorate is gratefully acknowledged. We acknowledge the support of the local institutes that realized the instrument. Also supported by NASA JPL contract 1267923 to the University of Maryland (M.F.A'H.), grant NSC 101-2111-M-008-016 from the Ministry of Science and Technology, Taiwan (W.H.I.), grant 2011/01/B/ST9/05442 from the Polish National Science Center and grant 74/10:2 of the Swedish National Space Board (H.R.), and travel grants from the Swiss National Science Foundation (N.T.). All data presented in this paper will be delivered to ESA's Planetary Science Archive and NASA's Planetary Data System in accordance with the schedule established by the Rosetta project and will be available on request prior to that archiving. We thank the Rosetta Science Ground Segment at ESAC, the Rosetta Mission Operations Centre at ESOC, and the Rosetta Project at ESTEC for their outstanding work enabling the science return of the Rosetta Mission. The OSIRIS team thanks the MIRO team and its principal investigator, S. Gulkis, for providing prepublication data and helpful discussion regarding use of their limb scans to determine the shape.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/aaa1044/suppl/DC1
Supplementary Text
Figs. S1 to S7
Table S1
References (*30*–*42*)

17 October 2014; accepted 30 December 2014
10.1126/science.aaa1044

COMETARY SCIENCE

Dust measurements in the coma of comet 67P/Churyumov-Gerasimenko inbound to the Sun

Alessandra Rotundi,[1,2]* Holger Sierks,[3] Vincenzo Della Corte,[1] Marco Fulle,[4] Pedro J. Gutierrez,[5] Luisa Lara,[5] Cesare Barbieri,[6] Philippe L. Lamy,[7] Rafael Rodrigo,[8,9] Detlef Koschny,[10] Hans Rickman,[11,12] Horst Uwe Keller,[13] José J. López-Moreno,[5] Mario Accolla,[1,2] Jessica Agarwal,[3] Michael F. A'Hearn,[14] Nicolas Altobelli,[15] Francesco Angrilli,[16] M. Antonietta Barucci,[17] Jean-Loup Bertaux,[18] Ivano Bertini,[19] Dennis Bodewits,[14] Ezio Bussoletti,[2] Luigi Colangeli,[20] Massimo Cosi,[21] Gabriele Cremonese,[22] Jean-Francois Crifo,[18] Vania Da Deppo,[23] Björn Davidsson,[11] Stefano Debei,[16] Mariolino De Cecco,[24] Francesca Esposito,[25] Marco Ferrari,[1,2] Sonia Fornasier,[17] Frank Giovane,[26] Bo Gustafson,[27] Simon F. Green,[28] Olivier Groussin,[7] Eberhard Grün,[29] Carsten Güttler,[3] Miguel L. Herranz,[5] Stubbe F. Hviid,[30] Wing Ip,[31] Stavro Ivanovski,[1] José M. Jerónimo,[5] Laurent Jorda,[7] Joerg Knollenberg,[30] Rainer Kramm,[3] Ekkehard Kührt,[30] Michael Küppers,[15] Monica Lazzarin,[6] Mark R. Leese,[28] Antonio C. López-Jiménez,[5] Francesca Lucarelli,[2] Stephen C. Lowry,[32] Francesco Marzari,[33] Elena Mazzotta Epifani,[25] J. Anthony M. McDonnell,[28,32,34] Vito Mennella,[25] Harald Michalik,[35] Antonio Molina,[36] Rafael Morales,[5] Fernando Moreno,[5] Stefano Mottola,[30] Giampiero Naletto,[37] Nilda Oklay,[3] José L. Ortiz,[5] Ernesto Palomba,[1] Pasquale Palumbo,[1,2] Jean-Marie Perrin,[18,38] Julio Rodríguez,[5] Lola Sabau,[39] Colin Snodgrass,[3,28] Roberto Sordini,[1] Nicolas Thomas,[40] Cecilia Tubiana,[3] Jean-Baptiste Vincent,[3] Paul Weissman,[41] Klaus-Peter Wenzel,[10] Vladimir Zakharov,[17] John C. Zarnecki[9,28]

Critical measurements for understanding accretion and the dust/gas ratio in the solar nebula, where planets were forming 4.5 billion years ago, are being obtained by the GIADA (Grain Impact Analyser and Dust Accumulator) experiment on the European Space Agency's Rosetta spacecraft orbiting comet 67P/Churyumov-Gerasimenko. Between 3.6 and 3.4 astronomical units inbound, GIADA and OSIRIS (Optical, Spectroscopic, and Infrared Remote Imaging System) detected 35 outflowing grains of mass 10^{-10} to 10^{-7} kilograms, and 48 grains of mass 10^{-5} to 10^{-2} kilograms, respectively. Combined with gas data from the MIRO (Microwave Instrument for the Rosetta Orbiter) and ROSINA (Rosetta Orbiter Spectrometer for Ion and Neutral Analysis) instruments, we find a dust/gas mass ratio of 4 ± 2 averaged over the sunlit nucleus surface. A cloud of larger grains also encircles the nucleus in bound orbits from the previous perihelion. The largest orbiting clumps are meter-sized, confirming the dust/gas ratio of 3 inferred at perihelion from models of dust comae and trails.

Although accurate measurements of the gas loss rate from comets are possible under favorable conditions even from Earth, estimates of the dust loss rate have been more uncertain thus far. Multiparametric models are needed to extract global dust parameters from the dust features of comets (coma, tail, and trail) observed from the ground and Earth-orbiting telescopes, and it is often difficult to establish the uniqueness of these model results. Past space missions had onboard instruments devoted to the measurement of the dust flux. Because all of these missions were fast flybys, it was impossible to disentangle the dust grains coming directly from the nucleus from those reflected back by solar radiation pressure (*1*, *2*). The latter component could explain part of the excess of millimeter-sized particles, showing as a bulge of the size distribution above the fitting power law derived by the observations performed during flybys at 1P/Halley (*3*) and at short-period comets 26P/Grigg-Skjellerup, 81P/Wild 2, and 9P/Tempel 1 (*4*–*6*).

An even more severe bias could affect all estimates of dust/gas ratio obtained so far in comets. The dust/gas ratio measured in 1P/Halley was close to 2 (*5*), but this number is valid up to the largest mass of ~1 g observed by the DIDSY (Dust Impact Detection System) detector (actually, this largest-mass grain was invoked to explain the spacecraft precession-inducing impact that occurred just before closest approach). Theoretical models predict that 1P/Halley was then ejecting larger masses, and the dust/gas ratio strongly depends on the actual largest grain ejected in the coma. Because it was impossible to fix the size distribution between 1 g and the unknown largest ejected mass, we cannot exclude dust/gas values a factor of 10 higher. In this paper, we show that for comet 67P/Churyumov-Gerasimenko (67P hereafter), we can disentangle the two families of ejected grains (direct and reflected) and extract the dust size distribution up to the largest ejected grain, obtaining for the first time an accurate estimate of the dust/gas ratio.

[1]Istituto di Astrofisica e Planetologia Spaziali, Istituto Nazionale di Astrofisica (INAF), Via Fosso del Cavaliere, 100, 0133 Rome, Italy. [2]Università degli Studi di Napoli "Parthenope," Dipartimento di Scienze e Tecnologie, CDN IC4, 80143 Naples, Italy. [3]Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg, 3, 37077 Göttingen, Germany. [4]Osservatorio Astronomico, INAF, Via Tiepolo 11, 34143 Trieste, Italy. [5]Instituto de Astrofísica de Andalucía, Consejo Superior de Investigaciones Científicas (CSIC), P.O. Box 3008, 18080 Granada, Spain. [6]Department of Physics and Astronomy, Padova University, Vicolo dell'Osservatorio 3, 35122 Padova, Italy. [7]Laboratoire d'Astrophysique de Marseille, UMR 7326, CNRS and Aix-Marseille Université, 13388 Marseille, France. [8]Centro de Astrobiologia (Instituto Nacional de Tecnica Aerospacial–CSIC), 28691 Villanueva de la Canada, Madrid, Spain. [9]International Space Science Institute, Hallerstrasse 6, CH-3012 Bern, Switzerland. [10]Scientific Support Office, European Space Agency, 2201 Noordwijk, Netherlands. [11]Department of Physics and Astronomy, Uppsala University, Box 516, 75120 Uppsala, Sweden. [12]Polish Academy of Sciences Space Research Center, Bartycka 18A, PL-00716 Warszawa, Poland. [13]Institute for Geophysics and Extraterrestrial Physics, Technische Universität Braunschweig, Braunschweig 38106, Germany. [14]Department of Astronomy, University of Maryland, College Park, MD 20742-2421, USA. [15]European Space Agency (ESA)–European Space Astronomy Center (ESAC), Camino Bajo del Castillo, s/n, 28692 Villanueva de la Cañada, Madrid, Spagna. [16]Department of Mechanical Engineering, University of Padova, via Venezia 1, 35131 Padova, Italy. [17]Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique, Observatoire de Paris, CNRS, Université Paris 06, Université Paris-Diderot, 5 place Johannes Janssen, 92195 Meudon, France. [18]Laboratoire Atmosphères, Milieux, Observations Spatiales, CNRS/Université de Versailles Saint-Quentin-en-Yvelines/Institut Pierre-Simon Laplace, 11 boulevard d'Alembert, 78280 Guyancourt, France. [19]University of Padova, Centro Interdipartimentale di Studi e Attività Spaziali (CISAS), via Venezia 15, 35100 Padova, Italy. [20]ESA, European Space Research and Technology Centre (ESTEC), Keplerlaan 1, 2201 AZ Noordwijk, Netherlands. [21]Selex-ES, Via Albert Einstein, 35, 50013 Campi Bisenzio, Firenze, Italy. [22]Osservatorio Astronomico di Padova, INAF, Vicolo dell'Osservatorio 5, 35122 Padova, Italy. [23]Consiglio Nazionale delle Ricerche–Istituto di Fotonica e Nanotecnologie–Unità Operativa di Supporto Padova LUXOR, via Trasea 7, 35131 Padova, Italy. [24]Università di Trento, via Mesiano, 77, 38100 Trento, Italy. [25]Osservatorio Astronomico di Capodimonte, INAF, Salita Moiariello, 16, 80133 Naples, Italy. [26]Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. [27]University of Florida, Gainesville, FL 32611, USA. [28]Planetary and Space Sciences, Department of Physical Sciences, The Open University, Milton Keynes MK7 6AA, UK. [29]Max-Planck-Institut fuer Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany. [30]Institute of Planetary Research, Deutsches Zentrum für Luft- und Raumfahrt (DLR), Rutherfordstrasse 2, 12489 Berlin, Germany. [31]Institute for Space Science, National Central University, 300 Chung Da Road, 32054 Chung-Li, Taiwan. [32]The University of Kent, School of Physical Sciences, Canterbury, Kent CT2 7NZ, UK. [33]Department of Physics, University of Padova, 35131 Padova, Italy. [34]Unispacekent, Canterbury CT2 8EF, UK. [35]Institut für Datentechnik und Kommunikationsnetze, 38106 Braunschweig, Germany. [36]Departamento de Física Aplicada, Universidad de Granada, Facultad de Ciencias, Avenida Severo Ochoa, s/n, 18071 Granada, Spain. [37]Department of Information Engineering, Padova University, via Gradenigo 6, 35131 Padova, Italy. [38]Observatoire de Haute Provence OSU Pythéas UMS 2244 CNRS-AMU, 04870 Saint Michel l'Observatoire, France. [39]Instituto Nacional de Tecnica Aeroespacial, Carretera de Ajalvir, p.k. 4, 28850 Torrejon de Ardoz, Madrid, Spain. [40]Physikalisches Institut, Sidlerstrasse 5, University of Bern, 3012 Bern, Switzerland. [41]Planetary Science Section, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, USA.
***Corresponding author. E-mail: rotundi@uniparthenope.it**

OSIRIS (Optical, Spectroscopic, and Infrared Remote Imaging System) (*7*) has detected a dust coma since an outburst that occurred between 27 and 30 April 2014, and possibly since the first images taken on 23 March (*8*). We started to detect coma features in long wide-angle camera (WAC) exposures at the end of July 2014 at 3.7 astronomical units (AU) from the Sun. Most of the activity was coming out of the transition region between the small and large lobes of the nucleus, the Hapi region (*9*). This major coma feature showed a diurnal variation of intensity, and its brightness contribution was measured in WAC images taken on 10 August 2014 at 110 km from the nucleus. This single jet-like feature contributed roughly 10% of the total dust brightness: Such a dominance is rare among active comets closer to the Sun but may be true for activity as it first turns on. The background dust brightness may come from smaller features and/or isotropic emission from the nucleus surface.

GIADA (Grain Impact Analyzer and Dust Accumulator) consists of three subsystems: (i) the grain-detection system (GDS) to detect dust grains as they pass through a laser curtain; (ii) the impact sensor (IS) to measure grain momentum derived from the impact on a plate connected to five piezoelectric sensors; and (iii) the microbalances system (MBS), five quartz crystal microbalances in roughly orthogonal directions providing the cumulative dust flux of grains smaller than 10 μm. GDS provides data on grain speed and its optical cross section. The IS momentum measurement, when combined with the GDS detection time, provides a direct measurement of grain speed and mass (*10*, *11*). Using these combined measurements, we are able to characterize the dynamics of single grains.

GIADA began operating in continuous mode on 18 July 2014 when the comet was at a heliocentric distance of 3.7 AU. The first grain detection occurred when the spacecraft was 814 km from the nucleus on 1 August 2014. Three additional grains were detected at distances of 603, 286, and 179 km from the nucleus during the spacecraft approach. Once in orbit about the nucleus, 19 impacts were recorded by the IS at distances ranging from 90 to 30 km, while during the same period GDS detected seven particles. A total of nine combined GDS+IS events (i.e., measured by both the GDS and IS detectors) allow us to extract the complete set of dust grain parameters: mass, speed, velocity vector, and optical cross section. From these GDS+IS detections, we combined the measured optical cross sections (GDS) with the momenta (IS) to constrain the bulk density range to $(1.9 \pm 1.1) \times 10^3$ kg m^{-3}, assuming spherical grains (i.e., the lowest cross section/mass ratio). Grain cross sections (Table 1) are retrieved by means of two GIADA calibration

Table 1. GIADA GDS+IS detections made at heliocentric distances between 3.6 and 3.4 AU. The grain momentum and speed (v) are direct GIADA measurements from which the mass (m) is calculated. Grain geometrical cross sections are derived by the GDS calibration curves obtained with amorphous carbon and with a mixture of alkali-feldspar, melilite, and anorthite grains. R is the detection distance from the nucleus. The IS-sensitive area* is 10^{-2} m^2. Q_m is the dust mass loss rate per grain. The error of the velocity measurements is 6%.

Event no.	Date [dd/mm/yyyy]	Momentum [kg m s^{-1}]	v [m s^{-1}]	m [kg]	Cross section [m^2]	R [km]	Q_m [kg s^{-1}]
5	11/08/2014	$3.1 \pm 1.3 \times 10^{-8}$	3.2	$9.8 \pm 4.0 \times 10^{-9}$	$5.6 \pm 1.9 \times 10^{-8}$	92.9	1.9×10^{-2}
16	23/08/2014	$2.1 \pm 0.8 \times 10^{-7}$	2.6	$8.0 \pm 3.2 \times 10^{-8}$	$1.4 \pm 0.4 \times 10^{-7}$	61.7	6.8×10^{-2}
21	29/08/2014	$2.3 \pm 0.9 \times 10^{-8}$	4.7	$4.9 \pm 2.0 \times 10^{-9}$	$2.3 \pm 1.2 \times 10^{-8}$	54.0	3.2×10^{-3}
23	30/08/2014	$2.0 \pm 0.8 \times 10^{-7}$	2.6	$7.8 \pm 3.2 \times 10^{-8}$	$9.2 \pm 3.0 \times 10^{-8}$	57.4	5.8×10^{-2}
26	01/09/2014	$2.0 \pm 0.8 \times 10^{-8}$	2.9	$6.9 \pm 2.9 \times 10^{-9}$	$2.7 \pm 1.9 \times 10^{-8}$	53.2	4.4×10^{-3}
27	01/09/2014	$3.9 \pm 1.6 \times 10^{-8}$	2.8	$1.4 \pm 0.5 \times 10^{-8}$	$7.4 \pm 2.3 \times 10^{-8}$	51.8	8.4×10^{-3}
33	12/09/2014	$1.6 \pm 1.0 \times 10^{-9}$	3.0	$5.3 \pm 3.3 \times 10^{-10}$	$3.1 \pm 0.3 \times 10^{-8}$	29.7	2.1×10^{-3}*
34	12/09/2014	$1.3 \pm 0.5 \times 10^{-7}$	2.8	$4.6 \pm 1.9 \times 10^{-8}$	$7.4 \pm 0.6 \times 10^{-8}$	29.7	9.1×10^{-3}
35	13/09/2014	$8.3 \pm 3.7 \times 10^{-9}$	3.6	$2.3 \pm 1.0 \times 10^{-9}$	$3.1 \pm 1.6 \times 10^{-8}$	30.1	4.7×10^{-4}

*For momentum values close to the IS detection limit (6.5×10^{-10} kg m s^{-1}), a correction factor to the IS-sensitive area has to be taken into account; e.g., only 30% of the total IS-sensitive area has a sufficient sensitivity to detect values of momentum as small as 2.5×10^{-9} kg m s^{-1}. For the grain density estimate, this detection was not taken into account because the derived density was far off the density of all the materials used to retrieve the calibration curves.

Table 2. Impacts on GIADA, detected by IS only, from 3.6 to 3.4 AU. Dust speed is assumed from (*18*) for a dust bulk density of 1×10^3 (grain a) and 3×10^3 kg m^{-3} (grain b). The IS-sensitive area* is 10^{-2} m^2.

Event no.	Date [dd/mm/yyyy]	Momentum [kg m s^{-1}]	A [m^2]	v (a) [m s^{-1}]	v (b) [m s^{-1}]	m (a) [kg]	m (b) [kg]	R [km]	Q_m (a) [kg s^{-1}]	Q_m (b) [kg s^{-1}]
9	11/08/2014	$1.8 \pm 1.6 \times 10^{-9}$	8.0×10^{-4}	10.0	6.6	1.8×10^{-10}	2.7×10^{-10}	92.9	4.4×10^{-3}	6.5×10^{-3}
10	11/08/2014	$1.2 \pm 0.5 \times 10^{-8}$	1.0×10^{-2}	6.6	4.7	1.8×10^{-9}	2.6×10^{-9}	91.1	3.4×10^{-3}	4.8×10^{-3}
11	12/08/2014	$2.2 \pm 1.7 \times 10^{-9}$	2.0×10^{-3}	10.0	6.6	2.2×10^{-10}	3.3×10^{-10}	93.9	2.2×10^{-3}	3.3×10^{-3}
12	15/08/2014	$5.3 \pm 2.1 \times 10^{-8}$	1.0×10^{-2}	4.7	3.2	1.1×10^{-8}	1.7×10^{-8}	92.3	2.1×10^{-2}	3.2×10^{-2}
13	19/08/2014	$2.8 \pm 1.1 \times 10^{-8}$	1.0×10^{-2}	6.6	4.7	4.2×10^{-9}	6.0×10^{-9}	78.9	5.9×10^{-3}	8.4×10^{-3}
14	22/08/2014	$1.4 \pm 0.6 \times 10^{-8}$	1.0×10^{-2}	6.6	4.7	2.2×10^{-9}	3.0×10^{-9}	63.4	2.0×10^{-3}	2.7×10^{-3}
17	26/08/2014	$2.5 \pm 1.9 \times 10^{-9}$	3.0×10^{-3}	10.0	6.6	2.5×10^{-10}	3.8×10^{-10}	55.0	5.7×10^{-4}	8.6×10^{-4}
18	26/08/2014	$6.6 \pm 3.1 \times 10^{-9}$	9.8×10^{-3}	6.6	4.7	1.0×10^{-9}	1.4×10^{-9}	61.9	8.8×10^{-4}	1.2×10^{-3}
19	29/08/2014	$1.8 \pm 0.8 \times 10^{-8}$	1.0×10^{-2}	6.6	4.7	2.7×10^{-9}	3.8×10^{-9}	53.7	1.7×10^{-3}	2.5×10^{-3}
20	29/08/2014	$6.2 \pm 3.0 \times 10^{-9}$	9.7×10^{-3}	6.6	4.7	9.4×10^{-10}	1.3×10^{-9}	53.8	6.3×10^{-4}	8.7×10^{-4}
22	30/08/2014	$9.4 \pm 4.1 \times 10^{-9}$	1.0×10^{-2}	6.6	4.7	1.4×10^{-9}	2.0×10^{-9}	56.0	9.8×10^{-4}	1.4×10^{-3}
24	30/08/2014	$1.2 \pm 0.5 \times 10^{-8}$	1.0×10^{-2}	6.6	4.7	1.8×10^{-9}	2.6×10^{-9}	58.6	1.4×10^{-3}	2.0×10^{-3}
28	02/09/2014	$2.0 \pm 0.8 \times 10^{-8}$	1.0×10^{-2}	6.6	4.7	3.0×10^{-9}	4.3×10^{-9}	54.2	2.0×10^{-3}	2.8×10^{-3}
30	09/09/2014	$1.1 \pm 0.4 \times 10^{-7}$	1.0×10^{-2}	4.7	3.2	2.3×10^{-8}	3.4×10^{-8}	29.3	4.4×10^{-3}	6.5×10^{-3}
31	09/09/2014	$9.6 \pm 4.2 \times 10^{-9}$	1.0×10^{-2}	6.6	4.7	1.5×10^{-9}	2.0×10^{-9}	29.2	2.9×10^{-4}	3.8×10^{-4}

*For momentum values close to the IS detection limit (6.5×10^{-10} kg m s^{-1}), a correction factor to the IS-sensitive area has to be taken into account; e.g., only 30% of the total IS-sensitive area has a sufficient sensitivity to detect values of momentum as small as 2.5×10^{-9} kg m s^{-1}, and only 8% has a sufficient sensitivity to detect values of momentum as small as 1.8×10^{-9} kg m s^{-1}.

curves, GDS signal versus the grain measured geometrical cross section, obtained using grains of amorphous carbon and grains of a mixture of alkali-feldspar, melilite, and anorthite to get the extreme cases of dark and bright materials. The GIADA estimates of grain bulk densities fit the range of the measured bulk densities of interplanetary dust particles, some of which have cometary origins and sizes comparable to the grains detected by GIADA (*12*). They are also compatible with those inferred for the collected comet 81P/Wild 2 particles (*13*). The Stardust mission caused a dramatic paradigm shift in our understanding of the transport of debris from the evolving asteroid belt to the heliocentric distances of the present-day Kuiper belt (*14*). It was found that the minerals, chondrule fragments, and calcium-aluminum–rich inclusions in comet 81P/Wild 2 dust have their counterparts in chondritic meteorites (*15*), which is confirmed by the still ongoing laboratory analyses of Wild 2 samples (*16*).

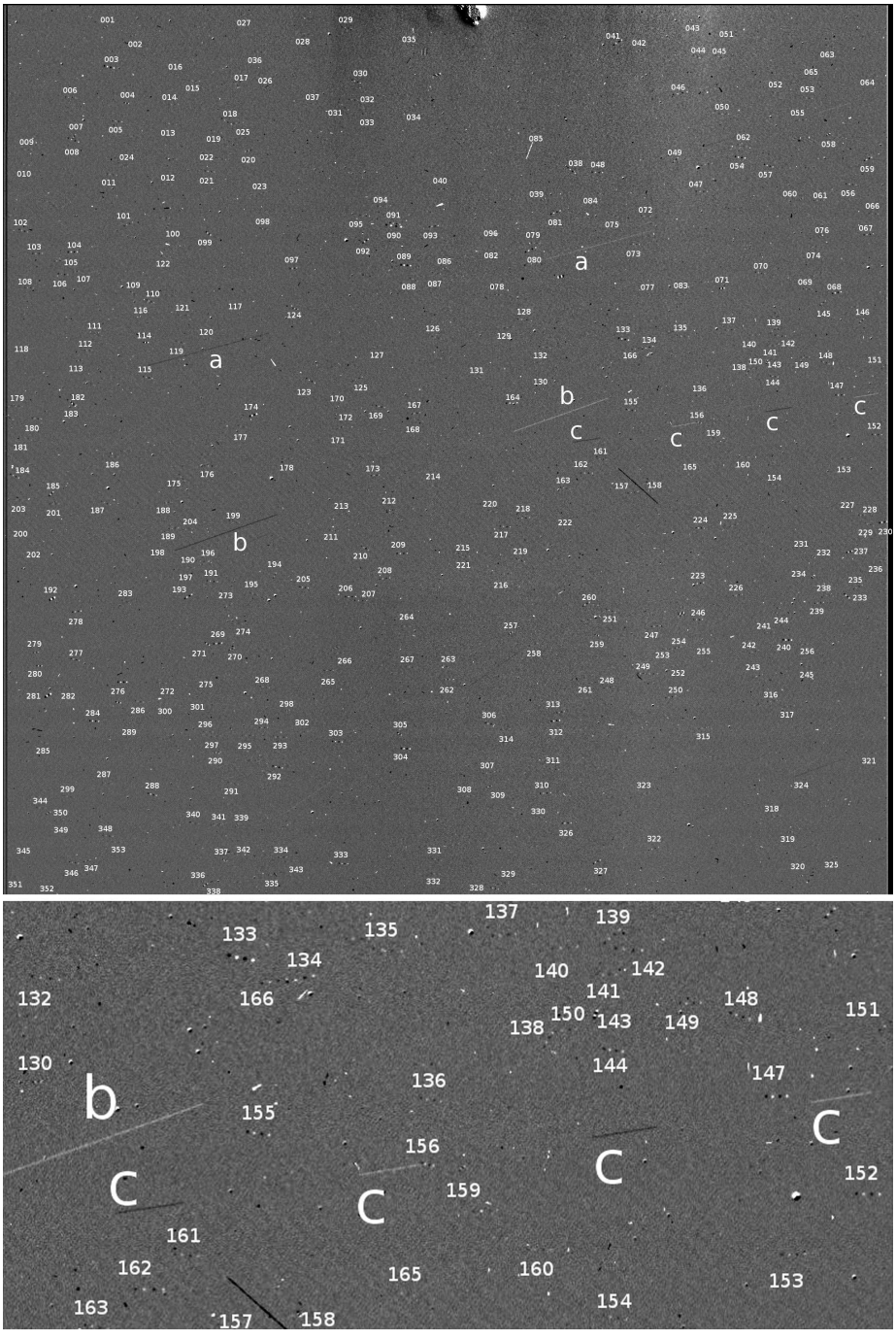


Fig. 1. Detections of single grains by OSIRIS. (**Top**) Bound grains detected in the first set of NAC images (each labeled by a number from 001 to 353) and the brightest three tracks of the 48 outflowing grains (each labeled by letters a, b, and c). The 67P nucleus is visible at the top of the panel. (**Bottom**) Zoom centered on track c. The images were shifted to align stars on the same pixels. The spacecraft motion is to the right, with an angle of 6° counterclockwise from the +x image axis. Grain c shows all of the expected four white and black tracks; grains a and b have two tracks outside the image. The detections of bound grains (white-black-white-black dots) show a uniform space distribution. Further details and explanations are available in the supplementary materials.

To extract the dust/gas mass ratio from our data, we compare our estimated dust mass loss rates to the water production rate measured by the MIRO (Microwave Instrument for the Rosetta Orbiter) instrument on board Rosetta [1.2 kg s^{-1} at 3.5 AU with diurnal variations of 50% (*17*)]. MIRO adopted a hemispherical gas ejection model from the nucleus surface. To retrieve a meaningful dust/gas ratio, we assume the same hemispherical dust ejection. If both gases and dust are ejected within a smaller solid angle, we will overestimate both the gas and the dust mass loss rates, but we will still obtain the correct dust/gas mass ratio. We discuss the dust size distribution between 3.6 and 3.4 AU. Table 1 lists the nine GDS+IS events with measured momentum and speed from which we can calculate the grain mass (m) directly. We compute the derived mass loss rate per grain at the nucleus

$$Q_m = 2\pi m R^2 (A\Delta t)^{-1} \quad (1)$$

where $\Delta t = 2.8 \times 10^6$ s is the collection time, R is the distance of the spacecraft from the nucleus center of mass, and A is the effective IS collecting area. Because the spacecraft radial velocity (always lower than 0.1 m s^{-1}) is much smaller than the dust velocity, the conversion from observed impacts to dust loss rate at the nucleus does not depend on the dust velocity. Table 2 completes the list of IS detections. The grain momentum is measured but not the speed. The grain speed (v) is therefore derived from a model applied to ground-based observations of 67P at 3.2 AU (*18*), for two assumed bulk densities of 1×10^3 kg m^{-3} (for grain a) and 3×10^3 kg m^{-3} (for grain b). The nine dust velocities measured by the GDS+IS match the predictions by Fulle *et al.* (*18*) at 3.2 AU very well for a dust bulk density of 3×10^3 kg m^{-3}. We cannot exclude higher dust velocities

Table 3. Radial velocity distribution in the comet reference frame of the 48 outflowing grains detected by OSIRIS.

Velocity bin (m s^{-1})	Number of grains
0–1	2
1–2	10
2–3	5
3–4	9
4–5	7
5–6	1
6–7	2
7–8	4
8–9	3
>9	5

at lower dust masses and bulk densities, as predicted by all models of dust drag by gas.

To complement the GIADA data and to estimate the largest dust mass ejected from 67P, we consider the detections of single grains by the OSIRIS narrow-angle camera (NAC) (*7*) on 4 August 2014 at a distance of 3.6 AU from the Sun and 275 km from the nucleus (Fig. 1). These observations allowed us to detect ~350 grains in bound orbits around the comet nucleus and 48 fast and outflowing grains that were ejected about a day before the observations and had diameters up to 17 mm (see details in the supplementary materials). We compute the velocity of the outflowing grains assuming that the grain motion is strictly radial from the nucleus. The velocity distribution in the nucleus reference frame is shown in Table 3. The wide distribution is also due to a nonperfectly radial velocity of the grains, because a grain proper transverse speed of 0.18 m s^{-1} changes the extracted radial velocity by a factor 2. Because proper transverse speeds are expected to be random, averages of the extracted radial velocities cancel out the effects of individual transverse speeds. We conclude that, on average, the median of the radial velocity is 3.5 m s^{-1}, in agreement with measurements of other grains by GIADA. The observed velocities do not show any evident trend with respect to the grain size.

Regarding bound grains, the radius of 67P's Hill sphere (318 km) corresponds to a coma volume of 10^{14} m^3 in NAC images, providing the observed space density of 3.5×10^{-12} m^{-3}. This value fits the space density predicted by models of bound dust clouds around comet 46P/Wirtanen (*19*), assuming a nucleus with the same mass as 67P (*9*, *19*). The predicted cloud density ranges from 5×10^{-13} m^{-3} to 2×10^{-11} m^{-3}, taking into account the difference between the dust mass loss of 46P and 67P (*18*, *19*) and a possible range of the dust bulk density from 4×10^2 kg m^{-3} (*9*) to 3×10^3 kg m^{-3} (this paper). Photometry of the bound grains provides diameters ranging from 4 cm, in agreement with dynamical constraints (*20*), to ~2 m. Models of bound grain orbits (*19*, *21*) show that an isotropic bound cloud needs years to be built up. The observed bound grains could not be ejected during the outburst observed at the end of April 2014 (*8*). They were probably put in bound orbits just after the previous perihelion, when the gas loss rate was decreasing and no longer perturbing the bound orbits (*19*). The same gas perturbations, during the current increase of the gas production rate, may dissipate the bound cloud in a few months.

The two families of detected grains (outflowing and bound) do not overlap in space. Outflowing grains are not detected farther than 20 km from the spacecraft, whereas bound grains are not detected closer than 130 km from the spacecraft. The space density of bound grains is at least 100 times lower than that of outflowing grains: Detections closer than 130 km are improbable (at most, three bound grains in a coma volume 100 times lower than up to 600 km). Taking into account the different pixel areas covered by the faintest bound (dots) and outflowing grains (tracks), the ratio of sizes between farthest outflowing and closest bound grains becomes consistent with the ratio of respective detection distances (20 versus 130 km). The track of the farthest outflowing grain is 16 NAC pixels long. This proves that, if larger and more distant young grains were outflowing in 67P's coma, they would have been easily detected. The size distribution of outflowing grains bigger than 2 cm has a large drop, probably due to the low gas density at the nucleus surface, still lifting very few larger grains at 3.6 AU. This conclusion is not affected by the possibility that some of the grains identified as bound are, in fact, outflowing. Their number and space density would be so low as to not change the cut-off of the size distribution above a diameter of 2 cm. This largest size is therefore well constrained and allows us to combine results from GIADA and OSIRIS to reach a precise measurement of the dust/gas ratio.

Tables 4 and 5 and Fig. 2 summarize our results for both OSIRIS (the three upper mass bins) and GIADA (the three lower-mass bins) and for dust bulk densities of 1×10^3 and 3×10^3 kg m^{-3}. For the GIADA detections, the dust mass loss rate per mass bin (Q_m) is the sum of all Q_m values in Tables 1 and 2. In addition, we performed a similar sum for three different subsets of Tables 1 and 2, namely dividing the total 24 samples into three sets of 8 samples each, one following the other in time. All the results were consistent (within 20%) with those shown in Tables 4 and 5. This shows that (i) the 24 samples provide a sufficient sampling for the 67P dust size distribution, and (ii) the dust mass loss rate remained approximately constant from 3.6 to 3.4 AU. Diurnal variations of the water loss rate mask any gas loss rate trend versus the heliocentric distance. For the OSIRIS detections, the dust loss rate was extracted from the dust space density ρ at the spacecraft, which was computed on coma volumes defined by the distance D from the spacecraft to the faintest (just above the detection limit) grain in each mass bin (i.e., the coma volume surveyed by the OSIRIS-NAC). N is the number of detections (for the OSIRIS data, N was divided by the number of images providing independent detections), v is the mean dust radial velocity (thus canceling out transverse speed effects), and Q_n is the number loss rate for hemispherical dust ejection

$$Q_n = 2\pi\rho v R^2 \quad (2)$$

Tables 4 and 5 and Fig. 2 show the quantity

$$Af\rho = 2A_p Q_n \sigma / v \quad (3)$$

for each mass bin, where σ is the mean dust cross section in each mass bin, and the geometric albedo A_p = 5% (*18*, *22*). $Af\rho$ provides the dust mass dominating the dust coma optical flux and shows a clear maximum at half-millimeter sizes. The total $Af\rho$ reaches 8 ± 4 cm after we fill

Table 4. GIADA and OSIRIS data on the 67P dust environment, assuming a dust bulk density of 1×10^3 kg m^{-3}. D is the distance of the grains observed (just above the detection limit) by OSIRIS-NAC to the spacecraft; N is the number of detections (for the OSIRIS data, N was divided by the number of images providing independent detections); ρ is the dust space density derived from N; v is the mean dust radial velocity (for the 48 OSIRIS detections, v has been computed assuming that the motion is radial from the nucleus); Q_n is the number loss rate for hemispherical dust ejection; Q_m is the dust mass loss rate per mass bin; and $Af\rho$ provides the dust brightness per mass bin. N/A, not applicable.

Mass bin [kg]	N	D [km]	ρ [m^{-3}]	v [m s^{-1}]	Q_n [s^{-1}]	Q_m [kg s^{-1}]	$Af\rho$ [m]
10^{-3}–10^{-2}	0.4	20	1.0×10^{-10}	7.2	3.9×10^2	0.9	1.8×10^{-3}
10^{-4}–10^{-3}	2.2	16	1.1×10^{-9}	5.7	3.9×10^3	1.2	5.0×10^{-3}
10^{-5}–10^{-4}	7	10	1.4×10^{-8}	4.4	3.6×10^4	1.5	1.3×10^{-2}
10^{-8}–10^{-7}	6	N/A	N/A	2.4	4.2×10^6	0.17	2.8×10^{-2}
10^{-9}–10^{-8}	13	N/A	N/A	4.7	1.4×10^7	4.6×10^{-2}	1.0×10^{-2}
10^{-10}–10^{-9}	5	N/A	N/A	6.9	2.4×10^7	9.9×10^{-3}	2.6×10^{-3}

Table 5. GIADA and OSIRIS data of the 67P dust environment, assuming a dust bulk density of 3×10^3 kg m^{-3}. See Table 4 for a description of parameters. N/A, not applicable.

Mass bin [kg]	N	D [km]	ρ [m^{-3}]	v [m s^{-1}]	Q_n [s^{-1}]	Q_m [kg s^{-1}]	$Af\rho$ [m]
10^{-3}–10^{-2}	1.2	20	3.0×10^{-10}	5.1	8.7×10^2	3.0	2.8×10^{-3}
10^{-4}–10^{-3}	5.2	16	2.6×10^{-9}	4.4	6.7×10^3	1.7	5.4×10^{-3}
10^{-5}–10^{-4}	3.2	10	6.4×10^{-9}	5.5	2.0×10^4	1.4	2.8×10^{-3}
10^{-8}–10^{-7}	6	N/A	N/A	1.9	4.2×10^6	0.18	1.7×10^{-2}
10^{-9}–10^{-8}	14	N/A	N/A	3.4	1.4×10^7	5.4×10^{-2}	6.8×10^{-3}
10^{-10}–10^{-9}	4	N/A	N/A	4.7	3.4×10^7	1.3×10^{-2}	2.6×10^{-3}

the gap from 10^{-7} to 10^{-5} kg with the observed lowest and largest $Af\rho$ values, respectively. Ground-based observations provided $Af\rho$ = 7 cm when the comet was at 5 AU, and $Af\rho$ = 20 cm at 3.4 AU inbound (*22*). Q_m is approximately constant at larger-than-millimeter sizes, suggesting a differential size distribution power law index = −4, in agreement with (*18*, *23*). At smaller-than-millimeter sizes, Q_m drops faster than predicted by (*18*), suggesting a shallower differential size distribution, with a power law index close to −2. The same model (*18*) predicts a MBS response 10 times higher than the MBS sensitivity. However, the MBS signal at these 67P heliocentric distances did not reach the minimum sensitivity level, suggesting no measurable accumulation of dust grains smaller than 10^{-11} kg and confirming that the power law index of the differential size distribution is close to −2 over the entire size spectrum, from submicrometer grains up to 10^{-7} kg. Assuming this size distribution (power index close to −2), the MBS flux becomes lower than the sensitivity limit, and $Af\rho$ has a sharp maximum at the upper mass; i.e., at 10^{-7} kg, as shown in Fig. 2. The number of micrometer-sized grains in 67P is too low to provide any substantial optical scattering with respect to that coming from millimeter-sized grains, a fact never confirmed before in any comet. The $Af\rho$ value observed at 3.4 AU (*22*), which is larger than our value of 8 ± 4 cm, may be explained by (i) larger $Af\rho$ values from 10^{-7} to 10^{-5} kg, (ii) a larger σ due to non-spherical dust shape, and (iii) A_p > 5%. Comet 67P's dust loss rate from 3.6 to 3.4 AU of 7 ± 1 kg s^{-1}, derived by interpolation between the GIADA and OSIRIS data to fill the gap from 10^{-7} to 10^{-5} kg, is consistent with an average of predicted dust loss rates at 3.2 AU [column for 3.2 AU in tables 3 and 4 of (*18*)].

Taking into account the MIRO measurements of the water production rate and the results obtained by the Rodionov-Zakharov-Crifo (RZC) coma model (*24*), applied to 67P to forecast the lander descent environment, the derived dust/gas mass ratio is 6 ± 2 for water only. The RZC model fits ROSINA-COPS (Rosetta Orbiter Spectrometer for Ion and Neutral Analysis–Comet Pressure Sensor) data, assuming the out-flow of water and of other gases in proportion to the local illumination. The obtained water loss rate fits that directly provided by MIRO data within the diurnal oscillations. Computations are ongoing to further constrain the whole three-dimensional and time-dependent coma by means of other Rosetta data. The dust/gas mass ratio decreases to 4 ± 2 if we use a total gas loss rate that includes CO and CO_2 [~50% in mass with respect to water according to ROSINA data (*25*)]. An estimated dust/gas ratio of 3 for 67P at perihelion was predicted assuming that meter-sized clumps are ejected at perihelion (*15*). This assumption is supported by the observations of the bound cloud of grains observed by OSIRIS-NAC. The dust number loss rate (Q_n) shows a systematic increase from the largest to the smallest measured dust masses, as is expected in a power law distribution. This is true also at the lowest mass bin measured by GIADA, showing that the effective IS collecting area (A) in Tables 1 and 2 compensates the bias of the low number of grain detections (N), due to the momentum sensitivity limit, in this mass bin. Therefore, the Q_m and $Af\rho$ values in the lowest mass bin (Table 3 and Fig. 2), should be considered a good estimate of the real values. The very good match of the two size distributions obtained by two independent instruments and techniques (in situ dust collection by GIADA versus remote optical detection by OSIRIS) suggests that the same dust bulk density range estimated from the GIADA data is a good assumption also for the single grains detected by OSIRIS.

In conclusion, at heliocentric distances between 3.7 and 3.4 AU: (i) GIADA on Rosetta is working well in characterizing the dust flux and, in particular, for carrying out measurements including dust grain speed, momentum, mass, and optical cross section and furthermore constraining derived values of the grain bulk densities. (ii) The dust/gas mass ratio, averaged over the sunlit nucleus surface, is 4 ± 2 but may change as the comet approaches closer to the Sun. The estimate of the diameter of the largest clumps in a bound cloud (probably built-up just after the previous perihelion passage) confirms the dust/gas ratio of 3 inferred at perihelion (*18*). (iii) The dust optical scattering is dominated by 100-μm to millimeter-sized grains, not by micrometer-sized grains [this is expected to change approaching to perihelion (*18*)]. (iv) The GIADA dust model predictions at 3.2 AU (*18*) are consistent with the measured data. (v) The detection by OSIRIS-NAC of a cloud of grains in bound orbits confirms predictions of its density (*19*) and sizes of the orbiting grains (*20*).

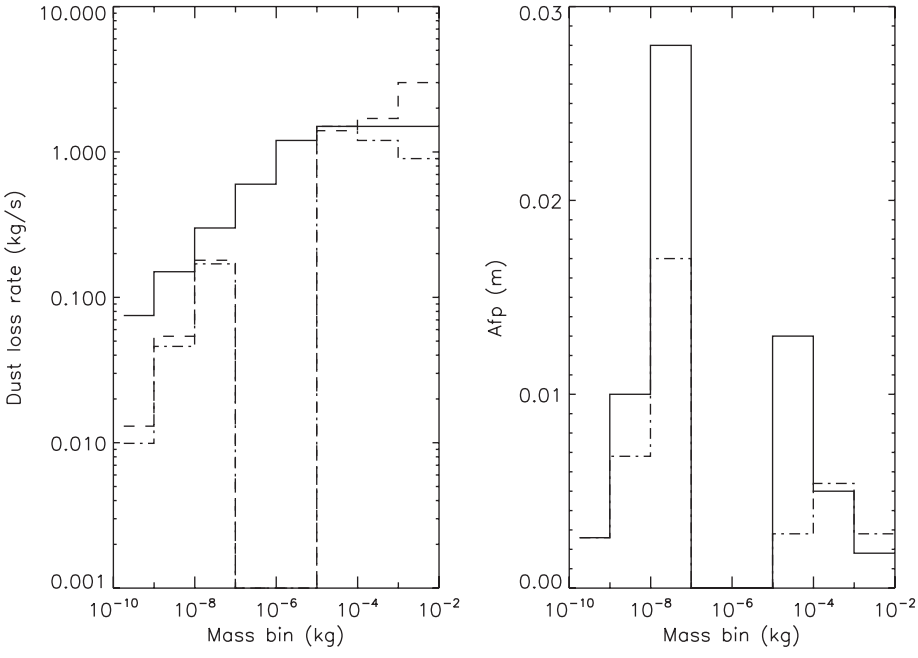


Fig. 2. 67P grain mass distribution. (**Left**) Mass distribution of 67P dust grains as measured by GIADA and OSIRIS. Dot-dashed line: dust loss rate by GIADA (three lower-mass bins) and OSIRIS (three higher-mass bins) at bulk density 1×10^3 kg m^{-3}. Dashed line: dust loss rate by GIADA (three lower-mass bins) and OSIRIS (three higher-mass bins) at bulk density 3×10^3 kg m^{-3}. Both the mass distributions have differential size distribution power law indices close to −2 in the three lower-mass bins. Solid line: average dust loss rate predicted by (*18*) at 3.2 AU. Differential size distribution power index = −4 in the three higher-mass bins and −3 elsewhere. (**Right**) Solid line: dust coma brightness-equivalent size $Af\rho$ derived from GIADA (three lower-mass bins) and OSIRIS (three higher-mass bins) data with bulk density 1×10^3 kg m^{-3}. Dot-dashed line: $Af\rho$ derived from GIADA (three lower-mass bins) and OSIRIS (three higher-mass bins) data with bulk density 3×10^3 kg m^{-3}. Both of these distributions have a clear maximum in the 10^{-8}- to 10^{-7}-kg mass bin, corresponding to a size of hundreds of micrometers.

REFERENCES AND NOTES

1. M. Fulle, L. Colangeli, V. Mennella, A. Rotundi, E. Bussoletti, The sensitivity of the size distribution to the grain dynamics: Simulation of the dust flux measured by GIOTTO at P/Halley. *Astron. Astrophys.* **304**, 622–630 (1995).
2. M. Fulle, A. C. Levasseur-Regourd, N. McBride, E. Hadamcik, In situ dust measurements from within the coma of 1P/Halley: First-order approximation with a dust dynamical model. *Astron. J.* **119**, 1968–1977 (2000). doi: 10.1086/301285
3. J. A. M. McDonnell, G. S. Pankiewicz, P. N. W. Birchley, S. F. Green, C. H. Perry, "The comet nucleus: Ice and dust morphological balances in a production surface of comet P/Halley," in *Proceedings of the 20th Lunar and Planetary Science Conference*, 13 to 17 March 1989 (Lunar and Planetary Institute, Houston, TX, 1990), pp. 373–378.
4. J. A. M. McDonnell *et al.*, Dust particle impacts during the Giotto encounter with Comet Grigg-Skjellerup. *Nature* **362**, 732–734 (1993). doi: 10.1038/362732a0
5. S. F. Green *et al.*, The dust mass distribution of comet 81P/Wild 2. *J. Geophys. Res.* **109**, E12S04 (2004). doi: 10.1029/2004JE002318
6. T. E. Economou, S. F. Green, D. E. Brownlee, B. C. Clark, Dust flux monitor instrument measurements during Stardust-NExT flyby of comet 9P/Tempel 1. *Icarus* **222**, 526–539 (2013). doi: 10.1016/j.icarus.2012.09.019

7. H. U. Keller *et al*., OSIRIS - The scientific camera system onboard Rosetta. *Space Sci. Rev.* **128**, 433–506 (2007). doi: 10.1007/s11214-006-9128-4
8. C. Tubiana *et al*., 67P/Churyumov-Gerasimenko: Activity between March and June 2014 as observed from Rosetta/OSIRIS. *Astron. Astrophys.* **573**, A62 (2015). doi: 10.1051/0004-6361/201424735
9. H. Sierks *et al*., On the nucleus structure and activity of comet 67P/Churyumov-Gerasimenko. *Science* **347**, aaa1044 (2015). doi: 10.1126/science.aaa1044
10. V. Della Corte *et al*., GIADA: Its status after the Rosetta cruise phase and on-ground activity in support of the encounter with comet 67P/Churyumov-Gerasimenko. *J. Astron. Instrum.* **03**, 1350011 (2014). doi: 10.1142/S2251171713500116
11. L. Colangeli *et al*., The Grain Impact Analyser and Dust Accumulator (GIADA) experiment for the Rosetta mission: Design, performances and first results. *Space Sci. Rev.* **128**, 803–821 (2007). doi: 10.1007/s11214-006-9038-5
12. F. J. M. Rietmeijer, "Interplanetary dust particles," in *Planetary Materials, Reviews in Mineralogy*, J.J. Papike, Ed. (Mineralogical Society of America, Chantilly, VA, 1998), vol. 36, pp. 2-1–2-95.
13. A. T. Kearsley *et al*., Dust from comet Wild 2: Interpreting particle size, shape, structure, and composition from impact features on the Stardust aluminum foils. *Meteorit. Planet. Sci.* **43**, 41–73 (2008). doi: 10.1111/j.1945-5100.2008.tb00609.x
14. D. Brownlee *et al*., Comet 81P/Wild 2 under a microscope. *Science* **314**, 1711–1716 (2006). doi: 10.1126/science.1135840; pmid: 17170289
15. M. Zolensky *et al*., Comparing Wild 2 particles to chondrites and IDPs. *Meteorit. Planet. Sci.* **43**, 261–272 (2008). doi: 10.1111/j.1945-5100.2008.tb00621.x
16. A. Rotundi *et al*., Two refractory Wild2 terminal particles from a carrot-shaped track characterized combining MIR/FIR/Raman micro-spectroscopy and FE-SEM/EDS analyses. *Meteorit. Planet. Sci.* **49**, 550–575 (2014).
17. S. Gulkis *et al*., Subsurface properties and early activity of comet Churyumov-Gerasemenko. *Science* **347**, aaa0709 (2015). doi: 10.1126/science.aaa0709
18. M. Fulle *et al*., Comet 67P/Churyumov-Gerasimenko: The GIADA dust environment model of the Rosetta mission target. *Astron. Astrophys.* **522**, A63–A80 (2010). doi: 10.1051/0004-6361/201014928
19. M. Fulle, Injection of large grains into orbits around comet nuclei. *Astron. Astrophys.* **325**, 1237–1248 (1997).
20. K. Richter, H. U. Keller, On the stability of dust particle orbits around cometary nuclei. *Icarus* **114**, 355–371 (1995). doi: 10.1006/icar.1995.1068
21. D. J. Scheeres, F. Marzari, Temporary orbital capture of ejecta from comets and asteroids: Application to the Deep Impact experiment. *Astron. Astrophys.* **356**, 747–756 (2000).
22. C. Snodgrass *et al*., Beginning of activity in 67P/Churyumov-Gerasimenko and predictions for 2014-2015. *Astron. Astrophys.* **557**, A33 (2013). doi: 10.1051/0004-6361/201322020
23. J. Agarwal *et al*., The dust trail of comet 67P/Churyumov-Gerasimenko between 2004 and 2006. *Icarus* **207**, 992–1012 (2010). doi: 10.1016/j.icarus.2010.01.003
24. J. F. Crifo, G. A. Lukyanov, V. V. Zakharov, A. V. Rodionov, "Physical model of the coma of Comet 67P/Churyumov–Gerasimenko," in *The New Rosetta Targets. Observations, Simulations and Instrument Performances*, L. Colangeli, E. Mazzotta Epifani, P. Palumbo, Eds. (Kluwer Academic, Dordrecht, 2004), pp. 119–131.
25. M. Hässig *et al*., Time variability and heterogeneity in the coma of 67P/Churyumov-Gerasimenko. *Science* **347**, aaa0276 (2015). doi: 10.1126/science.aaa0276

ACKNOWLEDGMENTS

GIADA was built by a consortium led by the Università degli Studi di Napoli "Parthenope" and INAF—Osservatorio Astronomico di Capodimonte, in collaboration with the Instituto de Astrofisica de Andalucia, Selex-ES, FI, and SENER. GIADA is presently managed and operated by Istituto di Astrofisica e Planetologia Spaziali–INAF, Italy. GIADA was funded and managed by the Agenzia Spaziale Italiana, with the support of the Spanish Ministry of Education and Science Ministerio de Educacion y Ciencias (MEC). GIADA was developed from a Principal Investigator proposal from the University of Kent; science and technology contributions were provided by CISAS, Italy; Laboratoire d'Astrophysique Spatiale, France, and institutions from the UK, Italy, France, Germany, and the USA. Science support was provided by NASA through the U.S. Rosetta Project managed by the Jet Propulsion Laboratory/California Institute of Technology. We would like to thank A. Coradini for her contribution as a GIADA Co-Investigator. GIADA calibrated data will be available through ESA's Planetary Science Archive (PSA) Web site (www.rssd.esa.int/index.php?project=PSA&page=index). All data presented here are available on request before archival in the PSA. OSIRIS was built by a consortium led by the Max-Planck-Institut für Sonnensystemforschung, Katlenburg-Lindau, Germany, in collaboration with CISAS, University of Padova, Italy; the Laboratoire d'Astrophysique de Marseille, France; the Instituto de Astrofísica de Andalucia, CSIC, Granada, Spain; the Scientific Support Office of the European Space Agency, Noordwijk, The Netherlands; the Instituto Nacional de Técnica Aeroespacial, Madrid, Spain; the Universidad Politéchnica de Madrid, Spain; the Department of Physics and Astronomy of Uppsala University, Sweden; and the Institut für Datentechnik und Kommunikationsnetze der Technischen Universität Braunschweig, Germany. The support of the national funding agencies of Germany (DLR), France (Centre National d'Etudes Spatiales), Italy (Italian Space Agency), Spain (MEC), Sweden (Sweden National Science Board), and the ESA Technical Directorate is gratefully acknowledged. We thank the MIRO and ROSINA teams for sharing their early results with us. We thank the Rosetta Science Ground Segment at ESAC, the Rosetta Mission Operations Centre at ESOC (European Space Operations Center), and the Rosetta Project at ESTEC (European Space Technology Center) for their outstanding work enabling the science return of the Rosetta Mission. We gratefully acknowledge the referees for very constructive comments.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/aaa3905/suppl/DC1
Supplementary Text
Fig. S1
Tables S1 to S3
References

29 November 2014; accepted 1 January 2015
10.1126/science.aaa3905

REVIEW SUMMARY

SEA LEVEL CHANGE

Inherited landscapes and sea level change

Sierd Cloetingh and Bilal U. Haq*

BACKGROUND: Knowledge of past sea level fluctuations is fundamental to geosciences and for exploration of Earth-bound resources. Recent years have seen a convergence of views between stratigraphers (who measure past sea level changes in marine strata) and geodynamicists (who investigate surficial expressions of lithospheric and mantle processes) with the realization that without understanding inherited topographies, their causal mechanisms and operative time scales, palinspastic (pre-diastrophic) reconstructions of the past landscapes and seascapes remain inchoate at best.

Sequestrations of seawater on land or its subduction-related entrainment in the mantle are two direct means of lowering global sea level. Sea level can also be changed by modifying the container capacity of the ocean through numerous interconnected solid-Earth processes. Some of these can only refashion landscapes regionally, thus affecting local measures of sea level change. Disentangling these processes to uncover the likely cause(s) for the resultant topography poses considerable challenges.

ADVANCES: Recent developments in seismic tomography and high-speed computing that allow detailed forward and inverse modeling, combined with new concepts in stratigraphy and geophysics that permit envisioning large-scale transfers of material among depositional centers, have brought us closer to understanding factors that influence landscapes and sea levels and the complex feedbacks. As a result, estimates of the amplitude of long-term eustatic changes have converged using different data sets. We have learned that solid-Earth processes operating on decadal to multi-million-year time scales are all responsible for retaining lithospheric memory and its surface expression: On time scales of tens to hundreds of years, glacial isostatic adjustments cause local topographic anomalies, whereas postglacial rebound can be enhanced by viscous mantle flow on time scales of thousands to hundreds of thousands of years; on time scales of more than 1 million years, oceanic crustal production variations, plate reorganizations, and mantle-lithosphere interactions (e.g., dynamic topography) become more influential in altering the longer-wavelength surface response. Additionally, the lithosphere's rheological heterogeneity, variations in its strength, and changes in intraplate stress fields can also cause regional topographic anomalies, and syn-rift volcanism may be an important determinant of the long-term eustatic change on time scales of 5 to 10 million years.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1258375

OUTLOOK: Despite these remarkable advances, we remain far from resolving the causes for third-order quasi-cyclic sea level changes (~500,000 to 3 million years in duration). Ascertaining whether ice volume changes were responsible for these cycles in the Cretaceous will require discerning the potential for extensive glaciation at higher altitudes on Antarctica by modeling topographic elevation involving large-scale mantle processes. Extensive sea floor volcanism, plate reorganizations, and continental breakup events need to be better constrained if causal connections between tectonics and eustasy have to be firmly established. Another promising avenue of inquiry is the leads and lags between entrainment and expulsion of water within the mantle on third-order time scales. Future geodynamic models will also need to consider lateral variations in upper mantle viscosity and lithosphere rheology that require building on current lithospheric strength models and constructing global paleorheological models. Deep-drilling efforts will be of crucial importance for achieving the integrative goals. ■

PHOTO COURTESY HOWARD SPERO, UNIVERSITY OF CALIFORNIA, DAVIS

Earth's sea level history is archived in sedimentary sections. This example is from the Umbrian Apennines near Gubbio, Italy, that was a part of the ancient Tethys Ocean. It preserves a continuous and nearly complete Cretaceous pelagic record, and its contained microfossils and stable isotopes provide valuable clues about paleoceanography, paleoclimatology, and sea level history of the region. The finer-grained sediment near the base marks the boundary between the Cenomanian and Turonian stages, at which time the highest sea levels of the Cretaceous have been documented around the world. Evidence and measure of the amplitude of this eustatic high based on geodynamics and stratigraphic data have converged.

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: bilhaq@gmail.com
Cite this article as S. Cloetingh, B. U. Haq, *Science* **347**, 1258375 (2015). DOI: 10.1126/science.1258375

REVIEW

SEA LEVEL CHANGE

Inherited landscapes and sea level change

Sierd Cloetingh[1] and Bilal U. Haq[2,3]*

Enabled by recently gained understanding of deep-seated and surficial Earth processes, a convergence of views between geophysics and sedimentary geology has been quietly taking place over the past several decades. Surface topography resulting from lithospheric memory, retained at various temporal and spatial scales, has become the connective link between these two methodologically diverse geoscience disciplines. Ideas leading to the hypothesis of plate tectonics originated largely with an oceanic focus, where dynamic and mostly horizontal movements of the crust could be envisioned. But when these notions were applied to the landscapes of the supposedly rigid plate interiors, there was less success in explaining the observed anomalies in terrestrial topography. Solid-Earth geophysics has now reached a developmental stage where vertical movements can be measured and modeled at meaningful scales and the deep-seated structures can be imaged with increasing resolution. Concurrently, there have been advances in quantifying mechanical properties of the lithosphere (the solid outer skin of Earth, usually defined to include both the crust and the solid but elastic upper mantle above the asthenosphere). The lithosphere acts as the intermediary that transfers the effects of mantle dynamics to the surface. These developments have allowed us to better understand the previously puzzling topographic features of plate interiors and continental margins. On the sedimentary geology side, new quantitative modeling techniques and holistic approaches to integrating source-to-sink sedimentary systems have led to clearer understanding of basin evolution and sediment budgets that allow the reconstruction of missing sedimentary records and past geological landscapes.

The interdisciplinary dissension between solid-Earth geophysics and soft-rock geology was at least partly due to the prevalent view within the sedimentological community that post-rift tectonic processes are normally too slow to contribute to punctuated stratigraphy. This was a likely outgrowth of the first generation of basin evolution models that implied steady long-term thermal subsidence of rifted margins. Another reason may have been that the solid-Earth geophysics community did not always appreciate variable time scales of basin-fill processes or lacked awareness of advances in sedimentary geology that now allowed the deciphering of the record of mass transfer at the scale of an entire continental margin or interior basin. In the past few decades, conceptual models as well as empirical evidence indicate that post-rift tectonic processes act on multiple scales. These lead to episodic vertical movements on shorter time scales, as well as far-field effects that may transform local topographies on longer spatial and temporal wavelengths. This newfound cognizance of the vertical movements of landscapes and seascapes at various time scales has indeed been a bridging factor between these two communities.

[1]Utrecht University, Utrecht, Netherlands. [2]National Science Foundation, Arlington, VA, USA. [3]Sorbonne, Pierre & Marie Curie University, and CNRS, UMR 7193, ISTeP, F-75005 Paris, France.
*Corresponding author. E-mail: bilhaq@gmail.com

Sequence stratigraphy that facilitates stratigraphic estimates of past sea level changes (*1*, *2*) and geodynamic lithospheric models of continental margins and interiors [see, e.g., (*3*)] have had parallel histories of advancement over the past four decades, which in turn have led to a number of major breakthroughs in basin analysis. Indeed, this trend toward increasing synergy between stratigraphy and tectonics, crustal geophysics, paleoclimatology, and numerical modeling, where stratigraphic data had begun to play an important role in providing constraints for geodynamic models, was already alluded to by Miall (*4*). In the past two decades, progressively improving resolution of seismic tomographic images of the mantle (*5*–*8*) has provided better constraints for absolute plate motions (*9*), as well as revolutionary insights into the causes for dynamic vertical movements of continental and oceanic lithosphere (Fig. 1) at wavelengths that can affect both eurybatic (regional) and eustatic (global) sea level signals (*10*). [Eurybatic is a term introduced by Haq (*10*) to distinguish local or regional measures of sea level change from eustatic (global) changes.] This in turn has reinforced the notion of intimate coupling between Earth's mantle and surface processes (*11*–*13*). In addition, the holistic source-to-sink approach and process-based studies on several active and passive continental margins have also helped to reveal the connection of surface topography and basin subsidence [see, e.g., (*14*, *15*)]. At the same time, numerical simulations of the viscous response of the upper mantle to crustal loading and unloading [i.e., ice (*16*, *17*), sediments (*18*, *19*), and subduction of oceanic slabs underneath the continental crust (*20*)] have led to new insights into how surface topography reacts dynamically to both isostasy and mantle flow, producing regionally extensive landscape anomalies at both shorter and longer time scales. This response of the landscape to mantle flow has been broadly termed as dynamic topography (*11*), which thwarts straightforward measures of sea level changes based on stratal patterns and thickness of stratigraphic sequences along continental margins, even when corrected for thermal and flexural subsidence and other local factors (*21*). Also important to this discussion is the inherited rheological heterogeneity of lithosphere (*22*, *23*) due to intrinsic variations in the density and thermal regimes (*24*, *25*) (Fig. 2).

Below, we review the relationship between inherited topography, caused both by deep-Earth and surficial processes, and sea level change, using the Cretaceous Period as the key example for several reasons: The Cretaceous is known to have been a period of warm climate when Earth was largely ice-free, with the possible exception of Antarctica (*26*). This makes the Cretaceous suitable for testing various tectonic models without the complications of ice cover in the Northern Hemisphere. The Cretaceous was also a period of extensive sea floor volcanism (*27*, *28*) associated with upper mantle instabilities and rifting and continental breakup (*29*) that is often accompanied by volcanic activity (*30*). Additionally, not all of the Cretaceous sea floor has been subducted, and the missing fragments can be reconstructed (*31*) or alternatively, some of the subducted slabs of this age can be imaged and modeled in mantle graveyards (*20*). The Cretaceous is also of considerable paleoclimatological interest, as it may be an appropriate analog for a greenhouse climate, a direction in which our planet currently seems to be headed. The presence of Cretaceous petroleum systems, both in terrestrial and offshore plays, also makes this interval of high interest to the hydrocarbon industry, where knowledge of the chronology of eustatic variations and their amplitudes is considered an indispensable component of the explorationists' toolbox, allowing them to broadly anticipate several exploration criteria (*10*). In this Review, we examine the record of both horizontal and vertical motions of the lithosphere with a focus on the Cretaceous record that could have influenced far-field sea level changes. We conclude by identifying unresolved questions and areas of future research.

Mechanisms and measures of sea level change

Known mechanisms capable of generating a global signature in the stratigraphic record fall into two categories: those caused by substantial changes in the volume of water in the oceans, and those driven by changes in the volume of the ocean basins, (i.e., their container capacity) [see also (*32*) for a recent review, and Table 1 for

a summary of known operative mechanisms of sea level change].

Water sequestration on land

The simplest means of transferring a substantial amount of seawater to land (so as to be observable in the sedimentary record) is by sequestering it as ice on the continents during glacial episodes. [Smaller amounts of water can also be sequestered on shorter time scales (hundreds to thousands of years) in terrestrial aquifers and lakes (*33*–*36*), but the current consensus dictates that this does not produce an appreciable global stratigraphic signal, although it could be relevant to higher-frequency Milankovitch-scale cyclicity reflecting swings between dry and wet periods.] Our knowledge of Earth's glacial history is still fragmentary; nonetheless, glacial episodes with varying degrees of ice cover are known to have occurred in the Late Ordovician–Early Silurian interval, in the Late Devonian, through much of the Carboniferous and Permian, and in the Late Paleogene–Neogene, in addition to the development of bipolar ice cover in the Late Neogene [see, e.g., (*37*)]. In the Late Cretaceous, ephemeral Antarctic glacial episodes have been hypothesized (*38*–*41*).

The second process capable of redistributing a substantial volume of free water in the oceans over relatively longer time scales (on the order of millions of years) is the potential of imbalance of ocean's water exchange with deep mantle. This is not a well understood process (*32*), and in the absence of reliable models or the knowledge of operative time scales, this exchange (i.e., water subducted down with hydrated minerals versus water released from mantle into the ocean by degassing) is expediently considered to be in balance over very long time scales (10 million years or more). But it stands to reason that during intervals of anomalously high oceanic crustal production, as in the case in the Late Cretaceous, there would be a net gain of deep-sourced water in the oceans on the relevant time scales, which would balance out by net loss following increased subduction, presumably after a time lag (*10*). This is a promising area for future investigation.

Variations in container capacity of the oceans

The geologically rapid elastic isostatic response (on time scales of tens to hundreds of years) to ice unloading during deglaciations can be distinguished from what follows (*32*) as the viscous post-deglacial response (on time scales of thousands to hundreds of thousands of years) of mass redistribution and vertical displacements along continental margins (*42*–*44*). These effects do not influence the ocean volume on the broader scale per se, but can cause regional anomalies in sea level measurements along a particular margin. Even on the somewhat longer time scales (million years or more) of third-order sea level changes, the long-wavelength warping effect due to mantle convection flow patterns in response to upwellings and downwellings and movements of plates over subduction graveyards [i.e., dynamic topography (*20*, *45*, *46*)] can cause regional anomalies along affected continental margins. It should be noted, however, that numerical modeling studies of glacial isostatic adjustment (GIA) after deglaciation show that the initial elastic rebound and the consequent viscous mantle flow under deglaciated areas are essentially completed within thousands to hundreds of thousands of years post-glacial (*16*, *17*). On these time scales, GIA also shows a strong dependence on distance from the ice sheet. Thus, by ~100,000 years post-glacial, the effects of GIA largely reach equilibrium, and for third-order cyclicity they become part of the steady-state background.

Factors that can modify the ocean's container capacity with the potential to cause eustatic variations on relatively longer time scales of several million years include variations in oceanic crustal production rates (including volume of mid-ocean ridges and extruded igneous material on the sea floor) and net dynamic uplift and subsidence of the ocean floor due to long-wavelength dynamic topographic effects. It has long been reasoned that variability in rates of oceanic crustal formation at mid-ocean ridge systems (i.e., rates of spreading and total ridge length) would affect the total volume of these subsea features and thus eustatic sea level on relatively longer time scales (~10 million years) (*47*, *48*). Estimating ridge volume from older, now-subducted sea floor poses considerable challenges. Nevertheless, ridge volume variability from the Cretaceous to the present has been modeled (*49*) and, despite some difference of opinion (*50*), a number of independent estimates have reached the conclusion that there has been a slowdown in ridge spreading rates (and thus effective ridge volume) of 60 to 80% since the Late Cretaceous (*31*, *32*, *51*). Other ocean floor volcanic activities that can effect change in the ocean's container capacity include emplacements of large igneous provinces (LIP) and hotspot activity on the sea floor. In the Early Cretaceous, LIP activity was particularly pronounced, which manifested itself in the emplacement of Shatsky Rise [~137 to 131 million years ago (Ma)] and Ontong-Java and Manihiki Plateaus (episodically from ~122 to 90 Ma) in the Pacific that are interpreted to have been emplaced over relatively long durations (*28*). The Deccan Traps region in the Indian Ocean, on the other hand, was emplaced in a relatively short time (several hundred thousand years) in the latest Cretaceous, both on the ocean floor and along the western Indian margins (*52*).

Mantle-lithosphere interactions

An important development in geosciences in recent years has been the recognition that mantle-lithosphere interactions can have important consequences for long-wavelength surface topography on the sea floor and the continents, with implications for long-term (time scales of >5 million years) sea level variations. This applies to zones of mantle upwelling as well as downwelling, as documented by high-resolution seismic tomography (*7*, *8*). Continental interior and marginal surface topography can be modified when large features, such as a subducting slab, are traversed over by an overriding continent. This has

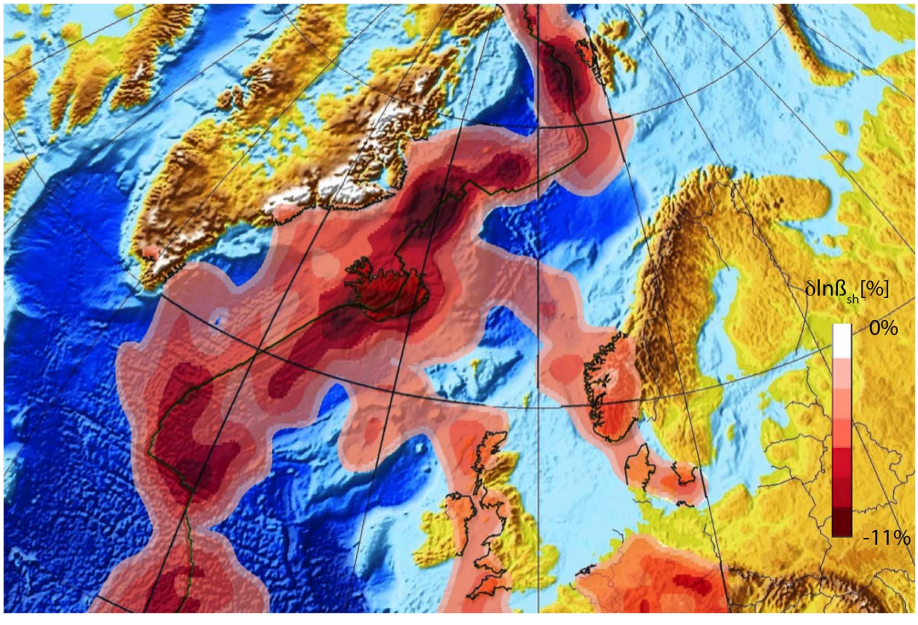


Fig. 1. A recent example of the resolving power of seismic tomography delineating a zone of mantle upwelling under the North Atlantic and adjacent continental areas [adapted from (*8*)]. Seismic velocity perturbations from full waveform inversion are shown for a depth slice at 120 km. Red areas depict location of Iceland plume and its side lobes extending into southern Norway, the British Isles, and central Europe. All of these areas are undergoing recent uplift in an intraplate tectonic setting, far from plate boundaries, thus affecting any measure of sea level change gauged along these margins. Perturbations are with respect to the reference velocity indicated in the figure.

been modeled and interpreted to be the case for the westward movement of North America over the remnants of the Farallon subduction slab (Fig. 3). This was inferred to have induced substantial far-field subsidence along New Jersey and adjacent margins since the Late Cretaceous, as modeled by a number of dynamic topographic studies (*20*, *31*, *53*–*55*). Although dynamic topography can produce both local and global effects, it is mainly relevant to long-term sea level trends and less so for the punctuated shorter-term third-order events, as long as it is recognized that the far-field warping effect that it can produce needs to be taken into account when measuring sea level change locally.

Thermomechanical models of mantle-lithosphere interactions predict vertical crustal movements within the same amplitude range as the third-order cycles, but they are not likely to be repetitive (*56*, *57*). For example, plume emplacement, slab detachment, or folding instabilities within the lithosphere are not likely to repeat in a strictly cyclic manner. However, one intriguing dynamic that seems worthy of consideration is the direct correlation between strong lithosphere and higher continental topography, and the association of low-strength lithosphere with lower topography that may allow access to oceanic flooding (*58*, *59*). The Cretaceous Period had inherently weaker lithosphere in the broad sense, due to widespread continental breakup and high oceanic crustal production rates. If this is combined with the potential for higher water expulsion rates from the mantle in the Late Cretaceous, weakening the lithosphere still further by water entrainment, the Late Cretaceous topography in general would have been lower and the global sea levels higher (*58*) (Fig. 4). Paleorheology varies in the Cretaceous, and the lithosphere as a whole becomes relatively stronger near the end of the Cretaceous. In addition, there is a feedback between stress and rheology (stresses weaken plates), thereby indirectly affecting topography. And as noted earlier, plume emplacements and their timing also play a role is resetting the thermomechanical structure. The plume emplacement effect dictates that there would be uplift in the earlier stages, followed by differential motions, including subsidence on the continents and margins (*56*). According to this tenet, we should expect shorter-wavelength (and higher-amplitude) changes in the earlier Cretaceous, and longer-wavelength (and lower-amplitude) changes in the later Cretaceous. On the ocean floor, swelling occurs during plume emplacement and subsidence follows only when the plume dies out (Fig. 5).

In this context, it is becoming evident that the assumption of rigid lithospheric plates may be an oversimplification, and the isostatic response to loading and unloading would depend strongly on rheological stratification of the lithosphere (*56*). As a result of the large number of rifted margin studies carried out in recent decades, it is becoming clear that traditional concepts resulting from the early stretching models (*60*), implying instantaneous rifting followed by a post-rift phase of steadily decaying thermal subsidence, require some additional thought. The duration of syn-rift phases can last up to hundreds of millions

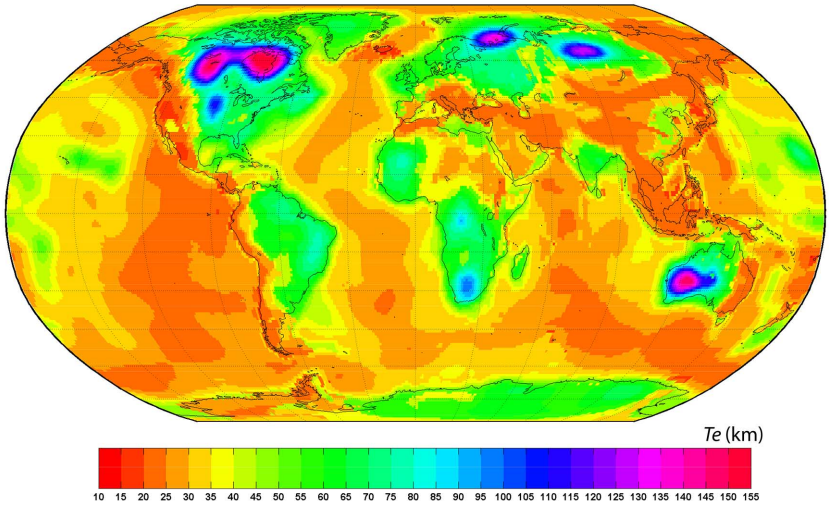


Fig. 2. Map of global modeled lithospheric rigidity (strength) expressed in term of effective elastic thickness (T_e) of oceans and continents, showing pronounced spatial variations from strong continental cratonic areas to weak alpine belts and young oceanic lithosphere [adapted from (*23*)]. Such variations have strong impact on the loading capacity of the lithosphere to support topography, ice caps, and sedimentary loads. Weak areas frequently manifest lithospheric domains (e.g., Iceland in the North Atlantic) that are subject to upwelling, whereas strong cratonic areas (such as the Canadian Shield) retain long-term memory of relative tectonic quiescence and the presence of a cold upper mantle underneath. Weak areas are also locus of intraplate deformation and associated vertical motions affecting inherited landscapes.

Table 1. Various mechanisms that can modify local/regional (eurybatic) and global (eustatic) sea levels, their time scales, their magnitude of change, their extent, and the mechanism's relevance to the Cretaceous Period under discussion in this Review. Data are from several sources, including (*32*, *47*, *48*, *84*); GIA, glacial isostatic adjustment; LIPS, large igneous provinces; My, millions of years.

Mechanism	Operative time scale	Order of magnitude of change	Potential extent	Cretaceous relevance
Water sequestration on land and in mantle				
Terrestrial aquifers and lakes	<0.01 My	Up to 10 m (? higher)	Global	?
Glaciations/deglaciations	0.01 to 0.1 My	100 to 250 m	Global	?
Water exchange with mantle	0.1? to 1.0 My	Unknown	Global?	++
Changes in container capacity of oceans				
GIA: Elastic rebound	0.00001 My	Up to 100 m	Regional	?
GIA: Viscous mantle flow	0.0001 to 0.1 My	Up to 100 m	Regional	?
Mean age of oceanic crust	50 to 100 My	100 to 300 m	Global	+++
Ridge production rate changes	50 to 100 My	100 to 300 m	Global	+++
Ocean floor volcanic activity (LIPS)	1 to 10 My	500 to 1000 m	Global	+++
Mantle/lithosphere interactions	1 to 10 My	10 to 1000 m	Regional/global	+++
Intraplate deformation	1 to 10 My	10 to 1000 m	Regional/global	+++
Dynamic topography	>5 My	Up to 1000 m	Regional/global	+++
Sedimentation	1 to 10 My	50 to 100 m	Global	+

of years (29); in the case of the Norwegian-Greenland Sea, the syn-rift phase lasted more than 200 million years, followed by a post-rift phase of 65 million years, leading to a complex pattern of subsidence and basin migration. Tectonic subsidence in the post-rift phase in many marginal systems also does not conform with the prediction of a steady decrease with time (61–65), but is instead often anomalous, implying that punctuated subsidence is not necessarily only a response to sea level fluctuations. Another factor that may be important in determining inherited landscape is the role of intraplate deformation and associated differential vertical motions. The wavelength and magnitude of lithospheric warping due to intraplate deformation and lithosphere-mantle interactions depend strongly on thermal age and rheology of the lithosphere and should not be ignored in geodynamic models (66, 67). In addition, better quantitative understanding of global effects of seemingly regional factors, such as stress field changes operating on plate scales and their effect on eustasy, is also needed. The same applies to the effects of major opening and closing of tectonic gateways (e.g., closure of neo-Tethys, closure of mid-America Seaway, breakup of North Atlantic, opening of Scotia Arc, etc.).

Dynamic topography

In practice, dynamic topography is what remains after shorter-term isostatic effects have been removed from the observed topography (68). Thus, dynamic topography complicates physical measures of sea level change based on stratal patterns, even when corrections for local factors have been applied. The long-wavelength warping effect of dynamic topography becomes notably apparent on the relatively longer time scales (multiple millions of years) and can dampen or enhance any underlying global signal of the third-order sea level events that occur on shorter time scales, but it rarely wipes out this signal completely. The U.S. East Coast provides an appropriate example of the influence of this factor because of the numerous modeling studies that followed the backstripped stratigraphic amplitude estimates along this margin. The modeled migration of the remnants of the Farallon Plate beneath North America (20) shows the effects of this dynamic passage on surface morphology since the Late Cretaceous. Assuming these models to be valid, this dynamic warping effect may have led the New Jersey margin to subside between 105 and 180 m since the late Cretaceous [see (10) for a discussion]. This example illustrates the considerable challenge involved in extracting a eustatic long-term sea level signal from local stratigraphic measurements.

Disentangling measures and mechanisms of Cretaceous sea level changes

It has become apparent that a measure of the amplitude of eustatic events cannot be gleaned from any one location, as discussed above. Instead, we have to resort to averaging it from global data after chronological consistency of events has been established. Nonetheless, the issue of the causes of third-order events (other than glacial sequestration of water on land) remains conjectural at this point. As mentioned earlier, at least in the Cretaceous, net gain and loss of water to the oceans from mantle sources due to the high crustal production rates remains a likely possibility that may have contributed to the long-term sea level highs and lows of that period.

Recently, an updated model of Cretaceous eustatic sea level variations based on a synthesis of broadly distributed stratigraphic database has been presented (10) (Fig. 6). This synthesis also took into account most of the published Cretaceous stable oxygen-isotopic data, plus an additional isotopic data set for a nearly complete Cretaceous pelagic carbonate section in Italy. Although such isotopic data (69) were not helpful in the estimation of amplitude of sea level changes [see discussion in (10)], major isotopic excursions aided in verifying the timing of the prominent shifts that were consistent with the stratigraphic data.

Long- and short-term trends

Constraints for the Cretaceous long-term eustatic curve were originally provided by continental flooding data and ridge volume estimates (47, 70). However, now that links with sea floor activities and ocean container capacity changes have been established, model reconstructions of the mean age of the oceanic crust and emplacement of substantial igneous accumulations on the sea floor (31, 71, 72) can provide additional constraints, but they do not change the overall shape of the original long-term curve (1) for the Cretaceous. Most recent estimates derived from geophysical models have now converged and are within the same range of amplitudes (e.g., the Cenomanian-Turonian all-time high in the Cretaceous is estimated at between 170 and 250 m above present-day mean sea level). A noticeable feature of the long-term envelope is that the mean sea level throughout the Cretaceous remained higher than at present, with a low in the Early Cretaceous and peak high in the late mid-Cretaceous. The updated Cretaceous short-term sea level curve is based on the premise that events that can be documented to be geologically synchronous in several, preferably noncontiguous, basins represent global rather than eurybatic events [see discussion in (10)].

Tectonics and eustasy: temporal connections

Figure 6 illustrates the revised eustatic curve for the Cretaceous Period (10) and the various potential mechanisms that could have been operative collectively to produce the resulting long-term sea level trends and possibly also contribute toward the short-term changes. Although the long-term trends can be explained by a number of these processes, individually or in combination, the reasons for shorter-term (third-order) cyclicity remain elusive. A simple comparison of the swing points (where the trends in sea level change) on the long-term eustatic curve with the timing of major tectonic events in the Atlantic domain (73) seems

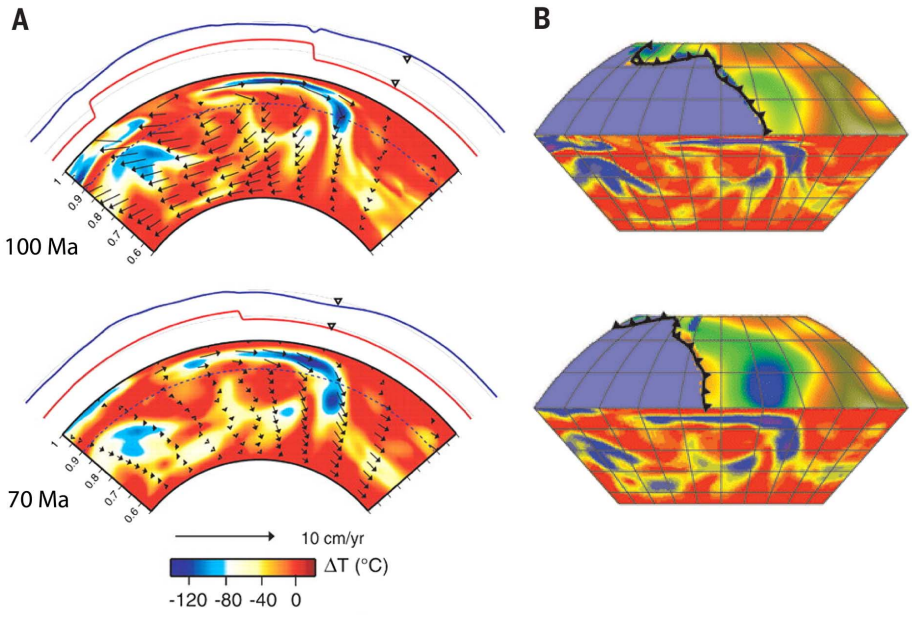


Fig. 3. Modeled temporal evolution of Farallon slab (shown in blue) subducting under North America leading to differential dynamic topography patterns, with down-warping of overlying continental lithosphere. (**A**) Results of a forward model, with a cross section showing temperature anomalies in the mantle, dynamic topography (blue line), and longitudinal component of plate motion (red line) during the Middle (100 Ma) and Late (70 Ma) Cretaceous [adapted from (20)]. (**B**) The same time slices with North America superimposed [adapted from (83)]. Oceanic lithosphere is subducted along a trench (solid line with triangles) on the western margin of the North American continent. During the Cretaceous, the topographic low (areas in green and blue) migrates east together with the remnants of the subducted plate.

to suggest a connection between major plate reorganizations and initiation of rifting with these swings. Large-scale volcanic activity and continental breakup events (*29*) also seem to complement the overall trends. It should be noted that the onset of rifting in the Atlantic was followed by a prolonged sea floor–spreading phase that occurred over a wide realm that must have affected the Cretaceous record. In contrast to major plate reorganizations (and their related stress fields) with global implications, most stress regime fluctuations and their effect on the sea level were more likely of regional nature.

Discussion: Causal mechanisms operative in the Cretaceous

Multiple interrelated factors can account for the long-term changes, and some of them hold promise to shed light on the shorter-term sea level variations.

Water sequestration on land

During the Cretaceous, the Northern Hemisphere was ice-free while the Southern Hemisphere had the large continent of Antarctica stationed in polar position. Because ice sheets are very sensitive to altitude, it remains to be determined whether Antarctica displayed enough temporal variations in topography to cause substantial variability in water sequestration on land to influence third-order cyclicity. This will have to wait better geodynamic models of this continent. Currently, however, collective opinion is leaning toward accepting potentially enough accumulation on Antarctica to cause ice sheets at least in the Late Cretaceous (*38*, *39*). If this mechanism was indeed operative, the extent of such transient ice sheets would have to have been more pronounced in the earliest Cretaceous (Berriasian to Hauterivian) than in the Late Cretaceous (Santonian through Maastrichtian). Evidence for this, of course, remains to be obtained.

Periodic water sequestration in the mantle

By any measure, the Cretaceous was a period of high oceanic crustal production rates and recycling in the mantle. As alluded to earlier, this could mean high water expulsion rates from the mantle sources into the ocean as well as high water entrainment rates with subduction. For this mechanism to produce a net increase (or decrease) of free water, the duration of the time lag between expulsion and entrainment would be of crucial importance. Also, the total volume of water in this exchange would have to be considerable to cause observable deviations from the long-term steady state. The relevance of this mechanism for the shorter third-order cycles as well as long-term trends remains undetermined at this time and needs to be explored. For this mechanism to be wholly or partially applicable to the Cretaceous eustatic record, more mantle-sourced water would have to be flushed into the oceans during late early and late mid-Cretaceous intervals (Fig. 6).

Variations in the container capacity of the oceans

In the Cretaceous, a large component of the long-term eustatic trends can be explained by changes in the container capacity of the oceans. Higher sea floor spreading rates producing greater amounts of new oceanic crust and larger ridge lengths all contributed to elevated bathymetry. This, combined with emplacement of LIPs, goes

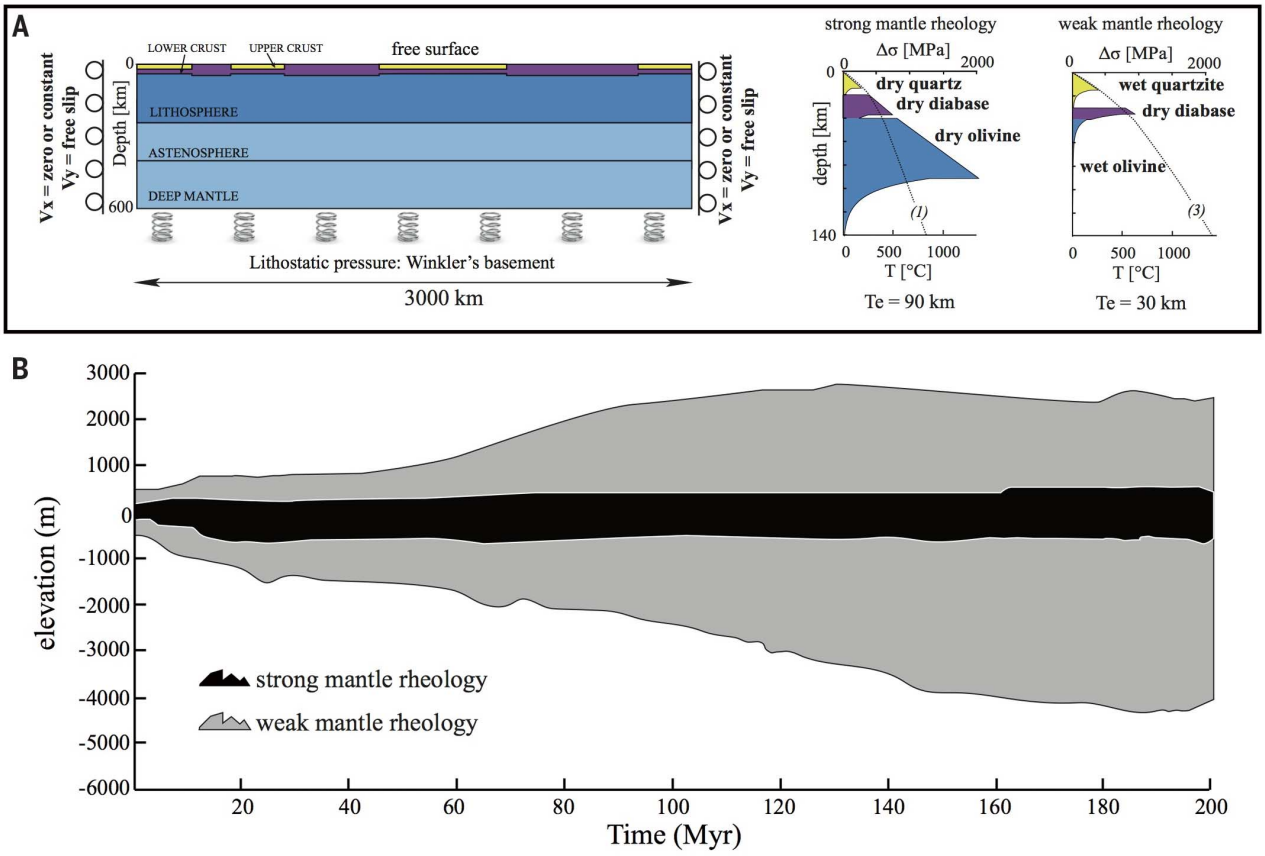


Fig. 4. Models showing capacity of continental lithosphere to support topographic and buried subsurface loads dependent on its strength. (**A**) Design of numerical model and two tested viscous-elasto-plastic yield strength rheological profiles with two end-member assumptions (strongest and weakest). (**B**) Predicted maximum amplitude of surface topography as a function of time [adapted from (*59*)]. During the Cretaceous, lithospheric strength has varied considerably, ranging from high production of ocean lithosphere to high plume activity. Strong surface amplitude variations in case of weak lithosphere with undepleted mantle result from Rayleigh-Taylor instabilities that develop in such lithosphere because of low viscosity of its negatively buoyant mantle part (the instabilities eventually initiate small-scale convection) and low flexural strength, which results in stronger surface response to the subsurface load variations associated with the R-T instabilities and mantle dynamics.

a long way toward explaining the overall high sea levels as well as variations in the long-term trends of this period. However, these could also contribute to some of the shorter-term variations, especially the rapid changes in container capacity such as those that occurred in the Late Cretaceous (e.g., rapid emplacement of the Deccan Traps, which were extruded within a few hundred thousand years, both on land and in the ocean) (52). Similarly, delivery of sediments to the ocean can overall enhance subsidence at the depocenters. This has not been systematically modeled, although available data (32) show that the effect of this factor was negative and did not vary much during the Cretaceous.

Mantle-lithosphere interaction and intraplate deformation

The Cretaceous was also a period of active plume emplacement in the oceanic settings that would have led to superswell development and rise of sea level to be later followed by subsidence. Plume emplacement is also often associated with continental breakup and the formation of volcanic rifted margins. A striking example of this is the emplacement of plumes at the end of the Cretaceous during the development of the Norwegian-Greenland Sea, while at the same time creating anomalous topography on the continental margins in the North Atlantic. The effects of the plumes on the surface are sensitive to inherited lithospheric structure and lead to spatial and temporal variations in vertical motions in continental interiors as well as along their margins. As shown by a recent study (74), intraplate stresses also have an impact on the long-wavelength surface response to mantle upwelling.

During the Cretaceous, much of the lithosphere was thermally rejuvenated at the onset of the period, and it subsequently aged and cooled. As a result, there was an overall long-term increase in plate rigidity. This increase in strength has in turn increased the capacity of the lithosphere to sustain greater topographic loads. At the same time, it led to flexural widening of the post-rift phase, resulting in apparent onlap in the longer term. As we have already observed, the Cretaceous was a time of major plate reorganizations with strong impact on stress fields in the lithosphere and resulting in intraplate deformation that affected differential vertical motions over wide areas. A characteristic feature of these stress-induced vertical motions is their short-term nature that might affect regional changes on the third-order time scales. For the Cretaceous record, it has been found that rift basins on the African plate underwent simultaneous anomalous differential motions, coinciding with the timing of plate reorganization in the adjacent Atlantic and Indian oceans (75).

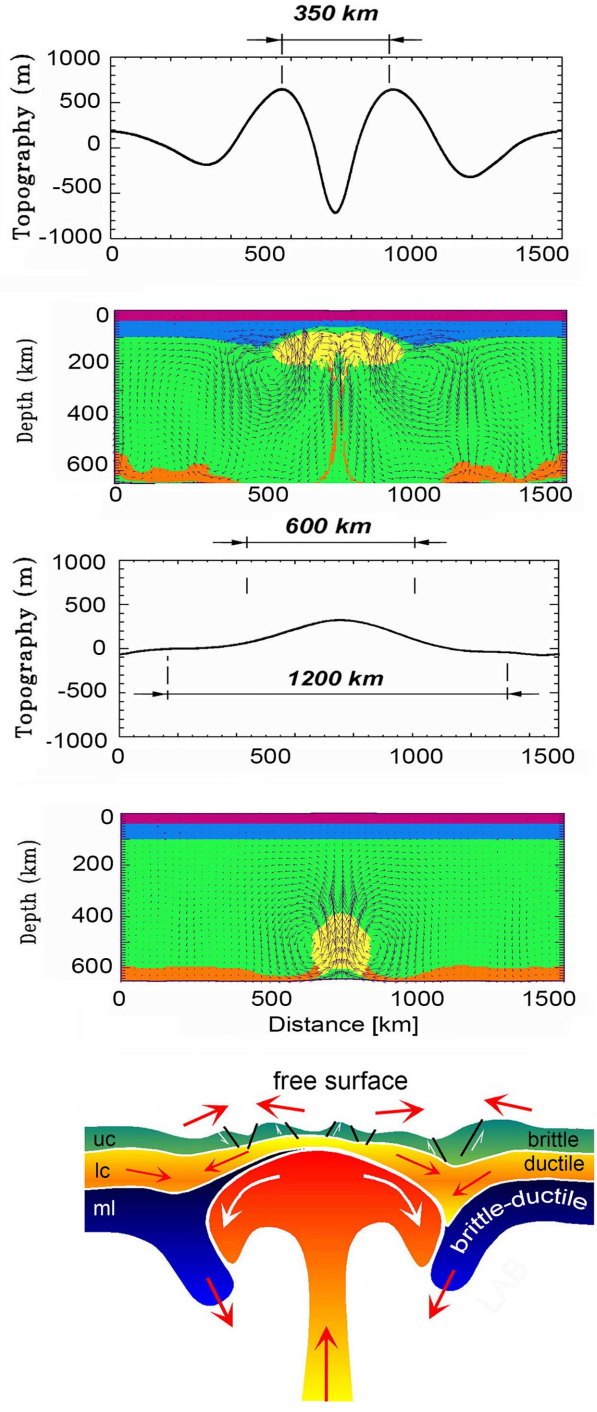


Fig. 5. The surface signature of mantle-lithosphere interactions is strongly dependent on the rheological stratification of the lithosphere. In contrast to the effect of ascending plume in oceanic domain leading to a topographic swell with a wavelength of several hundred kilometers, the interaction of a plume with the stratified continental lithosphere leads to differential uplift and subsidence becoming more pronouncedly differentiated with time [adapted from (56)]. uc, upper crust; lc, lower crust; ml, mantle lithosphere. A large part of the mantle impact on topography is dampened in the ductile lower crust and modulated by mechanical instabilities in the rheological layers.

Dynamic topography

Dynamic topographic predictions are very sensitive to assumptions about crustal thickness and rheological strength and the presumed characteristics of mantle upwelling and mode and geometry of subducting plates. The results of dynamic topographic models are thus prone to intrinsic uncertainties, requiring high-resolution tomographic imaging. Commonly, dynamic topographic models assume rigid plates, and for the Cretaceous most models focus on the effects of down-going slabs leading to overall downwarping and tilting of overlying continents. Dynamic topographic hindcasting rarely goes all the way back to the Early Cretaceous.

One such reconstruction is from Australia, where regional topography inferred is thought to have been influenced by an ancient subducting slab (76). In the Early Cretaceous, the eastern margin of Australia deflected downward in response to active subduction to its east. After subduction ceased in the Late Cretaceous and remnants of the subducted slab moved downward and westward relative to the eastward-drifting Australian Plate, the eastern half of Australia may have sunk as much as a kilometer, inverting only after the slab descended deeper into the mantle in the latest Cretaceous (76). These vertical motions would strongly influence all stratigraphic estimates

of sea level variations along this and other margins from Cretaceous strata in Australia.

The example of regional topographic anomalies related to the fragments of Farallon slab traversing underneath North America and its influence on the East Coast stratigraphic sea level estimates has already been mentioned (*20*) (Fig. 3). These reconstructions, which go back to the mid-Cretaceous, modeled regional subsidence and ensuing flooding, which shows sensitivity to changes in the assumed mantle viscosity structures through time. The authors used an inverse mantle convection model that simulates seismic structure and plate motion to reconstruct Farallon plate subduction back to the mid-Cretaceous with the objective to constrain depth dependence of mantle viscosity and buoyancy by comparing model prediction and stratigraphy. Another, younger than Cretaceous (Neogene), example is provided by a modeling study of African topographic anomalies that are interpreted to have been influenced by buoyant forces in the mantle for the past 30 million years (*77*).

Nevertheless, existing dynamic topography models indicate an overall negative global contribution by crustal topographic variations to the long-term sea level change (*32*, *72*). Models for present-day dynamic topography for the Mediterranean (*78*) show that mantle flow can produce strong differentiation in zones of positive and negative topography, linked to assumptions about crustal structure. Thus, there is considerable challenge involved in reconstructing detailed crustal structure in the geologic past (Fig. 7). It is also noteworthy that in the continents, dynamic topography can be substantially modulated as the result of a dampening effect of lithosphere-asthenosphere boundary undulations in the lower crust and short-wavelength modulation of the associated signal due to mechanical response of competent crustal and mantle layers (*56*).

Relative contribution of various mechanisms and their feedbacks and interconnections

Qualitative measures of contribution of various factors on the long-term sea level trends are relatively easy to envision. In contrast, quantitative assessments are fraught with difficulties. Most workers nevertheless agree that mean age of the oceanic crust and ocean floor productivity are the largest contributors to long-term sea level trends in the Cretaceous. Sea floor productivity contributes highest positive contribution, varying between 250 and 225 m of sea level in the Cretaceous (*32*). Volcanic activity becomes a relatively major positive contributor

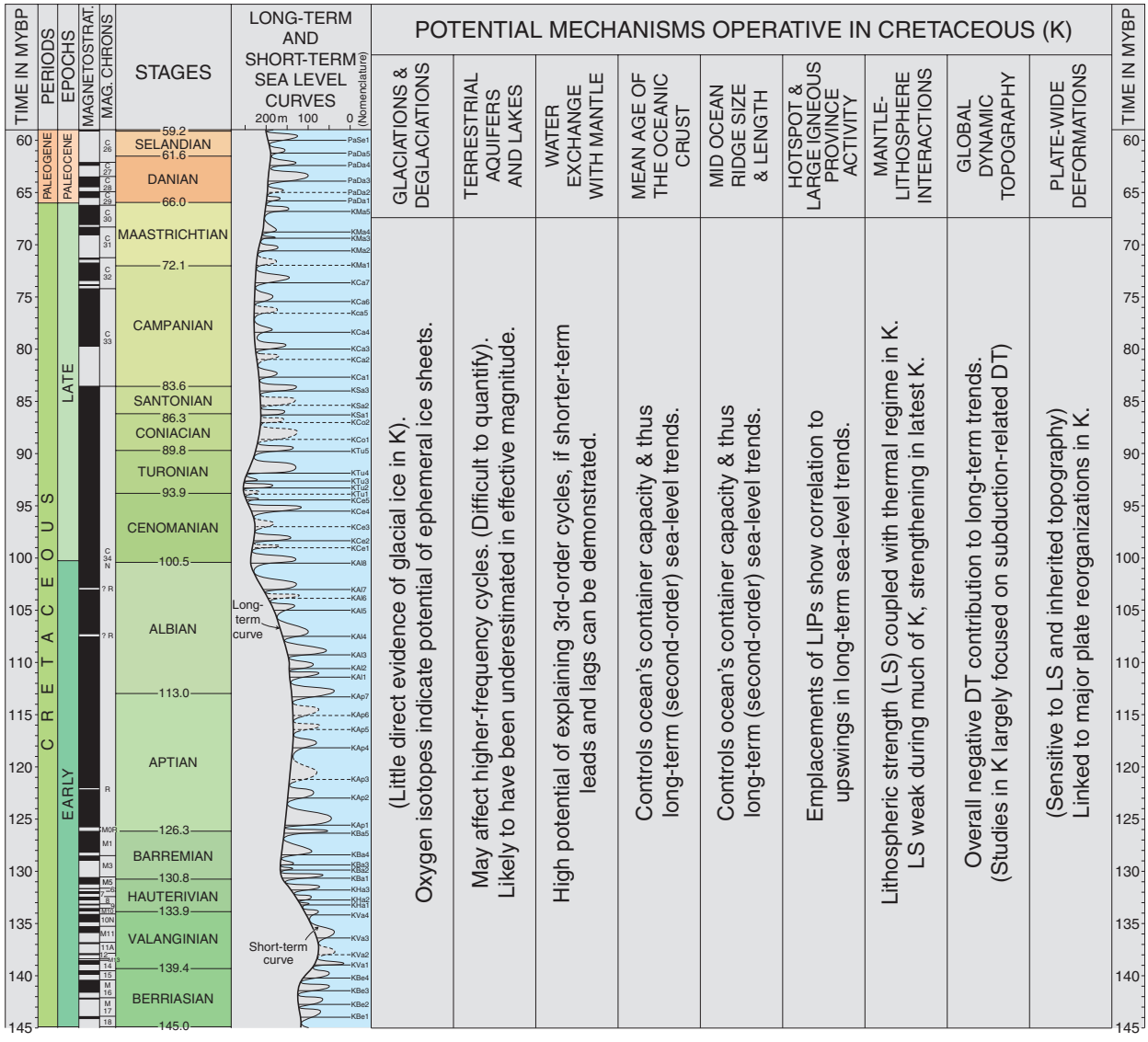


Fig. 6. An updated eustatic curve for the Cretaceous Period plotted against various potential mechanisms that could have been operative collectively in this period to produce the resulting long-term (and possibly also contribute toward the short-term) sea level trends [adapted from (*10*)]. Whereas the long-term trends can be explained by a number of these processes, individually or in combination, the reasons for shorter-term (third-order) cyclicity remain elusive. MYBP, millions of years before the present.

in the Late Cretaceous (~50 m). Sediments provide a relatively constant negative contribution overall of ~50 m through the same period.

Spasojevic and Gurnis (72) suggest that with the exception of Southeast Asia (and possibly Australia), in the Late Cretaceous time dynamic topography provides an overall negative contribution to sea level up to the order of 1 km, primarily due to slab subduction, creating broad topographic depressions at the surface of the overlying lithosphere. These models, however, could be effectively refined once the effects of internal deformation of the plates and their obvious dependence on rheology are incorporated. These deformations can lead to vertical motions on the time scale of third-order events, with a wide range of magnitudes that vary from a few tens of meters for strong plates to several hundreds of meters (up to 1 km) for weaker ones. As discussed earlier, the contribution of this factor was most likely higher in the Early Cretaceous than in the Late Cretaceous, depending on the rheological strength as well as level of tectonic activity.

We should once again underscore that most of the factors discussed above are interrelated and have feedbacks. Changes in mantle convection not only influence dynamic topography but could also have a direct effect on the stress field within the plate, leading to differential vertical motions at different spatial scales. Resultant topographies also generate stresses and deformation, whereas rigidity of the lithosphere is a key to sustained topography (75). During the Cretaceous, with relatively high and variable mantle activity and a peak in the crust generation in the Early Cretaceous and volcanism in the Late Cretaceous, the nature of feedbacks would have also changed. In all of these, inherited spatial variability in the lithosphere structure is the key to surface topography and its influence on sea level change.

Conclusions and future perspectives

Substantial progress has been made in our understanding of the connection between deep-Earth processes and their surface expression, which has brought about a desirable convergence of views between solid-Earth geophysics and sedimentary geology concerning eustasy. It is becoming obvious that knowledge of inherited topography is crucial to the disentangling of sea level changes at all spatial and temporal time scales and perhaps in the understanding of their causal mechanisms. Solid-Earth mechanisms that modify inherited landscapes and our measure of sea level are many and often interrelated. These include glacial isostatic adjustment, dynamic topography, mantle-lithosphere interactions, oceanic crustal production rate variations, and plate reorganizations and deformations, among others. Superimposed on these are the broad climatic influences, of which glaciations and deglaciations are most relevant to sea level change. We are now at a stage where we have begun to understand many of these causal mechanisms and their complex feedbacks for the long-term trends in eustatic changes, and estimates for the amplitude of these changes are also converging based on different data sets. Likewise, considerable headway has been made in deciphering the causes for short-term eurybatic changes on regional scales that are also influenced by lithospheric memory and its responses to coupled deep-Earth processes (79, 80). We have also become aware that the role of volcanism in rifting can be an important determinant of the long-term sea level change, which can produce punctuated although not cyclic changes with a frequency of 5 to 10 million years. Alternatively, the record of sea level change can provide an important constraint on upper mantle processes.

Despite this progress in understanding long-term eustasy, causes for third-order cyclicity remain enigmatic, although there are avenues that should be investigated further, such as periodic imbalance between water entrainment and expulsion within the mantle, and potentially longer-term residence of substantial amount of water on land during dry intervals (lowstands). The latter is especially relevant to sea level changes related to higher-frequency Milankovitch-scale oscillations. It is quite possible that we are misjudging both the time scales and magnitude of both these effects in influencing the stratigraphic record. Dynamic topography has come to the rescue in explaining away the disparities in the amplitudes of sea level measured stratigraphically at different places in spite of the underlying eustatic signal. But it does not shed light on the causal mechanisms for these changes or the occurrence of quasi-cyclic third-order changes on global scale. In addition, we have also learned that the role of volcanism in rifting can be a significant determinant of the long-term sea level change, which can produce punctuated although not cyclic changes with frequency of 5 to 10 million years.

As regards the causal mechanisms of quasi-cyclic third-order fluctuations, if we hold water sequestration on land as a major contributor, it would require ice volume to fluctuate with amplitudes that are large enough to drive several tens of meters of sea level change. Thus, it is imperative that we find direct evidence of significant glaciation in the Cretaceous. The most likely source of such glaciation would probably be located at high altitudes on Antarctica. This again necessitates inherited elevated topography on a scale requiring the involvement of large-scale mantle processes underneath Antarctica. The period was also characterized by large-scale volcanic activity associated with plate reorganizations and continental breakup events. However, the temporal evolution of Cretaceous volcanic activity associated with plate reorganizations and breakup events needs to be constrained in more detail if credible links must be established between tectonics and regional and global sea level changes in this period. Inherited land and sea floor topographies also set the stage for understanding the events to follow in the Cenozoic.

Lateral variations in upper mantle viscosity and lithosphere rheology that have implications for both glacial isostatic adjustment and dynamic topography must be further investigated. Glacial isostatic adjustment models have only recently begun to take into account mantle heterogeneity and associated lateral variation in lithospheric strength (81), but these need to incorporated into longer-term dynamic topographic

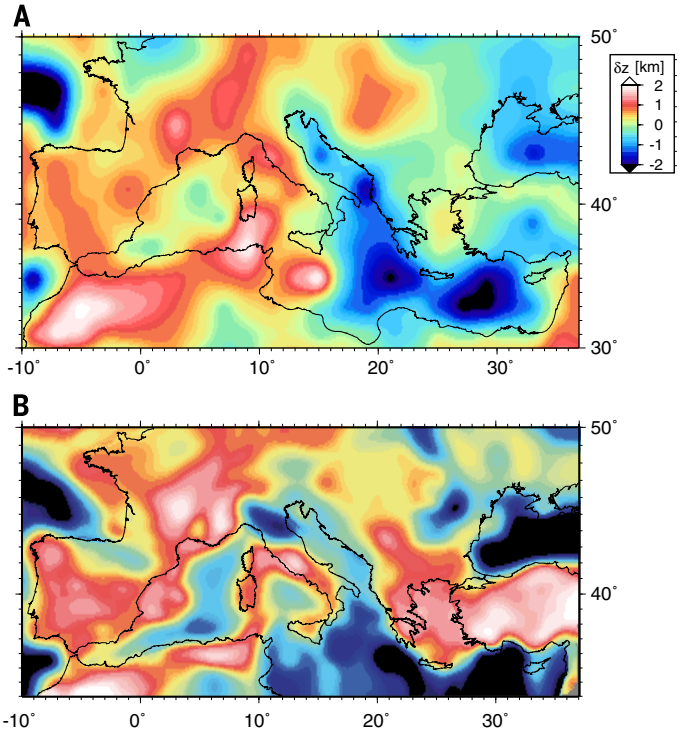


Fig. 7. Modeled prediction for dynamic topography for the Mediterranean area based on mantle flow patterns inferred from seismic tomography [adapted from (78)]. These models are also dependent on availability of constraints on crustal thickness. (**A**) Prediction adopting global crustal thickness model. (**B**) Vertical motions calculated adopting a higher resolution regional model the Mediterranean area based on available geophysical data in the region. This model predicts uplift in Anatolian plateau and Iberia, areas with elevated topography and subsidence in the eastern Mediterranean.

models as well. Recent model studies in the Mediterranean area have shown strong dependence of dynamic topography on lateral variations in crustal structure and rheology (*77*). Likewise, such coverage is needed on a global scale for crustal and lithosphere structure for the present day as well as the geologic past. A first-order global strength model for present-day oceanic and continental lithosphere has recently become available (*23*). We need to build on this initial effort to construct higher-resolution maps as well as past rheological models on a global scale.

It is evident that modern sea level–related research, through its intimate connection with inherited landscapes, is not only an integrative force between solid-Earth geophysics and sedimentary geology, but could also play a role in integrating continental and marine studies. In this connection, deep-drilling programs, both on land and at sea, will be of crucial importance for achieving the integrative goals. Finally, we note that a review by Schlanger (*82*) had already predicted in 1986 that higher-frequency eustatic variations will not be understood by a single discipline (stratigraphy) alone and may ultimately involve a geophysical solution.

REFERENCES AND NOTES

1. B. U. Haq, J. Hardenbol, P. R. Vail, Chronology of fluctuating sea levels since the triassic. *Science* **235**, 1156–1167 (1987). doi: 10.1126/science.235.4793.1156; pmid: 17818978
2. B. U. Haq, S. R. Schutter, A chronology of Paleozoic sea-level changes. *Science* **322**, 64–68 (2008). doi: 10.1126/science.1161648; pmid: 18832639
3. M. Gurnis, Phanerozoic marine inundation of continents driven by dynamic topography above subducting slabs. *Nature* **364**, 589–593 (1993). doi: 10.1038/364589a0
4. A. D. Miall, Whither stratigraphy? *Sediment. Geol.* **100**, 5–20 (1995). doi: 10.1016/0037-0738(95)00100-X
5. J. Ritsema, H. J. van Heijst, J. H. Woodhouse, Global transition zone tomography. *J. Geophys. Res.* **109**, B02302 (2004). doi: 10.1029/2003JB002610
6. J. Ritsema, A. K. McNamara, A. L. Bull, Tomographic filtering of geodynamic models: Implications for model interpretation and large-scale mantle structure. *J. Geophys. Res.* **112**, B01303 (2007). doi: 10.1029/2006JB004566
7. K. Sigloch, N. McQuarrie, G. Nolet, Two-stage subduction history under North America inferred from multiple-frequency tomography. *Nat. Geosci.* **1**, 458–462 (2008). doi: 10.1038/ngeo231
8. F. Rickers, A. Fichtner, J. Trampert, The Iceland–Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: Evidence from full-waveform inversion. *Earth Planet. Sci. Lett.* **367**, 39–51 (2013). doi: 10.1016/j.epsl.2013.02.022
9. D. G. van der Meer, W. Spakman, D. J. van Hinsbergen, M. I. Amaru, T. H. Torsvik, Towards absolute plate motions constrained by lower-mantle slab remnants. *Nat. Geosci.* **3**, 36–40 (2009). doi: 10.1038/ngeo708
10. B. U. Haq, Cretaceous eustasy revisited. *Global Planet. Change* **113**, 44–58 (2014). doi: 10.1016/j.gloplacha.2013.12.007
11. M. Gurnis, Rapid continental subsidence following the initiation and evolution of subduction. *Science* **255**, 1556–1558 (1992). doi: 10.1126/science.255.5051.1556; pmid: 17820168
12. C. Lithgow-Bertelloni, M. Gurnis, Cenozoic subsidence and uplift of continents from time-varying dynamic topography. *Geology* **25**, 735–738 (1997). doi: 10.1130/0091-7613(1997)025<0735:CSAUOC>2.3.CO;2
13. S. Cloetingh, S. D. Willett, TOPO-EUROPE: Understanding of the coupling between the deep Earth and continental topography. *Tectonophysics* **602**, 1–14 (2013). doi: 10.1016/j.tecto.2013.05.023
14. C. A. Nittrrouer, J. A. Austin, M. E. Field, J. H. Kravitz, J. P. Syvitski, P. L. Wiberg, Eds., *Intl. Assoc. Sedimentol. Spec. Pub. 37* (2007).
15. L. Matenco, P. Andriessen, Quantifying the mass transfer from mountain ranges to deposition in sedimentary basins: Source to sink studies in the Danube Basin–Black Sea system. *Global Planet. Change* **103**, 1–18 (2013). doi: 10.1016/j.gloplacha.2013.01.003
16. J. X. Mitrovica, W. R. Peltier, On postglacial geoid subsidence over the equatorial oceans. *J. Geophys. Res.* **96**, 20053–20071 (1991). doi: 10.1029/91JB01284
17. K. Lambeck, C. Smither, M. Ekman, Tests of glacial rebound models for Fennoscandinavia based on instrumented sea- and lake-level records. *Geophys. J. Int.* **135**, 375–387 (1998). doi: 10.1046/j.1365-246X.1998.00643.x
18. M. S. Steckler, A. B. Watts, Subsidence of the Atlantic-type continental margin off New York. *Earth Planet. Sci. Lett.* **41**, 1–13 (1978). doi: 10.1016/0012-821X(78)90036-5
19. A. B. Watts, Crust and lithosphere dynamics: An overview. In *Treatise on Geophysics, Vol. 6, Crust and Lithosphere Dynamics*, A. B. Watts, Ed. (Elsevier, Amsterdam, 2007), pp. 1–48.
20. L. Liu, S. Spasojevic, M. Gurnis, Reconstructing Farallon plate subduction beneath North America back to the Late Cretaceous. *Science* **322**, 934–938 (2008). doi: 10.1126/science.1162921; pmid: 18988850
21. M. Kominz *et al.*, Late Cretaceous to Miocene sea-level estimates from the New Jersey and Delaware coastal plain coreholes: An error analysis. *Basin Res.* **20**, 211–226 (2008). doi: 10.1111/j.1365-2117.2008.00354.x
22. E. B. Burov, Rheology and strength of the lithosphere. *Mar. Pet. Geol.* **28**, 1402–1443 (2011). doi: 10.1016/j.marpetgeo.2011.05.008
23. M. Tesauro, M. K. Kaban, S. Cloetingh, Global strength and elastic thickness of the lithosphere. *Global Planet. Change* **90-91**, 52–57 (2012). doi: 10.1016/j.gloplacha.2011.12.003
24. C. Jaupart, S. Labrosse, J. C. Mareschal, In *Treatise on Geophysics, Vol. 7, Mantle Dynamics*, D. Bercovici, Ed. (Elsevier, Amsterdam, 2007), pp. 253–303. doi: 10.1016/B978-044452748-6/00114-0
25. I. Artemieva, *The Lithosphere: An Interdisciplinary Approach* (Cambridge Univ. Press, Cambridge, 2011).
26. W. W. Hay, Evolving ideas about the Cretaceous climate and ocean circulation. *Cretac. Res.* **29**, 725–753 (2008). doi: 10.1016/j.cretres.2008.05.025
27. M. F. Coffin, O. Eldholm, Large igneous provinces: Crustal structure, dimensions, and external consequences. *Rev. Geophys.* **32**, 1–36 (1994). doi: 10.1029/93RG02508
28. G. Ito, P. D. Clift, Subsidence and growth of Pacific Cretaceous plateaus. *Earth Planet. Sci. Lett.* **161**, 85–100 (1998). doi: 10.1016/S0012-821X(98)00139-3
29. P. A. Ziegler, S. Cloetingh, Dynamic processes controlling evolution of rifted basins. *Earth Sci. Rev.* **64**, 1–50 (2004). doi: 10.1016/S0012-8252(03)00041-2
30. J. Skogseid *et al.*, NE Atlantic continental rifting and volcanic margin formation. *Spec. Pub. Geol. Soc. London* **167**, 295–326 (2000). doi: 10.1144/GSL.SP.2000.167.01.12
31. R. D. Müller, M. Sdrolias, C. Gaina, B. Steinberger, C. Heine, Long-term sea-level fluctuations driven by ocean basin dynamics. *Science* **319**, 1357–1362 (2008). doi: 10.1126/science.1151540; pmid: 18323446
32. C. P. Conrad, The solid Earth's influence on sea level. *Geol. Soc. Am. Bull.* **125**, 1027–1052 (2013). doi: 10.1130/B30764.1
33. W. W. Hay, M. A. Leslie, In *Studies in Geophysics: Sea-Level Change* (National Academy Press, Washington, DC, 1990), pp. 161–170.
34. D. K. Jacobs, D. L. Sahagian, Climate-induced fluctuations in sea level during non-glacial times. *Nature* **361**, 710–712 (1993). doi: 10.1038/361710a0
35. J. E. Wendler, S. R. Meyers, I. Wendler, J. Kuss, A million-year scale astronomical control on Late Cretaceous sea level. *Newslett. Stratigr.* **47**, 1–19 (2014). doi: 10.1127/0078-0421/2014/0038
36. M. Wagreich, R. Lein, B. Sames, Eustasy, its controlling factors and the Limno-eustatic hypothesis—concepts inspired by Eduard Suess. *Aust. J. Earth Sci.* **107**, 115–131 (2014).
37. M. Deynoux, J. M. G. Miller, E. W. Domack, Eds., *Earth's Glacial Record* (Cambridge Univ. Press, Cambridge, 1994).
38. R. K. Matthews, R. Z. Poore, Tertiary $\delta^{18}O$ record and glacio-eustatic sea-level fluctuations. *Geology* **8**, 501–504 (1980). doi: 10.1130/0091-7613(1980)8<501:TORAGS>2.0.CO;2
39. K. G. Miller *et al.*, Late Cretaceous chronology of large, rapid sea-level changes: Glacioeustasy during the greenhouse world. *Geology* **31**, 585–588 (2003). doi: 10.1130/0091-7613(2003)031<0585:LCCOLR>2.0.CO;2
40. H. M. Stoll, D. P. Schrag, Evidence of glacial control of rapid sea level changes in the Early Cretaceous. *Science* **272**, 1771–1774 (1996). doi: 10.1126/science.272.5269.1771; pmid: 8662480
41. H. M. Stoll, D. P. Schrag, High-resolution stable isotope records from the Upper Cretaceous rocks of Italy and Spain: Glacial episodes in a greenhouse planet? *Geol. Soc. Am. Bull.* **112**, 308–319 (2000). doi: 10.1130/0016-7606(2000)112<308:HSIRFT>2.0.CO;2
42. W. E. Farrell, J. A. Clark, On postglacial sea level. *Geophys. J. R. Astron. Soc.* **46**, 647–667 (1976). doi: 10.1111/j.1365-246X.1976.tb01252.x
43. C. P. Conrad, B. H. Hager, Spatial variations in the rate of sea level rise caused by the present-day melting of glaciers and ice sheets. *Geophys. Res. Lett.* **24**, 1503–1506 (1997). doi: 10.1029/97GL01338
44. J. X. Mitrovica, M. E. Tamisiea, J. L. Davis, G. A. Milne, Recent mass balance of polar ice sheets inferred from patterns of global sea-level change. *Nature* **409**, 1026–1029 (2001). doi: 10.1038/35059054; pmid: 11234008
45. M. Gurnis, Ridge spreading, subduction, and sea level fluctuations. *Science* **250**, 970–972 (1990). doi: 10.1126/science.250.4983.970; pmid: 17746922
46. J. Mitrovica, C. Beaumont, G. Jarvis, Tilting of continental interiors by the dynamical effects of subduction. *Tectonics* **8**, 1079–1094 (1989). doi: 10.1029/TC008i005p01079
47. W. C. Pitman III, X. Golovchenko, The effect of sea level change on the shelf edge and slope of passive margins. *SEPM Spec. Pub.* **33**, 41–58 (1983).
48. J. F. Dewey, W. C. Pitman, Sea-level changes: Mechanisms, magnitudes and rates. *SEPM Spec. Pub.* **58**, 95–127 (1997).
49. M. A. Kominz, Oceanic ridge volumes and sea-level change-an error analysis. *Am. Assoc. Petrol. Geol. Spec. Pub.* **A166**, 109–127 (1984).
50. D. B. Rowley, Rate of plate creation and destruction: 180 Ma to present. *Geol. Soc. Am. Bull.* **114**, 927–933 (2002). doi: 10.1130/0016-7606(2002)114<0927:ROPCAD>2.0.CO;2
51. R. W. Becker, C. P. Conrad, B. Buffett, R. D. Müller, Past and present seafloor age distributions and the temporal evolution of plate tectonic heat transport. *Earth Planet. Sci. Lett.* **278**, 233–242 (2009). doi: 10.1016/j.epsl.2008.12.007
52. V. Courtillot *et al.*, Deccan flood basalts at the Cretaceous/Tertiary boundary? *Earth Planet. Sci. Lett.* **80**, 361–374 (1986). doi: 10.1016/0012-821X(86)90118-4
53. S. Spasojević, L. Liu, M. Gurnis, R. D. Müller, The case for dynamic subsidence of the US east coast since the Eocene. *Geophys. Res. Lett.* **35**, L08305 (2008). doi: 10.1029/2008GL033511
54. R. Moucha *et al.*, Dynamic topography and long-term sea-level variations: There is no such thing as a stable continental platform. *Earth Planet. Sci. Lett.* **271**, 101–108 (2008). doi: 10.1016/j.epsl.2008.03.056
55. D. B. Rowley *et al.*, Dynamic topography change of the eastern United States since 3 million years ago. *Science* **340**, 1560–1563 (2013). doi: 10.1126/science.1229180; pmid: 23686342
56. E. Burov, S. Cloetingh, Controls of mantle plumes and lithospheric folding on modes of intraplate continental tectonics: Differences and similarities. *Geophys. J. Int.* **178**, 1691–1722 (2009). doi: 10.1111/j.1365-246X.2009.04238.x
57. T. Duretz, T. Gerya, Slab detachment during continental collision: Influence of crustal rheology and interaction with lithospheric delamination. *Tectonophysics* **602**, 124–140 (2013). doi: 10.1016/j.tecto.2012.12.024
58. E. Burov, The equivalent elastic thickness (T_e), seismicity and the long-term rheology of continental lithosphere: Time to burn-out "crème brûlée?: Insights from large-scale geodynamic modeling. *Tectonophysics* **484**, 4–26 (2010). doi: 10.1016/j.tecto.2009.06.013
59. T. François, E. Burov, B. Meyer, P. Agard, Surface topography as key constraint on thermo-rheological structure of stable cratons. *Tectonophysics* **602**, 106–123 (2013). doi: 10.1016/j.tecto.2012.10.009
60. D. P. McKenzie, Some remarks on the development of sedimentary basins. *Earth Planet. Sci. Lett.* **40**, 25–32 (1978). doi: 10.1016/0012-821X(78)90071-7
61. P. Unternehr, G. Péron-Pinvidic, G. Manatschal, E. Sutra, Hyper-extended crust in the South Atlantic: In search of a model. *Petrol. Geosci.* **16**, 207–215 (2010). doi: 10.1144/1354-079309-904
62. S. Cloetingh, P. Ziegler, In *Treatise on Geophysics, Vol. 6, Crust and Lithosphere Dynamics*, A. B. Watts, Ed. (Elsevier, Amsterdam, 2007), pp. 486–611.
63. H. Johnson, T. G. Doré, R. W. Gatliff, R. Holdsworth, E. Lundin, D. Ritchie, Eds., *Geol. Soc. London Spec. Pub. 206* (2008).

64. S. Dupré, G. Bertotti, S. Cloetingh, Tectonic history along the South Gabon Basin: Anomalous early post-rift subsidence. *Mar. Pet. Geol.* **24**, 151–172 (2007). doi: 10.1016/j.marpetgeo.2006.11.003
65. G. Bertotti, M. Gouiza, Post-rift vertical movements and horizontal deformations in the eastern margin of the Central Atlantic: Middle Jurassic to Early Cretaceous evolution of Morocco. *Int. J. Earth Sci.* **101**, 2151–2165 (2012). doi: 10.1007/s00531-012-0773-4
66. S. Cloetingh, H. McQueen, K. Lambeck, On a tectonic mechanism for regional sea level variations. *Earth Planet. Sci. Lett.* **75**, 157–166 (1985). doi: 10.1016/0012-821X(85)90098-6
67. S. Cloetingh, E. Burov, T. Francois, Thermo-mechanical controls on intra-plate deformation and the role of plume-folding interactions in continental topography. *Gondwana Res.* **24**, 815–837 (2013). doi: 10.1016/j.gr.2012.11.012
68. N. Flament, M. Gurnis, R. D. Müller, A review of observations and models of dynamic topography. *Lithosphere* **5**, 189–210 (2013). doi: 10.1130/L245.1
69. Bulk carbonate oxygen isotopic data from the Contessa Section in Italy that covers almost all of the Cretaceous (with the exception of early Barremian) was provided by Schrag [see (*10*)] and compared to the sea level curve. A significant correlation was seen between timing of positive oxygen isotopic excursions and eustatic falls of sea level recorded by Haq *et al*. (*1*).
70. A. B. Ronov, Phanerozoic transgressions and regressions on the continents; a quantitative approach based on areas flooded by the sea and areas of marine and continental deposition. *Am. J. Sci.* **294**, 777–801 (1994). doi: 10.2475/ajs.294.7.777
71. J. P. Cogné, E. Humler, Trends and rhythms in global seafloor generation rate. *Geochem. Geophys. Geosyst.* **7**, Q03011 (2006). doi: 10.1029/2005GC001148
72. S. Spasojević, M. Gurnis, Sea level and vertical motion of continents from dynamic Earth models since the Late Cretaceous. *Am. Assoc. Pet. Geol. Bull.* **96**, 2037–2064 (2012).
73. R. Zühlke, M.-S. Bouaouda, B. Ouajhain, T. Bechstädt, R. Leinfelder, Quantitative Meso-Cenozoic development of the eastern central Atlantic continental shelf, western High Atlas, Morocco. *Mar. Pet. Geol.* **21**, 225–276 (2004). doi: 10.1016/j.marpetgeo.2003.11.014
74. E. Burov, T. Gerya, Asymmetric three-dimensional topography over mantle plumes. *Nature* **513**, 85–89 (2014). doi: 10.1038/nature13703; pmid: 25186903
75. M. Janssen, R. Stephenson, S. Cloetingh, Temporal and spatial correlations between changes in plate motions and the evolution of rifted basins in Africa. *Geol. Soc. Am. Bull.* **107**, 1317–1332 (1995). doi: 10.1130/0016-7606(1995)107<1317:TASCBC>2.3.CO;2
76. M. Gurnis, R. D. Müller, L. Moresi, Cretaceous vertical motion of Australia and the Australian-Antarctic discordance. *Science* **279**, 1499–1504 (1998). doi: 10.1126/science.279.5356.1499; pmid: 9488643
77. R. Moucha, A. M. Forte, Changes in African topography driven by mantle convection. *Nat. Geosci.* **4**, 707–712 (2011). doi: 10.1038/ngeo1235
78. L. Boschi, C. Faccenna, T. Becker, Mantle structure and dynamic topography in the Mediterranean Basin. *Geophys. Res. Lett.* **37**, L20303 (2010). doi: 10.1029/2010GL045001
79. S. Liu, D. Nummedal, M. Gurnis, Dynamic versus flexural controls of Late Cretaceous Western Interior Basin, USA. *Earth Planet. Sci. Lett.* **389**, 221–229 (2014). doi: 10.1016/j.epsl.2014.01.006
80. M. S. Miller, T. W. Becker, Reactivated lithospheric-scale discontinuities localize dynamic uplift of the Moroccan Atlas Mountains. *Geology* **42**, (2013). doi: 10.1130/G34959.1
81. W. van der Wal *et al*., Glacial isostatic adjustment model with composite 3-D Earth rheology for Fennoscandia. *Geophys. J. Int.* **194**, 61–77 (2013). doi: 10.1093/gji/ggt099
82. S. Schlanger, High frequency sea-level fluctuations in Cretaceous time: An emerging geophysical problem. *Geodyn. Ser. Am. Geophys. Union* **15**, 61–74 (1986).
83. B. Steinberger, Reconstructing Earth history in three dimensions. *Science* **322**, 866–868 (2008). doi: 10.1126/science.1166395; pmid: 18988830
84. National Research Council, In *Studies in Geophysics: Sea-Level Change* (National Academy Press, Washington, DC, 1990), pp. 3–34.

ACKNOWLEDGMENTS

We thank E. Burov, T. Francois, and H. Thybo for valuable discussions and comments on this paper, and our colleagues T. Watts and the late P. Ziegler with whom our discussions over the years have evolved our views about the interaction of solid Earth with eustasy. B.U.H. acknowledges National Science Foundation for many years of unfettered access to research time as a part of his functions at the Foundation. S.C. acknowledges the generous support provided by the Netherlands Research Center for Integrated Solid Earth Science and Utrecht University.

10.1126/science.1258375

RESEARCH ARTICLE SUMMARY

PROTEOMICS

Tissue-based map of the human proteome

Mathias Uhlén,* Linn Fagerberg, Björn M. Hallström, Cecilia Lindskog, Per Oksvold, Adil Mardinoglu, Åsa Sivertsson, Caroline Kampf, Evelina Sjöstedt, Anna Asplund, IngMarie Olsson, Karolina Edlund, Emma Lundberg, Sanjay Navani, Cristina Al-Khalili Szigyarto, Jacob Odeberg, Dijana Djureinovic, Jenny Ottosson Takanen, Sophia Hober, Tove Alm, Per-Henrik Edqvist, Holger Berling, Hanna Tegel, Jan Mulder, Johan Rockberg, Peter Nilsson, Jochen M. Schwenk, Marica Hamsten, Kalle von Feilitzen, Mattias Forsberg, Lukas Persson, Fredric Johansson, Martin Zwahlen, Gunnar von Heijne, Jens Nielsen, Fredrik Pontén

INTRODUCTION: Resolving the molecular details of proteome variation in the different tissues and organs of the human body would greatly increase our knowledge of human biology and disease. Here, we present a map of the human tissue proteome based on quantitative transcriptomics on a tissue and organ level combined with protein profiling using microarray-based immunohistochemistry to achieve spatial localization of proteins down to the single-cell level. We provide a global analysis of the secreted and membrane proteins, as well as an analysis of the expression profiles for all proteins targeted by pharmaceutical drugs and proteins implicated in cancer.

RATIONALE: We have used an integrative omics approach to study the spatial human proteome. Samples representing all major tissues and organs (n = 44) in the human body have been analyzed based on 24,028 antibodies corresponding to 16,975 protein-encoding genes, complemented with RNA-sequencing data for 32 of the tissues. The antibodies have been used to produce more than 13 million tissue-based immunohistochemistry images, each annotated by pathologists for all sampled tissues. To facilitate integration with other biological resources, all data are available for download and cross-referencing.

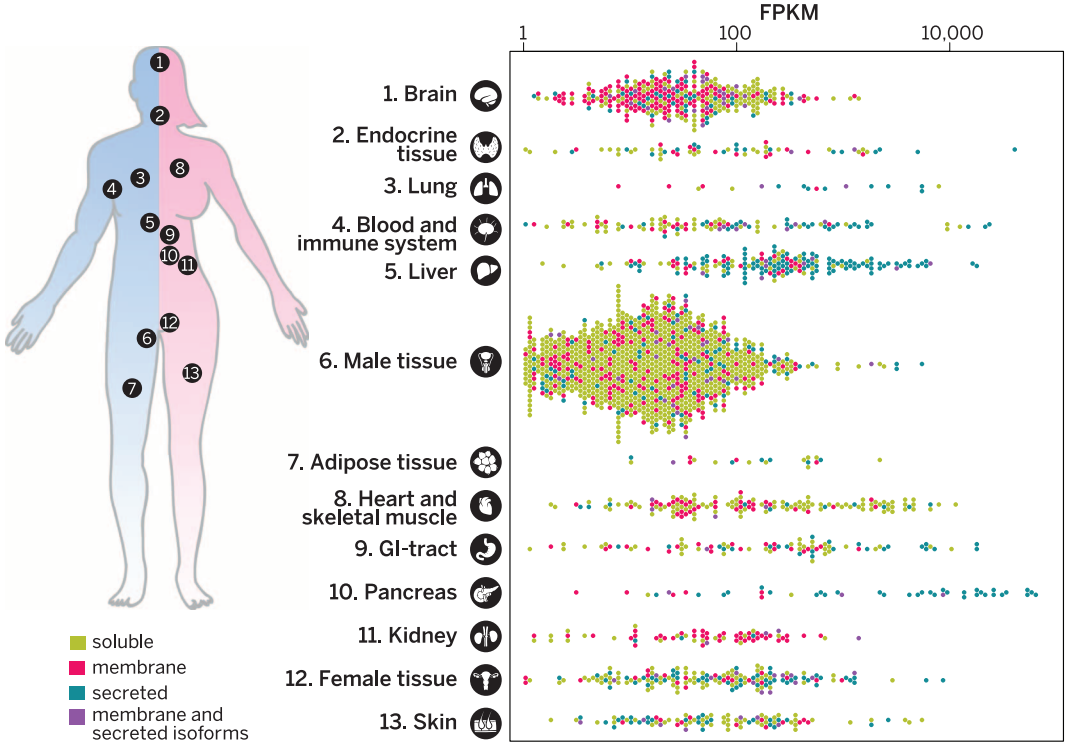
RESULTS: We report a genome-wide analysis of the tissue specificity of RNA and protein expression covering more than 90% of the putative protein-coding genes, complemented with analyses of various subproteomes, such as predicted secreted proteins (n = 3171) and membrane-bound proteins (n = 5570). The analysis shows that almost half of the genes are expressed in all analyzed tissues, which suggests that the gene products are needed in all cells to maintain "housekeeping" functions such as cell growth, energy generation, and basic metabolism. Furthermore, there is enrichment in metabolism among these genes, as 60% of all metabolic enzymes are expressed in all analyzed tissues. The largest number of tissue-enriched genes is found in the testis, followed by the brain and the liver. Analysis of the 618 proteins targeted by clinically approved drugs unexpectedly showed that 30% are expressed in all analyzed tissues. An analysis of metabolic activity based on genome-scale metabolic models (GEMS) revealed liver as the most metabolically active tissue, followed by adipose tissue and skeletal muscle.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1260419

CONCLUSIONS: A freely available interactive resource is presented as part of the Human Protein Atlas portal (www.proteinatlas.org), offering the possibility to explore the tissue-elevated proteomes in tissues and organs and to analyze tissue profiles for specific protein classes. Comprehensive lists of proteins expressed at elevated levels in the different tissues have been compiled to provide a spatial context with localization of the proteins in the subcompartments of each tissue and organ down to the single-cell level. ■

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: mathias.uhlen@scilifelab.se
Cite this article as M. Uhlén *et al.*, *Science* **347**, 1260419 (2015). DOI: 10.1126/science.1260419



The human tissue–enriched proteins. All tissue-enriched proteins are shown for 13 representative tissues or groups of tissues, stratified according to their predicted subcellular localization. Enriched proteins are mainly intracellular in testis, mainly membrane bound in brain and kidney, and mainly secreted in pancreas and liver.

RESEARCH ARTICLE

PROTEOMICS

Tissue-based map of the human proteome

Mathias Uhlén,[1,2,3]* Linn Fagerberg,[1] Björn M. Hallström,[1,2] Cecilia Lindskog,[4] Per Oksvold,[1] Adil Mardinoglu,[5] Åsa Sivertsson,[1] Caroline Kampf,[4] Evelina Sjöstedt,[1,4] Anna Asplund,[4] IngMarie Olsson,[4] Karolina Edlund,[6] Emma Lundberg,[1] Sanjay Navani,[7] Cristina Al-Khalili Szigyarto,[2] Jacob Odeberg,[1] Dijana Djureinovic,[4] Jenny Ottosson Takanen,[2] Sophia Hober,[2] Tove Alm,[1] Per-Henrik Edqvist,[4] Holger Berling,[2] Hanna Tegel,[2] Jan Mulder,[8] Johan Rockberg,[2] Peter Nilsson,[1] Jochen M. Schwenk,[1] Marica Hamsten,[2] Kalle von Feilitzen,[1] Mattias Forsberg,[1] Lukas Persson,[1] Fredric Johansson,[1] Martin Zwahlen,[1] Gunnar von Heijne,[9] Jens Nielsen,[3,5] Fredrik Pontén[4]

Resolving the molecular details of proteome variation in the different tissues and organs of the human body will greatly increase our knowledge of human biology and disease. Here, we present a map of the human tissue proteome based on an integrated omics approach that involves quantitative transcriptomics at the tissue and organ level, combined with tissue microarray–based immunohistochemistry, to achieve spatial localization of proteins down to the single-cell level. Our tissue-based analysis detected more than 90% of the putative protein-coding genes. We used this approach to explore the human secretome, the membrane proteome, the druggable proteome, the cancer proteome, and the metabolic functions in 32 different tissues and organs. All the data are integrated in an interactive Web-based database that allows exploration of individual proteins, as well as navigation of global expression patterns, in all major tissues and organs in the human body.

There is much interest in annotating all human genes at the level of DNA (*1*, *2*), RNA (*3*, *4*), and proteins (*5*, *6*), with the ultimate goal of defining structure, function, localization, expression, and interactions of all proteins. This has resulted in large-scale projects, such as ENCODE (*7*) and the Human Proteome Project (*8*), aimed to integrate results from many research groups and technical platforms to reach a detailed understanding of each of the ~20,000 human protein-coding genes predicted from the human genome and their corresponding protein isoforms. Recently, drafts of the human proteome based on proteogenomics efforts have been described (*9*, *10*), focusing on recent advances in mass spectrometry that allow comprehensive analyses using both isotope-labeled analysis systems (*11*) and deep proteomics methods (*12*) or genome-wide targeted proteomics efforts (*13*).

A complement to these efforts is the Human Protein Atlas program (*14*), which is exploring the human proteome using genecentric and genome-wide antibody-based profiling on tissue microarrays. This allows for spatial pathology-based annotation of protein expression, in combination with deep-sequencing transcriptomics of the same tissue types. The strategy is based on the quantitative assessment of transcript expression in complex tissue homogenates, involving a mixture of cell types combined with the precise localization of the corresponding proteins down to the single-cell level, using immunohistochemistry. Recently, we performed a transcriptomics study of 27 different tissues using this approach (*15*), followed by subsequent in-depth studies of the global proteome in a number of these tissues and organs, such as liver (*16*), testis (*17*), and the gastrointestinal (GI) tract (*18*). Here, we have used this approach and extended the analysis to 32 tissue types, representing all major tissues and organs in the human body, to create a genome-wide map of the human tissue–based proteome, with a focus on the analysis of the tissue-elevated proteins and all secreted and membrane proteins. Particular emphasis has been placed on analyses of proteins targeted by pharmaceutical drugs (*19*) and proteins implicated in cancer (*20*). We used the data to generate comprehensive metabolic maps for all 32 tissue types in order to identify differences in metabolism between tissues. In addition, new transcriptomics data from 36 human cell lines allowed us to compare the proteomes between cell lines and normal cells derived from the same tissue types. Finally, the protein isoforms generated by differential splicing between different tissues were studied with a focus on splice variants with predicted differential subcellular localization. All data are presented in an interactive database (www.proteinatlas.org).

[1]Science for Life Laboratory, KTH—Royal Institute of Technology, SE-171 21 Stockholm, Sweden. [2]Department of Proteomics, KTH—Royal Institute of Technology, SE-106 91 Stockholm, Sweden. [3]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK-2970 Hørsholm, Denmark. [4]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, SE-751 85 Uppsala, Sweden. [5]Department of Chemical and Biological Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden. [6]Leibniz Research Centre for Working Environment and Human Factors (IfADo) at Dortmund TU, D-44139 Dortmund, Germany. [7]Lab Surgpath, Mumbai, India. [8]Science for Life Laboratory, Department of Neuroscience, Karolinska Institute, SE-171 77 Stockholm, Sweden. [9]Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden.
***Corresponding author. E-mail: mathias.uhlen@scilifelab.se**

Results

Classification of all human protein-coding genes

Samples representing all major tissues and organs (n = 44) in the human body were analyzed (Fig. 1A) by using 20,456 antibodies generated "in-house," as well as 3572 antibodies provided by external suppliers. The antibodies have been used to produce more than 13 million tissue-based immunohistochemistry images, with each image annotated on the single-cell level for all sampled tissues by pathologists. The analysis was complemented with RNA sequencing (RNAseq) data for 32 out of the 44 tissue types. We investigated global expression profiles using hierarchical clustering based on the correlation between 122 biological replicates from the 32 organs and tissues (Fig. 1B and fig. S1). The results reveal testis and brain as outliers and a clear connectivity between the samples from the GI tract (stomach, duodenum, small intestine, colon, and rectum), the hematopoietic tissues (bone marrow, lymph node, spleen, tonsil, and appendix) and the two striated muscle samples (cardiac and skeletal muscle). A principal component analysis (fig. S2A) confirms a close resemblance between cardiac and skeletal muscle but also suggests similarities in global expression between pancreas and salivary gland, as well as differences between the primary lymphoid tissue (bone marrow) and the secondary lymphoid tissues, such as tonsil and spleen.

The transcriptomics study allowed us to refine the classification performed earlier (*15*) of all the 20,344 putative protein-coding genes with RNA-seq data into categories based on their expression across all 32 tissue types (Fig. 1C, Table 1, and tables S1 to S4). Indirectly, this also provides an estimate of the relative protein levels corresponding to each gene, because proteogenomics analyses have shown that the translation rate, in most cases, is constant for a specific transcript across different human cells and tissues at both a cellular level (*21*) and a tissue level (*9*). Although it is still a matter of scientific debate (*22*) whether protein degradation rates could, in some cases, vary for an individual protein in different tissues, an overall concurrence between mRNA and protein levels for a given gene product across various tissues is generally expected (*9*, *21*). A large fraction (44%) of the protein-coding genes were detected in all analyzed tissues, and these ubiquitously expressed genes include known "housekeeping" genes encoding mitochondrial proteins and proteins involved in overall cell structure, translation, transcription, and replication. Of all the protein coding genes, 34% showed an elevated expression in at least one of the analyzed tissues,

and these were further subdivided into (i) enriched genes with mRNA levels in one tissue type at least five times the maximum levels of all other analyzed tissues, (ii) group-enriched genes with enriched expression in a small number of tissues, and (iii) enhanced genes with only a moderately elevated expression. The use of the word "tissue-specific" has been avoided because this definition depends on arbitrary cut-off levels, and many proteins described in the literature as "tissue-specific" are here shown to be expressed in several tissues. This is exemplified by albumin, which we, as expected, identified as enriched in liver but also found at high levels, albeit much lower than for liver, in kidney and pancreas.

Evidence for the human protein-coding genes

We have determined the number of genes for which evidence is available at a protein level by combining our antibody-based data with the manual annotation of literature by the UniProt consortium (*5*) and the results from the recent mass spectrometry–based proteogenomics analyses (*9*, *10*, *12*). The analysis shows that there are 17,132 protein-coding genes with proteins identified from at least one of the three efforts and 13,841 genes with experimental evidence from at least two of the efforts (Fig. 1D). Furthermore, there is evidence, at the RNA level, for 2546 additional genes based on either our data or annotations by UniProt. Although proteins not yet detected by one of the three methods should be further investigated to establish them as true human proteins, it is noteworthy that out of the 20,356 putative protein-coding genes (in Ensembl release 75) there are only 677 genes (3.3%) for which there is no experimental evidence (table S5). Many of these genes were removed in the

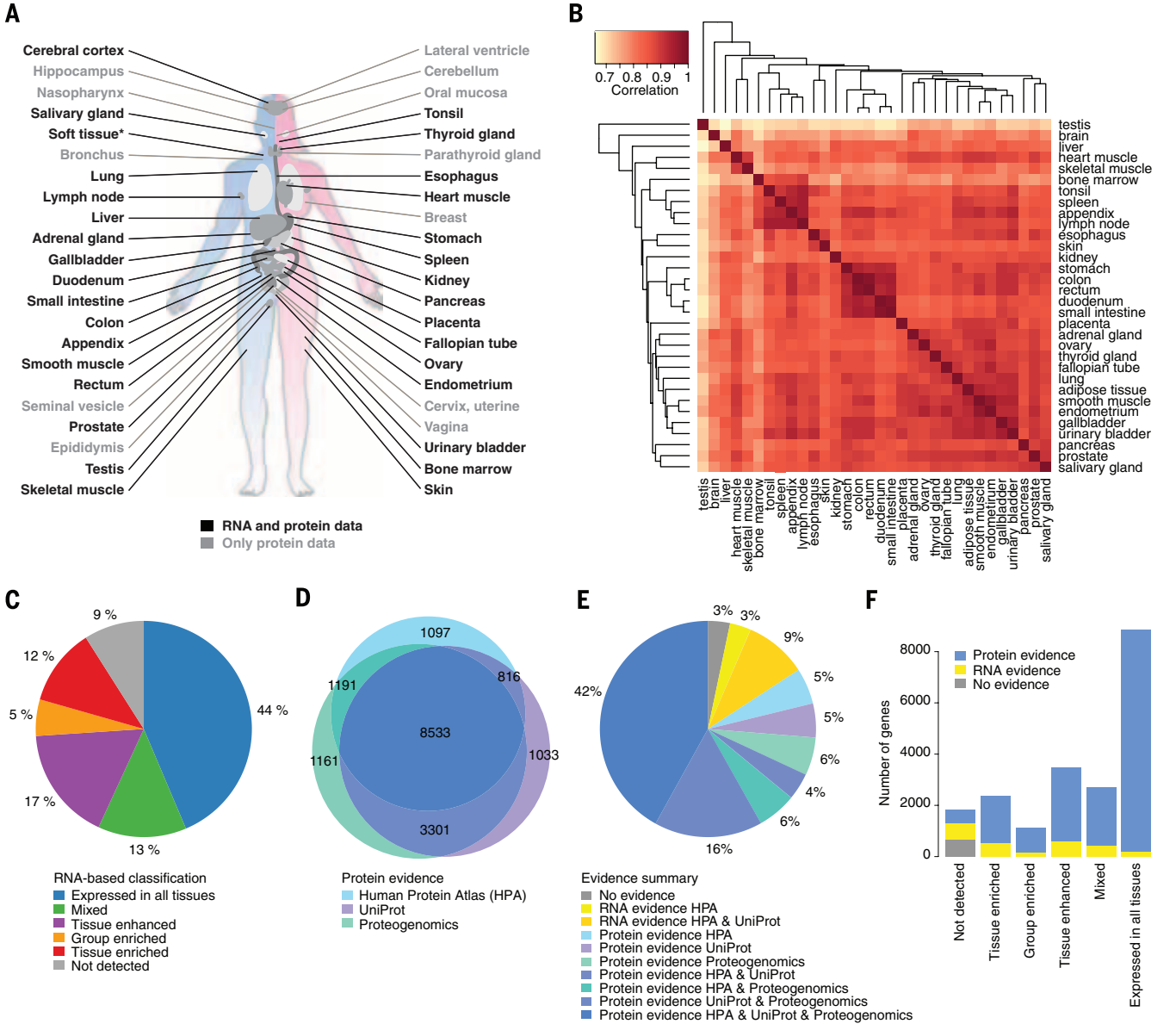


Fig. 1. Classification and protein evidence of the human protein-coding genes. (**A**) The tissues analyzed in this study, including tissues studied both by RNAseq and antibody-based profiling and those analyzed only by antibody-based profiling. For details see table S1. (**B**) Heat map showing the pairwise correlation between all 32 tissues based on transcript expression levels of 20,344 genes. The average FPKM values for each gene and tissue are used in the analysis. For correlation results of all individual samples, see fig. S1. (**C**) The number of genes classified in each expression category according to the definition stated in Table 1. (**D**) Venn diagram showing the overlap between protein evidence on the basis of three sources: Human Protein Atlas, UniProt, and Proteogenomics. (**E**) The distribution of genes classified as having protein evidence, evidence only at the transcript level, and genes without any experimental evidence. (**F**) The number of genes with protein evidence, RNA evidence, and no evidence stratified according to their transcriptomics-based classification into six categories.

later update of Ensembl (release 76) (fig. S2B), and others have been suggested to be noncoding genes on the basis of the lack of correlation in gene family age and cross-species conservation studies. Thus, it is possible that most of these "missing genes" will be removed from the list of protein-coding genes in the future. These genes and the genes with evidence only at the RNA level are obvious targets for more in-depth functional protein studies. A summary of the supporting data is shown in Fig. 1E. Few (2%) of the ubiquitously expressed genes lack protein evidence (Fig. 1F); however, protein evidence is lacking for 18% of the genes identified here by RNA analysis as elevated (tissue enriched, group enriched, or enhanced). Examples of genes with no previous evidence on the protein level according to UniProt, but now confirmed using antibody-based profiling and proteogenomics (*9*, *10*), are chromosome 2 open reading frame 57 (C2orf57), shown here with an enriched expression in testis localized to the sperm (Fig. 2A), and chromosome 8 open reading frame 47 (C8orf47), with expression in a subset of endocrine islet cells and ductal cells of the exocrine pancreas (Fig. 2B).

The tissue-elevated proteome

A network plot shows the number of tissue-enriched genes for each tissue type, as well as the number of genes enriched in different groups of tissues and organs (fig. S4). An analysis of selected tissues and organs (Fig. 2O) reveals a large number of elevated genes in male tissue, brain,

Table 1. Classification of all human protein-coding genes based on transcript expression levels in 32 tissues.

Category	Description	No. of genes	Fraction of genes (%)
Tissue enriched	mRNA levels in a particular tissue at least five times those in all other tissues	2,355	12
Group enriched	mRNA levels at least five times those in a group of 2–7 tissues	1,109	5
Tissue enhanced	mRNA levels in a particular tissue at least five times average levels in all tissues	3,478	17
Expressed in all	Detected in all tissues (FPKM > 1)	8,874	44
Mixed	Detected in fewer than 32 tissues but not elevated in any tissue	2,696	13
Not detected	FPKM < 1 in all tissues	1,832	9
Total	Total number of genes analyzed with RNAseq	20,344	100
Total elevated	Total number of tissue-enriched, group-enriched, and tissue-enhanced genes	6,942	34

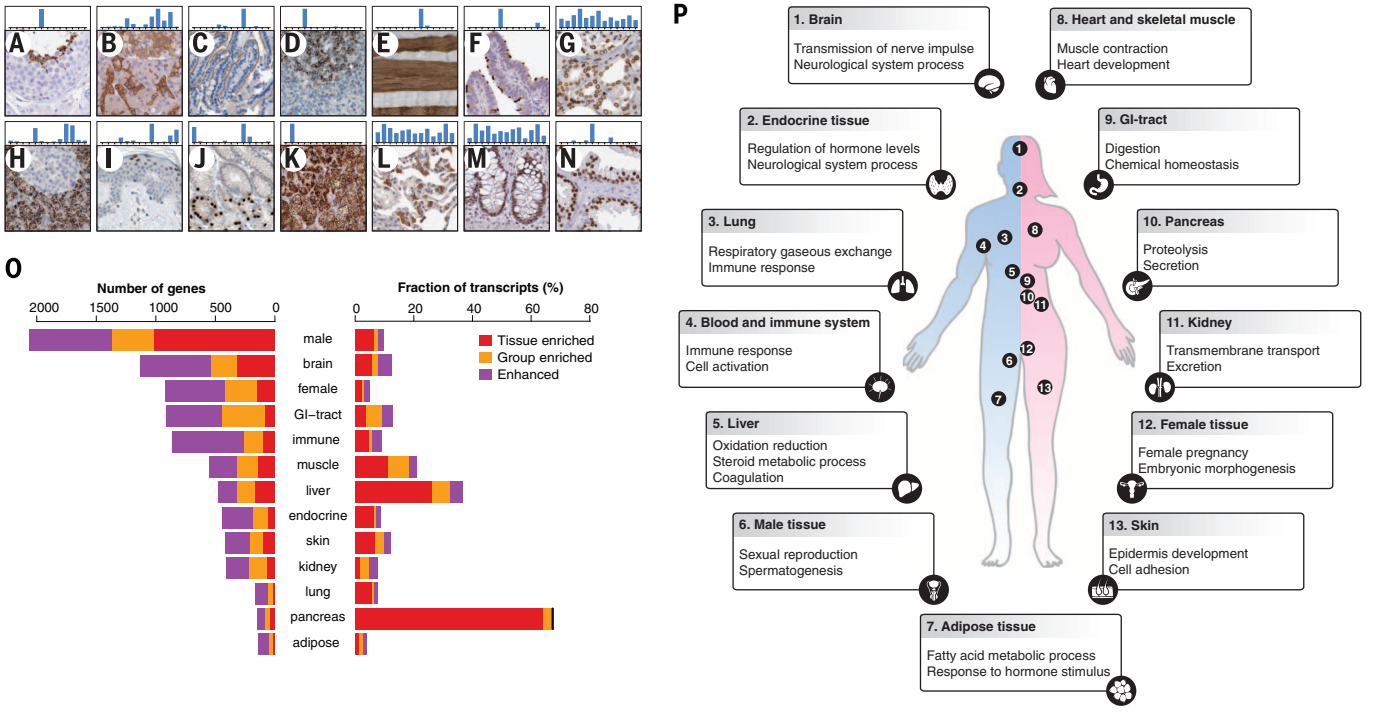


Fig. 2. Tissue microarray–based protein expression, and analysis of tissue-elevated genes in the different organ systems. (**A** to **N**) Tissue expression and localization for a selection of human proteins. Larger images corresponding to (A) to (N) of the figure are shown in fig. S3. The levels of the corresponding mRNA (FPKM) are displayed as bars for each of the 13 organ systems analyzed (from left: brain, endocrine tissue, lung, blood and immune system, liver, male tissue, adipose tissue, heart and skeletal muscle, GI tract, pancreas, kidney, female tissue, and skin). Examples include testis with C2orf57 expression in sperm (A), pancreas with cytoplasmic C8orf47 expression in both a subset of endocrine cells and ductal cells (B), duodenum with CDHR2 expression in microvilli (C), lymph node with cytoplasmic FCRLA expression in germinal center cells (D), skeletal muscle with cytoplasmic MYL3 expression in slow muscle fibers (E), fallopian tube with ROPN1L expression in cilia (F), kidney with SUN2 expression in all nuclear membranes (G), pancreas with GATM expression in mitochondria throughout the exocrine compartment (H), skin with GRHL1 expression in nuclei of the upper epidermal layer (I), stomach with nuclear PAX6 expression in endocrine cells (J), adrenal gland with cytoplasmic expression of CYP11B1 in cortical cells (K), lung with cytoplasmic COMT expression in a subset of pneumocytes and macrophages (L), colon with nuclear ATF1 expression in glandular cells (M), and prostate with nuclear FOXA1 expression in glandular cells (N). (**O**) The number of elevated genes in the 13 organ systems, as described in (P), and the fraction of all transcripts (FPKM) encoded by these elevated genes for each of these organ systems. (**P**) An analysis of major GO terms for each tissue on the basis of the tissue-elevated genes in 13 selected tissues or groups of tissues, as described in supplementary methods. For more details of the GO analysis, see table S6.

and liver and relatively few in lung, pancreas, and fat (adipose tissue). The transcriptomics analysis also allowed us to determine the fraction of elevated transcripts in each tissue (Fig. 2O). For most tissues, only ~10% of the transcripts are encoded by tissue-elevated genes, with the exception of pancreas and liver, where elevated genes encode 70% and 35% of the transcripts, respectively.

A functional Gene Ontology (GO) analysis for 13 tissues or groups of tissues is summarized in Fig. 2P (see table S6 for details), and the terms identified are consistent with the function of the respective tissues. The largest number of enriched genes is found in the testis (n = 999), with many of the corresponding testis-specific proteins involved in the reproductive process and spermatogenesis. It is not unlikely that many of these genes will show a shared expression with oocytes in the female ovaries, which are difficult to analyze because of the different kinetics of germ cell development, including first rounds of meiosis at the embryonic stages during female life. The tissue with the second largest number of enriched genes is the brain (n = 318). The number of genes with expression restricted to neuronal tissue is relatively small, but it is likely that more enriched genes would be added to the list if additional regions, such as the various specialized regions of the brain, were sampled. Genes elevated in liver encode secreted plasma and bile proteins, detoxification proteins, and proteins associated with metabolic processes and glycogen storage, whereas genes elevated in adipose tissue encode proteins involved in lipid metabolic processes, secretion, and transport. Genes elevated in skin encode proteins associated with functions related to the barrier function (squamous cell differentiation and cornification), skin pigmentation, and hair development. In the GI tract, elevated genes predominantly encode proteins involved in nutrient breakdown, transport, and metabolism; host protection; and tissue morphology maintenance.

As expected, many of the genes enriched in groups of tissues are common for the GI tract and the hematopoietic tissues, respectively, as exemplified on the protein level by cadherin-related family member 2 (CDHR2), expressed in the microvilli of duodenum and small intestine (Fig. 2C), and Fc receptor-like A (FCRLA), expressed in lymph node, tonsil, appendix, and spleen (Fig. 2D). A large number of group-enriched genes involved in contraction are observed in striated (cardiac and skeletal) muscle, as exemplified by the fiber type–specific expression of myosin light chain 3 (MYL3) (Fig. 2E), whereas many genes shared between testis and the fallopian tube, as well as testis and lung, are involved in cell motility, as exemplified by rophilin-associated tail protein–like (ROPN1L), which is expressed in

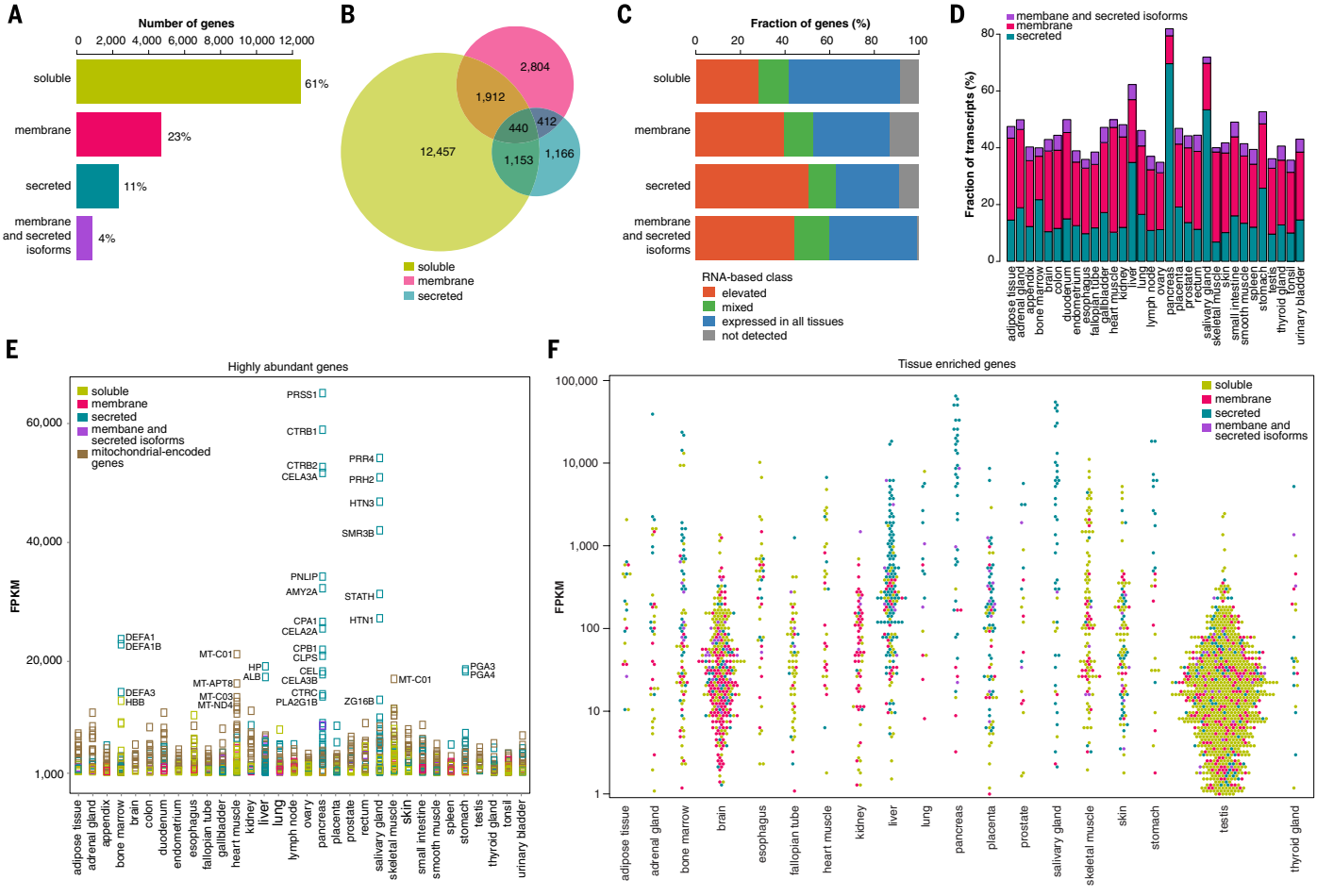


Fig. 3. Prediction and analysis of the human secreted and membrane-spanning proteins. (**A**) The number and fraction of all human genes (n = 20,356) classified into the categories soluble, membrane-spanning, and secreted, as well as genes with isoforms belonging to two or all three categories. (**B**) Venn diagram showing the number of genes in each of the three main subcellular location categories: membrane, secreted, and soluble. The overlap between the categories gives the number of genes with isoforms belonging to two or all three categories. (**C**) The fraction of genes in the various protein expression classes for the soluble, secreted, and membrane-spanning proteins, as well as genes with both secreted and membrane-spanning isoforms. (**D**) The fraction of transcripts based on FPKM values from each of the three secreted or membrane-spanning categories across the 32 analyzed tissues. (**E**) The 370 most-abundant genes (FPKM > 1000) in the different tissues, stratified according to their predicted localization on the basis of (C), as well as an additional category of the 13 genes encoded by the mitochondrial genome. The gene names for a selection of the most abundant genes are shown. (**F**) The transcript levels (FPKM) on a $\log_{10}$ scale for all genes identified as tissue-enriched are shown for a few selected tissues, with each gene stratified according to predicted localization.

sperm (testis), ciliated cells in respiratory epithelia (lung), and ciliated cells in the fallopian tube (Fig. 2F).

The human secretome and membrane proteome

Both secreted and membrane-bound proteins play crucial roles in many physiological and pathological processes. Important secreted proteins include cytokines, coagulation factors, hormones, and growth factors, whereas membrane proteins include ion channels or molecular transporters, enzymes, receptors, and anchors for other proteins. Here, we performed a whole-proteome scan to predict the complete set of human secreted proteins ("secretome") using three methods for signal-peptide prediction: SignalP4.0 (*23*), Phobius (*24*), and SPOCTOPUS (*25*). In addition, the human membrane proteome was predicted using seven membrane–protein topology prediction methods as described (*21*), which resulted in a majority decision–based method (MDM). For each protein-coding gene, all protein isoforms were annotated for predicted localization: secreted, membrane spanning, or soluble (intracellular proteins without a predicted signal peptide or membrane-spanning region) (table S1). Some of the proteins predicted to be membrane-spanning are intracellular, e.g., in the Golgi or mitochondrial membranes, and some of the proteins predicted to be secreted could potentially be retained in a compartment belonging to the secretory pathway, such as the endoplasmic reticulum (ER), or remain attached to the outer face of the cell membrane by a GPI anchor. About 3000 human genes are predicted to encode secreted proteins, with another 5500 encoding membrane-bound proteins (Fig. 3A). In the interactive database (www.proteinatlas.org), many of the secreted proteins are detected at the RNA level in tissues, but no protein expression is observed in the antibody-based analysis in the same tissue—most likely because the steady-state levels of proteins in the cell during the secretion process are too low to be detected.

A large fraction (72%) of human genes encode multiple splice variants with different protein sequences. In Fig. 3B, all genes have been classified according to the presence of protein isoforms that are intracellular, membrane-spanning, and/or secreted. Note that two-thirds of the genes encoding secreted proteins have at least one splice variant with alternative localization. All protein isoforms (n = 94,856) with their predicted localization based on the three signal-peptide–prediction methods, as well as the number of predicted transmembrane segments, are listed in table S7. An analysis across the 32 tissues (Fig. 3C) supports earlier suggestions (*21*, *26*) that a larger fraction of tissue-enriched proteins are secreted or membrane-spanning proteins than are intracellular proteins.

Furthermore, we investigated the fraction of the transcriptome that codes for each class of proteins across the 32 tissues (Fig. 3D and fig. S4). In most cases, the secreted proteins account for between 10 and 20% of the transcripts. In contrast, more than 70% of the transcripts from the pancreas and ~60% from the salivary gland encode secreted proteins. This demonstrates the extreme specialization of these two tissues for production of secreted proteins into the duodenum

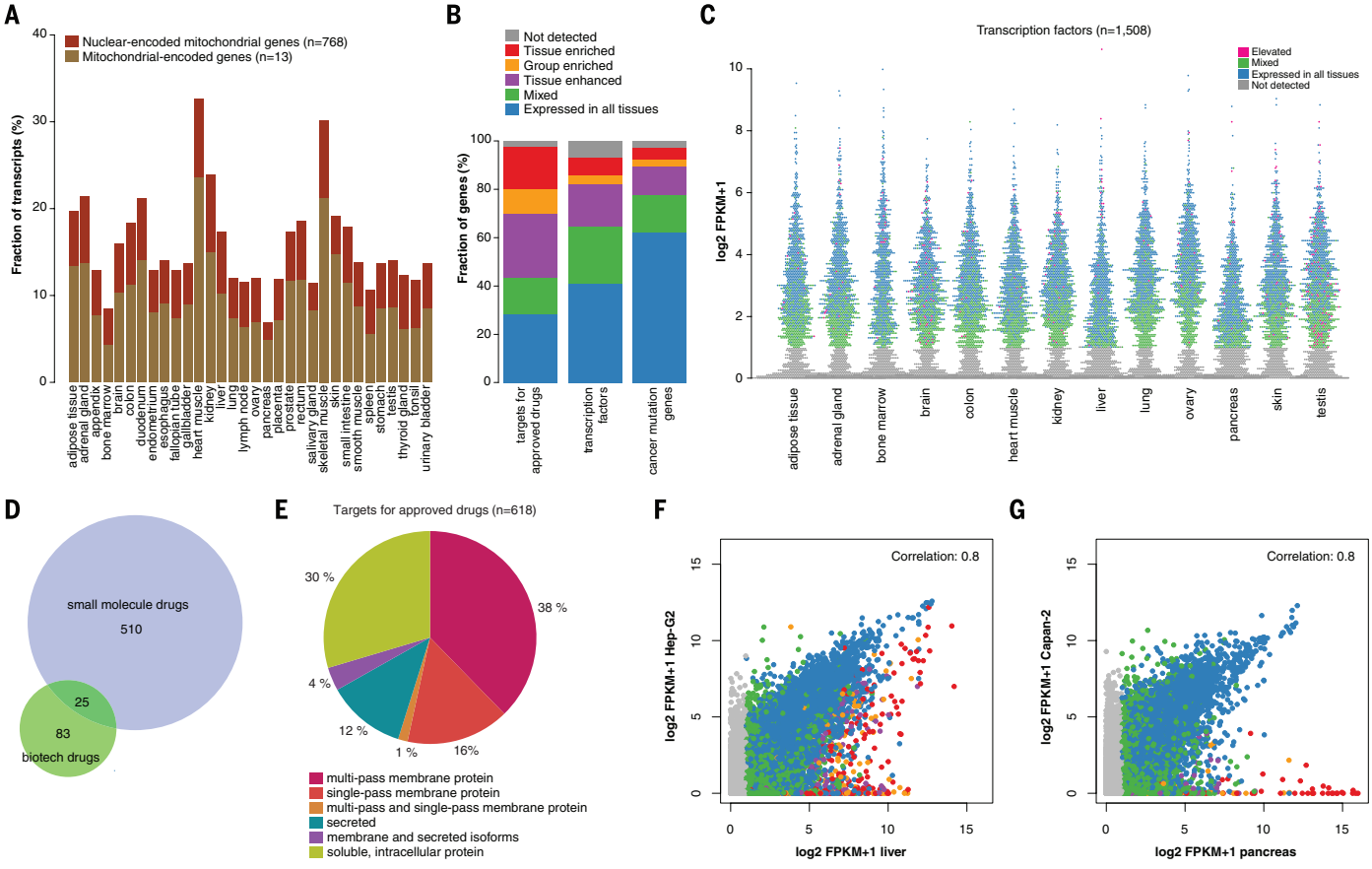


Fig. 4. The human transcriptome in different tissues and organs. (**A**) The fraction of transcripts encoded by mitochondrial genes for each of the different tissues and organs, subdivided by genes encoded by the mitochondrial genome and chromosomes, respectively. (**B**) The fraction of genes classified according to tissue expression pattern and analyzed for all targets of approved drugs (n = 618), all transcription factors (n = 1508), and proteins implicated in cancer (n = 525). (**C**) The transcript levels (FPKM) for all genes encoding transcription factors in some selected tissues, color-coded according to their global expression category. (**D**) The number of pharmaceutical drugs approved by FDA, according to Drugbank (*19*), that are chemical (small-molecule) or biotech drugs. (**E**) The number of pharmaceutical drugs approved by FDA (*19*) stratified according to the predicted localization of the target protein. (**F**) Pairwise comparison showing all genes expressed in liver tissue and the liver cell line Hep-G2, color-coded according to protein expression category as shown in (B). (**G**) Pairwise comparison showing all genes expressed in pancreas tissue and the pancreas cell line Capan-2, color-coded according to protein expression category as shown in (B).

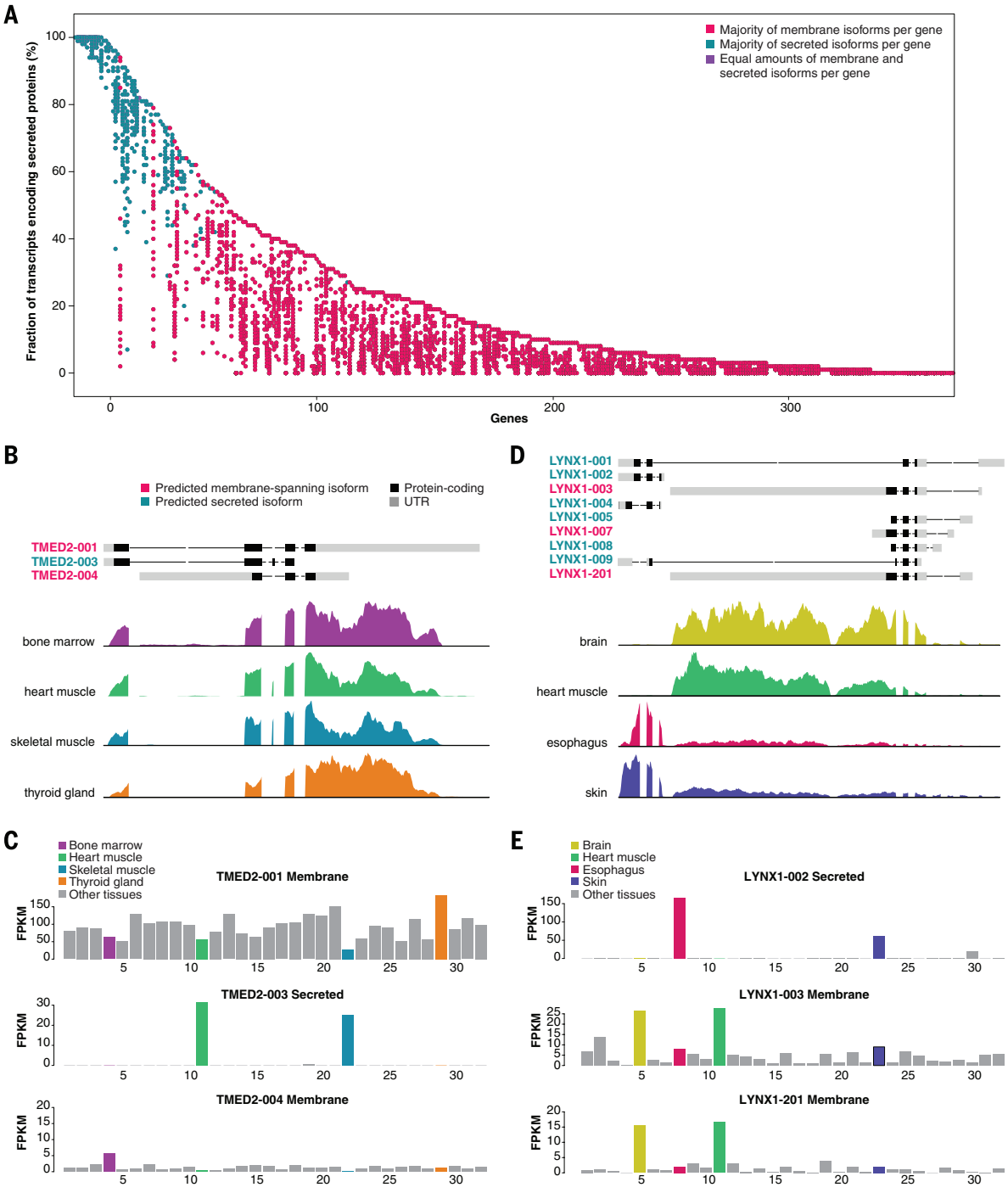


Fig. 5. Differential splicing analysis of transcripts. (**A**) Dot plot of genes with multiple isoforms, where at least one isoform is classified as membrane-spanning and another classified as secreted. The *x* axis shows 366 genes expressed at >5 FPKM in one or more tissues; the *y* axis shows the sum of FPKM values for all secreted isoforms divided by the total sum of FPKM values for each tissue expressed at >5 FPKM. For each gene, the number of tissues where the secreted transcripts are more abundant than the membrane-spanning transcripts is calculated to define a majority fraction-type as membrane (red), secreted (blue), or equal number for both categories (purple). Each tissue is represented by a circle, and the color is the same across all tissues for the same gene. (**B**) Example of differential splicing for the gene *TMED2*, with two isoforms predicted as membrane-spanning and one isoform predicted as secreted. The exon-intron structure (with pure intronic sites removed), as well as the location of the untranslated regions (UTR) of three splice variants of *TMED2*, are shown on top. Normalized read coverage plots for cardiac muscle, skeletal muscle, thyroid gland, and bone marrow highlight the differential use of exons in the selected tissues. (**C**) Transcript abundance (FPKM values) plotted across all 32 tissues for each isoform. The predicted membrane-spanning transcript (top) is expressed in all tissues, with thyroid gland as the most abundant tissue; a secreted isoform (middle) is only detected in cardiac and skeletal muscle; and a second membrane-spanning isoform (bottom) is expressed at very low levels, with bone marrow as most abundant. (**D**) Examples of differential splicing for the gene *LYNX1*, with three isoforms predicted as membrane-spanning and six isoforms predicted as secreted from the visualization used in (B). (**E**) Transcript abundance (FPKM values) for three isoforms of *LYNX1* detected at >5 FPKM. The secreted isoform (top) is expressed at high levels in esophagus and skin, whereas the two membrane-spanning isoforms (middle and bottom) are most abundant in brain and cardiac muscle.

and oral cavity, respectively. About 40% of the transcripts in liver encode secreted proteins. Other tissues with relatively high levels of transcripts encoding secreted proteins include gallbladder, bone marrow, placenta, and different parts of the GI tract, such as stomach, duodenum, and small intestine.

The most abundant genes, normalized as fragments per kilobase of exon per million fragments mapped (FPKM) with a value >1000, in the different tissues are shown in Fig. 3E, and the prediction of the localization of the corresponding proteins reveals that many (53%) are secreted proteins. Among the predicted membrane-spanning proteins, 13 proteins encoded in the mitochondrial genome are the most highly expressed. In Fig. 3F, tissue-enriched genes are shown stratified according to their predicted subcellular localization. Many of the tissue-enriched genes in testis are intracellular, whereas a large number of the tissue-enriched genes in brain and kidney are membrane-bound. In contrast, in many other tissues, such as pancreas, salivary gland, liver, stomach, and bone marrow, most tissue-enriched genes are secreted (fig. S5).

The housekeeping proteome

Transcriptomics analysis shows that close to 9000 genes (table S1) are expressed in all analyzed tissues, which suggests that the gene products are needed in all cells to maintain basic cellular structure and function. These housekeeping proteins include ribosomal proteins involved in protein synthesis, enzymes essential for cell metabolism and gene expression, and mitochondrial proteins needed for energy generation, as well as proteins responsible for the structural integrity of the cell. Most of these proteins are expressed at similar levels throughout the human body, as exemplified in kidney by the expression of the nuclear membrane protein SUN2 present in all cells (Fig. 2G), whereas a few proteins show great variability in expression levels—for example, the mitochondrial protein glycin amino transferase (GATM), with high expression in exocrine pancreas (Fig. 2H), kidney, and liver but relatively low expression levels in all other tissues. An interesting class of proteins is encoded by mitochondrial genes, and in Fig. 4A, the transcriptional load of these genes is shown across different tissues. The highest fractions of transcripts encoding mitochondrial proteins are found in cardiac muscle (32% of all transcripts) and skeletal muscle (28%), which demonstrates the importance of energy metabolism for striated muscle tissue.

The regulatory proteome

Transcription factors, of which ~1,500 have been identified in humans (*27*), comprise an important class of regulatory proteins as they function as on/off switches for gene expression. The fraction of transcription factor genes classified according to tissue specificity is shown in Fig. 4B, which suggests a tissue distribution similar to that of the complete proteome, with as many as 41% of the genes expressed in all tissues and only 29% identified as elevated (enriched, group enriched, or enhanced). Many of the more-abundantly expressed transcription factors are found in all tissues (Fig. 4C). However, there are examples of abundant transcription factors that belong to the tissue-elevated categories, such as (i) grainyhead-like 1 (GRHL1) with enhanced expression in esophagus and skin (squamous epithelia) and selective localization to the uppermost nucleated epidermal keratinocytes (Fig. 2I) and (ii) paired box 6 (PAX6) involved in eye and brain development and differentiation of pancreatic islet cells, with group-enriched expression in brain, pancreas, and stomach, selectively localized to a subset of glandular cells in the stomach mucosa (Fig. 2J) and to islet cells in the pancreas. The tissue-enriched transcription factors identified here (table S8) will enable new insights into the regulatory pattern of the different tissues.

The druggable proteome

Most pharmaceutical drugs act by targeting proteins and modulating their activity. Target proteins belong to four main families: enzymes, transporters, ion channels, and receptors. The U.S. Food and Drug Administration (FDA) has approved drugs targeting human proteins from 618 genes, according to Drugbank (*19*), with most drugs acting on signal transduction proteins that convert extracellular signals into intracellular responses. Antibody-based drugs are usually unable to penetrate the plasma membrane, and therefore, they target cell surface proteins, such

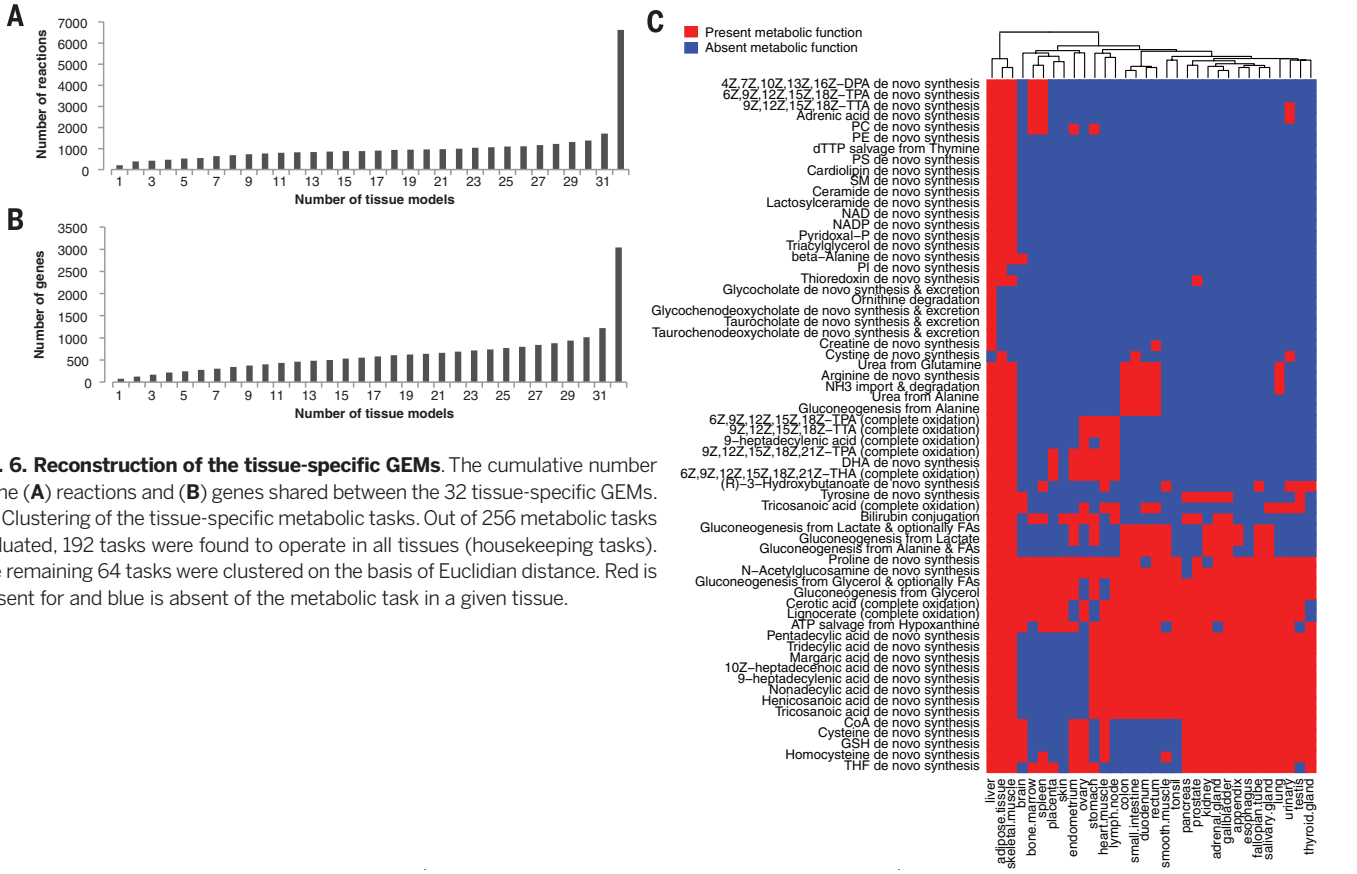


Fig. 6. Reconstruction of the tissue-specific GEMs. The cumulative number of the (**A**) reactions and (**B**) genes shared between the 32 tissue-specific GEMs. (**C**) Clustering of the tissue-specific metabolic tasks. Out of 256 metabolic tasks evaluated, 192 tasks were found to operate in all tissues (housekeeping tasks). The remaining 64 tasks were clustered on the basis of Euclidian distance. Red is present for and blue is absent of the metabolic task in a given tissue.

as receptors, whereas small-molecule drugs can diffuse into cells and act also on intracellular targets. An analysis of the proteins encoded from the 618 genes shows that 535 proteins are targeted by small chemical molecules, whereas 108 proteins are targeted by biotech drugs (Fig. 4D). The predicted subcellular localization (Fig. 4E) shows that 59% of the targets are predicted membrane proteins and that 16% are secreted, including those with both secreted and membrane-bound isoforms. The genes corresponding to these drug targets were classified according to tissue specificity, and the results (Fig. 4B and table S9) show a bias for tissue-elevated proteins (enriched, group enriched, or enhanced), although as many as 30% of the approved drugs target proteins expressed in all analyzed tissues. One example of a target with enriched expression is cytochrome P450 11B1 (CYP11B1), which is involved in the conversion of progesterone to cortisol in the adrenal gland (Fig. 2K), whereas a ubiquitously expressed protein is the catechol-*O*-methyltransferase (COMT), which is associated with degradation of neurotransmittors and is important in the metabolism of drugs used in treatment of Parkinson's disease. COMT displays cytoplasmic expression in all analyzed tissues, including lung (Fig. 2L). The ubiquitous expression may have implications for treatments using these proteins as drug targets.

The cancer proteome

Genes implicated in cancer are often essential for orderly growth, survival, and basic cell functions in normal cells and tissues, whereas overexpression, loss of expression, or expression of a mutated protein contributes to dysfunction and tumor growth. The number of genes implicated in cancer is dependent on definitions; however, 259 genes have been shown to be mutated across 21 tumor types (*28*); 290 genes have been reported as cancer driver genes across 12 tumor types (*29*); and 525 genes have been implicated in malignant transformation, according to a catalog of somatic mutations in cancer (COSMIC) (*20*). Expression analysis based on our transcriptomics data shows that a majority (60%) of these last-mentioned genes (Fig. 4B and table S10) is expressed in all tissues, with only a fraction of genes expressed in a tissue- or group-enriched manner. Examples are the activating transcription factor 1 (ATF1) (Fig. 2M), a protein expressed in all tissues with known translocations in sarcomas, and the forkhead box A1 (FOXA1) (Fig. 2N), a protein with enhanced expression where somatic mutations in a subset of prostate cancers have been reported (*30*). The lack of tissue specificity for many of these genes is not surprising because many of the corresponding proteins are involved in normal growth regulation and cell cycle control, but it also emphasizes the possible adverse effects of treatment with drugs targeting proteins expressed in all tissues.

Tissue versus cell lines

Human biology and diseases are often explored using cell lines as model systems. We compared the body-wide expression in human tissues with expression in cancer cell lines derived from corresponding tissue types. The transcriptomes for 11 cell lines were described earlier (*31*), whereas the transcriptomes for an additional 36 cell lines were generated as part of this study (see table S11). Genome-wide expression patterns comparing normal tissues with corresponding human cell lines are shown in fig. S6, as exemplified by the liver cancer–derived cell line Hep-G2 (Fig. 4F), and the pancreas cancer–derived cell line Capan-2 (Fig. 4G). Many of the tissue-enriched genes identified in normal tissues are down-regulated or completely "turned off" in the corresponding cell lines, and in contrast, the housekeeping proteins are expressed at the same level in both tissues and corresponding cell lines. These results support earlier studies (*32*) suggesting that cell lines are "dedifferentiated," with shared characteristics and lack of tissue-specific features due to down-regulation of tissue-enriched genes. This implies that conclusions from cell line studies should only be conferred on the corresponding tissue with caution.

The isoform proteome

Protein isoforms endow the structural space of the human proteome with breadth and complexity (*33*). Isoforms are produced through alternative splicing, posttranslational modifications, proteolytic cleavage, somatic recombination, or genetic variations in protein-coding regions. We explored genes encoding isoforms with different predicted localization (secreted or membrane spanning) (table S12). A large number of these genes (n = 366) are displayed together with the fraction of all transcripts (mRNA molecules) in Fig. 5A, with splice variants that yield secreted proteins. Most of the genes (67%) have more than 80% of the transcripts encoding only one of the two localizations across all 32 tissues, but there are some proteins for which the majority of the transcripts encode a secreted form in one tissue, whereas the majority of the transcripts encode a membrane protein in another tissue. As an example, the expression levels for different isoforms of the poorly understood transmembrane emp24 domain–trafficking protein 2 (TMED2) are shown in Fig. 5, B and C. Cardiac muscle has a tissue-specific expression of the secreted form, whereas the membrane-bound form is detected in all other tissue types, although at variable levels. Similarly, the protein Ly6 or neurotoxin 1 (LYNX1) shows a selective expression of the secreted isoform in the esophagus and the skin, whereas the membrane-bound form is found in other tissue types and is most abundantly expressed in the brain and the cardiac muscle (Fig. 5, D and E). The different localizations of the isoforms are consistent with the predicted functions of the different isoforms. In most cases, one of the isoforms dominates across all tissues, which is also consistent with earlier studies (*34*). These are starting points to explore the relation between tissue-specific expression and function.

Tissue-based map of human metabolism

Genome-scale metabolic models (GEMs) provide not only the best representation of the metabolic capabilities of cell and/or tissue types but also quantitative descriptions of the genotype-phenotype relationship (*35*). Using the RNAseq data, we reconstructed tissue-specific GEMs for 32 different tissues using the generic metabolic model, HMR2 (*36*), and generated a map of the complete human metabolism. All models were generated such that they can carry out 56 metabolic tasks identified to be present in all human cell types (*37*). The numbers of the reactions, metabolites, and genes incorporated into each tissue-specific GEM are presented (table S13), and the models are provided in SBML format at the Human Metabolic Atlas portal (*38*). In order to confirm that none of the models have futile cycles, we ensured that high-energy compounds cannot be generated from low-energy compounds using metabolic tasks including rephosphorylation of adenosine triphosphate or the generation of a proton gradient over the membranes (table S14).

A total of 6627 reactions, 3040 genes, and 4847 metabolites were present in at least one of the tissue models, and 4912 reactions, 1822 genes, and 3984 metabolites were present in all models. This shows that about 75% of all metabolic reactions in the human body are operating in all key tissues, which clearly illustrates the central role metabolism is playing for basic cellular function. At a gene level, the consensus expression in all tissues is, however, less (i.e., about 60%), which shows that, even though different tissues have the same metabolic reactions, it is different isoforms of the enzymes that are responsible for catalyzing these reactions. Our analysis is the first genome-wide illustration of this wide variation in enzyme usage for catalyzing the same reaction between human tissues.

We found that only 207 of the reactions (Fig. 6A) and 74 of the genes (Fig. 6B) were unique to any of the tissues, and notable differences between the genes (fig. S7) and reactions (fig. S8) based on pairwise comparisons of the various tissues were observed. Between 57 and 632 genes differed in these comparisons of the tissue models, representing 9 to 21% of the genes shared in all models. Bone marrow has the lowest number of genes and reactions, whereas liver has a large number of genes and reactions not present in any other tissue. Many of the metabolic reactions in liver involve specialized lipid metabolism, e.g., de novo synthesis and secretion of bile acids including glycocholate, taurocholate, glycochenodeoxycholate, and taurochenodeoxycholate, but there are also other metabolic functions specific to liver such as ornithine degradation. To further investigate the metabolic capability of each tissue-specific GEM, we defined 256 metabolic tasks (table S15) that are known to occur in humans. The analysis shows that 192 of these metabolic tasks can be performed in all analyzed tissues, whereas the remaining 64 metabolic tasks were performed by some GEMs and clustering of these 64 metabolic tasks is shown in Fig. 6C (see also table S16). The analysis demonstrates liver as the most metabolically active tissue, followed by adipose and skeletal

muscle. For all the remaining tissues, there are variations in the metabolic activities, but with clustering of activities in tissues with similar function and morphology, e.g., stomach, duodenum, and small intestine.

Discussion

Here, we present a tissue-based map of the human proteome from analyses of 32 tissues and 47 cell lines, with gene expression data on both the RNA and protein level and with supplementary analyses on the protein level for an additional 12 tissues. An interactive resource is presented as part of the Human Protein Atlas portal (www.proteinatlas.org). This allows exploration of the tissue-elevated proteomes in these tissues and organs and analysis of tissue profiles for specific protein classes, including proteins involved in housekeeping functions in the human body, such as cell growth, energy generation, and metabolic pathways; groups of proteins involved in diseases; and proteins targeted by pharmaceutical drugs. Comprehensive lists of genes expressed at elevated levels in these tissues have been compiled, with quantitative expression profiles provided by the deep-sequencing transcriptomics complemented with immunohistochemistry. This provides localization of the proteins in the subcompartments of each tissue and organ down to the single-cell level. To facilitate integration with other biological resources, all data are available for download and through collaborations cross-linked with efforts such as UniProt (*5*), NextProt (*6*), ProteomicsDB (*9*), Metabolic Atlas (*38*), and the pan-European ELIXIR project (*39*). An important short-term objective is to facilitate international efforts (*5*, *7*, *8*, *40*) to explore the "missing proteins," with the aim to provide a finite list of human protein-coding genes and to generate firm protein evidence and expression characteristics for all of these genes. In addition, the primary data here can be used to expand the analysis of the isoform proteome to better understand the role of this diverse proteome for the functional biology of humans.

REFERENCES AND NOTES

1. P. Flicek *et al*., Ensembl 2013. *Nucleic Acids Res.* **41** (Database), D48–D55 (2013). doi: 10.1093/nar/gks1236; pmid: 23203987
2. K. D. Pruitt, T. Tatusova, G. R. Brown, D. R. Maglott, NCBI Reference Sequences (RefSeq): Current status, new features and genome annotation policy. *Nucleic Acids Res.* **40** (Database), D130–D135 (2012). doi: 10.1093/nar/gkr1079; pmid: 22121212
3. H. Kawaji *et al*., Update of the FANTOM web resource: From mammalian transcriptional landscape to its dynamic regulation. *Nucleic Acids Res.* **39** (Database), D856–D860 (2011). doi: 10.1093/nar/gkq1112; pmid: 21075797
4. A. Brazma *et al*., ArrayExpress—a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res.* **31**, 68–71 (2003). doi: 10.1093/nar/gkg091; pmid: 12519949
5. M. Magrane, U. Consortium, UniProt Knowledgebase: A hub of integrated protein data. *Database (Oxford)* **2011**, bar009 (2011). doi: 10.1093/database/bar009; pmid: 21447597
6. P. Gaudet *et al*., neXtProt: Organizing protein knowledge in the context of human proteome projects. *J. Proteome Res.* **12**, 293–298 (2013). doi: 10.1021/pr300830v; pmid: 23205526
7. I. Dunham *et al*., An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012). doi: 10.1038/nature11247; pmid: 22955616
8. Y. K. Paik *et al*., The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome. *Nat. Biotechnol.* **30**, 221–223 (2012). doi: 10.1038/nbt.2152; pmid: 22398612
9. M. Wilhelm *et al*., Mass-spectrometry-based draft of the human proteome. *Nature* **509**, 582–587 (2014).pmid: 24870543
10. M. S. Kim *et al*., A draft map of the human proteome. *Nature* **509**, 575–581 (2014). doi: 10.1038/nature13302; pmid: 24870542
11. M. Mann, Functional and quantitative proteomics using SILAC. *Nat. Rev. Molec. Cell Biol.* **7**, 952–958 (2006). doi: 10.1038/nrm2067; pmid: 17139335
12. I. Ezkurdia *et al*., Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum. Mol. Genet.* **23**, 5866–5878 (2014). doi: 10.1093/hmg/ddu309; pmid: 24939910
13. V. Lange, P. Picotti, B. Domon, R. Aebersold, Selected reaction monitoring for quantitative proteomics: A tutorial. *Mol. Syst. Biol.* **4**, 222 (2008). doi: 10.1038/msb.2008.61; pmid: 18854821
14. M. Uhlen *et al*., Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **28**, 1248–1250 (2010). doi: 10.1038/nbt1210-1248; pmid: 21139605
15. L. Fagerberg *et al*., Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* **13**, 397–406 (2014). doi: 10.1074/mcp.M113.035600; pmid: 24309898
16. C. Kampf *et al*., The human liver-specific proteome defined by transcriptomics and antibody-based profiling. *FASEB J.* **28**, 2901–2914 (2014). doi: 10.1096/fj.14-250555; pmid: 24648543
17. D. Djureinovic *et al*., The human testis-specific proteome defined by transcriptomics and antibody-based profiling. *Mol. Hum. Reprod.* **20**, 476–488 (2014). doi: 10.1093/molehr/gau018; pmid: 24598113
18. G. Gremel *et al*., The human gastrointestinal tract-specific transcriptome and proteome as defined by RNA sequencing and antibody-based profiling. *J. Gastroenterol.* **50**, 46–57 (2014). doi: 10.1007/s00535-014-0958-7; pmid: 24789573
19. V. Law *et al*., DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res.* **42** (D1), D1091–D1097 (2014). doi: 10.1093/nar/gkt1068; pmid: 24203711
20. COSMIC catalogue of somatic mutations in cancer (2014); http://cancer.sanger.ac.uk/cancergenome/projects/census.
21. E. Lundberg *et al*., Defining the transcriptome and proteome in three functionally different human cell lines. *Mol. Syst. Biol.* **6**, 450 (2010). doi: 10.1038/msb.2010.106; pmid: 21179022
22. Y. Taniguchi *et al*., Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538 (2010). doi: 10.1126/science.1188308; pmid: 20671182
23. T. N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011). doi: 10.1038/nmeth.1701; pmid: 21959131
24. L. Käll, A. Krogh, E. L. Sonnhammer, Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res.* **35** (Web server), W429–W432 (2007). doi: 10.1093/nar/gkm256; pmid: 17483518
25. H. Viklund, A. Bernsel, M. Skwark, A. Elofsson, SPOCTOPUS: A combined predictor of signal peptides and membrane protein topology. *Bioinformatics* **24**, 2928–2929 (2008). doi: 10.1093/bioinformatics/btn550; pmid: 18945683
26. D. Ramsköld, E. T. Wang, C. B. Burge, R. Sandberg, An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLOS Comput. Biol.* **5**, e1000598 (2009). doi: 10.1371/journal.pcbi.1000598; pmid: 20011106
27. E. Wingender, T. Schoeps, J. Dönitz, TFClass: An expandable hierarchical classification of human transcription factors. *Nucleic Acids Res.* **41** (D1), D165–D170 (2013).
28. D. Tamborero *et al*., Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci. Rep.* **3**, 2650 (2013). pmid: 24084849
29. D. M. Muzny *et al*., Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012). doi: 10.1038/nature11252; pmid: 22810696
30. C. S. Grasso *et al*., The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239–243 (2012). doi: 10.1038/nature11125; pmid: 22722839
31. F. Danielsson *et al*., RNA deep sequencing as a tool for selection of cell lines for systematic subcellular localization of all human proteins. *J. Proteome Res.* **12**, 299–307 (2013). doi: 10.1021/pr3009308; pmid: 23227862
32. M. Schnabel *et al*., Dedifferentiation-associated changes in morphology and gene expression in primary human articular chondrocytes in cell culture. *Osteoarthritis Cartilage* **10**, 62–70 (2002). doi: 10.1053/joca.2001.0482; pmid: 11795984
33. E. Birney *et al*., Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816 (2007). doi: 10.1038/nature05874; pmid: 17571346
34. M. Gonzàlez-Porta, A. Frankish, J. Rung, J. Harrow, A. Brazma, Transcriptome analysis of human tissues and cell lines reveals one dominant transcript per gene. *Genome Biol.* **14**, R70 (2013). doi: 10.1186/gb-2013-14-7-r70; pmid: 23815980
35. A. Mardinoglu, J. Nielsen, Systems medicine and metabolic modelling. *J. Intern. Med.* **271**, 142–154 (2012). doi: 10.1111/j.1365-2796.2011.02493.x; pmid: 22142312
36. A. Mardinoglu *et al*., Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nat. Commun.* **5**, 3083 (2014). doi: 10.1038/ncomms4083; pmid: 24419221
37. R. Agren *et al*., Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol. Syst. Biol.* **10**, 721 (2014). doi: 10.1002/msb.145122; pmid: 24646661
38. Human Metabolic Atlas (2014); www.metabolicatlas.org/.
39. L. C. Crosswell, J. M. Thornton, ELIXIR: A distributed infrastructure for European biological data. *Trends Biotechnol.* **30**, 241–242 (2012). doi: 10.1016/j.tibtech.2012.02.002; pmid: 22417641
40. J. F. Rual *et al*., Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437**, 1173–1178 (2005).

ACKNOWLEDGMENTS

We acknowledge the entire staff of the Human Protein Atlas program; the Science for Life Laboratory; and the pathology team in Mumbai, India, for valuable contributions. We thank the Department of Pathology at the Uppsala Akademiska Hospital, Uppsala, Sweden, and Uppsala Biobank for kindly providing clinical diagnostics and specimens used in this study. We also acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure (NGI), and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure. Funding was provided by the Knut and Alice Wallenberg Foundation. The authors declare that they have no conflict of interest. Correspondence and requests for materials should be addressed to M.U. The mRNA levels of all genes in each tissue sample (n = 122) are available in table S18. The supplementary Excel tables are available in the supplementary material and at www.proteinatlas.org/about/publicationdata. The raw sequencing data are available at ArrayExpress (www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-2836/) and BioProject (NIH) (www.ncbi.nlm.nih.gov/bioproject/PRJNA183192). All Protein Atlas (protein) data are available in structured XML format and can be downloaded from www.proteinatlas.org/about/download.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/1260419/suppl/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S18
References (*41*–*61*)

25 August 2014; accepted 5 December 2014
10.1126/science.1260419

RESEARCH ARTICLES

HUMAN EVOLUTION

Human-like hand use in *Australopithecus africanus*

Matthew M. Skinner,[1,2,3,4]* Nicholas B. Stephens,[3] Zewdi J. Tsegai,[3] Alexandra C. Foote,[2] N. Huynh Nguyen,[3] Thomas Gross,[5] Dieter H. Pahr,[5] Jean-Jacques Hublin,[3] Tracy L. Kivell[1,3,4]*

The distinctly human ability for forceful precision and power "squeeze" gripping is linked to two key evolutionary transitions in hand use: a reduction in arboreal climbing and the manufacture and use of tools. However, it is unclear when these locomotory and manipulative transitions occurred. Here we show that *Australopithecus africanus* (~3 to 2 million years ago) and several Pleistocene hominins, traditionally considered not to have engaged in habitual tool manufacture, have a human-like trabecular bone pattern in the metacarpals consistent with forceful opposition of the thumb and fingers typically adopted during tool use. These results support archaeological evidence for stone tool use in australopiths and provide morphological evidence that Pliocene hominins achieved human-like hand postures much earlier and more frequently than previously considered.

The human hand is traditionally distinguished from that of other apes by a suite of morphological features that are considered advantageous for distinctive forceful precision and precision-pinch grips (*1*, *2*), as well as power "squeeze" grips, such as when grasping a hammer (*3*). In particular, the morphological configuration of the fingers and thumb is critical to the manipulative abilities of modern humans. However, when and why this morphology evolved is uncertain (*2*–*5*). Hand remains of *Orrorin tugenensis* [6 million years ago (Ma)] and fossil australopiths—including *Australopithecus afarensis* (~4 to 3 Ma), *A. africanus*, and *A. sediba* (1.98 Ma)—show some of these derived human-like features, such as a pollical distal phalanx with a broad apical tuft and a well-developed attachment for the flexor pollicis longus muscle to the thumb (i.e., a broad thumb tip capable of powerful flexion) (*6*–*9*), asymmetrical metacarpal heads (*9*–*11*), more proximodistally oriented joint configuration among the radial carpometacarpal joints (*5*, *8*–*11*), and/or a high thumb-to-finger ratio [(*4*, *8*, *11*–*13*), but see (*14*)], in combination with primitive, African ape–like features, such as curved fingers with robust digital flexor muscle attachments (*5*, *8*–*11*). This mosaic morphology in australopiths is often primarily attributed to removing the hands from the constraints of locomotion (*11*, *15*), with some increase in manipulative behaviors (*2*, *4*, *13*, *16*). However, based on morphology and the absence of directly associated archaeological evidence, most australopiths are generally considered to lack human-like forceful precision and power (squeeze) grips and a commitment to tool-related manipulative behaviors [(*2*, *3*, *5*, *11*, *15*), but see (*13*, *17*)].

The association of the Olduvai Hominid 7 (OH 7) hand bones with stone tools found in the late 1950s led to the long-standing assumption that *Homo habilis* was the first stone tool maker at as early as 2.4 Ma (*18*). This evidence coincides with the first recognizable stone tools in the archaeological record at ~2.6 Ma (*19*) and some of the earliest evidence for cut-marked bone at 2.5 Ma [(*20*), but see (*21*)]. However, in the early hominin fossil record there remain inherent challenges with drawing direct associations between the archaeological evidence and the species that may have produced it (*18*). Furthermore, clear evidence of the complete suite of derived human-like hand morphology does not appear until ~0.2 Ma in *H. neanderthalensis* (*5*), though fragmentary evidence suggests a generally human-like hand at 0.8 Ma in *H. antecessor* (*22*) and possibly at 1.4 Ma in *H. erectus* (*23*). The large gap in hominin fossil hand evidence associated with early *Homo* (*5*, *23*) has made it challenging to link the evolution of particular derived morphological features of the hand with tool-related behaviors (*2*, *3*, *5*). Recent evidence of stone tool cut marks at 3.4 Ma associated with *A. afarensis*–bearing sediments suggests that stone tool use evolved much earlier than traditionally thought [(*21*), but see (*24*)] and in a hominin whose hand morphology displays only a few derived, human-like features, such as asymmetrical metacarpal heads and possibly a long thumb (*4*, *13*). This evidence suggests that stone tool–use ability does not require the complete suite of derived later *Homo*-like hand morphology. However, it is uncertain whether early hominin tool-use behaviors were opportunistic and infrequent (*25*) and whether a *Homo*-like external hand morphology evolved only in response to habitual tool use and tool production (*5*). Consequently, we require alternative methods to determine how early hominins were using their hands.

The interpretation of external morphology can be ambiguous, as some features can be retentions from the ancestral condition and may not be functionally important (*26*). In contrast, trabecular bone remodels throughout an individual's life in response to mechanical loading, a concept known as Wolff's law or bone functional adaptation (*27*). Experimental studies have shown that altering loading direction is associated with corresponding changes in trabecular orientation and/or trabecular density (*28*). Comparative studies show that variation in trabecular structure of the hand correlates with differences in joint posture during locomotion in extant apes (*29*, *30*). Thus, comparative analyses of trabecular structure have the potential to provide a more informative signal of actual behavior during the life of individuals, rather than inferred behavior based on external skeletal morphology alone, in both extant and fossil taxa.

We use a "whole-epiphysis" method (*30*–*32*) to investigate trabecular structure in metacarpal hand bones in four successive and complementary steps. First, we assess bilateral asymmetry in recent *H. sapiens* (including 1st to 3rd century Romans, 4th to 6th century Egyptian Nubians, recent Europeans, and skeletally robust Tierra del Fuegians) and *Pan* (chimpanzees and bonobos) to more accurately interpret trabecular variation in isolated fossil hominin specimens (table S1). Second, we determine whether patterns of trabecular bone distribution in recent humans and other apes (*Pan*, *Gorilla*, *Pongo*, and hylobatids) match predictions of metacarpal head peak loading based on hand posture during the predominant locomotor and manipulative behaviors. Third, we investigate the trabecular patterns of Neandertal (Kebara 2, Amud 1) and early *H. sapiens* (Qafzeh 9) metacarpals as examples of fossil hominins that exhibit external morphology similar to that of recent *H. sapiens* and are known to be committed tool makers and tool users (table S1). Fourth, we examine *A. africanus* and Pleistocene South African hominin metacarpals to determine whether they present an ape-like, human-like, or distinct trabecular structure and, in turn, morphological evidence consistent with a particular locomotor and/or manipulative behavior.

Bilateral asymmetry

Bilateral asymmetry in cortical and trabecular bone morphology of the limbs (*33*), and specifically the hands, in modern humans and other apes is well documented (*33*, *34*) and can have potential confounding affects for interpreting isolated hominin fossils. Therefore, we first tested levels of absolute asymmetry (i.e., percentage

[1]School of Anthropology and Conservation, University of Kent, Canterbury CT2 7NR, UK. [2]Department of Anthropology, University College London, London WC1H 0BW, UK. [3]Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig Germany. [4]Evolutionary Studies Institute and Centre for Excellence in PalaeoSciences, University of the Witwatersrand, Private Bag 3, Wits 2050, South Africa. [5]Institute of Lightweight Design and Structural Biomechanics, Vienna University of Technology, Gusshausstrasse 27-29, 1040 Wien, Vienna, Austria.
*Corresponding author. E-mail: m.skinner@kent.ac.uk (M.M.S.); t.l.kivell@kent.ac.uk (T.L.K.)

of absolute difference between left and right pairs) and directional asymmetry [i.e., leftward bias (negative values) or rightward bias (positive values), calculated as percentage of directional asymmetry = (right-left)/(mean of left and right) × 100] in trabecular structure in the head (distal) and base (proximal) epiphyses of the first metacarpal (Mc1) in chimpanzees, bonobos (*Pan*), and recent *H. sapiens*. In trabecular bone volume fraction (bone volume/total volume), *Pan* exhibited low average absolute asymmetry (1.6% in the head, 1.5% in the base) and a leftward bias (89% in the head, 67% in the base) (table S2). *H. sapiens* exhibited only slightly greater average absolute asymmetry (4.1% in the head, 3.7% in the base) compared with *Pan*, as well as a rightward bias (72% in the head, 86% in the base), which is consistent with the 80 to 90% level of right-handedness across *H. sapiens* (*35*). The distribution of trabecular bone density throughout the Mc1 head and base (see below) did not differ markedly in left-right pairs (see fig. S1). These low levels of absolute asymmetry indicate that taxon-specific trabecular patterns can be detected in left or right isolated bones.

Behavioral signals of hand use in trabecular structure

We tested predictions (Fig. 1) linking the distribution of trabecular bone throughout the head of Mc1, Mc3, and Mc5 and throughout the base of Mc1 to joint posture at presumed peak loading during locomotor and manipulative behaviors in extant apes. We found that variation in trabecular bone volume fraction and trabecular bone distribution across apes was consistent with our predictions (*30*) and thus a useful method for inferring hand use and, specifically, metacarpal loading in fossil specimens.

Pan and *H. sapiens* were distinct in their Mc1 trabecular bone structure. *Pan* displayed high trabecular bone volume fraction in both Mc1 epiphyses (as well as the other metacarpals), with a distopalmar concentration of trabecular density in the Mc1 head and homogeneous trabecular organization throughout the Mc1 base (Table 1, Figs. 2 and 3, and fig. S2A). This trabecular pattern is consistent with inferred thumb use in *Pan*. Specifically, the *Pan* thumb is most often used (i) during vertical climbing and suspensory locomotion on large substrates, in which most of the thumb (as opposed to just the tip) grasps the branch in opposition to the fingers (*36*) and probably experiences high loads from superstrate reaction forces and muscle contraction, and (ii) during the manipulation of food or tools (*37*), in which the thumb is often opposed to the side of the index finger or abducted but with limited force [(*36*, *38*) but see (*39*)]. The *Pan* thumb is short compared with the length of the fingers (*8*, *11*) and is therefore not loaded during knuckle-walking (*40*) and is considered not to be capable of forcefully opposing the thumb and fingers in precision grips (*36*, *38*).

In contrast, humans showed low trabecular bone volume fraction in all metacarpals, with a palmar concentration of trabecular density in both the Mc1 head and base. This trabecular pattern is consistent with the forceful opposition of the thumb to the fingers in precision and precision-pinch grips (*1*, *2*, *38*) (Table 1, Figs. 2 and 3, and fig. S2A). In vivo kinematic analyses of the trapezium-Mc1 joint show flexion and palmar translation of the Mc1 during precision grip tasks (*41*) and that there is significantly increased trabecular bone (i.e., higher trabecular bone volume fraction, trabecular thickness, and connectivity) in the palmar portion of the corresponding Mc1 facet on the trapezium (*42*). This evidence is consistent with high compressive loading on the palmar region of the Mc1 base. Furthermore, cartilage loss in the trapezium-Mc1 joint is greatest in the palmar region (*43*), and the high prevalence of osteoarthritis in the human hand is at this joint (*44*), providing further support that the Mc1 base, and particularly the palmar region, is heavily loaded during manipulation. This palmar concentration of trabeculae in the Mc1 base is not found in *Pan* despite opposition of the thumb during arboreal locomotion and tool use, demonstrating that this is a distinct human pattern probably reflecting the distinctly human ability for forceful precision gripping (*1*, *2*, *38*) (Fig. 3).

Differences in the distribution of trabecular bone across apes and humans were also found in Mc3 and Mc5 (Fig. 2). Knuckle-walking *Pan* and *Gorilla* exhibit a dorsal concentration of trabecular bone in the Mc3 and Mc5 heads consistent with an extended metacarpophalangeal joint during knuckle-walking locomotion (*30*, *40*) (Fig. 2 and fig. S2B). In contrast, *Pongo* and hylobatids exhibit a distopalmar concentration consistent with palmar grasping and hook grips used during climbing and suspensory locomotion (*29*, *30*). *H. sapiens* also exhibit a distopalmar trabecular concentration in the Mc3 and Mc5 heads consistent with peak loading of the metacarpophalangeal joint during forceful flexed-finger grasping (*1*). However, the *H. sapiens* Mc3 and Mc5 heads also exhibit a distinct (relative to other apes) asymmetry in the concentration of trabecular bone on the palmar surface, extending proximally on the medial side. This asymmetric pattern is consistent with rotation (flexion and abduction) of the third and fifth proximal phalanges on the metacarpal heads during opposition to the thumb in precision grips (*2*), as well as during the distinctly human power (squeeze) grip of cylindrical objects (*3*). The *H. sapiens* Mc3 and Mc5 were also distinct in having significantly lower trabecular bone volume fraction compared with other apes, perhaps because the hands are rarely used for body weight support during locomotion (*29*, *30*) (Figs. 2 and 3, Table 1, and table S3). Thus, trabecular structure within *H. sapiens*

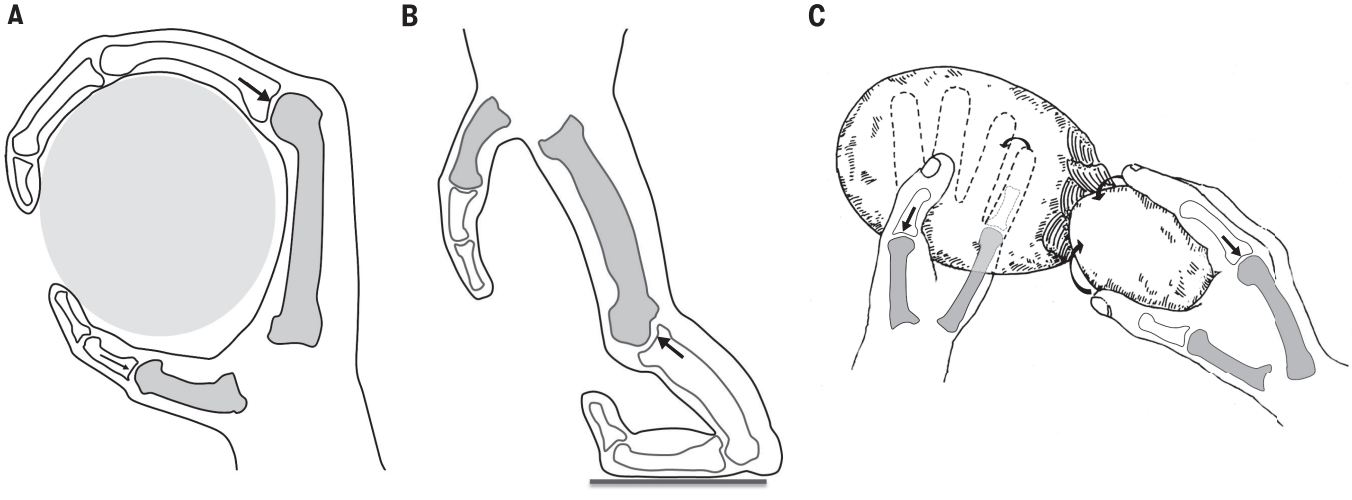


Fig. 1. Predictions of peak joint loading during the habitual hand postures used by apes. Hand postures are shown for (**A**) arboreal and (**B**) knuckle-walking locomotion and (**C**) human manipulation [adapted from (*2*)]. Straight arrows indicate the predicted direction of peak joint reaction force. Curved arrows in (C) help to illustrate opposition of the thumb and fingers during precision grasping.

Table 1. Trabecular bone volume fraction (percent) in the metacarpal sample for extant and fossil taxa. *H. sapiens* have significantly ($P < 0.001$) lower trabecular bone volume fraction than *Pan* in the Mc1 base and the Mc1, Mc3, and Mc5 heads (see table S3). Dashes indicate values not measured.

Taxon and specimen	Mc1 base		Mc1 head		Mc3 head		Mc5 head	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Hylobatids	–	–	–	–	18.02	4.94	22.21	1.56
Pongo	–	–	–	–	19.81	6.28	21.69	6.25
Gorilla	–	–	–	–	24.55	0.10	23.13	1.23
Pan	26.56	3.82	32.17	4.69	26.29	3.61	26.90	3.80
Recent *H. sapiens*	17.58	0.03	20.29	0.04	14.49	1.92	13.85	3.00
Early *H. sapiens*								
Qafzeh 9	18.30	–	18.00	–	23.80	–	20.42	–
H. neanderthalensis								
Amud 1	20.30	–	–	–	–	–	–	–
H. neanderthalensis								
Kebara 2	19.30	–	15.90	–	20.80	–	19.25	–
A. robustus/early *Homo*								
SKX 5020	26.30	–	–	–	–	–	–	–
A. robustus/early *Homo*								
SK 84	24.00	–	20.00	–	–	–	–	–
A. africanus								
StW 418	18.80	–	21.40	–	–	–	–	–
A. africanus								
StW 382 (Mc2)	–	–	–	–	25.60	–	–	–
A. africanus								
StW 394	–	–	–	–	25.90	–	–	–
A. africanus								
StW 552 (Mc4)	–	–	–	–	21.50	–	–	–

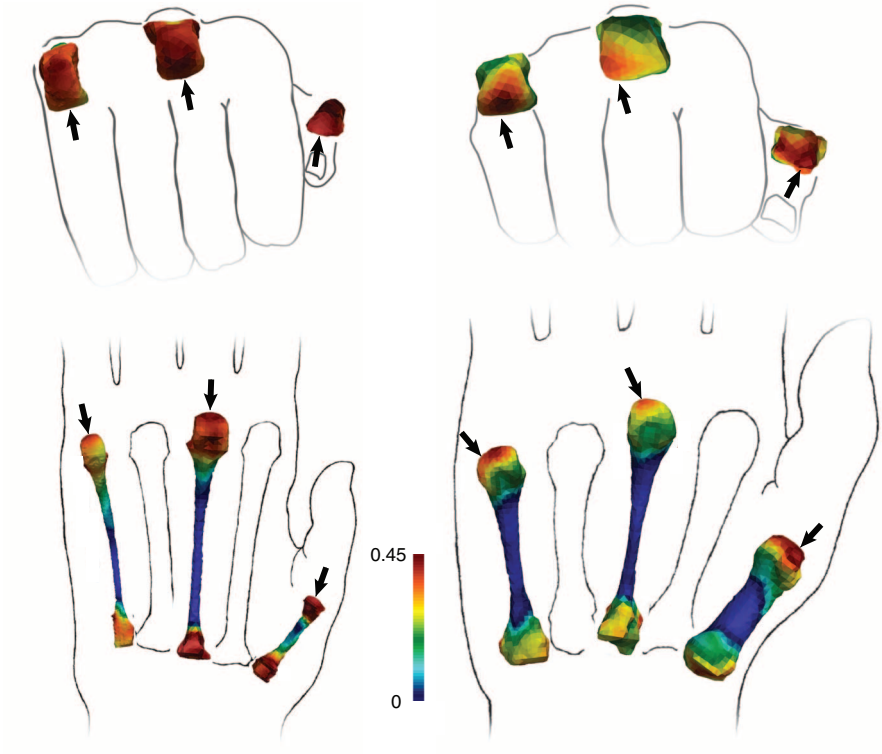


Fig. 2. Pattern of trabecular bone distribution in *Pan* (*P. paniscus*, left) and recent *H. sapiens* (right). Bones are scaled to trabecular bone volume fraction [bone volume/total volume (BV/TV)] of 0 to 45%. Black arrows point to the regions of highest BV/TV (red), demonstrating distopalmar and asymmetrical peak loading of the metacarpal heads in recent *H. sapiens* compared with the distodorsal and symmetrical loading of *Pan*.

Mc3 and Mc5 heads is distinct from that of other apes and is consistent with peak loading of the metacarpophalangeal joint during forceful precision or power (squeeze) gripping.

Pleistocene Neandertals and early *H. sapiens*

Neandertals and early *H. sapiens* were committed tool makers and tool users, although differences in tool kits (*45*) and hand morphology have been documented, with implications for potential subtle differences in hand biomechanics (*46*, *47*). The metacarpals of Neandertals (Kebara 2 and Amud 1) and early *H. sapiens* (Qafzeh 9) demonstrate the characteristic recent human-like pattern of trabecular bone (Figs. 3 and 4). The Mc1s of both fossil *Homo* taxa exhibit low trabecular bone volume fraction and a clear palmar concentration of trabeculae in the base like that of recent *H. sapiens* and distinctly unlike that of other apes (Table 1 and Fig. 3). All Mc3 and Mc5 Neandertal and early *H. sapiens* specimens showed the palmar concentration of the trabeculae with a proximally extended asymmetry on the medial portion in the head, although trabecular bone volume fraction is higher than that of recent *H. sapiens* (Table 1 and Fig. 4). Together, areas of peak loading of the palm in Pleistocene Neandertals and early *H. sapiens* individuals are similar in pattern, but perhaps higher in magnitude (at least in the Mc3 and Mc5), to that experienced by recent *H. sapiens* and are consistent with forceful opposition of the thumb and fingers during precision and power (squeeze) gripping.

A. africanus and other South African hominins

The external morphology of the *A. africanus* (3 to 2 Ma) and Swartkrans (1.9 to 1.8 Ma) hominin metacarpals indicates that they did not have the full suite of features associated with forceful precision gripping and committed tool use (*5*, *12*, *17*) and, at least in *A. africanus*, may still have used their hands for arboreal locomotion (*48*). Like the external morphology, the internal structure of earlier South African hominin metacarpals preserves a mosaic trabecular pattern between that of humans and other apes. All early hominin metacarpal specimens generally exhibit high trabecular bone volume fraction like that of *Pan* and unlike that of the lower trabecular density of recent *H. sapiens*. However, the distribution of trabecular bone is similar to the distinct patterns of committed tool users and tool makers, *H. sapiens* and Neandertals. The *A. africanus* Mc1 (StW 418) base exhibits a clear palmar concentration of trabecular bone like that found in *H. sapiens* and Neandertals, which is consistent with forceful opposition of the thumb in precision gripping (Fig. 3). The same palmar concentration of trabeculae is also found in the base of Swartkrans Mc1 specimen SK 84, although SKX 5020 displays a more homogeneous trabecular distribution. This variation in the Swartkrans hominins is perhaps not surprising given that the taxonomic affinity (either *A. robustus* or early

Homo) and potential tool-use capability (based on investigation of the external morphology) have been debated (*14*, *49*).

Furthermore, the one *A. africanus* Mc3 specimen that preserves trabeculae (StW 394), as well as a Mc2 (StW 382) and, less so, Mc4 (StW 552) (table S4), is similar to humans in expressing a distopalmar and asymmetrical concentration of trabecular bone in the metacarpal head (Fig. 4).

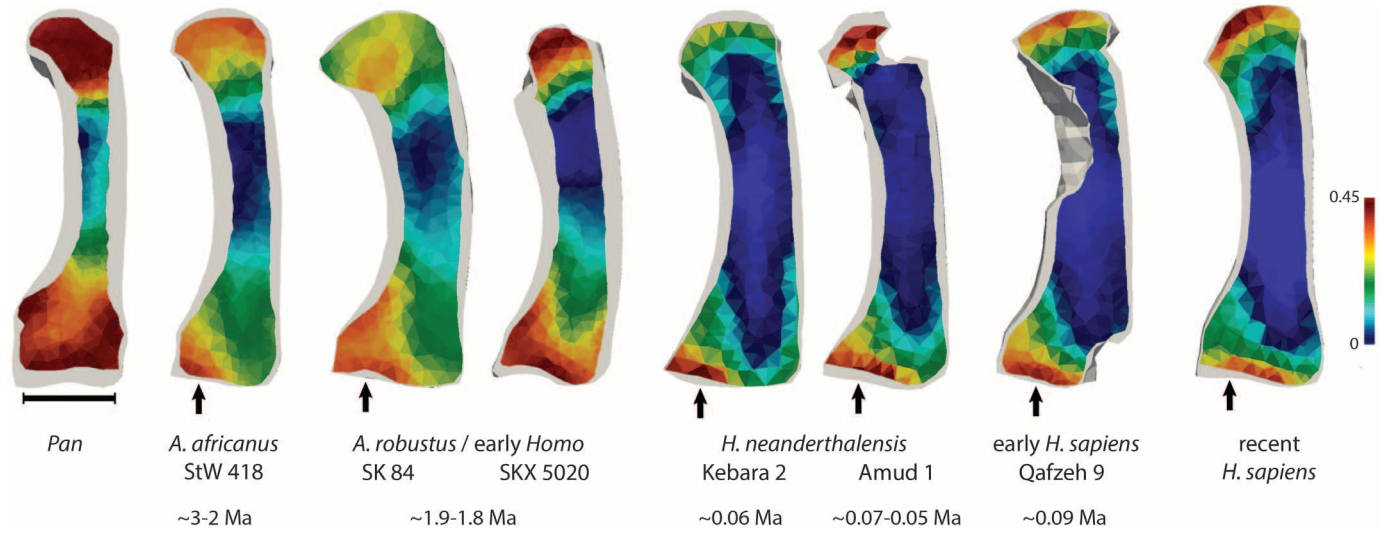


Fig. 3. Sagittal midline cross sections of the first metacarpal in *Pan*, recent *H. sapiens*, and fossil hominins. Recent and early *H. sapiens* demonstrate a palmar concentration of trabecular bone in the proximal base of the Mc1 consistent with forceful opposition of the thumb, which is distinct from the more homogeneous distribution of trabecular bone in *Pan*. *A. africanus*, Swartkrans specimen SK 84, and *H. neanderthalensis* also demonstrate the *H. sapiens* pattern, suggesting similar loading and use of the thumb. Swartkrans specimen SKX 5020 has a more homogeneous distribution of trabecular bone in the Mc1 base that is more similar to *Pan*. The distal heads of SKX 5020 and Amud 1, as well as the proximal base and palmar shaft of Qafzeh 9, are not fully preserved. Trabecular bone volume fraction is scaled to 0 to 45%.

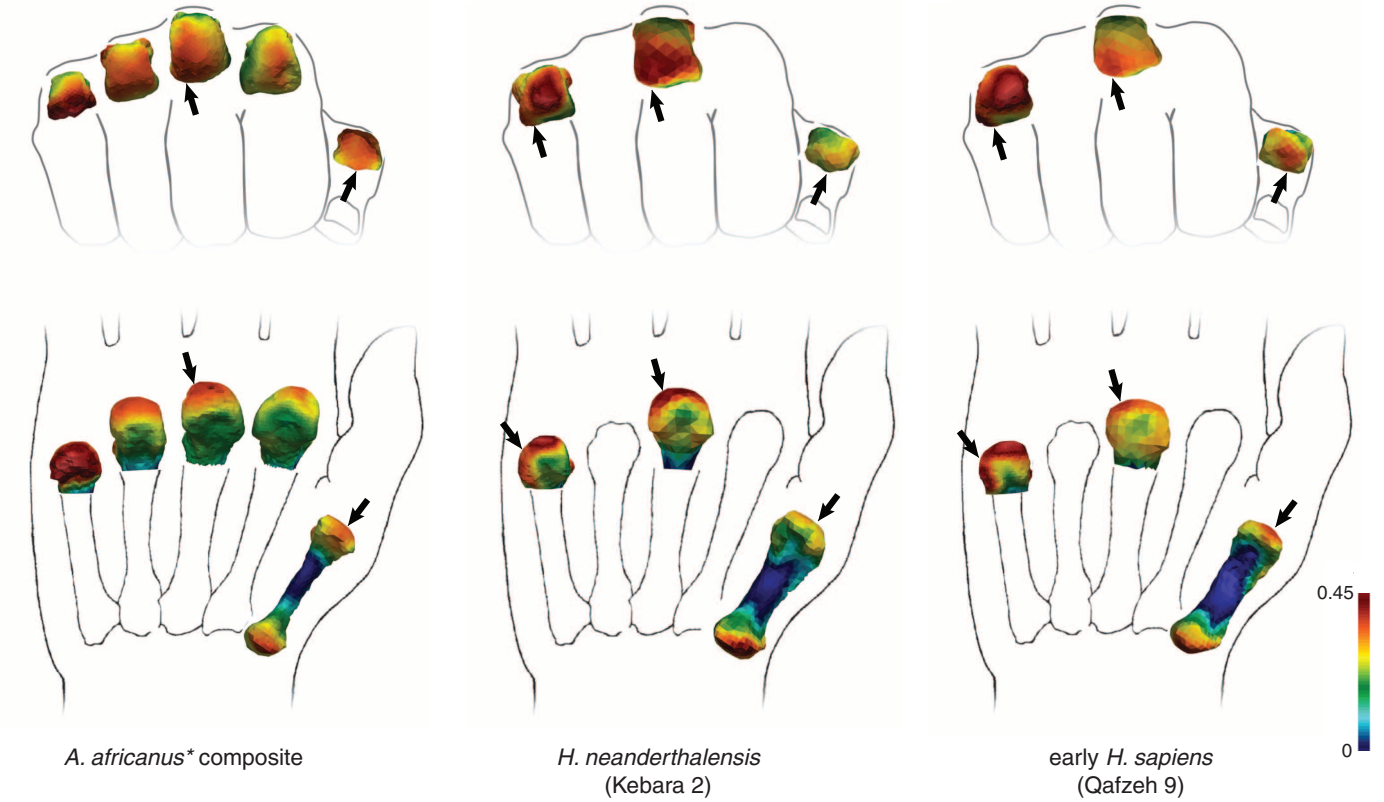


Fig. 4. Pattern of trabecular bone distribution in *A. africanus*, *H. neanderthalensis*, and early *H. sapiens*. The *A. africanus* composite includes StW 418 (left Mc1, mirrored), StW 382 (left Mc2, mirrored), StW 394 (left Mc3, mirrored), and StW 552 (right Mc4), as well as (*) Swartkrans specimen SK(W) 14147 (left Mc5, mirrored) associated with either *A. robustus* or early *Homo*. Arrows point to regions of highest BV/TV (red), indicating a recent *H. sapiens*–like distopalmar and asymmetrical peak loading of the metacarpal heads in all fossil hominins that is distinctly different from the pattern found in *Pan*. Trabecular bone volume fraction is scaled to 0 to 45%.

Swartkrans Mc5 specimen SK(W) 14147, attributed to either *A. robustus* or early *Homo*, also demonstrates a distopalmar concentration to trabecular bone; however, poor preservation of the trabeculae owing to matrix infilling along the medial side prohibits determination of the presence or absence of the asymmetry found in later *Homo*. Altogether, the trabecular structure preserved in the *A. africanus* and Swartkrans metacarpals shows a combination of *Pan*-like and human-like morphology. A generally higher trabecular bone volume fraction across all of the metacarpals compared with recent *H. sapiens* may suggest higher loading of the hand that occurs when using the hands during locomotion. However, the distinctly *H. sapiens*– and Neandertal-like distribution of trabeculae is consistent with forceful opposition of the thumb and fingers during precision and power (squeeze) gripping.

Discussion

Decades of investigation of the external morphology of early hominin hand fossils (*A. afarensis*, *A. africanus*, *A. sediba*, and *A. robustus*/early *Homo*) have stagnated in debates about the potential hand use behaviors of early hominins—in particular, the questions of whether australopiths were still using their hands for arboreal locomotion and what were their manipulative capabilities (*4*, *5*, *7*, *8*, *12*–*14*, *16*, *17*, *48*, *49*). These behavioral debates continue because of challenges associated with differentiating functionally important features from primitive, and potentially adaptively unimportant, retentions in the external morphology of the hominin skeleton (*26*). Our findings address this problem by providing morphological evidence from internal bone structure that can be linked to behavior and hand use during life.

The high *Pan*-like trabecular density in *A. africanus* and the Swartkrans hominin metacarpals suggests that the hands of these early hominins may still be used for arboreal locomotion. However, the higher trabecular bone volume also found in medial metacarpals of Neandertals and early *H. sapiens* may (i) indicate that the hands of both fossil *Homo* taxa were experiencing higher loads than those of recent *H. sapiens* or (ii) reflect a systemic pattern of high trabecular bone volume throughout the skeleton in all early hominins compared with recent *H. sapiens*. Nonetheless, the distribution of the trabecular structure in *A. africanus* and Swartkrans specimen SK 84 is similar to the characteristic pattern of committed tool makers, Neandertals and *H. sapiens*, and distinctly unlike that of other apes. The palmar concentration of trabecular bone in the Mc1 base and the asymmetrical distribution of trabeculae in the third and fifth metacarpal heads in all hominins, as well as the absence of this pattern in *Pan* and other apes, suggests that the hominin trabecular pattern reflects the forceful opposition of the thumb and fingers during human-like precision and power (squeeze) grips (*1*–*3*, *38*). Although hominins may have used forceful hand grips for any number of manipulative behaviors (*16*), tool-related behaviors probably generated some of the largest (nonlocomotor) forces on the hands and are thus considered to be a strong selective pressure on hominin hand morphology (*2*, *5*, *6*, *8*, *17*, *18*, *25*). Experimental evidence has demonstrated that forces are highest on the thumb during precision grips used during flake use rather than stone tool making (*50*), which is consistent with the earliest tool technologies (*25*) and, in particular, the archaeological evidence for use of stone flakes to remove meat from bone in early hominins (*20*, *21*). Thus, analyses of trabecular bone structure can provide direct morphological evidence of forceful precision and tool use in hominins prior to the first recognizable stone tools and may offer a method for linking archaeological evidence with particular hominin taxa. These results show that *A. africanus* was capable of habitual and forceful human-like opposition of the thumb and fingers during precision and power (squeeze) grips that are used during tool-related behaviors, providing morphological evidence of committed tool use in a hominin hitherto considered not to be capable of these behaviors and evolving ~500,000 years earlier than the first evidence of stone tool production.

REFERENCES AND NOTES

1. J. R. Napier, *J. Bone Joint Surg. Br.* **38-B**, 902–913 (1956).
2. M. W. Marzke, *Am. J. Phys. Anthropol.* **102**, 91–110 (1997).
3. M. W. Marzke, K. L. Wullstein, S. F. Viegas, *Am. J. Phys. Anthropol.* **89**, 283–298 (1992).
4. D. M. Alba, S. Moyà-Solà, M. Köhler, *J. Hum. Evol.* **44**, 225–254 (2003).
5. M. W. Tocheri, C. M. Orr, M. C. Jacofsky, M. W. Marzke, *J. Anat.* **212**, 544–562 (2008).
6. D. E. Ricklan, in *From Apes to Angels: Essays in Anthropology in Honor of Phillip V. Tobias*, G. H. Sperber, Ed. (Wiley-Liss, New York, 1990), pp. 171–183.
7. S. Almécija, S. Moyà-Solà, D. M. Alba, *PLOS ONE* **5**, e11727 (2010).
8. T. L. Kivell, J. M. Kibii, S. E. Churchill, P. Schmid, L. R. Berger, *Science* **333**, 1411–1417 (2011).
9. C. V. Ward, W. H. Kimbel, E. H. Harmon, D. C. Johanson, *J. Hum. Evol.* **63**, 1–51 (2012).
10. M. E. Bush, C. O. Lovejoy, D. C. Johanson, Y. Coppens, *Am. J. Phys. Anthropol.* **57**, 651–677 (1982).
11. M. W. Marzke, *J. Hum. Evol.* **12**, 197–211 (1983).
12. D. J. Green, A. D. Gordon, *J. Hum. Evol.* **54**, 705–719 (2008).
13. S. Almécija, D. M. Alba, *J. Hum. Evol.* **73**, 88–92 (2014).
14. C. Rolian, A. D. Gordon, *Am. J. Phys. Anthropol.* **152**, 393–406 (2013).
15. B. Latimer, in *Origine(s) de la Bipédie Chez les Hominidés*, B. Senut, Y. Coppens, Eds. (CNRS, Paris, 1991), pp. 169–176.
16. S. Moyà-Solà, M. Köhler, D. M. Alba, S. Almécija, *Folia Primatol. (Basel)* **79**, 215–250 (2008).
17. R. L. Susman, *Science* **265**, 1570–1573 (1994).
18. J. R. Napier, *Nature* **196**, 409–411 (1962).
19. S. Semaw *et al.*, *J. Hum. Evol.* **45**, 169–177 (2003).
20. J. de Heinzelin *et al.*, *Science* **284**, 625–629 (1999).
21. S. P. McPherron *et al.*, *Nature* **466**, 857–860 (2010).
22. C. Lorenzo, J. L. Arsuaga, J. M. Carretero, *J. Hum. Evol.* **37**, 501–522 (1999).
23. C. V. Ward, M. W. Tocheri, J. M. Plavcan, F. H. Brown, F. K. Manthi, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 121–124 (2014).
24. M. Domínguez-Rodrigo, T. R. Pickering, H. T. Bunn, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20929–20934 (2010).
25. M. A. Panger, A. S. Brooks, B. G. Richmond, B. Wood, *Evol. Anth.* **11**, 235–245 (2002).
26. C. V. Ward, *Am. J. Phys. Anthropol.* **119** (suppl. 35), 185–215 (2002).
27. C. Ruff, B. Holt, E. Trinkaus, *Am. J. Phys. Anthropol.* **129**, 484–498 (2006).
28. M. M. Barak, D. E. Lieberman, J.-J. Hublin, *Bone* **49**, 1141–1151 (2011).
29. A. Zeininger, B. G. Richmond, G. Hartman, *J. Hum. Evol.* **60**, 703–710 (2011).
30. Z. J. Tsegai *et al.*, *PLOS ONE* **8**, e78781 (2013).
31. T. Gross, T. L. Kivell, M. M. Skinner, N. H. Nguyen, D. H. Pahr, *Paleontol. Electron.* **17**, 17.3.33A (2014).
32. Materials and methods are available as supplementary materials on *Science* Online.
33. J. T. Stock, M. K. Shirley, L. A. Sarringhaus, T. G. Davies, C. N. Shaw, *Ann. N. Y. Acad. Sci.* **1288**, 86–99 (2013).
34. R. A. Lazenby, D. M. L. Cooper, S. Angus, B. Hallgrímsson, *J. Hum. Evol.* **54**, 875–885 (2008).
35. J. Steele, *Laterality* **5**, 193–220 (2000).
36. M. W. Marzke, K. L. Wullstein, *Int. J. Primatol.* **17**, 117–139 (1996).
37. C. Boesch, H. Boesch, in *Hands of Primates*, H. Preuschoft, D. J. Chivers, Eds. (Springer, Vienna, 1993), pp. 31–43.
38. M. W. Marzke *et al.*, *Am. J. Phys. Anthropol.* **110**, 163–178 (1999).
39. M. W. Marzke, L. F. Marchant, W. C. McGrew, S. P. Reece, *Am. J. Phys. Anthropol.* 10.1002/ajpa.22651 (2014).
40. F. A. Jenkins, J. G. Fleagle, in *Primate Functional Morphology and Evolution*, R. H. Tuttle, Ed. (Mouton, The Hague, Netherlands, 1975), pp. 213–227.
41. E. Halilaj *et al.*, *Clin. Orthop. Relat. Res.* **472**, 1114–1122 (2014).
42. A. T. Lee *et al.*, *J. Hand Surg. Am.* **38**, 309–315 (2013).
43. M. F. Koff *et al.*, *J. Hand Surg. Am.* **28**, 597–604 (2003).
44. M. M. Haara *et al.*, *J. Bone. Joint. Surg.* **86-A**, 1452–1457 (2004).
45. A. S. Brooks, J. E. Yellen, L. Nevell, G. Hartman, in *Transitions Before the Transition: Evolution and Stability in the Middle Paleolithic and Middle Stone Age*, E. Hovers, S. Kuhn, Eds. (Kluwer, New York, 2005), pp. 233–255.
46. E. Trinkaus, I. Villemeur, *Am. J. Phys. Anthropol.* **84**, 249–260 (1991).
47. W. A. Niewoehner, A. H. Weaver, E. Trinkaus, *Am. J. Phys. Anthropol.* **103**, 219–233 (1997).
48. H. M. McHenry, L. R. Berger, *J. Hum. Evol.* **35**, 1–22 (1998).
49. E. Trinkaus, J. C. Long, *Am. J. Phys. Anthropol.* **83**, 419–424 (1990).
50. E. M. Williams, A. D. Gordon, B. G. Richmond, *J. Hum. Evol.* **62**, 520–532 (2012).

ACKNOWLEDGMENTS

Tomographic scans of material from Sterkfontein were produced through a collaborative project between the University of the Witwatersrand and the Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology. We thank R. Lazenby for fruitful discussions and S. McPherron for helpful feedback on previous versions of this manuscript. We thank the following institutions and researchers for access to fossils and comparative specimens in their care: University of Witwatersrand (B. Zipfel), Berlin Museum für Naturkunde (F. Mayer, S. Jancke), Max Planck Institute for Evolutionary Anthropology (C. Boesch), Musée Royal de l'Afrique Centrale (E. Gilissen, W. Wendelen), Museo Nazionale Preistorico Etnografico "Luigi Pigorini" (R. Macchiarelli), Naturhistorisches Museum Wien (M. Teschler-Nicola, R. Muehl), Senckenberg Museum of Frankfurt (V. Volpato), Tel Aviv University (Y. Rak, I. Herskovitz), and University of Florence (J. Moggi Cecchi, S. Bortoluzzi). We thank C. Moore, D. Reinhardt, P. Schoenfeld, and H. Temming for scanning and technical assistance. This research is supported by the Max Planck Society (M.M.S., T.L.K., Z.J.T., N.B.S., J.-J.H.) and the European Research Council Starting Grant 336301 (T.L.K. and M.M.S.). Author contributions: M.M.S., T.L.K., N.B.S., Z.J.T., A.C.F., and J.-J.H. collected data; D.H.P. developed MedTool software; D.H.P., T.G., and N.H.N. provided analytical support; N.B.S., Z.J.T. and A.C.F. ran analyses; M.M.S., T.L.K., N.B.S., Z.J.T., and A.C.F. interpreted results; M.M.S., T.L.K., N.B.S., Z.J.T., T.G., D.H.P., and J.-J.H. wrote the paper. We confirm no competing financial interests. The data reported in this paper are tabulated in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/395/suppl/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S5
References (*51*–*59*)

25 September 2014; accepted 1 December 2014
10.1126/science.1261735

T CELL IMMUNITY

Functional heterogeneity of human memory CD4$^+$ T cell clones primed by pathogens or vaccines

Simone Becattini,[1,2] Daniela Latorre,[1] Federico Mele,[1] Mathilde Foglierini,[1] Corinne De Gregorio,[1] Antonino Cassotta,[1] Blanca Fernandez,[1] Sander Kelderman,[3] Ton N. Schumacher,[3] Davide Corti,[1] Antonio Lanzavecchia,[1,2] Federica Sallusto[1]*

Distinct types of CD4$^+$ T cells protect the host against different classes of pathogens. However, it is unclear whether a given pathogen induces a single type of polarized T cell. By combining antigenic stimulation and T cell receptor deep sequencing, we found that human pathogen- and vaccine-specific T helper 1 (T_H1), T_H2, and T_H17 memory cells have different frequencies but comparable diversity and comprise not only clones polarized toward a single fate, but also clones whose progeny have acquired multiple fates. Single naïve T cells primed by a pathogen in vitro could also give rise to multiple fates. Our results unravel an unexpected degree of interclonal and intraclonal functional heterogeneity of the human T cell response and suggest that polarized responses result from preferential expansion rather than priming.

The functional diversity of CD4 T helper (T_H) cells has evolved as a means to provide the immune system with the capacity to mount the appropriate type of defense against different classes of pathogens (*1*, *2*). The response to viruses and intracellular bacteria is dominated by T_H1 cells that produce the cytokine interferon (IFN)–γ and activate macrophage effector function. T_H17 cells that produce the cytokines interleukin (IL)–17 and IL-22 and activate neutrophils dominate the response to fungi and extracellular bacteria. The response to parasites and venoms is dominated by T_H2 cells that produce IL-4, IL-5, and IL-13 and activate eosinophils and other innate cells. Classical observations in mice and humans with defects in T cell polarization demonstrate the importance of generating the right type of response against each class of pathogen (*3*–*7*).

T_H1, T_H2, and T_H17 are considered alternative fates of differentiating CD4 T cells, which are instructed by signals from the T cell receptor (TCR), costimulatory molecules, and cytokines (*8*–*10*). However, it is increasingly appreciated that most polarized T cells are not terminally differentiated but rather maintain flexibility to change the type, or increase the range, of cytokines produced (*11*). Although there is evidence for such cytokine flexibility from in vitro experiments and animal models (*12*–*14*), it is unclear whether and how this mechanism contributes to the heterogeneity of the human memory T cell pool.

In principle, a functionally heterogeneous population of memory T cells may derive from the priming of multiple clones, each polarized to a distinct fate, or from the generation of multiple fates within the progeny of the same T cell clone (*15*). The "one cell, multiple fate" model has been documented in mouse experiments: Single naïve CD8 T cells were found to give rise to a heterogeneous progeny comprising T central memory (T_{CM}) and T effector memory (T_{EM}) cells (*16*–*18*), and single naïve CD4 T cells were found to give rise to T_H1 and T follicular helper (T_{FH}) cells (*19*). However, both T_{CM} and T_{FH} cells can be considered intermediate stages of differentiation rather than alternative fates. Therefore, it remains to be determined whether the "one cell, multiple fate" paradigm might apply to alternative differentiation states, such as T_H1, T_H2, and T_H17.

Functional heterogeneity and clonal composition of *Candida albicans*–specific CD4 T cells

To dissect the heterogeneity of human memory CD4 T cells, we isolated four memory T_H cell subsets, distinguished by chemokine receptor expression, from the peripheral blood mononuclear cells (PBMCs) of healthy donors (fig. S1) (*20*, *21*): T_H1 ($CXCR3^+CCR4^-CCR6^-$), T_H2 ($CCR4^+CXCR3^-CCR6^-$), T_H17 ($CCR4^+CCR6^+CXCR3^-$), and nonconventional T_H1 ($CXCR3^+CCR6^+CCR4^-$, hereafter defined as T_H1*) that produced IFN-γ and low levels of IL-17 (*22*). Cells were labeled with carboxyfluorescein succinimidyl ester (CFSE) and stimulated with *C. albicans* in the presence of autologous monocytes. In all donors analyzed, $CFSE^{lo}$ proliferating cells were found in high proportions in the T_H17 and T_H1* subsets but were also detected in the T_H1 and T_H2 subsets (Fig. 1, A and B) (*23*), consistent with previous reports (*22*, *24*, *25*). The pattern of cytokines produced in *C. albicans*–stimulated cultures was distinct and characteristic of each subset. *C. albicans*–specific T_H17 cells produced IL-17A, IL-22, and IFN-γ; T_H1* cells produced high amounts of IFN-γ and low amounts of IL-17A and IL-4; T_H2 cells produced IL-4, IL-5, and IL-13; and T_H1 cells produced IFN-γ (Fig. 1C). These findings indicate that human memory CD4 T cells responding to *C. albicans* are functionally heterogeneous in terms of migratory properties and effector function.

To analyze the TCR repertoire of antigen-specific memory cells within each T_H subset, we isolated $CFSE^{lo}$ cells from *C. albicans*–stimulated cultures from five healthy donors and used multiplex polymerase chain reaction (PCR) assays on genomic DNA for quantitative assessment of TCRβ clonotypes by deep sequencing [quantification and correction procedures are described in fig. S2 and (*20*, *26*)]. Even though the numbers of T cells recovered from the T_H17 and T_H1* subsets were higher than for T_H1 or T_H2 by a factor of 5 to 20 (table S1), the number of clonotypes detected in the four subsets was comparable within each donor tested, with an average of 976, 595, 830, and 696 unique clonotypes in T_H1, T_H2, T_H1*, and T_H17 subsets, respectively (Fig. 1D). Furthermore, the frequency distribution of individual clonotypes was comparable among the four T_H subsets and varied widely, ranging from 20% to 0.01% of total reads (Fig. 1E). In all subsets and donors, fewer than 20 clonotypes (representing 3 to 4.5% of total clonotypes) were present at frequencies above 1% and together contributed 34 to 43% of total reads (Fig. 1, F and G, and table S1).

To investigate whether the most frequent TCRβ clonotypes detected after in vitro stimulation were already expanded in vivo, we compared the clonotype composition of in vitro expanded *C. albicans*–specific CD4 T cells and total memory CD4 T cells immediately after isolation from peripheral blood. In one donor, we detected 300,945 clonotypes in 5×10^6 total memory cells, whereas we detected 2350 clonotypes in *C. albicans*–specific cells (Fig. 1H). Of the latter, 1426 (corresponding to 60.7% of unique clonotypes and 86.7% of total reads) were also found in unstimulated total memory cells. Furthermore, among the shared clonotypes, we observed a significant positive correlation between the frequency in the *C. albicans*–specific cell line and that in the total memory pool. Collectively, the above analysis indicates that memory CD4 T cells specific for *C. albicans* are highly polyclonal and that individual clonotypes show a broad range of frequency distribution that is comparable in different memory subsets and reflects their frequency in the total memory pool.

Unique and shared clonotypes among *C. albicans*–specific memory T cell subsets

The diversity and functional heterogeneity of the *C. albicans*–specific memory pool could be due to different clones each polarized to a single fate, or to clones that had acquired multiple fates through a process of intraclonal diversification. A comparative analysis of the TCRβ repertoires

[1]Institute for Research in Biomedicine, Bellinzona, Università della Svizzera Italiana, Lugano, Switzerland. [2]Institute of Microbiology, ETH Zürich, Zürich, Switzerland. [3]Division of Immunology, Netherlands Cancer Institute, Amsterdam, Netherlands.
*Corresponding author. E-mail: federica.sallusto@irb.usi.ch

of *C. albicans*–specific T cells from a representative donor (Fig. 2A) showed that of the 395 clonotypes detected in the T_H17 subset, 143 were also found in the T_H1^* subset, 53 in the T_H2 subset, and 8 in the T_H1 subset. Furthermore, among the 143 clonotypes shared between T_H17 and T_H1^*, 27 were also found in the T_H2 subset, and 3 in both T_H2 and T_H1 subsets (Fig. 2B). Several shared clonotypes were found to be highly expanded in both T_H17 and T_H1^* subsets. Moreover, several expanded clonotypes were found only in one subset, excluding the possibility that clonotype sharing was due to contamination during cell sorting. Similar patterns of clonotype sharing among memory subsets were found in the additional four donors analyzed (fig. S3). From the compiled analysis of the five donors, the highest sharing of clonotypes and reads was found between the T_H17 and T_H1^* subsets, followed by the T_H17 and T_H2 subsets, whereas the lowest was between the T_H17 and T_H1 subsets (Fig. 2C). Collectively, the above findings indicate that the functional heterogeneity of *C. albicans*–specific memory T cells rests both on clones that are uniquely polarized to a single fate and on clones that have generated multiple and distinct T cell fates through a process of intraclonal diversification.

Different patterns of clonotype sharing among *Mycobacterium tuberculosis*–specific T cell subsets

The high extent of clonotype sharing between T_H17 and T_H1^* observed in the response to *C. albicans* could have been predicted on the basis of some commonalities between these two subsets as well as the reported plasticity of T_H17 cells (*12*–*14*). To investigate whether clonotype sharing is a general property of these

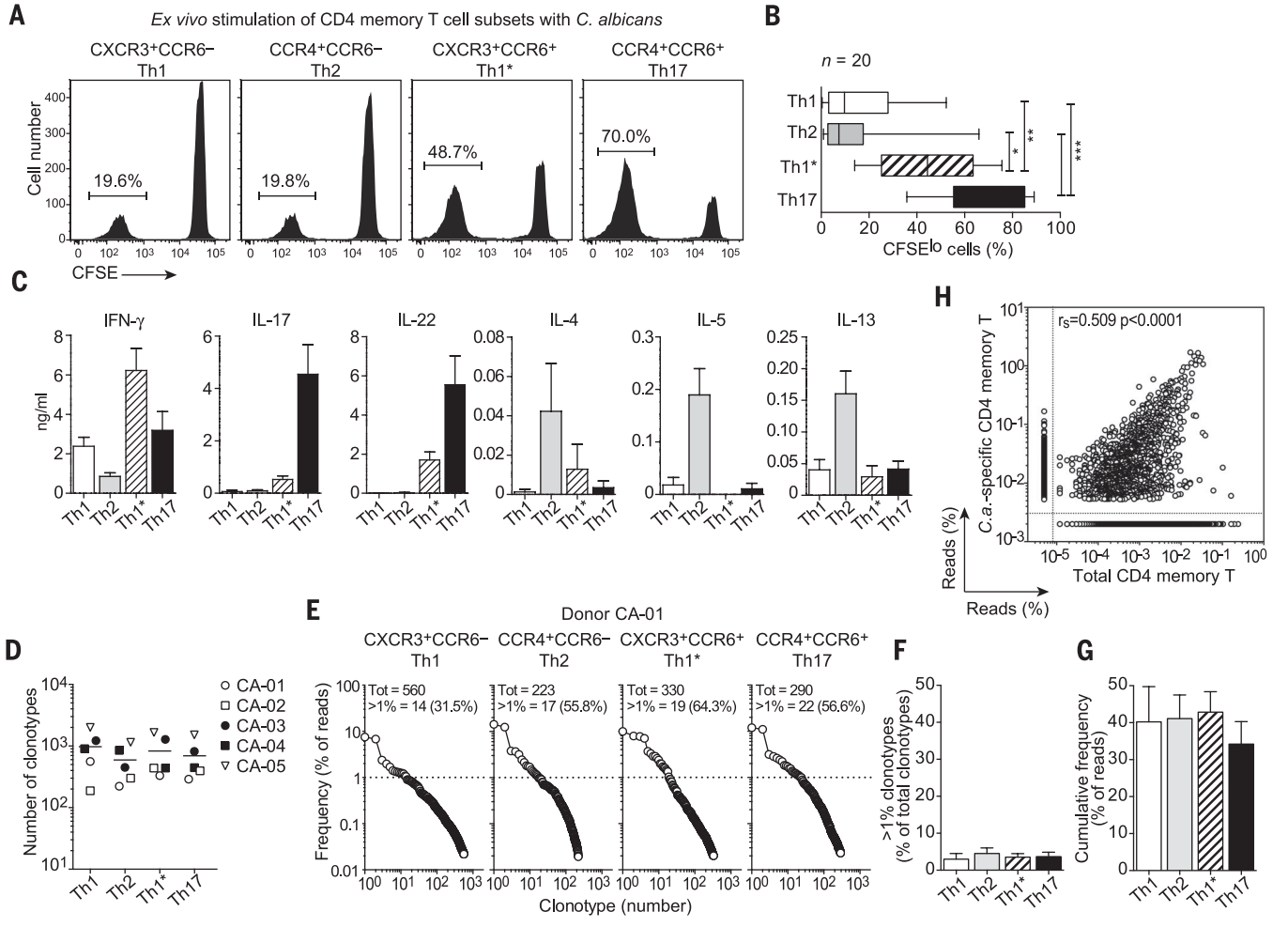


Fig. 1. Functional heterogeneity and clonal composition of *C. albicans*–specific CD4 T cells. Human memory CD4 T cell subsets were isolated according to the expression of chemokine receptors, labeled with CFSE, and stimulated with heat-killed *C. albicans* in the presence of autologous monocytes. (**A**) CFSE profiles on day 6 and percentage of $CFSE^{lo}$ proliferating cells in a representative donor. (**B**) Median values with 25th and 75th percentiles of $CFSE^{lo}$ cells in 20 donors, with whiskers representing the highest and lowest values. ***$P < 0.0001$, **$P < 0.001$, *$P < 0.05$, as determined by nonparametric Friedman test. (**C**) Cytokine production measured on day 6 in the supernatants of the *C. albicans*–stimulated cultures. Values represent means ± SEM ($n = 5$). For comparison, the amounts of cytokines produced by polyclonally stimulated naïve T cells are reported in (*39*). (**D**) Number of unique TCRβ clonotypes detected by deep sequencing in *C. albicans*–specific T cells isolated from the four memory subsets. Each symbol represents a different donor. Lines represent mean values. (**E**) Frequency distribution of clonotypes in *C. albicans*–specific T cells from the four subsets isolated from donor CA-01. Dotted line represents the 1% frequency threshold. Numbers indicate total clonotypes in the subset and clonotypes present at frequencies >1% with their cumulative frequency (percentage of reads). Disparity index was close to 1 in all subsets. (**F** and **G**) Percent of clonotypes with frequencies >1% (F) and their cumulative frequencies (percentage of reads) (G) in five donors (means ± SEM). (**H**) Comparison of clonotype frequency distribution of *C. albicans*–specific memory CD4 T cells (*y* axis) and total memory CD4 T cells sequenced directly ex vivo after sorting (*x* axis). Dots outside the dotted lines represent clonotypes that were found in only one subset and that were assigned an arbitrary frequency value for graphical purposes. Shown is the Spearman correlation and the significance analyzed by paired *t* test.

memory subsets, we analyzed memory T cells specific for a different pathogen. Three CCR6$^+$ T cell subsets were sorted from PBMCs of healthy donors (fig. S4) (*27*). *M. tuberculosis*–specific T cells were found in high numbers in the T_H1^* subset (CXCR3$^+$CCR4$^-$) and at lower numbers in the T_H17 subset (CCR4$^+$CXCR3$^-$) (*28*), as well as in a subset of CXCR3$^+$CCR4$^+$ T cells, which, similarly to T_H1^*, produced IFN-γ and low amounts of IL-17 (fig. S5, A and B). As observed in the response to *C. albicans*, despite the difference in proliferating cells (*27*), the TCRβ diversity of *M. tuberculosis*–specific T cells was comparable in the three subsets, with an average of 152, 158, and 142 clonotypes in T_H1^*, T_H17, and CXCR3$^+$CCR4$^+$ T cell subsets, respectively (fig. S5C). However, in contrast to what we observed for *C. albicans*, very few clonotypes were shared between *M. tuberculosis*–specific T_H1^* and T_H17, whereas several were shared between T_H1^* and CXCR3$^+$CCR4$^+$ T cells (fig. S5, C and D). These findings indicate that the main patterns of clonotype sharing between different T_H subsets are characteristic for different antigens and suggest that T_H1^* cells may be generated indirectly from plastic T_H17 cells as well as directly without transition through a T_H17 stage.

Extensive clonotype sharing among TT-specific memory T cell subsets

The heterogeneity of memory T cells observed for *C. albicans* and *M. tuberculosis* could be attributed to the multiplicity of antigens and innate receptor ligands present in these microbes. To address whether a single protein antigen would induce a less diverse and more polarized response, we analyzed memory T cells primed by the tetanus toxoid (TT) vaccine, which is given with alum as adjuvant. PBMC-sorted, CFSE-labeled T_H1, T_H2, T_H1^*, and T_H17 cells were stimulated with TT in the presence of autologous monocytes (*29*). Among the four subsets, the extent of cell proliferation was comparable but the pattern of cytokine production was distinct (Fig. 3, A to C). For example, TT-specific T_H1 cells produced IFN-γ, as expected, but also low levels of IL-4, IL-5, and IL-13.

TCRβ sequencing revealed a high and comparable number of clonotypes in the four subsets, with a broad distribution of frequencies and a few dominant clonotypes accounting for a large fraction of total reads (Fig. 3, D to F). Surprisingly, there was a high level of clonotype sharing among all four TT-specific subsets (Fig. 3, G and H), with several clonotypes being present in three and even all four subsets (fig. S6). These results indicate that, in contrast to our expectation, the response to TT is characterized by a higher TCR diversity and clonotype sharing as compared to the response to complex microbes.

Isolation and characterization of sister clones from different memory subsets

To provide further evidence that memory T cells expressing the same αβ TCR can acquire in vivo different phenotypes and functional properties, we characterized T cell clones isolated from the CFSE^{lo} *C. albicans*–specific cells that were frozen as backup. From donor CA-01 we isolated 242 *C. albicans*–specific T cell clones, of which 80 were sister clones found in more than one subset (Fig. 4A). Consistent with the deep sequencing data, sister clones were frequently isolated from the T_H17 and T_H1^* subsets and from the T_H17 and T_H2 subsets (Fig. 4B). Sequencing of the TCRα chain confirmed that the sister clones isolated from different subsets carried the same αβ TCR (Table 1).

The sister clones isolated were further characterized phenotypically and functionally. With the exception of IFN-γ, which was produced by virtually all clones as a consequence of the in vitro expansion, cytokine production was consistent with the origin of the cells. IL-4 was produced at high levels by sister clones isolated from the T_H2 subset, whereas IL-17 and IL-22 were produced by sister clones isolated from the T_H17 and T_H1^* subsets (Fig. 4C). In addition, CCR6 expression and RORγt mRNA were detected only in sister clones derived from T_H17 and T_H1^* (Fig. 4D). As examples, sister clones of clonotype 11, isolated from the T_H17 and T_H1^* subsets, and of clonotype 64, isolated from the T_H17 and T_H2 subsets, showed the characteristic polarized cytokine production and expression of transcription factors (Fig. 4, E and F). Sister clones with different functional properties (i.e., T_H1 and T_H2, or T_H2 and T_H17) were also isolated from different subsets of TT-specific memory cells.

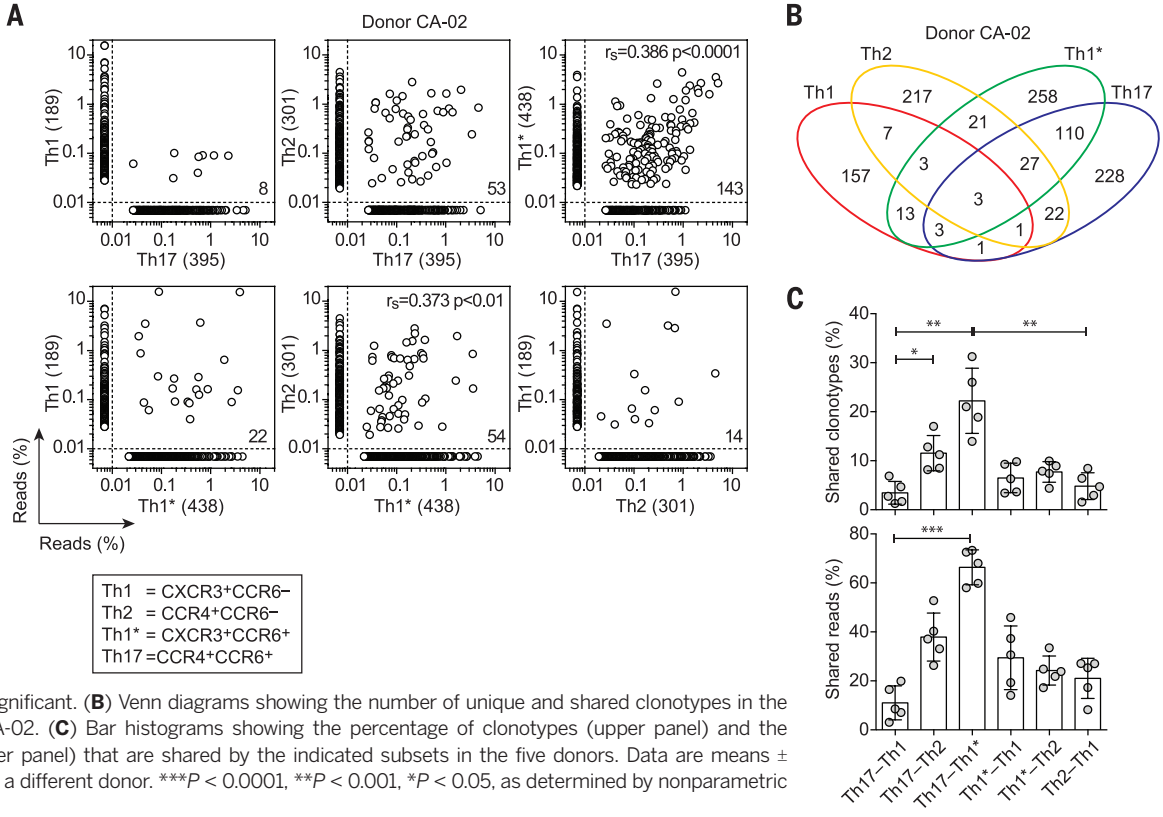


Fig. 2. Unique and shared clonotypes among *C. albicans*–specific memory T cell subsets. (**A**) Comparison of clonotype frequency distribution in samples of T cells isolated from *C. albicans*–stimulated T_H1, T_H2, T_H1^*, or T_H17 subsets from donor CA-02. Frequencies are shown as percentage of total reads. The total number of clonotypes in each sample is indicated on *x* and *y* axis. Values in the lower right corner represent the number of shared clonotypes between the two samples. The Spearman correlation and paired *t* test values are shown when significant. (**B**) Venn diagrams showing the number of unique and shared clonotypes in the four subsets of donor CA-02. (**C**) Bar histograms showing the percentage of clonotypes (upper panel) and the percentage of reads (lower panel) that are shared by the indicated subsets in the five donors. Data are means ± SEM. Each dot represents a different donor. ***$P < 0.0001$, **$P < 0.001$, *$P < 0.05$, as determined by nonparametric Friedman test.

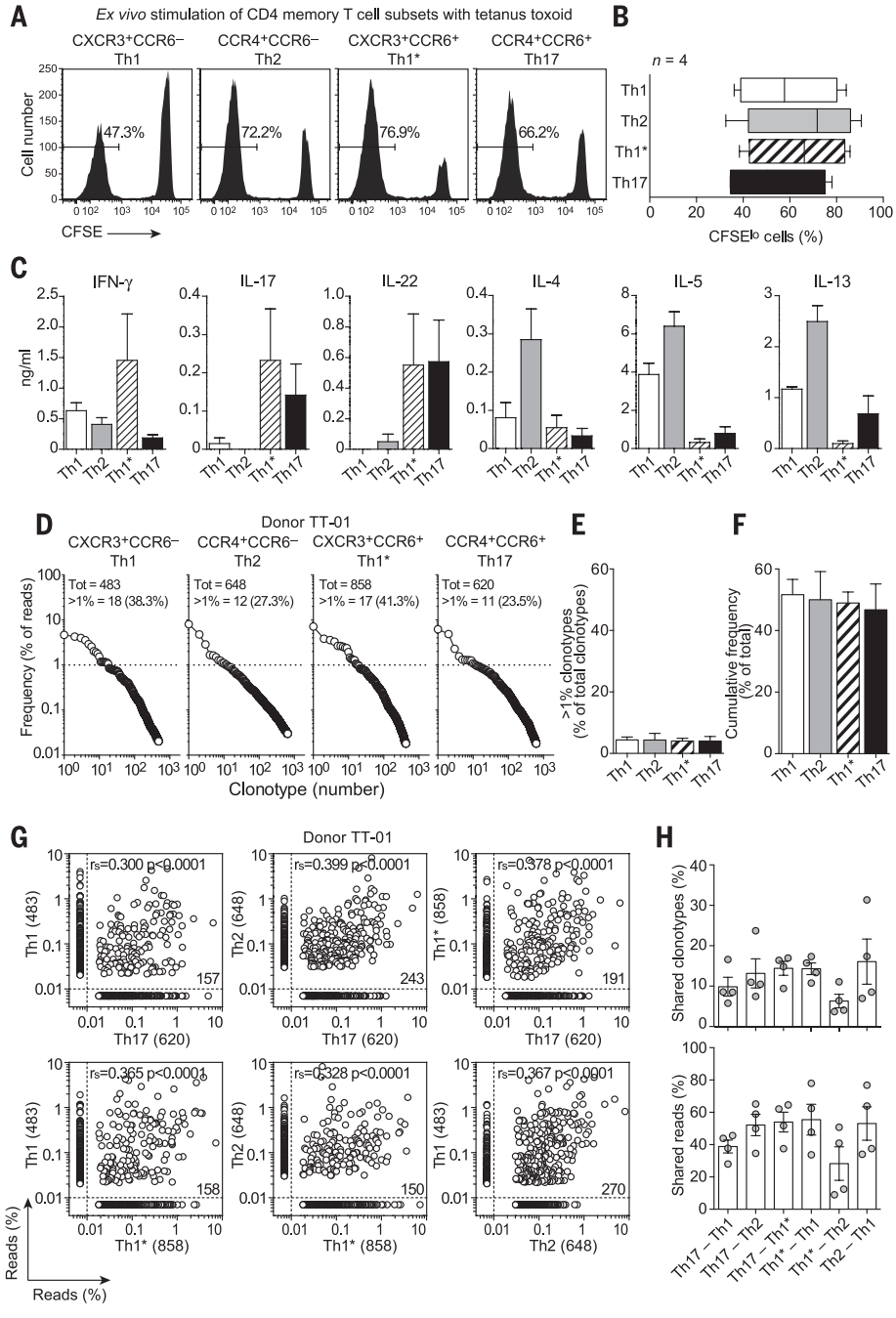


Fig. 3. Extensive clonotype sharing among TT-specific CD4 T cell subsets. (**A**) CFSE profiles and percentage of $CFSE^{lo}$ proliferating cells in memory T cell subsets stimulated with TT in one representative donor. (**B**) Median values with 25th and 75th percentiles in four donors; whiskers represent the highest and lowest values. (**C**) Cytokine production in day 6 supernatants of TT-stimulated cultures. Values represent means ± SEM (n = 4). (**D**) Frequency distribution of clonotypes in TT-specific T cells from the four subsets isolated from donor TT-01. Dotted line represents the 1% frequency threshold. Numbers indicate total clonotypes in the subset and clonotypes present at frequencies >1% with their cumulative frequency (percentage of reads). Disparity index was close to 1 in all subsets. (**E** and **F**) Percent of clonotypes with frequencies >1% (E) and their cumulative frequencies (percentage of reads) (F) in four donors (mean ± SEM). (**G**) Comparison of clonotype frequency distribution in samples of T cells isolated from TT-stimulated T_H1, T_H2, T_H1*, or T_H17 subsets from donor TT-01. Frequencies are shown as percentage of total reads. The total number of clonotypes in each sample is indicated on the x and y axis. Shown are the number of clonotypes shared between the two samples and the Spearman correlation and paired t test value. (**H**) Bar histograms showing for four donors the percentage of clonotypes (upper panel) and percentage of reads (lower panel) that are shared by the indicated subsets. Data are means ± SEM. Each dot represents a different donor.

Generation of multiple fates by in vitro priming of naïve CD4 T cells

The above analysis shows that memory CD4 T cells primed in vivo by pathogens or vaccines are highly heterogeneous, both at the population level and at the clonal level. Given the generation of multiple naïve T cells with identical TCRs in the thymus and their possible expansion in the periphery, one cannot rule out the possibility that some of the heterogeneity observed at the clonal level may reflect spatially and temporally distinct priming events of sister naïve T cells. Furthermore, heterogeneous fates may derive from the differential secondary stimulation of memory T cell clones. To test whether one round of stimulation could imprint heterogeneous fates within the progeny of a single naïve T cell, we primed in vitro a relatively small number of highly purified naïve CD4 T cells with *C. albicans* in the presence of autologous monocytes (*30*). As shown in Fig. 5 for one representative donor, proliferating $CFSE^{lo}$ T cells recovered on day 15 produced IFN-γ, IL-17, IL-22, and IL-4 in various combinations (Fig. 5A). Using the cytokine secretion assay, we sorted IL-17^+ (IFN-γ^-, IL-4^-), IFN-γ^+ (IL-17^-, IL-4^-), and IL-4^+ (IL-17^-, IFN-γ^-) T cells (*31*), which were directly cloned by limiting dilution and further expanded in bulk cultures with IL-2 for clonotypic analysis (Fig. 5B). As expected, the expanded populations produced primarily the original cytokine (i.e., IL-17, IFN-γ, or IL-4) (fig. S7). However, some IFN-γ^+ and some IL-17^+ sorted cells acquired the capacity to produce IL-4, consistent with flexibility in cytokine gene expression. TCRβ deep sequencing performed on the expanded cultures revealed a strikingly high level of clonotype sharing among the three cytokine-secreting cell populations (Fig. 5C). High frequencies of shared clonotypes and shared reads were found in three independent experiments performed with naïve T cells from different donors (Fig. 5D).

To further corroborate the deep sequencing data, we analyzed 205 T cell clones isolated from the different cytokine-secreting populations from one donor. Among the 88 unique TCRβ clonotypes identified, 21 were found in clones isolated from two and even three cytokine-secreting populations (table S2). In particular, six clonotypes (*8*, *32*, *37*, *41*, *43*, and *59*) were found in clones isolated from IL-17–, IFN-γ–, and IL-4–secreting cells. Collectively, these findings provide evidence that a single human naïve CD4 T cell can generate a heterogeneous progeny, even in a single round of antigenic stimulation.

Discussion

By combining antigenic stimulation with deep sequencing and isolation of T cell clones, we unraveled an unpredicted degree of heterogeneity in the human CD4 T cell response to microbes and vaccines. This heterogeneity is the result of two distinct phenomena: (i) interclonal heterogeneity (i.e., different clones that give rise to a homogeneous polarized progeny), and (ii) intraclonal heterogeneity (i.e., clones that give rise

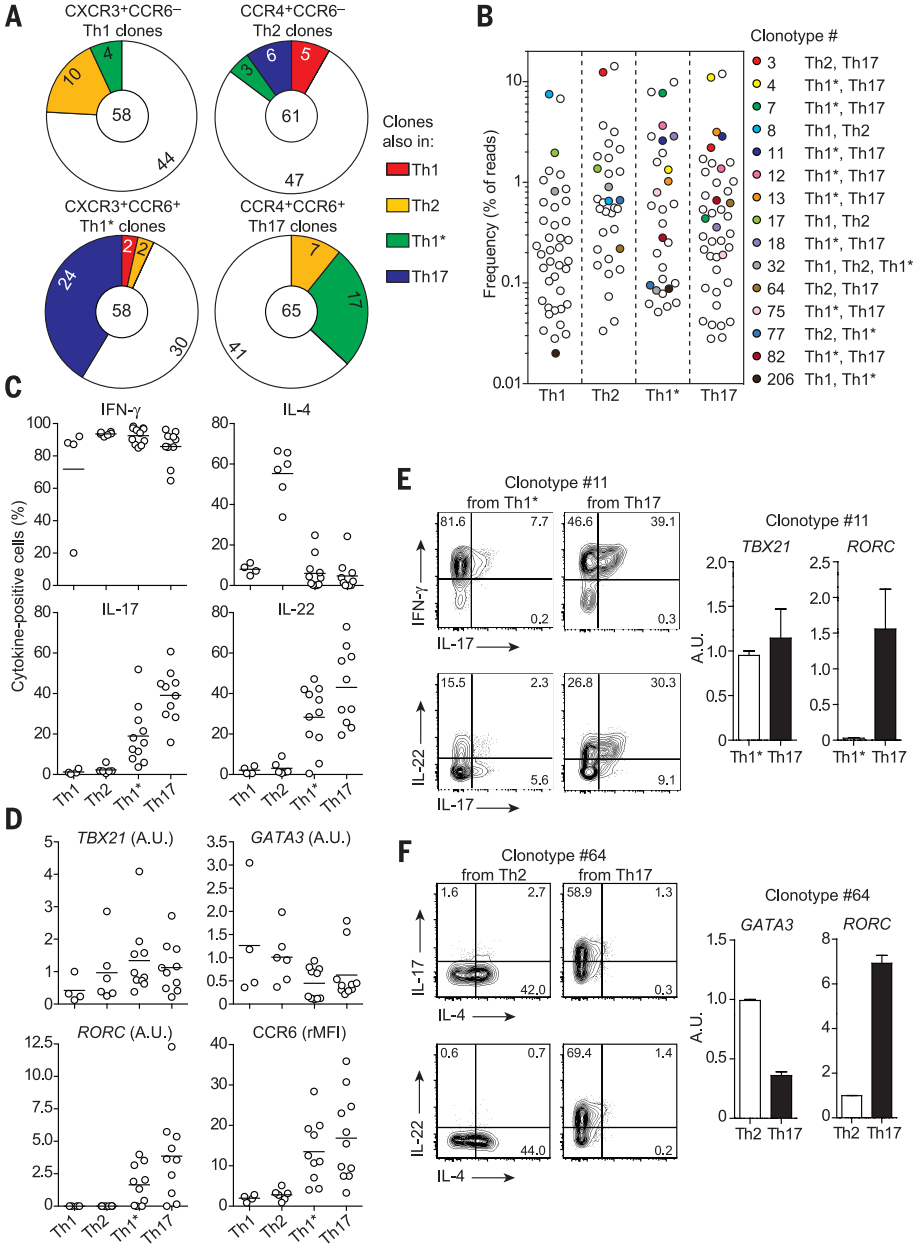


Fig. 4. Isolation of sister clones from different memory subsets. T cell clones were isolated from cryopreserved samples of *C. albicans*–stimulated CFSE^{lo} T_H1, T_H2, T_H1*, and T_H17 cells from donor CA-01. TCRβ sequences were obtained by Sanger sequencing on amplified cDNA. (**A**) For each subset the plot indicates the total number of clones isolated (center), the number of clones unique to that subset (white sector), and the number of clones whose TCRβ was found also in a sister clone isolated from another subset (color-coded sectors). (**B**) Frequency distribution (based on deep sequencing data) of the clonotypes corresponding to the clones that were found in a single subset (white circles) and in multiple subsets (colored circles). (**C** and **D**) Characterization of sister clones isolated from different memory subsets. Shown are cytokine production (C) and expression of transcription factors and CCR6 (D). (**E** and **F**) Characterization of two pairs of sister clones isolated from T_H1* and T_H17 subsets (E) and from T_H2 and T_H17 subsets (F).

to daughter cells that adopt different fates). The high level of intraclonal heterogeneity observed in this study supports the "one cell, multiple fates" model of CD4 T cell differentiation and provides evidence for T cell plasticity in the context of the human immune response. Possible mechanisms that lead to interclonal and intraclonal heterogeneity can involve the spectrum of pathogen- and damage-associated molecular patterns associated with a given pathogen or vaccine, the nature of the antigen-presenting cells in different tissues, the strength of TCR stimulation and costimulation, and the stochastic stimulation and restimulation of individual T cells within a proliferating clone (*10*, *32*–*34*). Another mechanism, which may be particularly relevant for complex pathogens (*35*), involves the stimulation of T cells by cross-reactive antigens that trigger heterologous immunity (*36*, *37*).

Intraclonal heterogeneity was observed in all responses analyzed, but the pattern and extent of clonotype sharing appeared to be characteristic for each antigen. In the response to *C. albicans*, the extensive clonotype sharing between T_H17 and T_H1* could be explained by the plasticity of T_H17 cells that have been shown to acquire IFN-γ production in appropriate stimulatory conditions (*12*–*14*). However, clonotype sharing was not observed between *M. tuberculosis*–specific T_H17 and T_H1* cells of healthy donors—a finding that would be consistent with different pathways leading to the generation of T_H1*. In contrast, there are no precedents to explain the substantial degree of clonotype sharing between *C. albicans*–specific T_H17 and T_H2 subsets. However, a similar pattern of intraclonal diversification was found in clones of naïve T cells primed in vitro by *C. albicans*, indicating that multiple fates can be rapidly induced within a single round of stimulation through mechanisms that remain to be defined. Finally, it should be noted that the clonal differentiation along multiple fates does not apply to all clones. Several expanded clones were found uniquely in a single T cell subset.

The extensive clonotype sharing among all TT-specific memory subsets was surprising and may be related to repeated exposure to the vaccine in booster immunizations. In all cases, this finding supports the notion that intraclonal heterogeneity can be generated even in response to a single protein antigen and under defined, albeit non-physiological, stimulatory conditions. We suggest that in this case intraclonal heterogeneity may result from the initial priming of nonpolarized memory T cells that are subsequently driven to differentiate toward distinct fates by homeostatic or tissue-specific mechanisms, a phenomenon that has been shown to occur in human T_{CM} cells in vitro (*38*). All considered, the extent of intraclonal heterogeneity found ex vivo and after in vitro priming suggest that intraclonal diversification is a fundamental property of the CD4 T cell response.

The finding that memory T cells are present in higher numbers in particular subsets associated with a protective response (for instance, in T_H17 for *C. albicans* and in T_H1* for *M. tuberculosis*) is at odds with our finding that the diversity of antigen-specific clonotypes was comparable in all subsets analyzed. This observation would be compatible with a model whereby polarized T cell responses result from the selective expansion, rather than priming, of T cell clones stimulated by the pathogen in a particular microenvironment. Accordingly, a broad priming of effector T cells endowed with different functions and migratory capacities would ensure a range of differentiated precursors to be recruited and expanded where necessary.

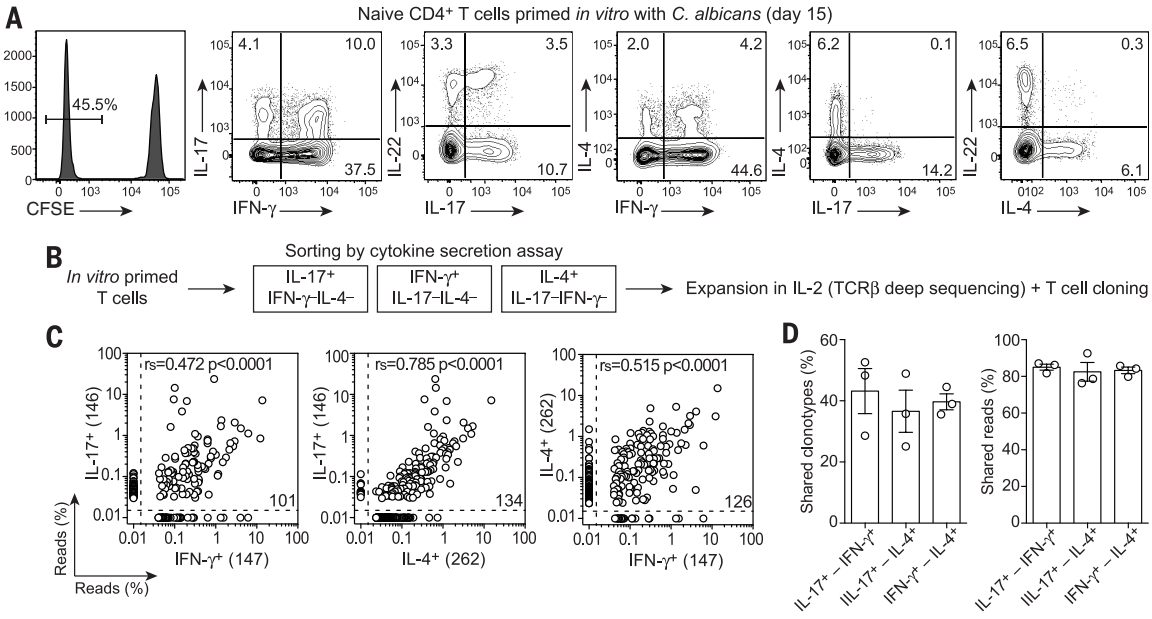


Fig. 5. Generation of multiple fates by in vitro priming of naïve T cells. (**A**) Left: CFSE profile of CFSE-labeled naïve CD4 T cells primed in vitro by *C. albicans* in the presence of autologous monocytes and analyzed on day 15. Right: Intracellular cytokine staining on $CFSE^{lo}$ cells stimulated with phorbol myristate acetate and ionomycin. (**B**) Schematic outline of the experimental approach: After priming in vitro, viable cytokine-secreting cells were enriched using the cytokine secretion assay (Miltenyi). Three populations sorted as IL-17$^+$ (IFN-γ$^-$ IL-4$^-$), IFN-γ$^+$ (IL-17$^-$ IL-4$^-$), IL-4$^+$ (IL-17$^-$ IFN-γ$^-$) were further expanded in IL-2 and used for TCRβ deep sequencing and T cell cloning. (**C**) Frequency distribution of TCRβ clonotypes in the indicated cytokine-secreting T cell populations from donor N-01. Frequencies are shown as percentage of total reads. The total number of clonotypes in each population is indicated in parenthesis on *x* and *y* axis. Shown are the number of clonotypes shared between two samples and the Spearman correlation and paired *t* test values. (**D**) Bar histograms showing the percentage of clonotypes (left) and percentage of reads (right) that are shared between the indicated cytokine-secreting T cell populations. Data are means ± SEM (*n* = 3). Each dot represents a different donor.

Table 1. T cell clones with identical TCR isolated from different memory T cell subsets. Complementarity-determining region 3 (CDR3) sequences were obtained for both TCRα and TCRβ chains. Amino acid abbreviations: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

Clonotype	TCR	CDR3	Memory subset (number of clones)			
			T_H1	T_H2	T_H1*	T_H17
3	Vα	CAVTRMDSSYKLIF	–	4	–	2
	Vβ	CASSYVSQDKNEAFF				
4	Vα	CILRDLNTGTASKLTF	–	–	1	4
	Vβ	CASRAGRYNEQFF				
7	Vα	CATCHSGYSTLTF	–	–	12	1
	Vα	CAGLFYNQGGKLIF				
	Vβ	CASSPGGSGGWNEQFF				
8	Vα	CAVRPISGSARQLTF	6	2	–	–
	Vα	CAVGRQGQAGTALIF				
	Vβ	CASSSTFNGRGTDTQYF				
11	Vα	CATDASSGANSKLTF	–	–	2	3
	Vβ	CASSYHRGAFF				
12	Vα	CAIIQGAQKLVF	–	–	1	1
	Vβ	CASRDRGVGRSNEKLFF				
13	Vα	CATPRGQAGTALIF	–	–	1	1
	Vβ	CASSKAGSSYNEQFF				
17	Vα	CAVRPGGYQKVTF	1	1	–	–
	Vβ	CASSPTDRGAFF				
18	Vα	CVVSEERDDKIIF	–	–	4	1
	Vβ	CSAQQVHEQFF				
32	Vα	CATTGGGNKLTF	3	2	1	–
	Vβ	CSARPSGANLQDTQYF				
64	Vα	CAVTPSGFQKLVF	–	2	–	5
	Vβ	CATSRQTQGVGETQYF				
75	Vα	CAAWGSGDKLTF	–	–	2	1
	Vβ	CASKIGGGETQYF				
77	Vα	CAPSWGKLQF	–	1	1	–
	Vα	CALNTDKLIF				
	Vβ	CASSPEGSGTKNYGYTF				
82	Vα	CAYRGPRDDKIIF	–	–	1	5
	Vβ	CSAQQVHEQYF				
206	Vα	CGADNTGNQFYF	1	–	1	–
	Vβ	CASSLPQGYNEQFF				

REFERENCES AND NOTES

1. R. Medzhitov, *Nature* **449**, 819–826 (2007).
2. F. Sallusto, A. Lanzavecchia, *Eur. J. Immunol.* **39**, 2076–2082 (2009).
3. F. P. Heinzel, M. D. Sadick, B. J. Holaday, R. L. Coffman, R. M. Locksley, *J. Exp. Med.* **169**, 59–72 (1989).
4. M. Yamamura *et al*., *Science* **254**, 277–279 (1991).
5. C. S. Ma *et al*., *J. Exp. Med.* **205**, 1551–1557 (2008).
6. J. D. Milner *et al*., *Nature* **452**, 773–776 (2008).
7. A. Puel *et al*., *Science* **332**, 65–68 (2011).
8. A. Iwasaki, R. Medzhitov, *Science* **327**, 291–295 (2010).
9. J. Zhu, H. Yamane, W. E. Paul, *Annu. Rev. Immunol.* **28**, 445–489 (2010).
10. A. O'Garra, L. Gabryšová, H. Spits, *Nat. Immunol.* **12**, 288–294 (2011).
11. J. J. O'Shea, W. E. Paul, *Science* **327**, 1098–1102 (2010).
12. L. Cosmi *et al*., *J. Exp. Med.* **205**, 1903–1916 (2008).
13. D. Bending *et al*., *J. Clin. Invest.* **119**, 565–572 (2009).
14. K. Hirota *et al*., *Nat. Immunol.* **12**, 255–263 (2011).
15. S. L. Reiner, F. Sallusto, A. Lanzavecchia, *Science* **317**, 622–625 (2007).
16. J. T. Chang *et al*., *Science* **315**, 1687–1691 (2007).
17. C. Gerlach *et al*., *J. Exp. Med.* **207**, 1235–1246 (2010).
18. C. Stemberger *et al*., *Immunity* **27**, 985–997 (2007).
19. N. J. Tubo *et al*., *Cell* **153**, 785–796 (2013).
20. See supplementary materials on *Science* Online.
21. T cell subsets were isolated according to the differential expression of chemokine receptors, as previously described (*22*). Cells in the $CXCR3^+CCR4^-CCR6^-$ subset (defined as T_H1) produced IFN-γ and expressed T-bet mRNA; cells in the $CCR4^+CXCR3^-CCR6^-$ subset (T_H2) produced IL-4 and expressed high level of GATA3 mRNA; cells in the $CCR4^+CCR6^+CXCR3^-$ subset (T_H17) produced IL-17A and expressed RORγt; cells in the $CXCR3^+CCR6^+CCR4^-$ (nonconventional T_H1, defined as T_H1*) produced IFN-γ and low levels of IL-17A and expressed both T-bet and RORγt mRNA.
22. E. V. Acosta-Rodriguez *et al*., *Nat. Immunol.* **8**, 639–646 (2007).
23. In all experiments, the specificity of proliferating memory T cells was assessed by sorting $CFSE^{lo}$ cells and isolating T cell clones by limiting dilution. On average, more than 95% of the T cell clones responded to the antigen used in the initial stimulation.
24. C. E. Zielinski *et al*., *Nature* **484**, 514–518 (2012).
25. P. Bacher *et al*., *J. Immunol.* **190**, 3967–3976 (2013).
26. H. S. Robins *et al*., *Blood* **114**, 4099–4107 (2009).
27. T cell subsets from two healthy donors were isolated from PBMCs, labeled with CFSE, and stimulated with *M. tuberculosis*.

For donor MT-01, input cell numbers were T_H1^*, 1.5×10^6; T_H17, 1.5×10^6; $CXCR3^+CCR4^+$ T_H, 1.2×10^6. For donor MT-02, input cell numbers were T_H1^*, 1.5×10^6; T_H17, 1.4×10^6; $CXCR3^+CCR4^+$ T_H, 1.1×10^6. Recovered $CFSE^{lo}$ cells on day 6 for donor MT-01 were T_H1^*, 2.2×10^6; T_H17, 0.95×10^6; $CXCR3^+CCR4^+$ T_H, 2.0×10^6. For donor MT-02, these numbers were T_H1^*, 2.4×10^6; T_H17, 0.7×10^6; $CXCR3^+CCR4^+$ T_H, 1.7×10^6.

28. C. S. Lindestam Arlehamn *et al.*, *PLOS Pathog.* **9**, e1003130 (2013).
29. T cell subsets from four donors were isolated and stimulated with TT. Input cell number for each subset was 2.5×10^6 (donor TT-01), 1.5×10^6 (donors TT-02 and TT-03), and 1.1×10^6 (donor TT-04).
30. Naïve T cells were isolated from three donors. Input cell numbers were 3.2×10^6 (donor N-01), 6.6×10^6 (donor N-02), and 8.8×10^6 (donor N-03).
31. In control experiments, the purity of the sorted populations was confirmed by measuring *IFNG*, *IL4*, and *IL17* mRNAs by quantitative PCR immediately after sorting.
32. B. Pulendran, *J. Immunol.* **174**, 2457–2465 (2005).
33. H. Sigmundsdottir, E. C. Butcher, *Nat. Immunol.* **9**, 981–987 (2008).
34. C. R. Plumlee, B. S. Sheridan, B. B. Cicek, L. Lefrançois, *Immunity* **39**, 347–356 (2013).
35. M. E. Birnbaum *et al.*, *Cell* **157**, 1073–1087 (2014).
36. R. M. Welsh, L. K. Selin, *Nat. Rev. Immunol.* **2**, 417–426 (2002).
37. D. Mason, *Immunol. Today* **19**, 395–404 (1998).
38. L. Rivino *et al.*, *J. Exp. Med.* **200**, 725–735 (2004).
39. Production of cytokines was measured on day 6 in the supernatants of 5×10^4 naïve CD4 T cells stimulated with plate-bound CD3 and CD28 antibodies. Mean values and SEM ($n = 3$) were as follows: IFN-γ, 414 ± 270.2; IL-17, 30.7 ± 3.3; IL-22, 391.6 ± 129.2; IL-4, 3.6 ± 3.6; IL-5, 9.6 ± 5.4; IL-13, 91 ± 20.5.

ACKNOWLEDGMENTS

We thank D. Jarrossay for cell sorting; O. Petrini, C. Fragoso, and A. Sette for providing valuable reagents; the Adaptive Biotechnologies team for technical support; and S. Monticelli for critical reading of the manuscript. The data presented in this manuscript are tabulated in the main paper and in the supplementary materials. All TCR sequences are available in the supplementary materials as .txt files. FASTA format files are available at downloads.adaptivebiotech.com (username: SallustoSciDec2014data; password: di2PhaiT). The sequences have been deposited in the Gene Expression Omnibus (accession numbers xxxxxx). Supported by European Research Council grant 323183 PREDICT (F.S.), Swiss National Science Foundation grants 149475 (F.S.) and 147662 (A.L.), and European Commission grants FP7-HEALTH-2011-280873, ADITEC (A.L. and F.S.) and FP7-HEALTH-2013-601958, SUPERSIST (T.N.S.). The Institute for Research in Biomedicine is supported by the Helmut Horten Foundation.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/400/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 and S2
References (*40*–*44*)
Data files

1 September 2014; accepted 25 November 2014
Published online 4 December 2014;
10.1126/science.1260668

REPORTS

ASTROPHYSICS

The exceptionally powerful TeV γ-ray emitters in the Large Magellanic Cloud

The H.E.S.S. Collaboration*†

The Large Magellanic Cloud, a satellite galaxy of the Milky Way, has been observed with the High Energy Stereoscopic System (H.E.S.S.) above an energy of 100 billion electron volts for a deep exposure of 210 hours. Three sources of different types were detected: the pulsar wind nebula of the most energetic pulsar known, N 157B; the radio-loud supernova remnant N 132D; and the largest nonthermal x-ray shell, the superbubble 30 Dor C. The unique object SN 1987A is, unexpectedly, not detected, which constrains the theoretical framework of particle acceleration in very young supernova remnants. These detections reveal the most energetic tip of a γ-ray source population in an external galaxy and provide via 30 Dor C the unambiguous detection of γ-ray emission from a superbubble.

The origin of cosmic rays (CRs), the very high energy (VHE) ($\gtrsim 10^{11}$ eV) and ultrahigh energy ($\gtrsim 10^{18}$ eV) particles that bombard Earth, has puzzled us for more than a century. Much progress has been made during the past decade due to the advent of VHE γ-ray telescopes. These telescopes detect ~10^{11} to 10^{14} eV γ rays from atomic nuclei (hadronic CRs) collisions with local gas or from ultrarelativistic electrons/positrons (leptonic CRs), which produce γ-ray emission by upscattering low-energy background photons (*1*). Indeed, a survey of the inner part of the Milky Way with the High Energy Stereoscopic System (H.E.S.S.), an array of imaging atmospheric Cherenkov telescopes (*2*), revealed a population of supernova remnants (SNRs) and pulsar wind nebulae (PWNe) emitting γ rays with energies in excess of 100 GeV (*3*).

Here, we report on VHE γ-ray sources detected outside the Milky Way, namely in the Large Magellanic Cloud (LMC). This satellite galaxy of the Milky Way has a stellar mass of about 4% of the Milky Way (*4*, *5*). Located at a distance of ~50 kpc (*6*), it is an irregular galaxy seen almost face-on (*7*). Consequently, source confusion is much less of a problem than for the inner Milky Way, and there is less uncertainty in the distances of the sources. The LMC stands out among nearby galaxies for its high star formation rate per unit mass, which is about a factor of five higher than in the Milky Way (*8*, *9*), and contains the best example of a local starburst, the Tarantula Nebula. The LMC also harbors numerous massive stellar clusters and SNRs. Among the SNRs is a unique source, SN 1987A, the remnant of the nearest supernova observed in modern times (*10*).

High-energy γ-ray emission from the LMC was detected by the Energetic Gamma Ray Experiment Telescope (EGRET) (*11*) and, more recently, by the Fermi Large Area Telescope (LAT) (*12*), which revealed diffuse emission with an extension of several degrees in diameter, tracing massive star-forming regions. VHE γ-ray telescopes, like H.E.S.S., besides providing information on much higher energy CRs, have an angular resolution of a few arcminutes, which is substantially better than Fermi-LAT's resolution at γ-ray energies <10 GeV. The good angular resolution allows H.E.S.S. to identify individual sources in the LMC. As we will detail below, a deep H.E.S.S. observation revealed three luminous sources in the LMC: the superbubble 30 Dor C, the energetic PWN N 157B, and the radio-loud SNR N 132D. Of these sources, only N 157B was detected previously in a 47-hour exposure (*13*). The observations extend the scope of VHE γ-ray astronomy by providing examples of sources from a population outside the Milky Way. N 157B and N 132D belong to known γ-ray source classes, but both have distinguishing characteristics, N 157B being powered by the most energetic young pulsar, and N 132D being one of the oldest VHE γ-ray emitting SNRs. The superbubble 30 Dor C, however, provides an unambiguous detection of a superbubble in VHE γ rays. Conspicuously absent from our list of three sources is SN 1987A, despite predictions that it should be a bright γ-ray source (*14*, *15*).

H.E.S.S. observations

We report on a deep, 210-hour H.E.S.S. exposure, targeted at the region of the Tarantula nebula—corresponding to 30 Doradus (30 Dor)—the largest star-forming region in the Local Group of galaxies. We reconstructed γ-ray showers with an image-fitting analysis (*16*) and cross-checked with a multivariate analysis based on image parameterization (*17*, *18*), with consistent results. In both analyses, a cut on the uncertainty of the reconstructed γ-ray direction indicated an angular resolution of ~0.05°.

Figure 1A shows an optical image of the LMC overlaid with TeV γ-ray point-source significance contours. In this data set, 613 γ rays with a statistical significance of 33 SD are detected from the PWN N 157B. Figure 1B provides a close-up view of the γ-ray emission from N 157B. The diameter

*The full list of authors and affiliations is at the end of the paper.
†Corresponding author. E-mail: nukri.komin@wits.ac.za (N.K.); chia-chun.lu@mpi-hd.mpg.de (C.-C.L.); michael.mayer@physik.hu-berlin.de (M.M.); stefan.ohm@desy.de (S.O.); j.vink@uva.nl (J.V.)

of N 157B of 100″ (*19*) is of the order of the H.E.S.S. angular resolution. Further significant γ-ray emission is detected to the southwest of N 157B.

A likelihood fit of a model of two γ-ray sources to the on-source and background sky maps establishes the detection of a second source at an angular distance of 9′ (corresponding to 130 pc at a distance of 50 kpc) from N 157B. The model consisting of two sources is preferred by 8.8 SD over the model of one single source. Figure 1C shows an x-ray image with overlaid contours of confidence of the source position. The position of the second source [right ascension = $5^h35^m(55 \pm 5)^s$, declination = −69°11′(10 ± 20)″, equinox J2000, 1 SD errors] coincides with the superbubble 30 Dor C, the first such source detected in VHE γ rays, and thus represents an additional source class in this energy regime. A γ-ray signal around the energetic pulsar PSR J0540-6919 is not detected, despite the presence of an x-ray luminous PWN (*20*). A flux upper limit (99% confidence level) is derived at F_γ (>1 TeV) < 4.8×10^{-14} ph cm^{-2} s^{-1}.

Along with the clear detection of N 157B and 30 Dor C, evidence for VHE γ-ray emission is observed from the prominent SNR N 132D (Fig. 1D). The emission peaks at a significance of about 5 SD above a background that is estimated from a ring around each sky bin. At the nominal position of the SNR, 43 γ rays with a statistical significance of 4.7 SD are recorded.

The γ-ray spectra of all three objects are well described by a power law in energy, $\Phi(E) = d^3N/(dE\,dt\,dA) = \Phi_0\,(E/1\text{ TeV})^{-\Gamma}$ (where E is energy, t is time, and A is detector area) (Fig. 2). The best-fit spectral indices and integral γ-ray luminosities are summarized in Table 1.

Even with a deep exposure of 210 hours, significant emission from SN 1987A is not detected, and we derive an upper limit on the integral γ-ray flux of F_γ(>1TeV) < 5.6×10^{-14} ph cm^{-2} s^{-1} at a 99% confidence level.

Discussion of individual sources

The three VHE emitters belong to different source classes, and their energy output exceeds or at least

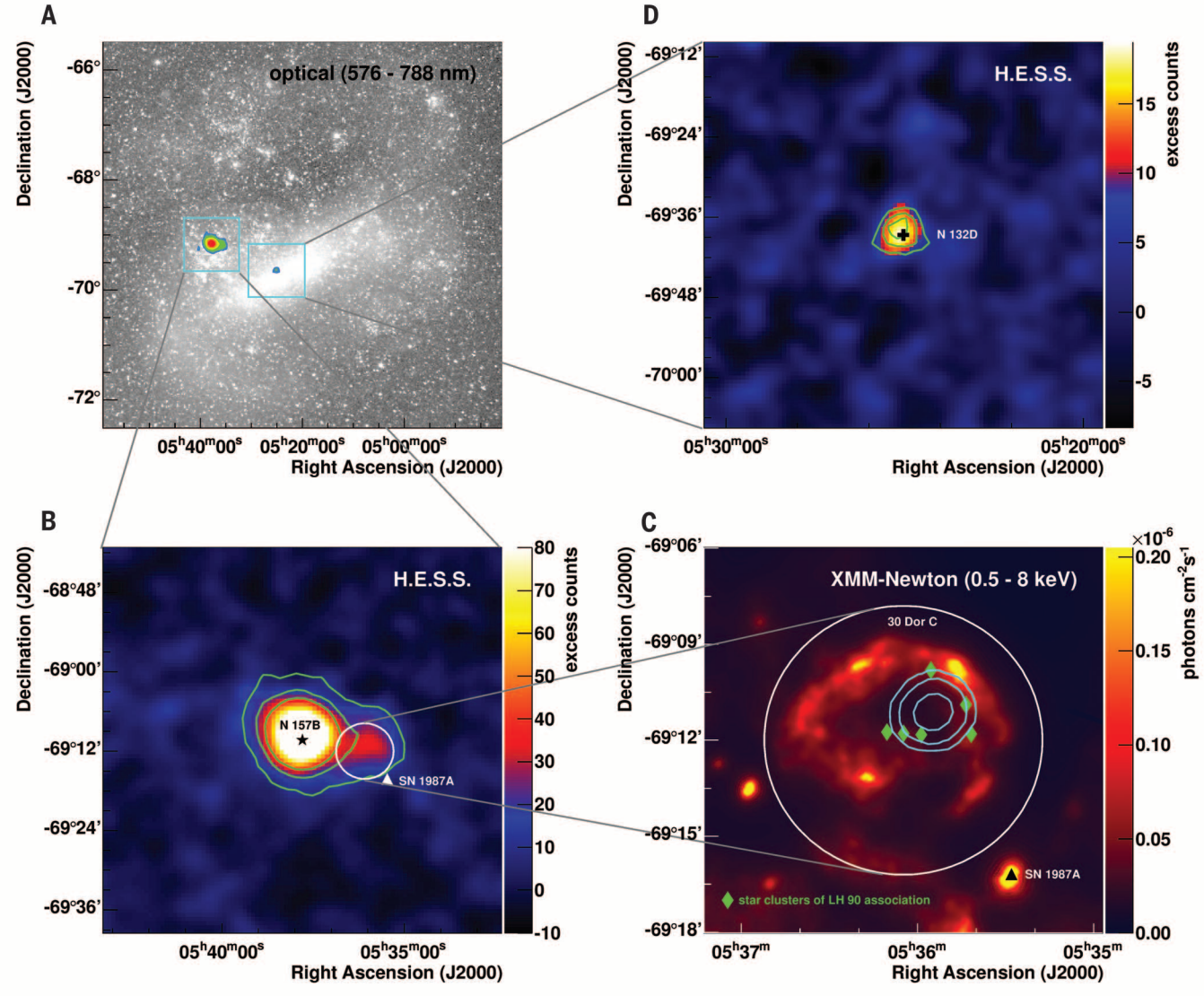


Fig. 1. Sky maps of the LMC. (**A**) Optical image of the entire LMC (*55*). The boxes denote the regions of interest discussed in this paper. Colors denote levels of 3, 5, 10, and 20 SD statistical significance of the γ-ray signal. (**B**) VHE γ-ray emission in the region around N 157B. The green lines represent contours of 5, 10, and 15 SD statistical significance of the γ-ray signal. (**C**) X-Ray Multi-Mirror Mission (XMM-Newton) x-ray flux image of the region of 30 Dor C. The superimposed cyan lines represent contours of 68, 95, and 99% confidence level of the position of the γ-ray source. Diamonds denote the positions of the star clusters of the LH 90 association. See the supplementary materials for details on the x-ray analysis. (**D**) VHE γ-ray emission in the region around N 132D. The green lines represent contours of 3, 4, and 5 SD statistical significance. The background of the γ-ray emission [in (B) and (D)] was obtained using the ring background method (*56*). The resulting excess sky map is smoothed to the angular resolution of the instrument.

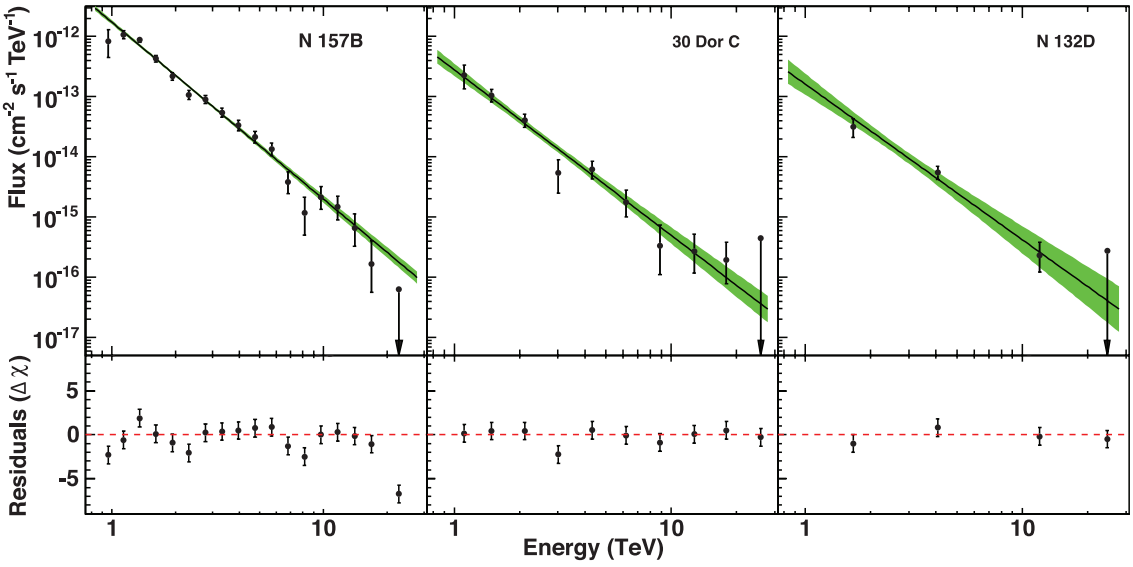


Fig. 2. Gamma-ray spectra of N 157B, 30 Dor C, and N 132D. The spectral points of 30 Dor C are not corrected for spill-over emission from N 157B (see the supplementary materials). Data points have ±1-SD error bars; upper limits are at the 99% confidence level. The bottom panels show the residuals of the data points compared with the best-fit model.

Table 1. Statistics and spectral parameters of the three sources. The exposure time is corrected for the acceptance differences due to different offsets from the camera center. The significances of N157B and N132D are statistical significances of the γ-ray emission obtained by using formula 17 of (*58*). The background was estimated from regions with similar offsets from the camera center as the on-source region. The significance of 30 Dor C is the significance by which a two-source morphology (N 157B and 30 Dor C) is preferred over a single-source morphology (N 157B only). The γ-ray count and the flux of 30 Dor C are corrected for spill-over emission from N 157B (see the supplementary materials). Γ is the photon index and Φ (1 TeV) the differential flux at 1 TeV of a power law fit to the energy spectrum. The luminosity L_γ is calculated for an assumed distance of 50 kpc (*6*). The listed errors are statistical, ±1-SD errors. Systematic errors are estimated to be ±0.3 for Γ and ±30% for Φ (1 TeV) (*13*).

Source H.E.S.S. identifier	N 157B J0537-691	30 Dor C J0535-691	N 132D J0525-696
Exposure time (hours)	181	183	148
γ rays	613	74	43
Significance (SD)	33.0	8.8	4.7
Photon index Γ	2.8 ± 0.1	2.6 ± 0.2	2.4 ± 0.3
Φ (1 TeV) (10^{-12} cm^{-2} s^{-1} TeV^{-1})	1.3 ± 0.1	0.16 ± 0.04	0.13 ± 0.05
L_γ (1 – 10 TeV) (10^{35} erg s^{-1})	6.8 ± 0.3	0.9 ± 0.2	0.9 ± 0.2

equals that of their most powerful relatives in the Milky Way.

30 Dor C

The superbubble 30 Dor C stands out in x-rays, as it contains, in the western part, an x-ray synchrotron-emitting shell with a radius of 47 pc, which makes it the largest known x-ray synchrotron shell (*21*–*23*). X-ray synchrotron emission, which indicates the presence of VHE electrons, is usually associated with 100- to 2000-year-old SNRs with radii smaller than 25 pc. In addition, the x-ray synchrotron luminosity of 30 Dor C is 10 times that of the archetypal young SNR SN 1006 (*21*). The 30 Dor C shell also emits radio and optical radiation (*24*) and appears to have been produced by the stellar winds and supernovae in the OB association LH 90 (NGC 2044) (*25*).

The measured H.E.S.S. flux of 30 Dor C corresponds to a 1 – 10 TeV γ-ray luminosity of (0.9 ± 0.2) × 10^{35} erg s^{-1}, with the best-fit position of the γ-ray emission lying in between the six identified subclusters (*25*). The TeV emission can be explained by the production of neutral pions due to collisions of hadronic CRs with the background plasma. Alternatively, the so-called leptonic emission scenario may apply, in which case the TeV emission is the result of Compton upscattering of low-energy photons to γ-ray energies, by the same population of electrons that is responsible for the x-ray synchrotron radiation (*1*).

For the hadronic scenario, a combination of energy in CRs (assumed to be protons) and density of hydrogen atoms, n_H, of W_{pp} = (0.7 – 25) × 10^{52} (n_H/1 cm^{-3})$^{-1}$ erg is required (see S1.3). 30 Dor C probably experienced ~5 supernova explosions (*22*), which likely provided ~5 × 10^{50} erg in CR energy. Hence, the average gas density should be $n_H \gtrsim 20$ cm^{-3}, which is higher than the density estimate of $n_H \approx 0.1 - 0.4$ cm^{-3} based on the x-ray thermal emission in the southwest (*21*, *26*). However, locations of high densities may be present if the x-ray thermal emission comes from smaller radii than the dense outer shell or if cool, dense, clumped gas survived inside the otherwise rarified interior of the bubble (*27*).

This hadronic scenario puts constraints on the CR diffusion coefficient, because the diffusion length scale should be smaller than the radius of the shell: $l_{diff} = \sqrt{2Dt} \lesssim 47$ pc for CRs around 10 TeV. Therefore, D(10 TeV) $\lesssim 3.3 \times 10^{26}(t/10^6$ year$)^{-1}$ cm^2 s^{-1}, which, given an age for the superbubble of a few million years, gives a much smaller diffusion coefficient than the typical Galactic diffusion coefficient of D(10 TeV) $\gtrsim 5 \times 10^{29}$ cm^2 s^{-1} (*28*). This small diffusion coefficient requires magnetic-field amplification combined with turbulent magnetic fields, as hypothesized by (*29*).

X-ray synchrotron emission from 30 Dor C requires large shock velocities, $v_{shock} \gtrsim 3000$ km s^{-1} (*30*). Assuming that this shock originates from an explosion centered at the superbubble, we obtain a rough estimate of the age of the western x-ray shell of $t = 0.4R/v_{shock} \approx 6000$ years, assuming a Sedov expansion model [$R = 2.8 \times 10^8 (Et^2/n_H)^{1/5}$ cm]. Because the OB association is much older, this age most likely refers to a recent supernova explosion, whose remnant evolves in the rarified medium of the superbubble. The Sedov expansion model then gives us a very low estimate for the density of $n_H \approx 5 \times 10^{-4}$ cm^{-3} for an explosion energy of $E = 10^{51}$ erg. Although this is very low, it can occur under certain conditions (*31*). This model for the x-ray synchrotron shell can even be reconciled with the hadronic model of the TeV emission, if the rarified medium also contains dense clumps. For the leptonic scenario for the TeV emission, the broad spectral energy distribution (SED) (Fig. 3) requires an energy in accelerated electrons of ~4 × 10^{48} erg and average magnetic-field strength of 15 μG, low compared to most young SNRs (*32*) but a factor of three to four higher than the average magnetic field in the LMC (*33*).

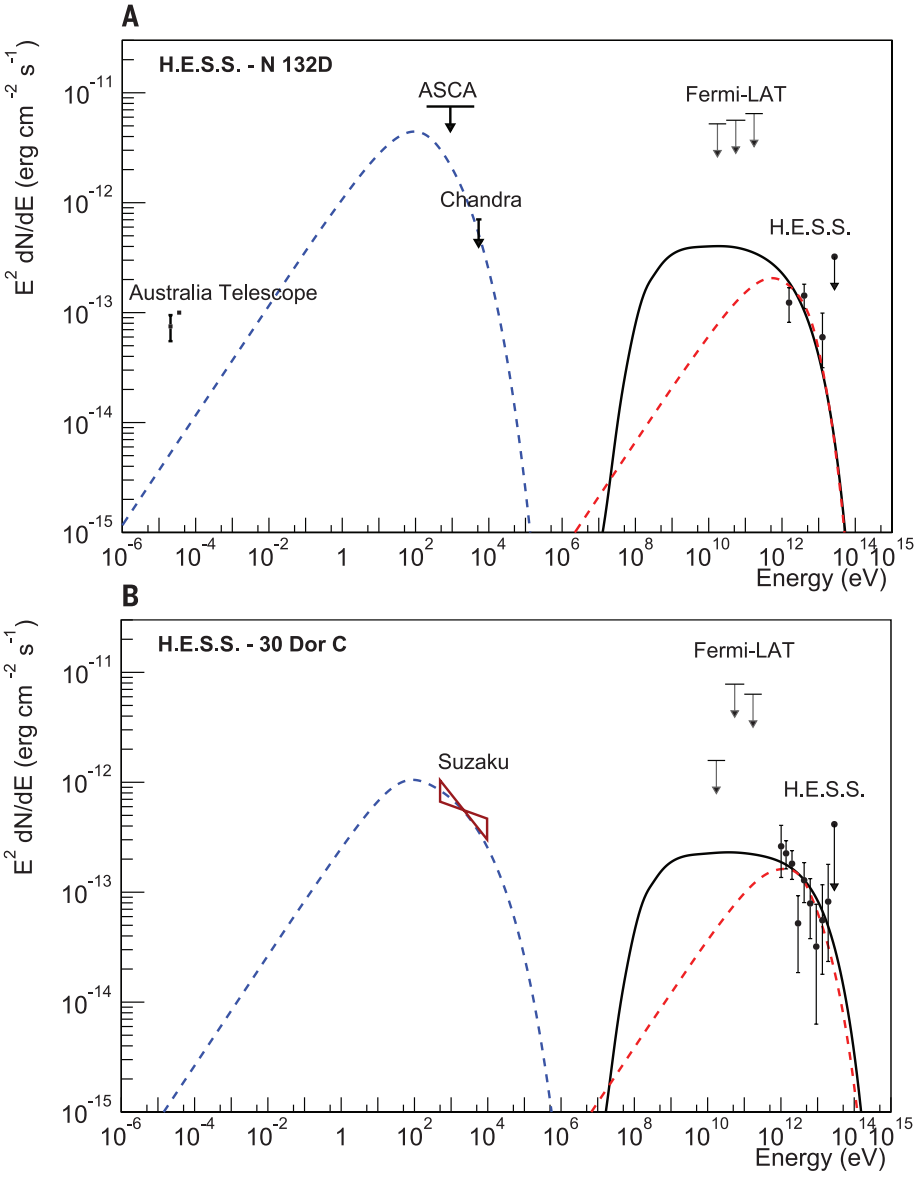


Fig. 3. Spectral energy distribution of 30 Dor C and N 132D. (**A**) For N 132D, radio data are from (*44*) and x-ray limits are from (*42*) and from reanalyzed Chandra data. (**B**) The 30 Dor C x-ray data are from (*23*). Both leptonic (dashed lines) and hadronic (solid lines) models are shown. For further details on Fermi-LAT data and spectral modeling, see S1.2 and S1.3.

Although at this stage we cannot rule out either the leptonic or the hadronic scenario, the H.E.S.S. observations reveal that the conditions inside the superbubble must be extreme: The hadronic scenario requires locations with high densities and a high degree of magnetic turbulence, whereas the leptonic scenario requires the stellar cluster to be extremely rarified. Moreover, the γ-ray and x-ray observations suggest active particle acceleration by a very large, fast-expanding shell. This may provide the right conditions for accelerating some protons to energies exceeding 3×10^{15} eV, which is the maximum energy detected for Galactic CRs. These observations, therefore, lend support to the view expressed in (*29*, *34*, *35*) that superbubbles may provide the right conditions for particle acceleration to very high energies, because they are thought to contain very turbulent magnetic fields and they are large enough to contain VHE particles for up to millions of years.

In the Milky Way, the most closely related object to 30 Dor C is the stellar cluster Westerlund 1 (*36*), which, however, has a completely different x-ray morphology. More important, it is not clear whether the γ rays originate from the cluster wind itself, a PWN, or the numerous supernovae that exploded inside Westerlund 1 in the recent past. Because a large fraction of supernovae are thought to go off in superbubbles, this first unambiguous detection of VHE γ rays from a superbubble may have broad implications for the circumstances in which a large fraction of CRs are accelerated.

N 157B

The source HESS J0537-691 is coincident with the PWN N 157B, which surrounds the pulsar PSR J0537-6910. PWNe are nebulae of ultrarelativistic particles driven by highly magnetized, fast-rotating neutron stars that convert a considerable amount of their spin-down energy into a particle wind. The archetypal Crab nebula is one of the brightest sources of nonthermal radiation in the sky and is powered by the pulsar with the highest spin-down energy known in the Milky Way (*37*). With comparably extreme rotational energy loss rates, N 157B ($\dot{E} = 4.9 \times 10^{38}$ erg s^{-1}) and the Crab nebula ($\dot{E} = 4.6 \times 10^{38}$ erg s^{-1}) appear to be twins. The study of N 157B thus provides an opportunity to compare two extreme PWNe and to disentangle object-specific and generic properties.

Given a population of ultrarelativistic electrons and positrons forming the PWN, the x-ray luminosity is determined by the strength of the magnetic field and the γ-ray luminosity by the intensity of radiation fields, which serve as targets for the inverse Compton upscattering. If the radiation fields are known, the magnetic field can be inferred from the combination of x-ray and γ-ray measurements. N 157B is likely associated with the LH 99 star cluster (*19*, *38*, *39*) and therefore is embedded in strong infrared radiation fields (see S1.5). In this environment, the magnetic field in the PWN must be rather weak, not larger than 45 μG, in order to explain the multiwavelength data (Fig. 4). When considering the region from which the hard x-ray emission is coming, the total energy in the magnetic field is $W_{B,\mathrm{tot}} = 1.4 \times 10^{47}$ erg, an order of magnitude smaller than the energy in >400 GeV electrons. The derived maximum magnetic field is also much lower than that inferred for the Crab nebula [~124 μG (*40*)] and suggests at least a factor ~7 lower magnetic pressure. Because most of the electrons that radiate in the Chandra, x-ray, and H.E.S.S., TeV domains have very short lifetimes (≤300 years), the energy in ultrarelativistic particles in N 157B can be inferred independently of the spin-down evolution of the pulsar. For the model shown in Fig. 4, a constant fraction of 11% of the current spin-down power of N 157B needs to be injected into the nebula in the form of relativistic electrons (compared with 50% for the Crab nebula under the same model). This fraction converted into x-ray and TeV emission is rather insensitive to the spectral index of injected electrons and the spin-evolution or braking index of the pulsar and only relies on the association of N 157B with LH 99 (see S1.5 for more information).

In this high-radiation field scenario, the situation for the Crab nebula is very different from N 157B. Not only is the best-fit electron spectrum of N 157B harder ($\Gamma_e = 2.0$ versus 2.35), exhibiting a lower cut-off energy ($E_c = 100$ TeV versus 3.5 PeV), but much of the spin-down energy of N 157B is also hidden and is not carried by ultrarelativistic particles or magnetic fields. The remainder of the available rotational energy is likely to be fed into electrons with energies ≤400 GeV that radiate at lower photon energies, adiabatic expansion, and/or particles escaping into the interstellar medium

via diffusive escape [e.g., (*41*)]. It therefore appears that N 157B is such a bright γ-ray emitter because of the enhanced radiation fields, despite the fact that it is apparently a much less efficient particle accelerator than the Crab nebula.

N 132D

In addition to the two unambiguously detected sources, we find strong evidence for a third source at the position of the core-collapse SNR N 132D. N 132D is a SNR with strong thermal x-ray emission, which has been used to estimate a preshock density of $n_H \approx 2.6$ cm^{-3} (*42*), a high explosion energy of ~6×10^{51} erg (*42*), and an age of ~6000 years, based on a Sedov model. Such x-ray bright SNRs are predicted to be γ-ray emitters (*43*). N 132D is also luminous in the radio (*44*) and infrared bands (*45*). N 132D is often compared to the brightest radio source Cas A, which, like N 132D, is an oxygen-rich SNR. N 132D has a higher infrared luminosity (*45*), but its radio luminosity is 50% that of Cas A. This is still remarkable, given that N 132D has a kinematic age of ~2500 years (*46*), whereas Cas A is ~330 years old and declines in radio luminosity by about 0.8% per year. The radio properties have been used to infer a magnetic-field strength of ~40 μG (*44*). The discrepancy between the age estimate based on the x-ray emission and the kinematic age may indicate that the supernova exploded within a bubble created by the progenitor star's wind before encountering the high-density material it now interacts with.

The γ-ray flux measured by H.E.S.S. translates to a 1 – 10 TeV γ-ray luminosity of $(0.9 \pm 0.2) \times 10^{35}(d/50\ \mathrm{kpc})^2$ erg s^{-1}. Assuming that the γ-ray emission is caused by neutral-pion production, this luminosity implies an energy of $10^{52}(n_H/1\ \mathrm{cm}^{-3})^{-1}(d/50\ \mathrm{kpc})^{-2}$ erg in relativistic protons. A hadronic origin of the γ-ray emission, therefore, implies either a large CR-energy fraction of 17% of the explosion energy, for an estimated post-shock density of $n_H \approx 10$ cm^{-3} (*42*), or that the gas density is higher than the x-ray–based estimates. The latter is plausible, given that N 132D appears to interact with dense, shocked interstellar clouds, seen in the optical and the infrared bands (*45*). It is interesting to compare N 132D to the most luminous Galactic SNR detected at TeV energies, HESS J1640–465: Both SNRs are believed to interact with a wind-blown cavity wall, to possibly have similar ages and sizes (*46*, *47*), and to have transferred a large fraction of their explosion energies into CRs.

The bright radio synchrotron luminosity of N 132D and the tentative claim of x-ray synchrotron emission from this source (*48*) also raises the possibility that the γ-ray emission is caused by inverse Compton scattering of low-energy photons. In and around N 132D, the radiation energy density is dominated by the bright infrared flux from dust inside the SNR and can be roughly estimated to be at least $u_{rad} \approx 1.0$ eV cm^{-3}. This leptonic scenario requires that the average magnetic-field strength needs to be ~20 μG, somewhat lower but still consistent with the equipartition value (see S1.3). However, this leptonic scenario critically depends on whether the 4 to 6 keV x-ray continuum emission indeed contains a substantial synchrotron component.

Whatever the emission mechanism for the γ-ray emission from N 132D, it is an exciting new γ-ray–emitting SNR, because its age lies in the gap between young (<2000 years) TeV-emitting SNRs, and old (≳10000 years) TeV-quiet SNRs. The latter can be bright pion-decay sources, but their spectra appear to be cut off above ~10 GeV. N 132D provides, therefore, an indication of how long SNRs contain CRs with energies in excess of 10^{13} eV.

SN 1987A

SN 1987A, the only naked-eye SN event since the Kepler SN (AD 1604), has been extensively observed at all wavelengths from the radio to the soft γ-ray band, providing invaluable insights into the evolution of a core-collapse SNR in its early stage (*49*).

It has been suggested that even in the early stages of the SNR development, the shock wave, which is heating the dense circumstellar medium (CSM) structured by stellar winds of the progenitor star, should have led to efficient acceleration of VHE nuclear CRs, accompanied by strong magnetic-field amplification through CR-induced instabilities (*14*, *50*). In collisions of the CRs with CSM particles, γ rays are produced. Estimates for the γ-ray flux (*14*, *15*) strongly depend on the magnetic-field topology and on the properties of the nonuniform CSM (*51*), making flux estimates uncertain by at least a factor of 2 (*14*).

Based on a nonlinear kinetic theory of CR acceleration, successfully applied to several young Galactic SNRs, the volume-integrated γ-ray flux at TeV energies, F_γ(>1 TeV), from SN 1987A was predicted to be rising in time and to have reached a level of ≈2.5×10^{-13} ph cm^{-2} s^{-1} in the year 2010 (*14*). An analysis with different assumptions on CSM properties and a more phenomenological approach to CR acceleration resulted in a predicted flux of ~8×10^{-14} ph cm^{-2} s^{-1} in the year 2013 (*15*). The H.E.S.S. upper limit F_γ (>1 TeV) < 5×10^{-14} ph cm^{-2} s^{-1} at a 99% confidence level, obtained from observations made between 2003 and 2012, being below the aforementioned predictions and a factor of 3 below similar estimates for the year 2005, therefore places constraints on the models despite their uncertainties.

The H.E.S.S. upper limit on the γ-ray flux translates into an upper limit for the γ-ray luminosity of L_γ (>1 TeV) < 2.2×10^{34} erg s^{-1}, which can be used to derive an approximate upper limit on the energy of the accelerated particles, W_{pp}, for a given average target density. Multiwavelength studies of SN 1987A suggest that the shock at the current epoch has reached and is interacting with the so-called equatorial ring, for which gas densities ranging from 10^3 cm^{-3} to 3×10^4 cm^{-3} have been found (*52*). Thus, one finds a conservative upper limit, $W_{pp} \lesssim 1.4 \times 10^{48} f^{-1}$ erg, where $0 < f < 1$ is the fraction of accelerated particles that are interacting with the dense regions. This upper limit on the energy of accelerated CR particles corresponds to 0.15 f^{-1}% of the explosion energy of 10^{51} erg.

Assuming a spherically symmetric distribution of accelerated particles, one can estimate $f \sim 0.2$ with the geometry of the equatorial ring found in (*53*). This translates to $W_{pp} \lesssim 9 \times 10^{48}$ erg, implying that less than 1% of the explosion energy is carried by accelerated CR nuclei. This fraction is rather small compared with typical values of ~10% for young SNRs (of ages ~1000 to 2000 years) but

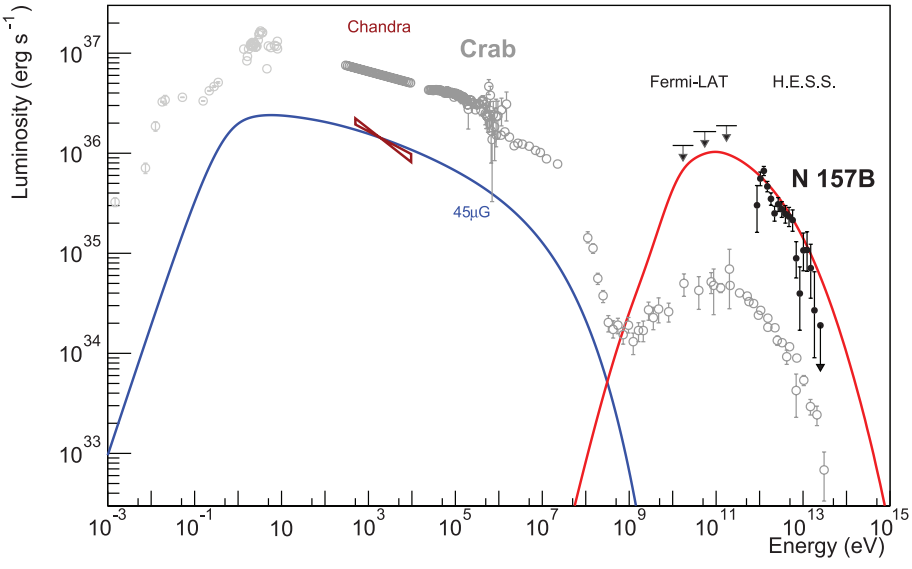


Fig. 4. Intrinsic SED of N 157B (black) and the Crab nebula (gray). The model shown assumes the same injection parameters as derived for the Crab nebula [E_{min} = 400 GeV, E_c = 3.5 PeV, Γ_e = 2.35 (*57*)]. The magnetic field required to explain the Chandra data of N 157B in the highest possible radiation fields is 45 μG. A significantly better fit to the Chandra data is obtained with Γ_e = 2.0 but requires a much lower cutoff of $E_c \lesssim 100$ TeV.

is not unreasonable for a very young object like SN 1987A.

Summary

With the deep H.E.S.S. observations of the LMC, we have detected three luminous examples of CR sources in an external galaxy. These sources detected in γ rays include a superbubble and counterparts to the most luminous sources in the Milky Way. N 157B provides a counterpart to the Crab Nebula, but its electron acceleration efficiency is five times less than for the Crab nebula, and its magnetic-field pressure is seven times less. N 132D has been long regarded as an older version of the brightest Galactic radio SNR Cas A and is one of the most radio-luminous SNRs known. N 132D is also notable in that it is one of the oldest VHE γ-ray–emitting SNRs. With the three detected sources, we increase our understanding of the variety of VHE γ-ray sources, which will likely require observations with the future Cherenkov Telescope Array (*54*), which should be an order of magnitude more sensitive than H.E.S.S.

REFERENCES AND NOTES

1. J. A. Hinton, W. Hofmann, *Annu. Rev. Astron. Astrophys.* **47**, 523–565 (2009).
2. F. Aharonian *et al*., *Astron. Astrophys.* **457**, 899–915 (2006).
3. F. Aharonian *et al*., *Science* **307**, 1938–1942 (2005).
4. P. J. McMillan, *Mon. Not. R. Astron. Soc.* **414**, 2446–2457 (2011).
5. R. P. van der Marel, *The Local Group as an Astrophysical Laboratory*, M. Livio, T. M. Brown, Eds. (Cambridge Univ. Press, Cambridge, 2006), pp. 47–71.
6. G. Pietrzyński *et al*., *Nature* **495**, 76–79 (2013).
7. S. Kim *et al*., *Astrophys. J.* **503**, 674–688 (1998).
8. J. Harris, D. Zaritsky, *Astron. J.* **138**, 1243–1260 (2009).
9. T. P. Robitaille, B. A. Whitney, *Astrophys. J.* **710**, L11–L15 (2010).
10. W. D. Arnett, J. N. Bahcall, R. P. Kirshner, S. E. Woosley, *Annu. Rev. Astron. Astrophys.* **27**, 629–700 (1989).
11. P. Sreekumar *et al*., *Astrophys. J.* **400**, L67 (1992).
12. A. A. Abdo *et al*., *Astron. Astrophys.* **512**, A7 (2010).
13. A. Abramowski *et al*., *Astron. Astrophys.* **545**, L2 (2012).
14. E. G. Berezhko, L. T. Ksenofontov, H. J. Völk, *Astrophys. J.* **732**, 58 (2011).
15. V. V. Dwarkadas, *Mon. Not. R. Astron. Soc.* **434**, 3368–3377 (2013).
16. M. de Naurois, L. Rolland, *Astropart. Phys.* **32**, 231–252 (2009).
17. S. Ohm, C. van Eldik, K. Egberts, *Astropart. Phys.* **31**, 383–391 (2009).
18. C.-C. Lu, H.E.S.S. Collaboration, 33rd International Cosmic Ray Conference, Rio de Janeiro, 2 to 9 July 2013.
19. Y. Chen *et al*., *Astrophys. J.* **651**, 237–249 (2006).
20. P. Kaaret *et al*., *Astrophys. J.* **546**, 1159–1167 (2001).
21. A. Bamba, M. Ueno, H. Nakajima, K. Koyama, *Astrophys. J.* **602**, 257–263 (2004).
22. D. A. Smith, Q. D. Wang, *Astrophys. J.* **611**, 881–891 (2004).
23. H. Yamaguchi, A. Bamba, K. Koyama, *Publ. Astron. Soc. Jpn.* **61**, 175 (2009).
24. D. S. Mathewson *et al*., *Astrophys. J. Suppl. Ser.* **58**, 197 (1985).
25. M. C. Lortet, G. Testor, *Astron. Astrophys.* **139**, 330 (1984).
26. P. J. Kavanagh *et al., Astron. Astrophys.*; available at http://arxiv.org/abs/1409.6547 (2014).
27. S. Gabici, F. A. Aharonian, *Mon. Not. R. Astron. Soc.* **445**, L70–L73 (2014).
28. A. W. Strong, I. V. Moskalenko, V. S. Ptuskin, *Annu. Rev. Nucl. Part. Sci.* **57**, 285–327 (2007).
29. A. M. Bykov, *Space Sci. Rev.* **99**, 317–326 (2001).
30. F. A. Aharonian, A. M. Atoyan, *Astron. Astrophys.* **351**, 330 (1999).
31. M. S. Oey, G. García-Segura, *Astrophys. J.* **613**, 302–311 (2004).
32. E. A. Helder *et al*., *Space Sci. Rev.* **173**, 369–431 (2012).
33. B. M. Gaensler *et al*., *Science* **307**, 1610–1612 (2005).
34. E. Parizot, A. Marcowith, E. van der Swaluw, A. M. Bykov, V. Tatischeff, *Astron. Astrophys.* **424**, 747–760 (2004).
35. G. Ferrand, A. Marcowith, *Astron. Astrophys.* **510**, A101 (2010).
36. A. Abramowski *et al*., *Astron. Astrophys.* **537**, A114 (2012).
37. R. Bühler, R. Blandford, *Rep. Prog. Phys.* **77**, 066901 (2014).
38. Y.-H. Chu, R. C. Kennicutt Jr., R. A. Schommer, J. Laff, *Astron. J.* **103**, 1545 (1992).
39. Q. D. Wang, E. V. Gotthelf, *Astrophys. J.* **494**, 623–635 (1998).
40. M. Meyer, D. Horns, H.-S. Zechlin, *Astron. Astrophys.* **523**, A2 (2010).
41. J. A. Hinton, S. Funk, R. D. Parsons, S. Ohm, *Astrophys. J.* **743**, L7 (2011).
42. J. P. Hughes, I. Hayashi, K. Koyama, *Astrophys. J.* **505**, 732–748 (1998).
43. B. Katz, E. Waxman, *J. Cosmol. Astropart. Phys.* **2008**, 018 (2008).
44. J. R. Dickel, D. K. Milne, *Astron. J.* **109**, 200 (1995).
45. A. Tappe, J. Rho, W. T. Reach, *Astrophys. J.* **653**, 267–279 (2006).
46. F. Vogt, M. A. Dopita, *Astrophys. Space Sci.* **331**, 521–535 (2011).
47. A. Abramowski *et al*., *Mon. Not. R. Astron. Soc.* **439**, 2828–2836 (2014).
48. X. Xiao, Y. Chen, *Adv. Space Res.* **41**, 416–423 (2008).
49. S. Immler, K. Weiler, R. McCray, Eds., *SUPERNOVA 1987A: 20 YEARS AFTER: Super-novae and Gamma-Ray Bursters*, vol. 937 of American Institute of Physics Conference Series (American Institute of Physics, College Park, MD, 2007).
50. A. R. Bell, K. M. Schure, B. Reville, G. Giacinti, *Mon. Not. R. Astron. Soc.* **431**, 415–429 (2013).
51. R. A. Chevalier, V. V. Dwarkadas, *Astrophys. J.* **452**, L45 (1995).
52. S. Mattila *et al*., *Astrophys. J.* **717**, 1140–1156 (2010).
53. C.-Y. Ng *et al*., *Astrophys. J.* **728**, L15 (2011).
54. M. Actis *et al*., *Exp. Astron.* **32**, 193–316 (2011).
55. D. Berge, S. Funk, J. Hinton, *Astron. Astrophys.* **466**, 1219–1229 (2007).
56. W. B. Atwood *et al*., *Astrophys. J.* **697**, 1071–1102 (2009).
57. P. L. Nolan *et al*., *Astrophys. J. Supp.* **199**, 31 (2012).
58. T. Kamae, N. Karlsson, T. Mizuno, T. Abe, T. Koi, *Astrophys. J.* **647**, 692–708 (2006).

ACKNOWLEDGMENTS

The support of the Namibian authorities and of the University of Namibia in facilitating the construction and operation of H.E.S.S. is gratefully acknowledged, as is the support by the German Ministry for Education and Research (BMBF), the Max Planck Society, the German Research Foundation (DFG), the French Ministry for Research, the CNRS-IN2P3 and the Astroparticle Interdisciplinary Programme of the CNRS, the U.K. Science and Technology Facilities Council (STFC), the Institute of Particle and Nuclear Physics (IPNP) of Charles University, the Czech Science Foundation, the Polish Ministry of Science and Higher Education, the South African Department of Science and Technology and National Research Foundation, and the University of Namibia. J.C. is aWallenberg Academy Fellow. M.D. and J.M. are funded by contract ERC-StG-259391 from the European Community. We appreciate the excellent work of the technical support staff in Berlin, Durham, Hamburg, Heidelberg, Palaiseau, Paris, Saclay, and Namibia in the construction and operation of the equipment. The H.E.S.S. Collaboration will make public the smoothed excess sky map and the associated correlated significance map as shown in Fig. 1, together with the source spectral points, on the H.E.S.S. website on the link to this publication: www.mpi-hd.mpg.de/hfm/HESS/pages/publications. We thank B. Katz, E. Waxman, and R. Budnik for their external proposal supporting observations of the SNR N 132D based on their work on γ-ray emission from shell-type SNRs (*43*).

The H.E.S.S. Collaboration

A. Abramowski,[1] F. Aharonian,[2,3,4] F. Ait Benkhali,[2] A.G. Akhperjanian,[5,4] E. O. Angüner,[6] M. Backes,[7] S. Balenderan,[8] A. Balzer,[9] A. Barnacka,[10,11] Y. Becherini,[12] J. Becker-Tjus,[13] D. Berge,[14] S. Bernhard,[15] K. Bernlöhr,[2,6] E. Birsin,[6] J. Biteau,[16,17] M. Böttcher,[18] C. Boisson,[19] J. Bolmont,[20] P. Bordas,[21] J. Bregeon,[22] F. Brun,[23] P. Brun,[23] M. Bryan,[9] T. Bulik,[24] S. Carrigan,[2] S. Casanova,[25,2] P.M. Chadwick,[8] N. Chakraborty,[2] R. Chalme-Calvet,[20] R.C.G. Chaves,[22] M. Chrétien,[20] S. Colafrancesco,[26] G. Cologna,[27] J. Conrad,[28] C. Couturier,[20] Y. Cui,[21] M. Dalton,[29] I. D. Davids,[18,7] B. Degrange,[16] C. Deil,[2] P. de Wilt,[30] A. Djannati-Ataï,[31] W. Domainko,[2] A. Donath,[2] L. O'C. Drury,[3] G. Dubus,[32] K. Dutson,[33] J. Dyks,[34] M. Dyrda,[25] T. Edwards,[2] K. Egberts,[35] P. Eger,[2] P. Espigat,[31] C. Farnier,[28] S. Fegan,[16] F. Feinstein,[22] M. V. Fernandes,[1] D. Fernandez,[22] A. Fiasson,[36] G. Fontaine,[16] A. Förster,[2] M. Füßling,[35] S. Gabici,[31] M. Gajdus,[6] Y. A. Gallant,[22] T. Garrigoux,[20] G. Giavitto,[37] B. Giebels,[16] J. F. Glicenstein,[23] D. Gottschall,[21] M.-H. Grondin,[2,27] M. Grudzińska,[24] D. Hadasch,[15] S. Häffner,[38] J. Hahn,[2] J. Harris,[8] G. Heinzelmann,[1] G. Henri,[32] G. Hermann,[2] O. Hervet,[19] A. Hillert,[2] J. A. Hinton,[33] W. Hofmann,[2] P. Hofverberg,[2] M. Holler,[16,35] D. Horns,[1] A. Ivascenko,[18] A. Jacholkowska,[20] C. Jahn,[38] M. Jamrozy,[10] M. Janiak,[34] F. Jankowsky,[27] I. Jung,[38] M. A. Kastendieck,[1] K. Katarzyński,[39] U. Katz,[38] S. Kaufmann,[27] B. Khélifi,[31] M. Kieffer,[20] S. Klepser,[37] D. Klochkov,[21] W. Kluźniak,[34] D. Kolitzus,[15] Nu. Komin,[*26] K. Kosack,[23] S. Krakau,[13] F. Krayzel,[36] P. P. Krüger,[18] H. Laffon,[29] G. Lamanna,[36] J. Lefaucheur,[31] V. Lefranc,[23] A. Lemière,[31] M. Lemoine-Goumard,[29] J.-P. Lenain,[20] T. Lohse,[6] A. Lopatin,[38] C.-C. Lu,[*2] V. Marandon,[2] A. Marcowith,[22] R. Marx,[2] G. Maurin,[36] N. Maxted,[30] M. Mayer,[*35] T. J. L. McComb,[8] J. Méhault,[29] P. J. Meintjes,[40] U. Menzler,[13] M. Meyer,[28] A. M. W. Mitchell,[2] R. Moderski,[34] M. Mohamed,[27] K. Morå,[28] E. Moulin,[23] T. Murach,[6] M. de Naurois,[16] J. Niemiec,[25] S. J. Nolan,[8] L. Oakes,[6] H. Odaka,[2] S. Ohm,[*37] B. Opitz,[1] M. Ostrowski,[10] I. Oya,[6] M. Panter,[2] R. D. Parsons,[2] M. Paz Arribas,[6] N. W. Pekeur,[18] G. Pelletier,[32] J. Perez,[15] P.-O. Petrucci,[32] B. Peyaud,[23] S. Pita,[31] H. Poon,[2] G. Pühlhofer,[21] M. Punch,[31] A. Quirrenbach,[27] S. Raab,[38] I. Reichardt,[31] A. Reimer,[15] O. Reimer,[15] M. Renaud,[22] R. de los Reyes,[2] F. Rieger,[2] L. Rob,[41] C. Romoli,[3] S. Rosier-Lees,[36] G. Rowell,[30] B. Rudak,[34] C. B. Rulten[19] V. Sahakian,[5,4] D. Salek,[42] D. A. Sanchez,[36] A. Santangelo,[21] R. Schlickeiser,[13] F. Schüssler,[23] A. Schulz,[37] U. Schwanke,[6] S. Schwarzburg,[21] S. Schwemmer,[27] H. Sol,[19] F. Spanier,[18] G. Spengler,[28] F. Spies,[1] Ł. Stawarz,[10] R. Steenkamp,[7] C. Stegmann,[35,37] F. Stinzing,[38] K. Stycz,[39] I. Sushch,[6,18] J.-P. Tavernet,[20] T. Tavernier,[31] A. M. Taylor,[3] R. Terrier,[31] M. Tluczykont,[1] C. Trichard,[36] K. Valerius,[38] C. van Eldik,[38] B. van Soelen,[40] G. Vasileiadis,[22] J. Veh,[38] C. Venter,[18] A. Viana,[2] P. Vincent,[20] J. Vink,[*9] H. J. Völk,[2] F. Volpe,[2] M. Vorster,[18] T. Vuillaume,[32] S. J. Wagner,[27] P. Wagner,[6] R. M. Wagner,[28] M. Ward,[8] M. Weidinger,[13] Q. Weitzel,[2] R. White,[33] A. Wierzcholska,[25] P. Willmann,[38] A. Wörnlein,[38] D. Wouters,[23] R. Yang,[2] V. Zabalza,[2,33] D. Zaborov,[16] M. Zacharias,[27] A. A. Zdziarski,[34] A. Zech,[19] H.-S. Zechlin[1]

[1]Universität Hamburg Institut, für Experimentalphysik, Luruper Chaussee 149, D 22761 Hamburg, Germany. [2]Max-Planck-Institut für Kernphysik, P.O. Box 103980, D 69029 Heidelberg, Germany. [3]Dublin Institute for Advanced Studies, 31 Fitzwilliam Place, Dublin 2, Ireland. [4]National Academy of Sciences of the Republic of Armenia, Marshall Baghramian Avenue, 24, 0019 Yerevan, Republic of Armenia. [5]Yerevan Physics Institute, 2 Alikhanian Brothers St., 375036 Yerevan, Armenia. [6]Institut für Physik, Humboldt-Universität zu Berlin, Newtonstrasse 15, D 12489 Berlin, Germany. [7]University of Namibia, Department of Physics, Private Bag 13301, Windhoek, Namibia. [8]University of Durham, Department of Physics, South Road, Durham DH1 3LE, UK. [9]Gravitation Astroparticle Physics Amsterdam (GRAPPA), Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands. [10]Obserwatorium Astronomiczne, Uniwersytet Jagielloński, Ulica Orla 171, 30-244 Kraków, Poland. [11]Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, MS-20, Cambridge, MA 02138, USA. [12]Department of Physics and Electrical Engineering, Linnaeus University, 351 95 Växjö, Sweden. [13]Institut für Theoretische Physik, Lehrstuhl IV: Weltraum und Astrophysik, Ruhr-Universität Bochum, D 44780 Bochum, Germany. [14]GRAPPA, Anton Pannekoek Institute for Astronomy and Institute of High-Energy Physics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands. [15]Institut für Astro- und Teilchenphysik, Leopold- Franzens-Universität Innsbruck, A-6020 Innsbruck, Austria. [16]Laboratoire Leprince-Ringuet, Ecole Polytechnique, CNRS/IN2P3, F-91128 Palaiseau, France. [17]Santa Cruz Institute for Particle Physics, Department of Physics, University of California at Santa Cruz, Santa Cruz, CA 95064, USA. [18]Centre for Space Research, North-West University, Potchefstroom 2520, South Africa. [19]Laboratoire Univers et Théories (LUTH), Observatoire de Paris, CNRS, Université Paris Diderot, 5 Place Jules Janssen, 92190 Meudon, France. [20]Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), Université Pierre et Marie Curie Paris 6, Université Denis Diderot Paris 7, CNRS/IN2P3, 4 Place Jussieu, F-75252, Paris Cedex 5, France. [21]Institut für Astronomie und Astrophysik,

Universität Tübingen, Sand 1, D 72076 Tübingen, Germany. [22] Laboratoire Univers et Particules de Montpellier, Université Montpellier 2, CNRS/IN2P3, CC 72, Place Eugène Bataillon, F-34095 Montpellier Cedex 5, France. [23]DSM/Irfu, CEA Saclay, F-91191 Gif-Sur-Yvette Cedex, France. [24]Astronomical Observatory, The University of Warsaw, Aleje Ujazdowskie 4, 00-478 Warsaw, Poland. [25]Instytut Fizyki Jadrowej PAN, Ulica Radzikowskiego 152, 31-342 Kraków, Poland. [26]School of Physics, University of the Witwatersrand, 1 Jan Smuts Avenue, Braamfontein, Johannesburg, 2050 South Africa. [27]Landessternwarte, Universität Heidelberg, Königstuhl, D 69117 Heidelberg, Germany. [28]Oskar Klein Centre, Department of Physics, Stockholm University, Albanova University Center, SE-10691 Stockholm, Sweden. [29]Université Bordeaux 1, CNRS/IN2P3, Centre d'Études Nucléaires de Borde H.E.S.S. Collaborationaux Gradignan, 33175 Gradignan, France. [30]School of Chemistry and Physics, University of Adelaide, Adelaide 5005, Australia. [31]APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, 10, Rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France. [32]Universite Grenoble Alpes, IPAG, F-38000 Grenoble, France; CNRS, IPAG, F-38000 Grenoble, France. [33]Department of Physics and Astronomy, The University of Leicester, University Road, Leicester, LE1 7RH, United Kingdom. [34]Nicolaus Copernicus Astronomical Center, Ulica Bartycka 18, 00-716 Warsaw, Poland. [35]Institut für Physik und Astronomie, Universität Potsdam, Karl-Liebknecht-Strasse 24/25, D 14476 Potsdam, Germany. [36]Laboratoire d'Annecy-le-Vieux de Physique des Particules, Université de Savoie, CNRS/IN2P3, F-74941 Annecy-le-Vieux, France. [37]Deutsches Elektronen-Synchrotron (DESY), D-15738 Zeuthen, Germany. [38]Universität Erlangen-Nürnberg, Physikalisches Institut, Erwin-Rommel-Strasse 1, D 91058 Erlangen, Germany. [39]Centre for Astronomy, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Grudziadzka 5, 87-100 Torun, Poland. [40]Department of Physics, University of the Free State, Post Office Box 339, Bloemfontein 9300, South Africa. [41]Charles University, Faculty of Mathematics and Physics, Institute of Particle and Nuclear Physics, V Holešovičkách 2, 180 00 Prague 8, Czech Republic. [42]GRAPPA, Institute of High-Energy Physics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands. *Corresponding author. E-mail: nukri.komin@wits.ac.za (N.K.); chia-chun.lu@mpi-hd.mpg.de (C.-C.L.); michael.mayer@physik.hu-berlin.de (M.M.); stefan.ohm@desy.de (S.O.); j.vink@uva.nl (J.V.)

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/406/suppl/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 and S2
References (*59*–*61*)

16 September 2014; accepted 16 December 2014
10.1126/science.1261313

MARS ATMOSPHERE

The imprint of atmospheric evolution in the D/H of Hesperian clay minerals on Mars

P. R. Mahaffy,[1]* C. R. Webster,[2] J. C. Stern,[1] A. E. Brunner,[1,3,4] S. K. Atreya,[5] P. G. Conrad,[1] S. Domagal-Goldman,[1] J. L. Eigenbrode,[1] G. J. Flesch,[2] L. E. Christensen,[2] H. B. Franz,[1,4] C. Freissinet,[1,6] D. P. Glavin,[1] J. P. Grotzinger,[7] J. H. Jones,[8] L. A. Leshin,[9] C. Malespin,[1,10] A. C. McAdam,[1] D. W. Ming,[8] R. Navarro-Gonzalez,[11] P. B. Niles,[8] T. Owen,[12] A. A. Pavlov,[1] A. Steele,[13] M. G. Trainer,[1] K. H. Williford,[2] J. J. Wray,[14] the MSL Science Team†

The deuterium-to-hydrogen (D/H) ratio in strongly bound water or hydroxyl groups in ancient martian clays retains the imprint of the water of formation of these minerals. Curiosity's Sample Analysis at Mars (SAM) experiment measured thermally evolved water and hydrogen gas released between 550° and 950°C from samples of Hesperian-era Gale crater smectite to determine this isotope ratio. The D/H value is 3.0 (±0.2) times the ratio in standard mean ocean water. The D/H ratio in this ~3-billion-year-old mudstone, which is half that of the present martian atmosphere but substantially higher than that expected in very early Mars, indicates an extended history of hydrogen escape and desiccation of the planet.

We obtained a reference point for the evolution of the martian atmosphere and loss of near-surface water by comparing deuterium-to-hydrogen (D/H) ratios from the atmosphere with those in ancient clay minerals, such as those in the Yellowknife Bay lake bed on the floor of Gale crater (*1*). Each process of atmospheric loss to the surface or to space can leave an isotopic imprint, so models of the evolution of the atmosphere and ancient climates (*2*–*4*) can be constrained by isotopic measurements as well as by geological studies (*5*–*7*). The Sample Analysis at Mars (SAM) experiment (*8*) on the Curiosity rover of the Mars Science Laboratory (MSL) mission has refined the atmospheric measurements not only of D/H in water (*9*) but also of $\delta^{13}C$ in CO_2 (*9*, *10*), $\delta^{15}N$ in N_2 (*11*), and $^{36}Ar/^{38}Ar$ (*12*). These combined measurements consistently support the paradigm of continued atmospheric loss to space over the course of martian history, with this process dominating over loss to surface reservoirs.

The mechanism for enriching D in water over the past billions of years entails the photolysis of water by solar ultraviolet (UV) radiation and the more rapid escape to space of the lighter H, leaving D behind to be incorporated again in surface water. The present loss rate is estimated to be at least 10^{26} atoms s^{-1} (*13*). For the water that formed the clays of the Sheepbed member at Yellowknife Bay, we have now measured the D/H ratio from water and hydrogen that was released in stepped heating to 950°C and analyzed with SAM's tunable laser spectrometer (TLS) and quadrupole mass spectrometer (QMS). The clays sampled in the Hesperian-age (*14*–*16*) Yellowknife Bay formation were likely formed during diagenesis before lithification (*17*) more than 3 billion years ago. Thus, the D/H value provides a reference point for estimates of surface water loss over geologic time scales. The Hesperian period in martian history is not sampled by the present meteorite record.

The ancient record of D/H in water is preserved in the crystal structure of the clay minerals formed by aqueous transformation of basaltic materials because fractionation of D/H between water and clay minerals during formation is no greater than a few 10s per mil (*18*). In the absence of high levels of heating or recrystallization, the structural OH in the octahedral layers of clay minerals thus preserves the isotopic value of the liquid water of formation. Predictions of the global loss of water from Mars have used the D/H ratio measured in Mars meteorites (*19*–*22*), with the earliest martian value presumed to be near that of terrestrial oceans. Recent analysis of D/H in the primitive martian mantle melt Yamato 980459 supports this hypothesis, with its D/H of ≤1.28 × standard mean ocean water (SMOW) (*21*) close to terrestrial, although contributions of cometary water to both early Earth and Mars have been postulated (*23*). The terrestrial SMOW D/H value of 1.558×10^{-4} is half that of several Oort-class comets (*24*–*26*) but closer to the value found in the Kuiper Belt comet Hartley 2 (*27*) and in a range of carbonaceous chondrites (*28*). The Hesperian period in martian history explored with the MSL measurements is not sampled by the present meteorite record.

The age of the 150-km-diameter Gale crater is estimated, by modeling of crater densities, to be

[1]Planetary Environments Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA. [3]School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85281, USA. [4]Center for Research and Exploration in Space Science and Technology, University of Maryland College Park, College Park, MD 20742, USA. [5]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI 48109-2143, USA. [6]NASA Postdoctoral Program, Oak Ridge Associated Universities, Oak Ridge, TN 37831, USA. [7]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA. [8]NASA Johnson Space Flight Center, Houston, TX 77058, USA. [9]Office of the President, Worcester Polytechnic Institute, Worcester, MA 01609, USA. [10]Goddard Earth Sciences Technology and Research (GESTAR)/Universities Space Research Association (USRA) NASA Goddard Space Flight Center Greenbelt, MD 20771, USA. [11]Universidad Nacional Autónoma de México, Ciudad Universitaria, México D.F. 04510, Mexico. [12]Institute for Astronomy, University of Hawaii, Honolulu, HI 96822, USA. [13]Geophysical Laboratory, Carnegie Institution of Washington, Washington, DC 20015, USA. [14]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA.
***Corresponding author. E-mail: paul.r.mahaffy@nasa.gov** †The MSL Science Team authors and affiliations are listed in the supplementary materials.

~3.6 billion years (*15*, *16*), whereas some regions on the crater floor are found to be younger at ~2.9 billion to 3.5 billion years old (*14*–*16*), having formed during the Hesperian period of martian history. Our analysis used samples from two drill holes designated John Klein and Cumberland from the Sheepbed member, the lowest stratigraphic unit in the Yellowknife Bay formation located ~500 m northeast of the Mars Curiosity rover landing site (*29*). These samples were analyzed by the full suite of instruments on the rover, including SAM and the Chemistry and Mineralogy (CheMin) instruments. These analyses revealed a mudstone containing a smectite clay mineral in addition to an amorphous component and basaltic minerals (*1*, *30*, *31*), with isochemical alteration that indicates authigenic smectite formation (*17*, *29*). The sandstones and mudstones of the Yellowknife Bay formation provide strong evidence for both flowing water in streams expressed in conglomerates and sandstones of the Gillespie member, and bodies of standing water as lakes that formed the Sheepbed member. The thickness of the Sheepbed member suggests that this body of water must have existed for at least hundreds to thousands of years (*1*). The existence of streams and stable lakes requires water production in the upper reaches of the Gale crater rim, which was a source for sediments that were transported to the Peace Vallis fan and its stratigraphic and facies equivalents. This is consistent with previous evidence for wetter Hesperian climates required by stream networks cut into bedrock (*32*) and deltas that built outward into bodies of standing water (*33*).

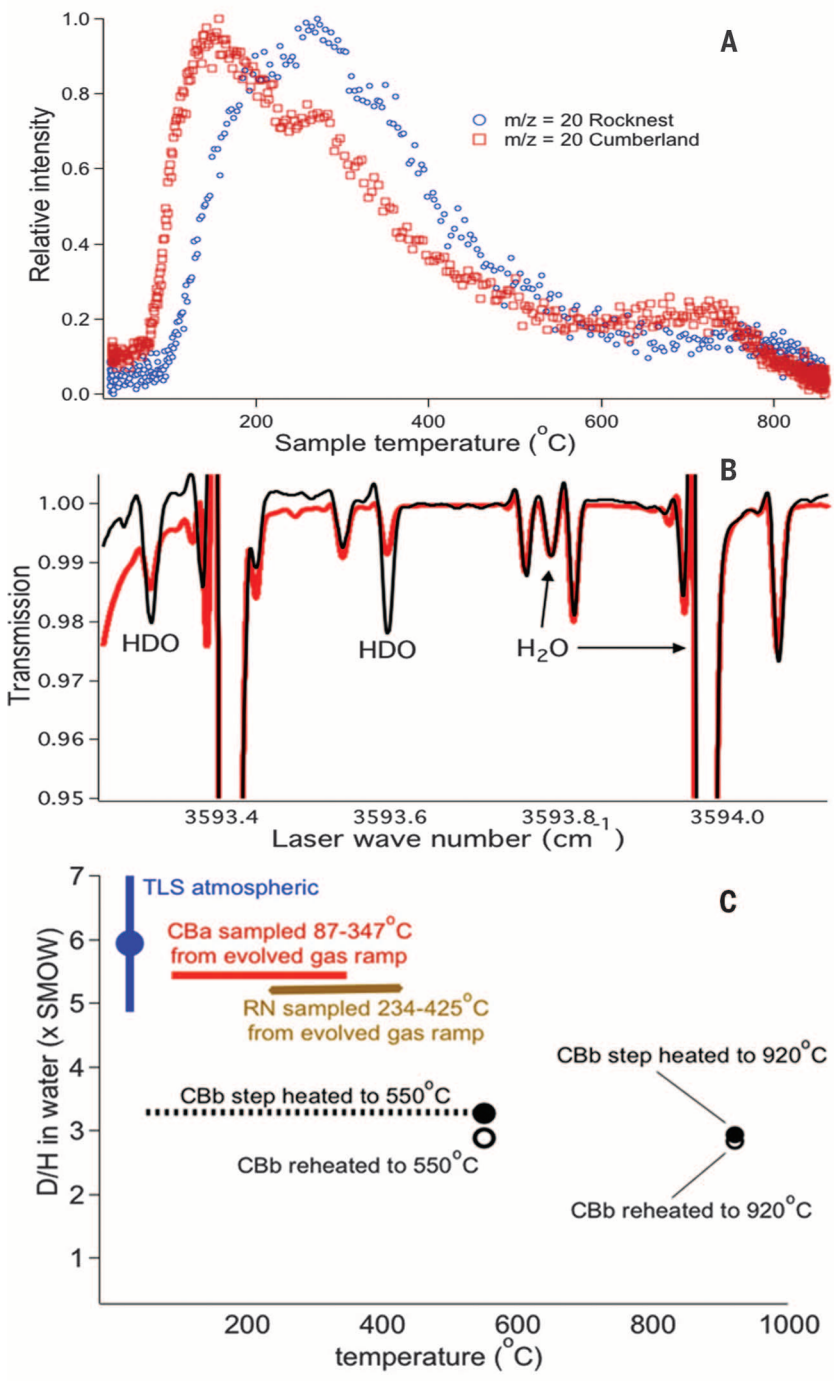


Fig. 1. Water released from martian samples. (**A**) Evolved water signature as monitored by the mass spectrometer at a mass-to-charge (*m*/*z*) ratio of 20 ($H_2{}^{18}O$) for both a Rocknest eolian deposit and the Cumberland mudstone. (**B**) TLS measurements showing the high-temperature water evolution from stepped heating in the black trace compared with the high-resolution transmission (HITRAN) spectrum in red normalized to the large $H_2{}^{16}O$ peak near 3594 cm^{-1}. (**C**) TLS measurements of D/H relative to terrestrial SMOW for water captured from a temperature ramp from Cumberland (CBa) and Rocknest (RN) samples, and water evolved at different temperatures in the step heating experiment from a second Cumberland sample (CBb). Rocknest D/H values in (*9*) and (*40*) are updated in (*41*).

A comparison of the evolved gas analysis (EGA) trace for H_2O from one of the Yellowknife Bay Cumberland drill hole samples (*30*) is shown in Fig. 1A with the same trace for a Rocknest aeolian deposit sample (*40*). The low-temperature water in the Cumberland sample, which peaks at ~200°C and extends to over 500°C, may have a number of sources, including water that is surface adsorbed, smectite interlayer, associated with amorphous materials, or structurally bound in compounds such as Fe-oxyhydroxides (*30*). The total water released from the Cumberland sample is ~2% by mass (*30*). The high-temperature water peak (600° to 800°C) evident in the Cumberland mudstone EGA trace is the signature of the structural hydroxyl hydrogen derived from the octahedral layer of a smectite component of this sample, which was definitively identified

The D/H ratio in martian atmospheric water derived from spectroscopic observations was reported in 1988 as (6 ± 3) × SMOW (*34*), and some years later as (5.5 ± 2) × SMOW (*35*). Higher-resolution spectroscopy has allowed spatial variations to be mapped (*36*, *37*), and values of 6.5 × and 7 × SMOW have recently been reported (*37*, *38*). Early results from the SAM TLS gave a value of (6 ± 1) × SMOW for a single atmospheric sample (*9*). The variability in D/H in the meteorite record may point to somewhat isolated reservoirs with different D/H ratios (*19*, *20*), but the atmospheric D/H ratios are generally consistent with those reported in meteorites with younger rock formation ages, such as the 6.08 × SMOW value (*21*) in the geochemically enriched 183-million-year-old (*39*) Larkman Nunatak 06319 shergottite and the ~5.6 × SMOW value reported (*20*) from two fracture-free zones in apatite grains of the (~170-million-year-old) Shergotty meteorite.

by the CheMin x-ray diffraction analysis of the Cumberland mudstone (*31*).

Most of the solid samples analyzed with SAM used the EGA measurement protocol (*8*) that identified a suite of volatiles, including the most abundant gas, H_2O, as well as CO_2, O_2, SO_2, H_2S, HCl, NO, H_2, H_2S, and several chlorinated organic compounds. The gas stream was diverted to the TLS in selected temperature intervals to measure C, O, and H isotopes in CO_2 and H_2O and the abundances of these molecules (Fig. 1B). However, the D/H in the high-temperature water component cannot be accurately isolated with the EGA experiment by using a continuous temperature ramp because of the mixing of the tail of the low-temperature-evolved water with the smaller high-temperature peak. In order to better separate the low- and high-temperature components, a specialized experiment was designed and implemented (*41*) that used stepped heating. The D content of the water released in the four steps of this experiment compared with EGA runs is given in Fig. 1C and table S1. Three high-temperature water releases give nearly identical values of δD (table S1) as measured with the TLS. The mole-weighted average δD that includes a small volume of H_2, and a small H_2 blank correction, gives a high-temperature D/H of (3.0 ± 0.2) × SMOW.

As evidenced by meteorite studies and this result, escape of hydrogen from the upper atmosphere of Mars has left its imprint in the D/H ratio in near-surface water over the past 3 billion to 4 billion years. The loss of water on Mars can be modeled by using single or multiple near-surface reservoirs, with the amount of surface water often expressed as a water-equivalent global layer (GEL). In the single-reservoir model, the entire near-surface H_2O inventory is exposed to atmospheric loss through time. Using the notation of Kurokawa (*22*), the amount of water R_{t1} in a near-surface reservoir at a time $t1$ with a D/H ratio of I_{t1} can be related to the amount R_{t2} at a later time $t2$ with a D/H ratio of I_{t2} by the expression $\frac{R_{t1}}{R_{t2}} = \left(\frac{I_{t2}}{I_{t1}}\right)^{\frac{1}{1-f}}$, where f is the fractionation factor. Thus, knowledge of a present and past D/H ratio and an estimate of the present GEL allows the volume of past water to be established.

In the multiple-reservoir models (*42*) invoked to explain variability in meteorite studies, a surficial D/H reservoir continuously participates in atmospheric cycling and is exposed to atmospheric loss, whereas another reservoir consists of deeply buried ice that is not exposed to atmospheric loss processes. Vigorous escape during early stages of planetary evolution when both reservoirs would have been rapidly exchanging would have provided an initial enrichment of the δD of both reservoirs to a value of 2 × to 4 × SMOW (*20*, *42*, *43*). Early catastrophic atmospheric loss (*44*) and global cooling separate a majority of the water as ice in the upper crust, isolating it from atmospheric processes and atmospheric loss. Because the amount of H_2O available to readily exchange with the atmosphere is small, only small amounts of atmospheric loss would be needed to create large D/H enrichments. However, short periods of warm, wet conditions on Mars could cause mobilization and exchange with this subsurface reservoir, providing surface water with a lower D/H ratio such as observed with our D/H measurement in Yellowknife Bay sediments.

In the single-reservoir model, Yellowknife Bay D/H values are compared with data from selected martian meteorites assuming a continuous D/H evolutionary timeline (*22*). This approach allows constraints on the amount of near-surface Hesperian water. The Amazonian features a high D/H of ~6 × SMOW (*21*) based on the <200-million-year-old meteorites LAR 06319 and Shergotty; the Noachian has a mid-range D/H value of 2.2 to 4 × SMOW based on the 4.1 Ga ALH 84001 pyroxenite (obtained from analyses of carbonates and magmatic apatite); and the pre-Noachian 4.5-billion-year-old primordial martian water exhibits a D/H of ≤1.28 × SMOW, based on melt inclusions in the Yamato 980459 shergottite meteorite (*21*). Under the assumptions of a fractionation factor for H and D loss in the range of 0.016 (*22*) to the diffusion limit of 0.4, the amount of water lost to space since the time of formation of the Yellowknife Bay clay minerals would be approximately equal to 1 to 1.5 times the amount of water in current-surface and near-surface reservoirs. Current estimates of these reservoirs that include the observable polar cap water reservoir (20 to 30 m) and other sources of cryospheric ice not bound in minerals is at least 50 m (*45*). Thus, the GEL at the time of the Cumberland mudstone formation would be 100 to 150 m or more. The current very limited set of exospheric H observations (*46*) predict substantially less loss and are not consistent with this measurement. Earlier loss including nonthermal escape processes (*47*, *48*) would push the predicted surface-water volume before this time higher as indicated by this ~3 × SMOW D/H value and supported by geological studies (*7*, *49*).

The D/H ratio of Hesperian water is preserved within a Yellowknife Bay mudstone, whose clay minerals were formed in an active lacustrine environment (*1*) several hundred million years after the Gale cratering event. This ratio established with in situ measurements from instruments on the Curiosity rover provides a new data point to help constrain the volume of water lost through escape processes over the past 3 billion years.

REFERENCES AND NOTES

1. J. P. Grotzinger *et al.*, *Science* **343**, 1242777 (2014).
2. R. O. Pepin, *Earth Planet. Sci. Lett.* **252**, 1–14 (2006).
3. H. Lammer *et al.*, *Space Sci. Rev.* **174**, 113–154 (2013).
4. B. M. Jakosky, R. O. Pepin, R. E. Johnson, J. L. Fox, *Icarus* **111**, 271–288 (1994).
5. J. W. Head *et al.*, *Space Sci. Rev.* **96**, 263–292 (2001).
6. M. H. Carr, J. W. Head III, *Earth Planet. Sci. Lett.* **294**, 185–203 (2010).
7. M. H. Carr, J. W. Head, *J. Geophys. Res.* **108** (E5), 5042 (2003).
8. P. R. Mahaffy *et al.*, *Space Sci. Rev.* **170**, 401–478 (2012).
9. C. R. Webster *et al.*, *Science* **341**, 260–263 (2013).
10. P. R. Mahaffy *et al.*, *Science* **341**, 263–266 (2013).
11. M. H. Wong *et al.*, *Geophys. Res. Lett.* **40**, 6033–6037 (2013).
12. S. K. Atreya *et al.*, *Geophys. Res. Lett.* **40**, 5605–5609 (2013).
13. D. M. Hunten, M. B. Mcelroy, *J. Geophys. Res.* **75**, 5989–6001 (1970).
14. J. A. Grant, S. A. Wilson, N. Mangold, F. Calef III, J. P. Grotzinger, *Geophys. Res. Lett.* **41**, 1142–1149 (2014).
15. B. J. Thomson *et al.*, *Icarus* **214**, 413–432 (2011).
16. L. L. Deit *et al.*, *J. Geophys. Res. Planets* **118**, 2439–2473 (2013).
17. T. F. Bristow *et al.*, *Am. Mineral.* (2015).
18. S. M. F. G. Sheppard, *Clay Miner.* **31**, 1–24 (1996).
19. L. Leshin, *Geophys. Res. Lett.* **27**, 2017–2020 (2000).
20. J. P. Greenwood, S. Itoh, N. Sakamoto, E. P. Vicenzi, H. Yurimoto, *Geophys. Res. Lett.* **35**, 5203 (2008).
21. T. Usui, C. M. O. D. Alexander, J. Wang, J. I. Simon, J. H. Jones, *Earth Planet. Sci. Lett.* **357**, 119–129 (2012).
22. H. Kurokawa *et al.*, *Earth Planet. Sci. Lett.* **394**, 179–185 (2014).
23. T. Owen, *Space Sci. Rev.* **138**, 301–316 (2008).
24. P. Eberhardt, M. Reber, D. Krankowsky, R. R. Hodges, *Astron. Astrophys.* **302**, 301 (1995).
25. D. Bockelée-Morvan *et al.*, *Icarus* **133**, 147–162 (1998).
26. R. Meier *et al.*, *Science* **279**, 1707–1710 (1998).
27. P. Hartogh *et al.*, *Nature* **478**, 218–220 (2011).
28. C. M. O. D. Alexander *et al.*, *Science* **337**, 721–723 (2012).
29. J. P. Grotzinger *et al.*, *Science* **343**, 1242777 (2014).
30. D. W. Ming *et al.*, *Science* **343**, 1245267 (2013).
31. D. T. Vaniman *et al.*, *Science* **343**, 1243480 (2014).
32. V. Ansan, N. Mangold, *Planet. Space Sci.* **54**, 219–242 (2006).
33. K. W. Lewis, O. Aharonson, *J. Geophys. Res. Planets* **111** (E6), 6001 (2006).
34. T. Owen, J. P. Maillard, C. de Bergh, B. L. Lutz, *Science* **240**, 1767–1770 (1988).
35. V. A. Krasnopolsky, G. L. Bjoraker, M. J. Mumma, D. E. Jennings, *J. Geophys. Res. Planets* **102** (E3), 6525–6534 (1997).
36. G. L. Villanueva *et al.*, *AGU Fall Meet. Abstr.* **52**, 05 (2010).
37. R. E. Novak, M. J. Mumma, G. L. Villanueva, *Planet. Space Sci.* **59**, 163–168 (2011).
38. G. L. Villanueva *et al.*, *J. Quant. Spectrosc. Radiat. Transf.* **113**, 202–220 (2012).
39. J. T. Shafer *et al.*, *Geochim. Cosmochim. Acta* **74**, 7307–7328 (2010).
40. L. A. Leshin *et al.*, *Science* **341**, 1238937 (2013).
41. Materials and methods are available as supplementary materials on *Science* Online.
42. T. Usui, C. M. O. D. Alexander, J. Wang, J. I. Simon, J. H. Jones, *44th Lunar Planet. Sci. Conf.* **1454**, abstract 1623 (2013).
43. L. A. Leshin, *Geophys. Res. Lett.* **27**, 2017–2020 (2000).
44. H. Lammer *et al.*, *Space Sci. Rev.* **174**, 113–154 (2013).
45. M. H. Carr, J. W. Head, in *45th Lunar Planet. Sci. Conf.*, abstr. 1427 (2014).
46. F. Tian *et al.*, in *Comparative Climatology of Terrestrial Planets*, S. J. Mackwell *et al.*, Eds. (Univ. of Arizona Press, Tucson, AZ, 2013), pp. 567-581.
47. H. Lammer *et al.*, *Icarus* **165**, 9–25 (2003).
48. R. O. Pepin, *Icarus* **111**, 289–304 (1994).
49. R. A. DiBiase, A. B. Limaye, J. S. Scheingross, W. W. Fischer, M. P. Lamb, *J. Geophys. Res. Planets* **118**, 1285–1302 (2013).

ACKNOWLEDGMENTS

This work was supported by NASA's Mars Exploration Program. The dedicated teams that developed the SAM suite of instruments are acknowledged. All data described can be found in NASA's the Planetary Data System archive pds.nasa.gov.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/412/suppl/DC1
Materials and Methods
SupplementaryText
Figs. S1 to S4
Tables S1 to S2
MSL Team Members and Affiliations

21 August 2014; accepted 1 December 2014
Published online 18 December 2014;
10.1126/science.1260291

Tree Physiology 16, 425–432

Effects of elevated CO_2 concentration on leaf characteristics and photosynthetic capacity of beech (*Fagus sylvatica*) during the growing season

DANIEL EPRON,[1,2] RODOLPHE LIOZON[1] and MARIANNE MOUSSEAU[1]

[1] *Laboratoire d'Ecologie Végétale, CNRS URA 1492, Université Paris Sud, Batîment 362, 91405 Orsay cedex, France*

[2] *Laboratoire de Sciences Végétales, Institut des Sciences et des Techniques de l'Environnement, Pôle Universitaire, BP 427, 25211 Montbéliard cedex, France*

Received April 13, 1995

Summary Two-year-old beech (*Fagus sylvatica* L.) saplings were planted directly in the ground at high density (100 per m^2), in an experimental design that realistically mimicked field conditions, and grown for two years in air containing CO_2 at either ambient or an elevated (ambient + 350 ppm) concentration. Plant dry mass and leaf area were increased by a two-year exposure to elevated CO_2. The saplings produced physiologically distinct types of sun leaves associated with the first and second growth flushes. Leaves of the second flush had a higher leaf mass per unit area and less chlorophyll per unit area, per unit dry mass and per unit nitrogen than leaves of the first flush. Chlorophyll content expressed per unit nitrogen decreased over time in plants grown in elevated CO_2, which suggests that, in elevated CO_2, less nitrogen was invested in machinery of the photosynthetic light reactions. In early summer, the photosynthetic capacity measured at saturating irradiance and CO_2 was slightly but not significantly higher in saplings grown in elevated CO_2 than in saplings grown in ambient CO_2. However, a decrease in photosynthetic capacity was observed after July in leaves of saplings grown in CO_2-enriched air. The results demonstrate that photosynthetic acclimation to elevated CO_2 can occur in field-grown saplings in late summer, at the time of growth cessation.

Keywords: acclimation, leaf, photosynthesis.

Introduction

As evidence of rising atmospheric CO_2 has accumulated during the last twenty years, the role of temperate forests in the global carbon cycle has been emphasized (Tans et al. 1990). The major uncertainty is the extent to which forest ecosystems will serve as a sink for extra atmospheric carbon dioxide. The responses of tree growth to elevated CO_2 are now well documented (Eamus and Jarvis 1989, Gunderson and Wullschleger 1994). However, there have been few studies concerned with the long-term effects of CO_2 enrichment on photosynthesis in broad-leaved tree species of European temperate forests (see review by Ceulemans and Mousseau 1994).

Photosynthesis is one of the main physiological processes thought to respond to increasing atmospheric CO_2 in C_3 plants, and growth enhancement of trees is primarily mediated by an increase in carbon assimilation. Short-term exposure to elevated CO_2 concentration increases net CO_2 assimilation because the CO_2 concentration at the catalytic site limits the activity of ribulose bisphosphate carboxylase oxygenase (Stitt 1991). The potential stimulation of net CO_2 assimilation due to atmospheric CO_2 enrichment is especially high for species like beech that have low stomatal and internal conductances to CO_2, and therefore low CO_2 concentrations at the chloroplast (Epron et al. 1995). However, this primary stimulation often diminishes with time of exposure to elevated CO_2. The nature of this so-called acclimation or down-regulation of photosynthesis continues to be debated and may sometimes reflect artefacts induced by the experimental conditions (Ceulemans and Mousseau 1994, Gunderson and Wullschleger 1994). Until recently, most studies on the responses of tree photosynthesis to elevated CO_2 have been done with seedlings grown for a few months under artificial climatic conditions with a restricted rooting volume. It is unclear, therefore, whether a decrease in photosynthetic capacity will occur under field conditions. However, in one field study, with yellow poplars and white oaks planted directly in the ground, no photosynthetic acclimation was observed (Gunderson et al. 1993).

Beech (*Fagus sylvatica* L.) is a major deciduous forest tree species in western Europe. Previously, it has been shown that beech seedlings grown in pots exhibit a large increase in total biomass (60%) and a sharp increase in whole-plant carbon assimilation after one growing season in CO_2 at twice the ambient atmospheric concentration (El Kohen et al. 1993). In the present study, beech saplings were planted directly in the ground at high density in an experimental design that realistically mimicked field conditions. Because the main characteristics of beech leaves are fixed during the early stages of leaf initiation (Eschrich et al. 1989, Thiebaut et al. 1990), measurements were carried out during the second season of growth in elevated or ambient CO_2. We examined the long-term effect of a doubling of atmospheric CO_2 concentration on seasonal

variation in photosynthesis of two physiologically distinct types of beech leaves. To separate possible stomatal effects from true photosynthetic acclimation, photosynthetic capacity was assessed by mean of oxygen evolution rates at saturating irradiance and CO_2. The effects of elevated CO_2 on leaf nitrogen content, chlorophyll content and leaf mass per unit area, which are some of the major leaf characteristics linked to photosynthetic processes, were also investigated. The aim of this study was to test whether the photosynthetic capacity of field-grown saplings remains unaffected after a two-year exposure to elevated CO_2.

Materials and methods

Plant material and experimental design

Beech seedlings were grown from seeds in a commercial forest nursery (Pépinière Bouchery et Fils, Crouy sur Cosson, France) for two years. The saplings were transplanted, while dormant, to trenches (1 m deep, 1 m wide and 2 m long) in a meadow of the Orsay University (48° N, 2° E). A cement wall was built around each trench, and the bottoms of the trenches were filled with a 15-cm layer of gravel. These physical barriers prevented roots from growing out of the trenches. The trenches were covered with small, naturally illuminated greenhouses (1.8 m high, 1 m wide and 2 m long) that were open sided and constantly ventilated with outside air at a flow rate of 5.7 $m^3\ min^{-1}$. The greenhouses were made from transparent polypropylene film (28 μm thickness) glued to aluminum frames.

Saplings were planted at high density (100 per m^2) in a sandy soil (clay 11%, silt 23%, sand 64%, soil organic matter 1.4% and total nitrogen 0.09%) without the addition of fertilizer. High densities (up to 300 per m^2) are common in naturally regenerating forests and may be maintained for more than three years (Oswald 1981). Plants were watered two or three times a week depending on weather conditions. Day and night temperatures inside the greenhouses ranged from 10 to 38 °C and 5 to 22 °C, respectively, during the growing season. High irradiances caused the air temperature inside the greenhouse to increase up to 4 °C above the outside air temperature, but temperature differences between the greenhouses were less than 0.5 °C. The photosynthetic photon flux density was attenuated by 15 to 20% by the polypropylene film, but the difference between the greenhouses never exceeded 10%.

Elevated CO_2 concentrations in the greenhouse were established by injecting pure industrial CO_2 at the blower inlet at a constant flow rate of 2 $dm^3\ min^{-1}$. Therefore, the concentrations of atmospheric CO_2 in this study were ambient and ambient + 350 ppm (elevated). The CO_2 enrichment was initiated when the saplings were transplanted (March 1992), and was maintained day and night for 20 months with regular monitoring with an infrared gas analyzer.

Leaf photosynthesis

During the second growing season, net CO_2 assimilation and transpiration rates of sun-exposed, attached, and fully expanded leaves were measured with a portable gas exchange system (Model CI301, CID Inc., Vancouver, WA). Gas exchange was measured before noon under saturating solar irradiance (> 1600 $\mu mol\ m^{-2}\ s^{-1}$) at ambient atmospheric CO_2 concentration. Stomatal conductance was derived from transpiration data according to standard equations (Von Caemmerer and Farquhar 1981).

Measurements of photosynthetic O_2 evolution rates were made at 30 °C in a leaf disc electrode chamber (LD2, Hansatech, Norfolk, England) filled with water-saturated air containing 5% CO_2 and 19% O_2 (it had previously been determined that 5% CO_2 was saturating for photosynthesis in well-watered beech leaves). Light was provided at the top of the chamber by a slide projector and a mirror. The photon flux density (PFD) at the leaf disc level was 1200 $\mu mol\ m^{-2}\ s^{-1}$. Preliminary light response curves indicated that photosynthesis was saturated at PFDs above 900 $\mu mol\ m^{-2}\ s^{-1}$. Measurements were carried out on leaf discs (9.8 cm^2) punched from sun-exposed, fully expanded leaves. Leaves were harvested at dawn and stored in darkness in a humidified petri dish until measurements were made (not more than 6 h after excision). It was determined that the potential rate of O_2 evolution did not change during the storage period. Leaf discs were illuminated for 30 to 40 min until a steady-state rate of O_2 evolution was observed, and then darkened for at least 5 min before dark respiration was measured.

Leaf characteristics

Following the measurement of maximal rates of photosynthetic O_2 evolution, chlorophyll content was determined. A small part of the leaf disc (1.13 cm^2) was extracted in 3 ml of dimethylsulfoxide at 65 °C for at least 90 min (Hiscox and Israelstam 1979). The remaining part of the leaf disc was weighed after oven drying at 60 °C for 48 h and dry mass was used to calculated leaf mass per unit area (LMA). Finally, the dry leaf material was ground and sent to the service central d'Analyse of the Centre National de la Recherche Scientifique, (Vernaison, France) for the catharometric determination of nitrogen content.

Plant growth

Stem height (h) and diameter (d) were measured periodically on 15 plants during the second season of growth. The same group of plants was used during the whole season. The volume index (hd^2) was computed and a highly significant linear regression found between the index and aboveground plant dry mass (DM): $DM = 0.23\ hd^2$ ($r = 0.96$, $P < 0.001$, data not shown). The linear regression did not differ between the treatments. The volume index and the shoot dry mass of all saplings were determined at the end of the growing season. Therefore, the volume index rather than the calculated dry mass is presented in Figure 1 (periodic measurements), because the regression was established at the final harvest. All leaves and roots were collected, and mean leaf area per plant and mean root dry mass per plant were calculated from the total harvest divided by the number of plants. Root collection was achieved by excavating a 2-m^3 soil volume from beneath each greenhouse.

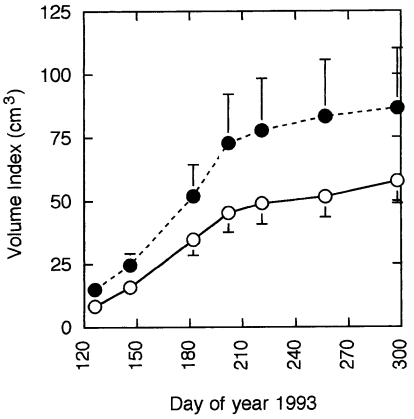


Figure 1. Volume index (hd^2) during the second growing season of beech saplings grown in ambient (○) or elevated (●) CO_2. Values are means of 15 replicates. The vertical bars are equal to one half of the standard error.

Data analysis

There was only one greenhouse for each CO_2 treatment, with the consequence that statistical analysis allowed comparisons only between greenhouses. However, initial soil characteristics, the irrigation procedure, as well as air temperature and PFD at plant level, were virtually the same in the two greenhouses. Therefore, differences between the greenhouses in plant characteristics are most probably attributable to differences in the CO_2 concentration of the air.

Means of six (O_2 evolution rates, LMA, nitrogen and chlorophyll contents) or 16 replicates (gas exchange measurements) were calculated with their standard errors. All data were subjected to one-way (CO_2 treatment) or two-way (CO_2 treatment and flush) analysis of variance. Separate analyses were performed for each sampling date. Differences were considered statistically significant when $P < 0.05$.

Results

At the beginning of the second growing season, bud break occurred at the end of April and was neither advanced nor delayed by elevated CO_2. The second flush of shoot growth began in the middle of June. At the end of the second growing season, leaves of the second flush represented 20% of mean plant leaf area in both the ambient and elevated CO_2 treatments. The mean leaf area per plant determined from litter collection at the end of the second growing season was increased by elevated CO_2 (53%, Table 1). This increase was due almost entirely to an increase in the number of leaves per plant. The increase with time in aboveground dry mass during the second year of the experiment was estimated from stem height and diameter measurements (volume index) of 15 plants per treatment (Figure 1). The growth of the aboveground parts of the plants was strongly reduced after July in both ambient and elevated CO_2. The saplings grown in elevated CO_2 always had a higher volume index than saplings grown in ambient CO_2, but the differences were not significant because of large between-plant differences and the small sample size. At the end of the second growing season, all plants were harvested for the determination of dry mass and, with this larger sample, shoot dry mass was significantly ($P < 0.001$) higher (90%) in the elevated CO_2 treatment than in the ambient CO_2 treatment (Table 1). Leaf and root dry mass were 67 and 124% higher, respectively, in the elevated CO_2 treatment than in the ambient CO_2 treatment (Table 1).

Table 1. Mean leaf area per plant and mean shoot, root and leaf dry mass at the end of the second growing season in beech saplings grown in ambient or elevated CO_2. Values are means and standard errors for all saplings. Different letters indicate significant differences between CO_2 treatments at $P < 0.001$. No statistics could be calculated for leaf area and leaf and root dry mass because leaves and roots were combined on collection.

	Ambient CO_2	Elevated CO_2
Shoot dry mass (g)	11.7 ± 0.54 a	22.4 ± 2.22 b
Root dry mass (g)	11.5	25.8
Leaf dry mass (g)	3.7	6.2
Leaf area (dm^2)	9.2	14.1

Periodic measurements throughout the second growing season indicated that leaf mass per unit area (LMA) was significantly ($P < 0.05$) higher in plants grown in elevated CO_2 than in plants grown in ambient CO_2, except in May (Day 125, Figure 2). Similarly, the leaves of the second flush had a significantly ($P < 0.05$) higher LMA than the leaves of the first flush in both ambient and elevated CO_2, except in July (Day 198, Figure 2). Leaf nitrogen content per unit area did not differ significantly with CO_2 treatment, except in June (Day 168) and July (Day 198) when saplings grown in elevated CO_2 had more leaf nitrogen than saplings grown in ambient CO_2 ($P < 0.01$, Figure 2). Leaves of the second flush of shoot growth had similar amounts of nitrogen to leaves of the first flush.

The chlorophyll content per unit area was slightly lower in saplings grown in elevated CO_2 than in saplings grown in ambient CO_2. The leaves of the second flush had less chlorophyll per unit area than the leaves of the first flush ($P < 0.001$, Figure 3). When expressed on a per unit nitrogen basis, the amount of chlorophyll in plants grown in elevated CO_2 was less than that in plants grown in ambient CO_2 ($P < 0.05$), except in August, Day 218, (Figure 3). The difference was less pronounced in leaves of the second flush than in leaves of the first flush. Leaves of the second flush also differed from leaves of the first flush in having less chlorophyll per unit nitrogen ($P < 0.001$, Figure 3).

Increasing the CO_2 concentration of the air had no significant effect on the maximal rate of net photosynthetic O_2 evolution measured at saturating irradiance and CO_2 (Figure 4). At the beginning of the growing season, the rates of oxygen evolution were slightly higher in saplings grown in elevated CO_2 than in saplings grown in ambient CO_2 ($P = 0.072$ in July, Day 198), but became slightly though not significantly lower after July in the leaves of both the first and second flushes ($P = 0.092$ in August and $P = 0.051$ in September, Days 218 and

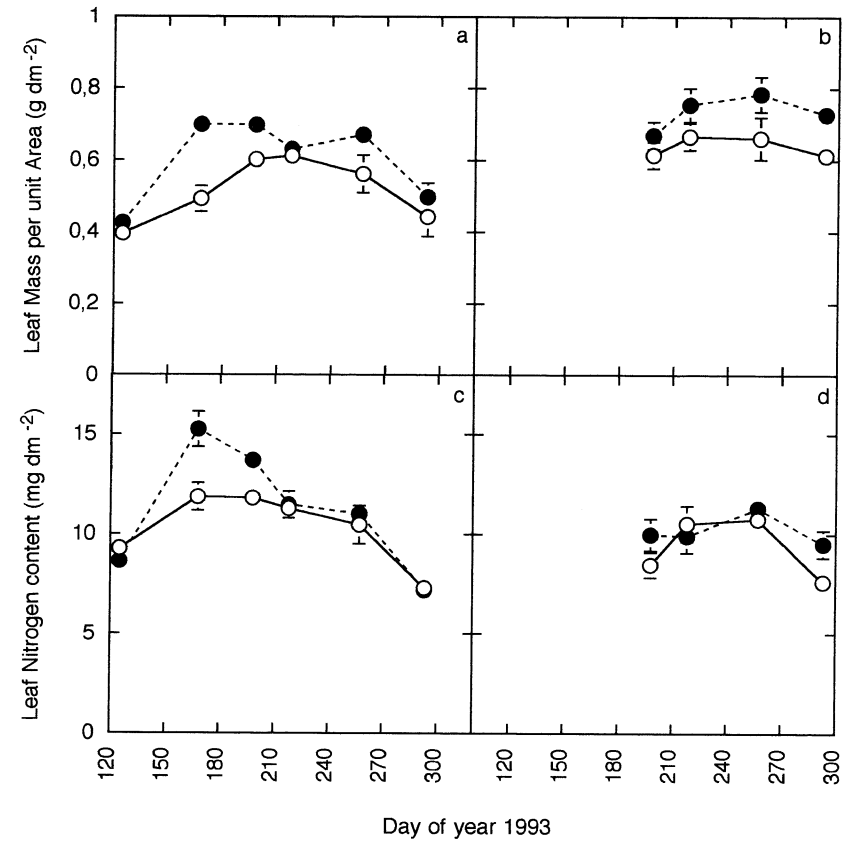


Figure 2. Leaf mass per unit area (a, b) and leaf nitrogen per unit area (c, d) in leaves of the first (a, c) and second (b, d) growth flush during the second growing season of beech saplings grown in ambient (○) or elevated (●) CO_2. Values are means of six replicates and vertical bars equal the standard errors.

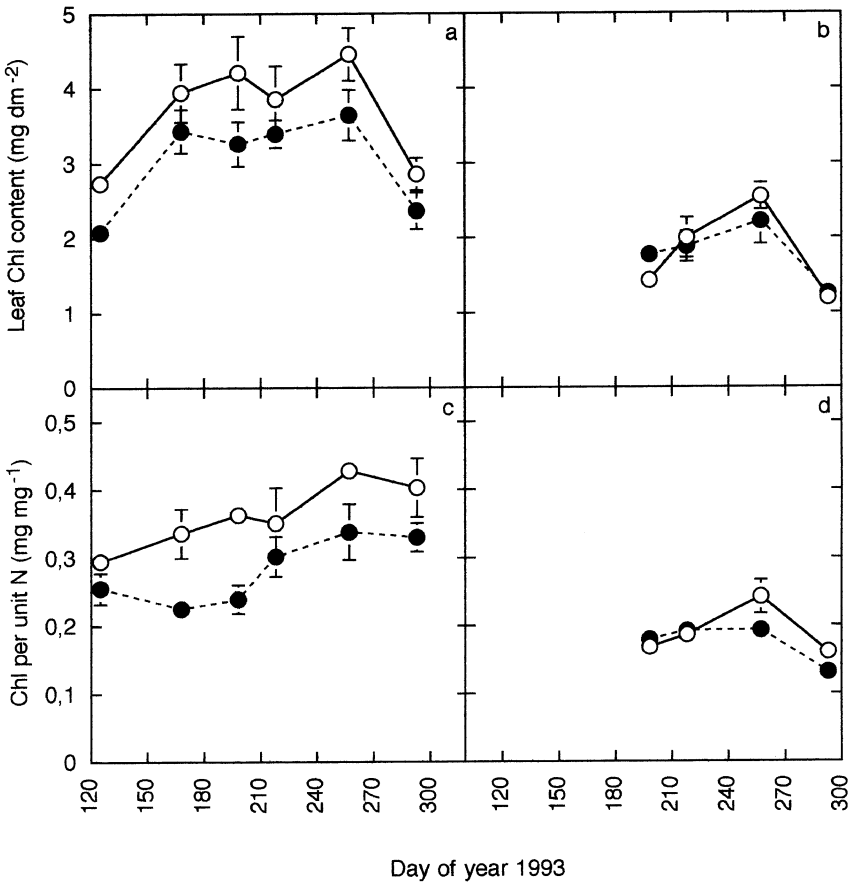


Figure 3. Amounts of chlorophyll per unit area (a, b) and per unit nitrogen (c, d) in leaves of the first (a, c) and the second (b, d) growth flush during the second growing season of beech saplings grown in ambient (○) or elevated (●) CO_2. Values are means of six replicates and vertical bars equal the standard error.

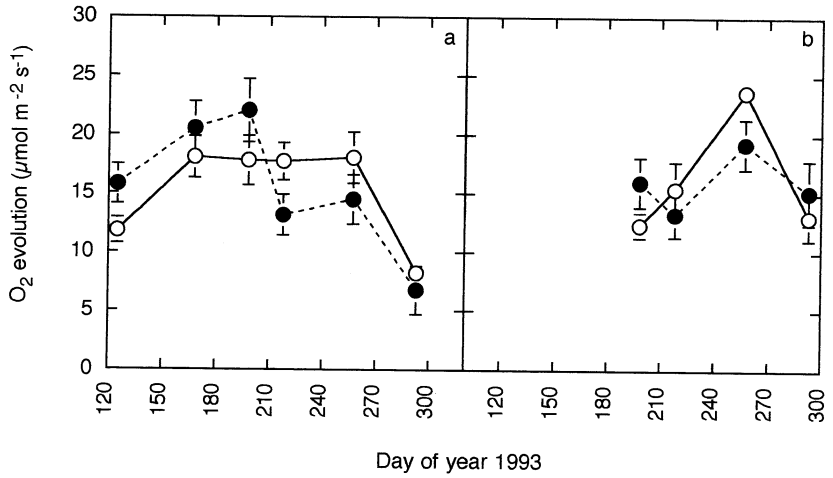


Figure 4. Maximal rate of photosynthetic O_2 evolution in leaves of the first (a) and second (b) growth flush during the second growing season of beech saplings grown in ambient (○) or elevated (●) CO_2. Values are means of six replicates and vertical bars equal the standard error.

257, respectively). The leaves of the second flush had slightly higher maximal rates of O_2 evolution than the leaves of the first flush (up to 20 μmol m^{-2} s^{-1}) later in the season. Thus, the leaves of the second flush had lower rates of O_2 evolution than the leaves of the first flush in July (Day 198, $P < 0.01$), but the reverse was observed in September and October ($P < 0.01$, Days 257 and 293, respectively). The rate of O_2 consumption in darkness was similar in both kinds of leaves during the whole growing season (about 2 μmol m^{-2} s^{-1}, data not shown). Thus, both the net and the gross O_2 evolution rates showed a similar trend.

Leaf gas exchange of sunlit leaves was measured before noon with saturating irradiance at ambient atmospheric CO_2 concentration. Under these conditions, both net CO_2 assimilation rates and stomatal conductances were unaffected by CO_2 concentration during growth at the beginning of July (Table 2). But in late August, second flush leaves of plants grown in elevated CO_2 had lower rates of net CO_2 assimilation than leaves of plants grown in ambient CO_2 ($P < 0.05$, Table 2). Measurements on sunlit leaves of the first flush were not made at this time of the year because of intermittent self-shading. The large difference in net CO_2 assimilation rates between leaves of the first and second flushes in early July ($P < 0.001$, Table 2), which was not a reflection of a difference in stomatal conductance, was in agreement with the lower photosynthetic capacity of young leaves of the second flush (Figure 4).

Discussion

Differences between leaves of the first and the second growth flush

Morphological differences between leaves of the first and second growth flush of beech have already been reported. Thicker leaves, mainly because of an increase in the thickness of both palisade (more cell layers) and spongy parenchyma, are characteristic of the second flush of beech (Roy et al. 1986, Thiebaut et al. 1990). In the present study, the higher leaf mass per unit area in leaves of the second flush compared to those of the first flush suggests a greater leaf thickness.

Higher mass and lower chlorophyll content per unit area in leaves of the second flush compared to leaves of the first flush led to much lower chlorophyll concentrations expressed on a dry mass basis. This may reflect physiological differences between leaves of the first and second flushes. A lower LMA and a higher chlorophyll concentration are among the characteristics that often distinguish sun leaves from shade leaves (Boardman 1977). The leaves of the first flush also had more chlorophyll per unit nitrogen than leaves of the second flush. The decrease in the ratio of nitrogen to chlorophyll, and hence in the allocation of nitrogen to light trapping (Seemann et al. 1987, Evans 1989, Ellsworth and Reich 1992), is another indication that leaves of the second flush exhibit more distinct sun-leaf characteristic than leaves of the first flush, although leaves of both flushes were exposed to full sun light.

Sun leaves are known to have a higher photosynthetic capacity than shade leaves (Boardman 1977, Seemann et al. 1987). It has also been reported that leaves of the second flush have a higher rate of photosynthesis at saturating irradiance than leaves of the first flush (Roy et al. 1986). A similar a difference

Table 2. Net CO_2 assimilation rate (A) and stomatal conductance (g_w) measured before noon with saturating irradiance at ambient atmospheric CO_2 concentration in sunlit leaves of beech saplings grown in ambient or elevated CO_2. Values are means and standard errors for 16 replicates. Different letters indicate significant differences between CO_2 treatments or growth flushes at $P \leq 0.05$. Separate analyses were performed for each sampling date.

	A (μmol m^{-2} s^{-1})		g_w (mmol m^{-2} s^{-1})	
	Flush 1	Flush 2	Flush 1	Flush 2
July 8 (Day 189)				
Ambient CO_2	11.5 ± 0.74 a	6.1 ± 0.62 b	127 ± 7.2 a	112 ± 3.8 a
Elevated CO_2	11.1 ± 0.64 a	6.8 ± 0.67 b	114 ±7.3 a	113 ± 6.1 a
August 20 (Day 232)				
Ambient CO_2		7.9 ± 0.49 a		71± 5.3 a
Elevated CO_2		6.6 ± 0.31 b		63 ± 3.2 a

between growth flushes in maximal rate of photosynthetic O_2 evolution at saturating irradiance and CO_2 was observed in the present study during September and October. However, the opposite was observed in July, and no difference was observed in August. Thus it appears that differences between growth flushes in maximal photosynthetic rate reflect differences in leaf age not intrinsic physiology.

Nevertheless, differences in leaf mass per unit area, chlorophyll concentration and amount of chlorophyll per unit nitrogen, the reported anatomical differences (Roy et al. 1986, Thiebaut et al. 1990), and the fact that the first flush is plagiotropic whereas the second flush is orthotropic (Thiebaut 1982), suggest that leaves of the second flush are better adapted to high irradiances than leaves of the first flush.

Effects of elevated CO_2 on leaf characteristics

In beech, morphological and anatomical differences between leaves of the first and the second flushes (Thiebaut et al. 1990) as well as differences between sun and shade leaves (Eschrich et al. 1989) appear to be related to differences in environmental conditions during leaf initiation rather than during leaf emergence. Because leaves harvested during the second growing season had been initiated and maintained throughout in either ambient or elevated atmospheric CO_2, we assume that both the short- and long-term effects of CO_2 concentration on leaf characteristics were fully expressed.

An increase in leaf mass per unit area is a common response to CO_2 enrichment (Mousseau and Saugier 1992). However, it is unclear whether this response is mainly attributable to accumulations of starch or other nonstructural carbohydrates (Delucia et al. 1985, Wong 1990), or to an increase in leaf thickness caused by an increase in the number of palisade cell layers (Thomas and Harvey 1983). In any case, the increase in LMA led to a slight decrease in nitrogen concentration per unit dry mass, which was mostly attributable to a dilution effect, because nitrogen content per unit area either increased or remained unchanged in response to elevated CO_2. In contrast, the decline in leaf chlorophyll per unit dry mass in response to CO_2 enrichment cannot be solely explained by a dilution effect because chlorophyll per unit area also decreased in response to CO_2 enrichment. Such a decrease in leaf chlorophyll has frequently been reported in young plants in response to elevated CO_2 (Delucia et al. 1985, Houpis et al. 1988, Wullschleger et al. 1992) and may have some physiological significance. The finding in this study that leaf chlorophyll per unit nitrogen was reduced by atmospheric CO_2 enrichment suggests that elevated CO_2 causes less nitrogen to be invested in the light reactions of photosynthesis. A decrease in nitrogen invested in Rubisco has frequently been reported in plants grown in elevated CO_2 (Sage et al. 1989, Tissue et al. 1993). It is unlikely, therefore, that the decrease in the proportion of nitrogen invested in light trapping was balanced by an increase in the proportion of nitrogen invested in Rubisco. We postulate, instead, that nitrogen is allocated to nonphotosynthetic functions of leaves, rather than to primary photosynthetic processes. In pea leaves exposed to high CO_2 concentrations, the relative abundance of the chloroplast phosphate translocator, starch phosphorylase and ADP-glucose pyrophosphorylase increased with atmospheric CO_2 enrichment (Rivière and Betsche 1994). This may reflect a higher requirement for proteins involved in carbohydrate metabolism and export arising from a higher rate of carbohydrate synthesis in plants grown in elevated CO_2.

Effects of elevated CO_2 on leaf photosynthesis

It has often been reported that trees exposed to elevated CO_2 for a period of weeks or months exhibit a decrease in photosynthesis, and this has been described as an acclimation process (for recent reviews see Ceulemans and Mousseau 1994 and Gunderson and Wullschleger 1994). Gunderson and Wullschleger (1994) show that, in most of the tree species tested (80%), there is a decline in net photosynthesis measured at ambient CO_2 when plants are grown in elevated CO_2. However, in some studies, photosynthetic capacity was unaffected by the growth concentration of CO_2 (Norby and O'Neill 1989, Curtis and Terry 1992, Epron et al. 1994). In the present study, early summer photosynthetic capacity measured at saturating irradiance and CO_2 was slightly but not significantly higher in beech saplings grown in elevated CO_2 than in ambient CO_2. The lack of photosynthetic acclimation during the first months of the second year was confirmed by measuring net CO_2 assimilation rates at ambient CO_2 concentration. As pointed out by Gunderson and Wullschleger (1994), comparing CO_2 assimilation rates of plants grown at different CO_2 concentrations at the ambient CO_2 concentration is more convenient for assessing photosynthetic acclimation than measuring CO_2 assimilation rate at the respective growth CO_2 concentrations. The lack of acclimation observed in the present study indicates that photosynthetic capacity can be maintained for two years in elevated CO_2, and is consistent with previous results on potted beech seedlings (El Kohen et al. 1993) and on branches of beech trees exposed to elevated CO_2 (Dufrêne et al. 1993).

However, evidence of photosynthetic acclimation during the second part of the summer is suggested in Figure 3. The decrease evident at the beginning of August in the maximal rate of photosynthetic O_2 evolution in second flush leaves of plants grown in elevated CO_2 was confirmed by measurements in ambient CO_2.

The acclimation process has often been related to source–sink imbalance (Arp 1991, Stitt 1991). Restricted rooting volume is one of the main factors in photosynthetic acclimation (Thomas and Strain 1991). However, such limitations were unlikely in this study because the experimental saplings were planted in the ground. It has been suggested that species able to initiate new sinks in response to an increase in carbohydrate availability will show no decrease in photosynthetic capacity. This was typically the case for *Quercus alba* (Norby and O'Neill 1989) and *Quercus petraea* (Epron et al. 1994). Because beech saplings are capable of more than one growth flush in a season, they have the potential to utilize a CO_2-induced surplus of carbohydrate in early summer. Because the increase in aboveground biomass slowed strongly at the end of July in beech, whether grown in ambient or elevated CO_2, we postulate that the late-season acclimation of photosynthesis was related to growth cessation.

Conclusion

Our results suggest that effects of elevated CO_2 on leaf photosynthetic capacity are time-dependent during the growing season and that a sustained increase in net CO_2 assimilation over the whole season is unlikely. Because we demonstrated that photosynthetic acclimation to elevated CO_2 may occur in field-grown trees, the time-dependent pattern of the process must be taken into account when attempting to predict ecosystem responses to global changes and the potential for carbon acquisition and sequestration by forest ecosystems.

Acknowledgments

This work was carried out within the framework of an E.C. project ECOCRAFT No. EV57CT920127. We thank Justine Harley for her help with the measurements and are grateful to Bernard Thiebaut for helpful discussions on beech growth and morphology.

References

Arp, W.J. 1991. Effects of source–sink relations on photosynthetic acclimation to elevated CO_2. Plant Cell Environ. 14:869–875.

Boardman, N.K. 1977. Comparative photosynthesis of sun and shade plants. Annu. Rev. Plant Physiol. 28:355–377.

Ceulemans, R. and M. Mousseau. 1994. Effects of elevated atmospheric CO_2 on woody plants. New Phytol. 127:425–446.

Curtis, P.S. and J.A. Teeri. 1992. Seasonal responses of leaf gas exchange to elevated carbon dioxide in *Populus grandidentata*. Can. J. For. Res. 22:1320–1325.

Delucia, E.H., T.W. Sasek and B.R. Strain. 1985. Photosynthesis inhibition after long-term exposure to elevated levels of atmospheric carbon dioxide. Photosynth. Res. 7:175–184.

Dufrêne, E., J.Y. Pontailler and B. Saugier. 1993. A branch bag technique for simultaneous CO_2 enrichment and assimilation measurements on beech (*Fagus sylvatica* L.). Plant Cell Environ. 16:1131–1138.

Eamus, D. and P.J. Jarvis. 1989. The direct effect of increase in the global atmospheric CO_2 concentration on natural and commercial temperate trees and forests. Adv. Ecol. Res. 19:1–55.

El Kohen, A., L. Venet and M. Mousseau. 1993. Growth and photosynthesis of two deciduous forest species at elevated carbon dioxide. Funct. Ecol. 7:480–486.

Ellsworth, D.S. and P.B. Reich. 1992. Leaf mass per area, nitrogen content and photosynthetic carbon gain in *Acer saccharum* seedlings in contrasting forest light environments. Funct. Ecol. 6:432–435.

Epron, D., E. Dreyer, C. Picon and G. Guehl. 1994. Relationship between CO_2 dependent O_2 evolution and photosystem II activity in oak (*Quercus petraea*) trees grown in the field and in seedlings grown in ambient or elevated CO_2. Tree Physiol. 14:725–733.

Epron, D., D. Godard, G. Cornic and B. Genty. 1995. Limitation of net CO_2 assimilation rate by internal resistances to CO_2 transfer in the leaves of two tree species (*Fagus sylvatica* L. and *Castanea sativa* Mill.). Plant Cell Environ. 18:43–51.

Eschrich, W., R. Burchardt and S. Essiamah. 1989. The induction of sun and shade leaves of the European beech (*Fagus sylvatica* L.): anatomical studies. Trees 3:1–10.

Evans, J.R. 1989. Partitioning of nitrogen between and within leaves grown under different irradiances. Austr. J. Plant Physiol. 16:533–548.

Gunderson, C.A., S.D. Wullschleger. 1994. Photosynthetic acclimation in trees to rising atmospheric CO_2: A broader perspective. Photosynth. Res. 39:369–388.

Gunderson, C.A., R.J. Norby and S.D. Wullschleger. 1993. Foliar gas exchange responses of two deciduous hardwoods during 3 years of growth in elevated CO_2: no loss of photosynthetic enhancement. Plant Cell Environ. 16:797–807.

Hiscox, J.D., G.F. Israelstam. 1979. A method for extraction of chlorophyll from leaf tissue without maceration. Can J. Bot. 57:1332–1334.

Houpis, J.L.J., K.A. Surano, S. Cowles and J.H. Shinn. 1988. Chlorophyll and carotenoid concentrations in two varieties of *Pinus ponderosa* seedlings subjected to long-term elevated carbon dioxide. Tree Physiol. 4:187–193.

Mousseau M. and B. Saugier. 1992. The direct effect of increased CO_2 on gas exchange and growth of forest tree species. J. Exp. Bot. 43:1121–1130.

Norby, R.J. and E.G. O'Neill. 1989. Growth dynamics and water use of seedlings of *Quercus alba* L. in CO_2-enriched atmospheres. New Phytol. 111:491–500.

Oswald, H. 1981. Régénération naturelle. Conditions de germination des faînes, de survie et de croissance des semis, densité de semis nécessaire à la réussite d'une régénération naturelle. *In* Le Hêtre. Eds. E. Tessier du Cros, F. Le Tacon, G. Nepveu, J. Pardé, R. Perrin and J. Timbal. INRA, Paris, France, pp 238–239.

Rivière, H. and T. Betsche. 1994. Effects of atmospheric CO_2 enrichment on protein levels, protein phosphorylation, and gene expression in leaves of C_3 plants. *In* Plant Sciences 1994. Eds. Y. de Kouchkovsky, F. Lahrer. SFPV, Paris, France, p 290.

Roy, J., B. Thiebaut and M.A. Watson. 1986. Physiological and anatomical consequences of morphogenetic polymorphism: leaf response to light intensity in young beech trees (*Fagus sylvatica* L.). *In* Colloque International sur l'Arbre, 9–14 Septembre 1985. Naturalia monspeliensia, Montpellier, France, pp 431–449.

Sage, R.F., T.D. Sharkey and J.F. Seemann. 1989. Acclimation of photosynthesis to elevated CO_2 in five C_3 species. Plant Physiol. 89:590–596.

Seemann J.R., T.D. Sharkey, J. Wang and C.B. Osmond. 1987. Environmental effects on photosynthesis, nitrogen use efficiency, and metabolite pools in leaves of sun and shade plants. Plant Physiol. 84:796–802.

Stitt, M. 1991. Rising CO_2 levels and their potential significance for carbon flow in photosynthetic cells. Plant Cell Environ. 14:741–762.

Tans, P.P, I.Y. Ting and T. Takahashi. 1990. Observational constraints on the global atmospheric CO_2 budget. Science 247:1431–1438.

Thiebaut, B. 1982. Observations sur le développement de plantules de Hêtre (*Fagus sylvatica*) cultivées en pépinière, orthotropie et plagiotropie. Can. J. Bot. 60:1292–1303.

Thiebaut, B, B. Comps and F. Plancheron. 1990. Anatomie des feuilles dans les pousses polycycliques du Hêtre européen (*Fagus sylvatica*). Can. J. Bot. 68:2595–2606.

Thomas, J.F. and C.N.B. Harvey. 1983. Leaf anatomy of four species grown under continuous long-term CO_2 enrichment. Bot. Gaz. 144:303–339.

Thomas, R.B. and B.R. Strain. 1991. Root restriction as a factor in photosynthetic acclimation of cotton seedlings grown in elevated carbon dioxide. Plant Physiol. 96:627–634.

Tissue, D.T., R.B. Thomas and B.R. Strain. 1993. Long-term effects of elevated CO_2 and nutrients on photosynthesis and Rubisco in loblolly pine seedlings. Plant Cell Environ. 16:859–865.

Von Caemmerer, S. and G.D. Farquhar. 1981. Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves. Planta 153:376–387.

Wong, S.C. 1990. Elevated atmospheric partial pressure of CO_2 and plant growth. II. Nonstructural carbohydrate content in cotton plants and its effect on growth parameters. Photosynth. Res. 23:171–180.

Wullschleger, S.D., R.J. Norby and D.L. Hendrix. 1992. Carbon exchange rates, chlorophyll content, and carbohydrate status of two forest tree species exposed to carbon dioxide enrichment. Tree Physiol. 10:21–31.

PLANETARY SCIENCE

Shock compression of stishovite and melting of silica at planetary interior conditions

M. Millot,[1,2]* N. Dubrovinskaia,[3] A. Černok,[4] S. Blaha,[4] L. Dubrovinsky,[4] D. G. Braun,[1] P. M. Celliers,[1] G. W. Collins,[1] J. H. Eggert,[1] R. Jeanloz[2]

Deep inside planets, extreme density, pressure, and temperature strongly modify the properties of the constituent materials. In particular, how much heat solids can sustain before melting under pressure is key to determining a planet's internal structure and evolution. We report laser-driven shock experiments on fused silica, α-quartz, and stishovite yielding equation-of-state and electronic conductivity data at unprecedented conditions and showing that the melting temperature of SiO_2 rises to 8300 K at a pressure of 500 gigapascals, comparable to the core-mantle boundary conditions for a 5–Earth mass super-Earth. We show that mantle silicates and core metal have comparable melting temperatures above 500 to 700 gigapascals, which could favor long-lived magma oceans for large terrestrial planets with implications for planetary magnetic-field generation in silicate magma layers deep inside such planets.

Understanding the structure, formation, and evolution of giant planets and extrasolar terrestrial planets (super-Earths) discovered to date requires knowledge of the properties of basic constituents such as iron, magnesium oxide, and silica at the relevant extreme conditions, including pressures of 100s to 1000s of GPa. Melting is arguably the most important process determining the physical and chemical evolution of planetary interiors, as differentiation of a terrestrial planet into a dense metallic core surrounded by rocky mantle and atmosphere proceeds by gravitational separation of a liquid phase (*1*). Moreover, giant impacts during the terminal stages of planetary formation can cause large-scale melting and generate a magma ocean encompassing much of the planet's rocky constituents (*2*, *3*). As mantle viscosity typically increases by more than 10 to 15 orders of magnitude upon solidification (*4*), the potential freezing of this magma ocean would greatly influence the planet's subsequent thermal evolution, geochemistry, and magnetic field.

We used shock compression of fused silica, α-quartz, and stishovite to document the pressure-density-temperature equation-of-state and optical properties (hence, electronic conductivity) of SiO_2. Stishovite's high initial density allowed us to access unprecedented high densities, which extended the experimental melting line of SiO_2 to more than 500 GPa. In combination with melting data for other oxides and iron, the high-pressure measurements provide constraints on the thermal structure and evolution of rocky planets and provide a benchmark for future theoretical (e.g., first-principles molecular dynamics), as well as experimental studies.

We used a TW-power laser pulse to send a strong, but decaying, shock through a planar target assembly (Fig. 1, A and B) (*5*). Nanosecond streaked optical pyrometry (SOP) and Doppler velocity interferometry (VISAR) recorded the shock-front velocity, reflectivity, and thermal emission as a function of time (Fig. 1, C and D). We applied impedance matching to obtain pressure-density data up to 2.5 TPa along the locus of shock (Hugoniot) states of stishovite (Fig. 2A and fig. S7). Then, following the decay of the compression wave as it traveled through the target, we obtained continuous measurements of the temperature T and reflectivity R as a function of the shock velocity U_S along the Hugoniot for three different starting materials: stishovite, α-quartz, and fused silica (initial densities of 4.29, 2.65, and 2.20 g/cm^3) (Fig. 2, B and C). Using pressure-density Hugoniot data to associate shock velocities with pressure, density, and internal energy, we can then relate the measured shock temperatures along the three Hugoniots with their respective pressure, density, and internal energy.

Shock temperature rises with increasing pressure to accommodate the increase of internal energy in excess of the compression work, in the absence of phase transitions or chemical reaction: Kinks in the temperature versus pressure are due to latent heat associated with phase changes. Our data for fused-silica and α-quartz (figs. S8 and S9 and Fig. 3A) match well with previous observations of temperature reversals (*6*–*9*) and sound speed discontinuities upon melting near 70 and 120 GPa (*10*). In addition, we observed a shock temperature plateau at 8300 ± 300 K and 500 ± 30 GPa along the stishovite Hugoniot (Figs. 2B and 3A) that we interpret as melting: an interpretation supported by previous shock data (*6*, *7*, *9*, *11*), static compression experiments at lower pressure (*12*, *13*), as well as qualitative velocimetry signatures that differentiate shock breakouts from solid and liquid silica (fig. S10). We fit the dynamic compression data and the coesite-stishovite–liquid silica triple point from static compression (*14*) (Fig. 3A), using Simon's equation (*5*) obtaining the temperature (T_m) (in Kelvins) as a function of the melting pressure (P_m) (in GPa), $T_m = 1968.5 + 307.8\ P_m^{0.485}$.

The onset of optical reflectivity when silica is shock compressed above ~10,000 K along the stishovite Hugoniot (Fig. 2C) reveals the transition to an electrically conducting state. Electronic conductivities estimated with a Drude model and assuming a strong-scattering regime (Ioffe-Regel

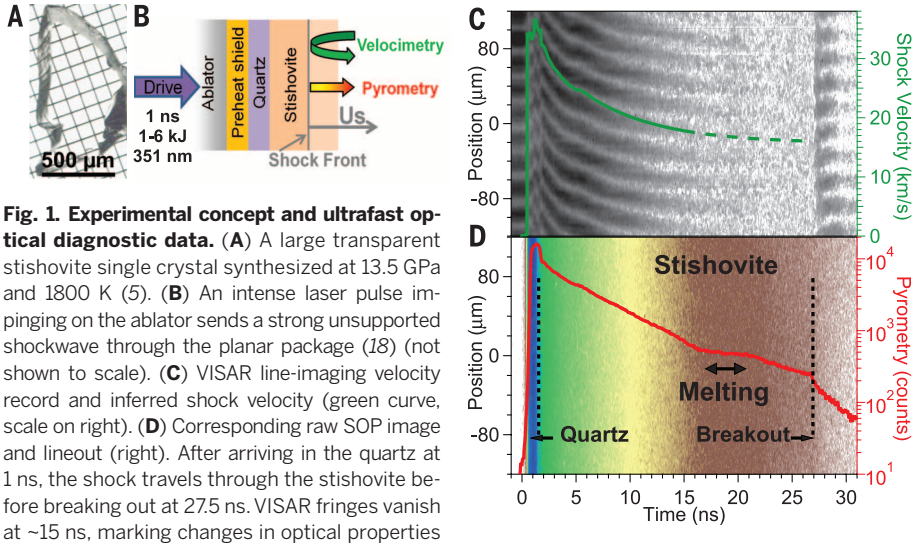


Fig. 1. Experimental concept and ultrafast optical diagnostic data. (**A**) A large transparent stishovite single crystal synthesized at 13.5 GPa and 1800 K (*5*). (**B**) An intense laser pulse impinging on the ablator sends a strong unsupported shockwave through the planar package (*18*) (not shown to scale). (**C**) VISAR line-imaging velocity record and inferred shock velocity (green curve, scale on right). (**D**) Corresponding raw SOP image and lineout (right). After arriving in the quartz at 1 ns, the shock travels through the stishovite before breaking out at 27.5 ns. VISAR fringes vanish at ~15 ns, marking changes in optical properties associated with the threshold in electrical conductivity. Melting is evident from a thermal anomaly attributed to latent heat and seen here as a change in the decay of the Hugoniot temperature and the recovery of VISAR fringes upon breakout into vacuum (*5*).

[1]Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. [2]University of California Berkeley, Berkeley, CA 94720, USA. [3]Material Physics and Technology at Extreme Conditions, Laboratory of Crystallography, University of Bayreuth, 95440 Bayreuth, Germany. [4]Bayerisches Geoinstitut, University of Bayreuth, 95440 Bayreuth, Germany.
*Corresponding author. E-mail: millot1@llnl.gov

limit) increase with temperature to ~5 × 10^5 S/m and approach predicted values for liquid iron at TPa pressures ~10^6 S/m (*15*) (fig. S12). Such high values of electrical conductivity imply that liquid silicates could contribute to dynamo generation of magnetic fields inside large planets (*16*).

Dense liquid silica is also found to have an anomalously high specific heat right above the melting point (fig. S13). The shock temperature along the stishovite Hugoniot increases at a slower rate than predicted by the ANEOS model, which is based on the classical Dulong-Petit limit for the specific heat: $C_V = 3nTk_B$, where n is the number of ions and k_B is the Boltzmann constant (*9*, *17*, *18*). This reveals additional degrees of freedom in the fluid, possibly including both configurational and electronic contributions. The emergence of a complex, polymeric fluid right above the melting temperature may offer one explanation (*9*, *18*).

In combination with other recent measurements (*19*), we therefore conclude that silica and magnesium oxide are solid in the deep interior of icy giants like Uranus and Neptune, as well as—with extrapolation—in the rocky core of Saturn and Jupiter (Fig. 3B). This is particularly interesting as silicates are expected to dissociate

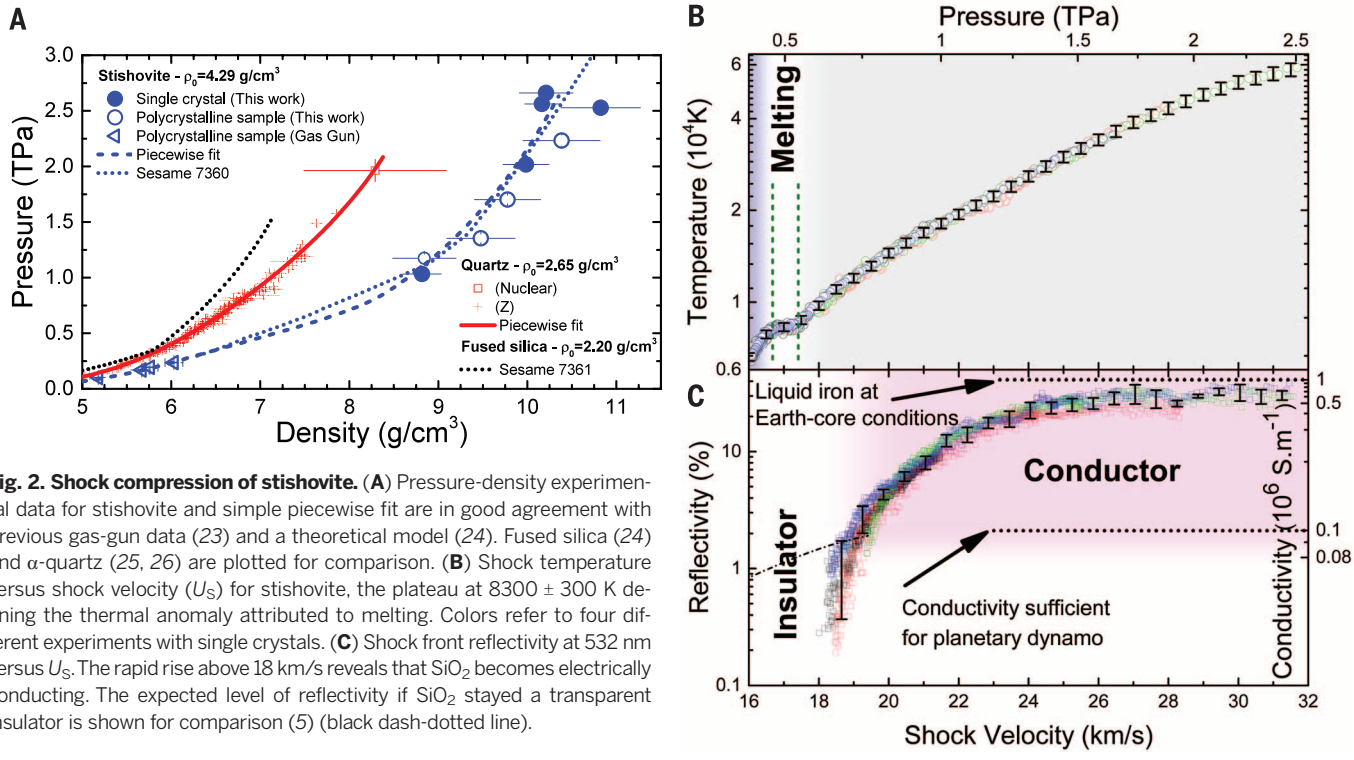


Fig. 2. Shock compression of stishovite. (**A**) Pressure-density experimental data for stishovite and simple piecewise fit are in good agreement with previous gas-gun data (*23*) and a theoretical model (*24*). Fused silica (*24*) and α-quartz (*25*, *26*) are plotted for comparison. (**B**) Shock temperature versus shock velocity (U_S) for stishovite, the plateau at 8300 ± 300 K defining the thermal anomaly attributed to melting. Colors refer to four different experiments with single crystals. (**C**) Shock front reflectivity at 532 nm versus U_S. The rapid rise above 18 km/s reveals that SiO_2 becomes electrically conducting. The expected level of reflectivity if SiO_2 stayed a transparent insulator is shown for comparison (*5*) (black dash-dotted line).

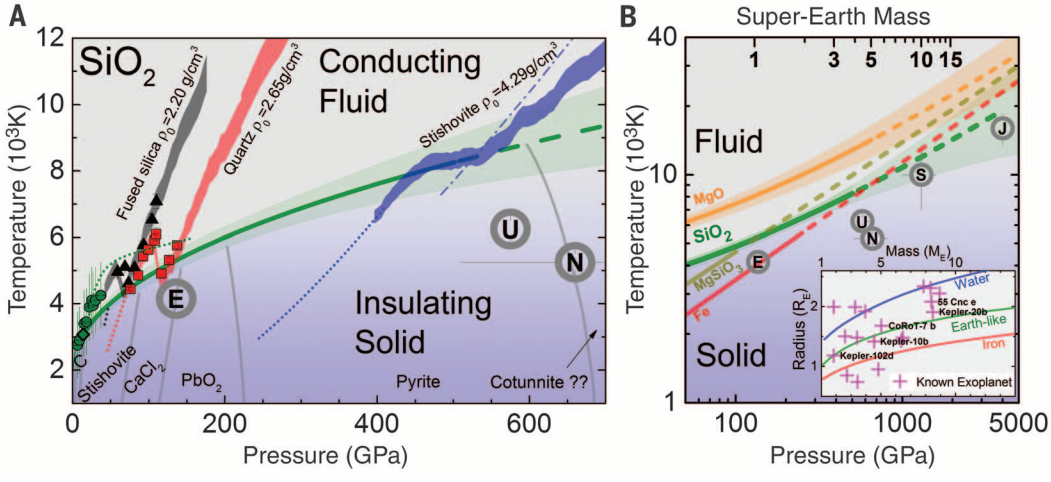


Fig. 3. Melting line of silica at planetary interior conditions determined by shock compression experiments. (**A**) The shock temperature plateau at 500 GPa reveals the crossing of the stishovite Hugoniot (blue) with the silica melting line. The data for α-quartz (red) and fused silica (gray) reproduce the strong temperature reversal upon melting observed with gas-gun experiments (*6*, *7*) (squares and triangles). A Simon fit of the shock and static compression data (*12*, *14*) (diamond, dots) provides a melting line up to TPa pressure (green line and shaded 1 σ uncertainty). Thick gray lines are known and predicted solid-phase boundaries (*27*). Dotted blue, red, and black lines are calculated Hugoniot curves below melting for stishovite (*5*), quartz (*8*), and fused silica (*24*). Predicted fluid properties are shown as the Hugoniot calculated from the ANEOS model (*17*) (blue dash-dotted) and the classical molecular dynamics melting line (*28*) (dotted green). (**B**) New experimental SiO_2 melting line and core mantle boundaries (gray circles) for Earth (E), Uranus (U) (*29*), Neptune (N) (*29*), Saturn (S) (*21*), and Jupiter (J) (*21*); melting lines of MgO, $MgSiO_3$, and Fe (dashed when extrapolated) (*5*) are shown for comparison. Top scale relates the pressure at the core-mantle boundary to the mass of Earth-like large exoplanets (*30*). (Inset) Discovered super-Earths and mass-radius relation for pure iron, water, or Earth-like structure that allow us to identify potentially rocky exoplanets (*5*).

into MgO and SiO_2 at TPa pressures (*20*). Solid cores in these planets would not participate in magnetic field generation and may be more resistant to erosion by the surrounding fluids (*21*).

Because rocky super-Earths are probably molten at the end of their accretion and many of them are expected to have high surface temperature, melting of silicates can regulate planetary evolution through control of internal heat flow. In particular, several mechanisms may tend to pin the core-mantle boundary temperature to the silicate melting line (*4*). If we overlook alloying effects and solidus-liquidus differences, our high-pressure measurements can thus be used to compare the melting behavior of silicates and iron alloys up to the terapascal range (Fig. 3 and fig. S14). As iron melts at lower temperature than silica up to ~0.6 TPa, an Earth-like structure with an iron-rich liquid outer core surrounded by a solid silicate mantle could subsist in a super-Earth up to ~5 Earth masses (M_E), such as exoplanets Kepler 102-d, Kepler 10 b, and Corot 7b (inset in Fig. 3B). In contrast, larger rocky exoplanets, like Kepler 20-b or 55 Cnc e, would likely have an extended basal magma ocean in the lower mantle that could affect their magnetic field and reinforce tidal response (*22*).

We find that molten silicate deep within large planetary mantles, in general, can contribute to magnetic field generation like the iron-rich liquid in the planetary cores. Magnetic fields provide information about the interior dynamics of planets and are potentially observable for exoplanets. Our discovery of conductive liquid silica at high pressures suggests a need to investigate planetary structure models that include multiple layers of conductive fluids.

REFERENCES AND NOTES

1. G. F. Davies, *Geophys. Res. Lett.* **9**, 1267–1270 (1982).
2. R. M. Canup, *Science* **338**, 1052–1055 (2012).
3. L. T. Elkins-Tanton, *Annu. Rev. Earth Planet. Sci.* **40**, 113–139 (2012).
4. L. Stixrude, *Philos. Trans. A. Math. Phys. Eng. Sci.* **372**, 20130076 (2014).
5. Materials and methods are available as supplementary materials on *Science* Online.
6. G. A. Lyzenga, T. J. Ahrens, A. C. Mitchell, *J. Geophys. Res.* **88** (B3), 2431 (1983).
7. M. B. Boslough, *J. Geophys. Res.* **93**(B6), 6477 (1988).
8. S.-N. Luo, T. J. Ahrens, P. D. Asimow, *J. Geophys. Res.* **108** (B9), 2421 (2003).
9. D. G. Hicks *et al.*, *Phys. Rev. Lett.* **97**, 025502 (2006).
10. R. McQueen, *Shock Compression of Condensed Matter–1991* (Elsevier, Amsterdam, 1992).
11. D. G. Hicks *et al.*, *Phys. Plasmas* **12**, 082702 (2005).
12. G. Shen, P. Lazor, *J. Geophys. Res.* **100** (B9), 17699 (1995).
13. F. Wang, Y. Tange, T. Irifune, K.-i. Funakoshi, *J. Geophys. Res.* **117**(B6) (2012) 10.1029/2011JB009100.
14. J. Zhang, R. C. Liebermann, T. Gasparik, C. T. Herzberg, Y. Fei, *J. Geophys. Res.* **98**(B11), 19785 (1993).
15. M. Pozzo, C. Davies, D. Gubbins, D. Alfè, *Nature* **485**, 355–358 (2012).
16. D. J. Stevenson, *Space Sci. Rev.* **152**, 651–664 (2010).
17. H. J. Melosh, *Meteorit. Planet. Sci.* **42**, 2079–2098 (2007).
18. J. H. Eggert *et al.*, *Nat. Phys.* **6**, 40–43 (2009).
19. R. S. McWilliams *et al.*, *Science* **338**, 1330–1333 (2012).
20. K. Umemoto, R. M. Wentzcovitch, P. B. Allen, *Science* **311**, 983–986 (2006).
21. F. González-Cataldo, H. F. Wilson, B. Militzer, *Astrophys. J.* **787**, 79 (2014).
22. Y. Harada *et al.*, *Nat. Geosci.* **7**, 569–572 (2014).
23. S.-N. Luo, *J. Geophys. Res.* **109**(B5), B05205 (2004).
24. G. I. Kerley, "Equations of state for composite materials" (Report KPS99-4, Kerley Publishing Services, 1999).
25. R. Trunin, *Physics-Uspekhi* **37**, 1123–1145 (1994).
26. M. D. Knudson, M. P. Desjarlais, *Phys. Rev. B* **88**, 184107 (2013).
27. T. Tsuchiya, J. Tsuchiya, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1252–1255 (2011).
28. A. B. Belonoshko, L. S. Dubrovinsky, *Geochim. Cosmochim. Acta* **59**, 1883–1889 (1995).
29. N. Nettelmann, R. Helled, J. Fortney, R. Redmer, *Planet. Space Sci.* **77**, 143–151 (2013).
30. F. W. Wagner, N. Tosi, F. Sohl, H. Rauer, T. Spohn, *Astron. Astrophys.* **541**, A103 (2012).

ACKNOWLEDGMENTS

Data are available in the supplementary materials and upon request to millot1@llnl.gov. Authors gratefully acknowledge assistance of A. Correa Barrios, W. Unites, R. Wallace, C. Davis, and J. Emig, Lawrence Livermore National Laboratory (LLNL); R. Paguio, M. Farrell, and A. Nikroo, General Atomics; M. Bonino, D. Harding, and the entire Omega Laser staff at the Laboratory for Laser Energetics (LLE); H. Schulze, T. Boffa Ballaran, A. Audetat, and H. Fischer (Bayreuth University); and discussions with D. K. Spaulding, T. R. Boehly, and J. R. Rygg. Prepared by LLNL under contract DE-AC52-07NA27344. Omega shots were allocated by LLE Laboratory Basic Science program. N.D. thanks the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Federal Ministry of Education and Research (BMBF, Germany) for financial support through the DFG Heisenberg Program and project no. DU 954-8/1 and the BMBF grant no. 5K13WC3 (Verbundprojekt 05K2013, Teilprojekt 2, PT-DESY). Partial support was provided by the U.S. Department of Energy and the University of California, including UC Berkeley's Miller Institute for Basic Research in Science.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/418/suppl/DC1
Materials and Methods
Figures S1 to S15
Tables S1 to S3
References (*31*–*75*)

22 September 2014; accepted 19 December 2014
10.1126/science.1261507

MAGNETIC MATERIALS

Tilt engineering of spontaneous polarization and magnetization above 300 K in a bulk layered perovskite

Michael J. Pitcher,[1] Pranab Mandal,[1] Matthew S. Dyer,[1] Jonathan Alaria,[2] Pavel Borisov,[1]* Hongjun Niu,[1] John B. Claridge,[1]† Matthew J. Rosseinsky[1]†

Crystalline materials that combine electrical polarization and magnetization could be advantageous in applications such as information storage, but these properties are usually considered to have incompatible chemical bonding and electronic requirements. Recent theoretical work on perovskite materials suggested a route for combining both properties. We used crystal chemistry to engineer specific atomic displacements in a layered perovskite, $(Ca_ySr_{1-y})_{1.15}Tb_{1.85}Fe_2O_7$, that change its symmetry and simultaneously generate electrical polarization and magnetization above room temperature. The two resulting properties are magnetoelectrically coupled as they arise from the same displacements.

For many technical applications, crystalline solids must combine distinct properties. For example, in thermoelectrics, thermal and electronic conductivity need to be optimized simultaneously (*1*). This level of structure-property-composition control is particularly difficult when the properties have antagonistic chemical requirements. For example, the synthesis of a single phase combining electrical polarization P and spontaneous magnetization M is challenging to achieve because of the distinct electronic structure requirements for the main mechanisms producing each property (*2*); for example, the closed-shell d^0 Ti^{4+} and s^2 Pb^{2+} cations, which produce polarization in the ferroelectric perovskite oxide $PbZr_{1-x}Ti_xO_3$, do not have the unpaired electrons needed for magnetization (*3*). Efficiently combining these two long-range orders could be useful for multiferroic or magnetoelectric (ME) information storage, which could overcome the drawbacks of ferroelectric memory (slow writing) and magnetic random access memory (high power density) and open the possibility of four-state memory (*4*, *5*) with reduced energy consumption. It is possible to combine these ground states by making composites over a range of length scales between phases that individually have the chemistry and, thus, properties required (*6*, *7*) or by lowering spatial symmetry through the onset of magnetic order at low temperatures (*8*). We use chemical control of the crystal structure of a single-phase material to generate atomic displacements that produce both properties simultaneously in a coupled manner above room temperature (RT).

The chemical incompatibility between P and M arises when the ferroelectricity is driven by a classical gamma point instability (*9*). Recent theoretical work (*10*, *11*) has identified that specific zone-boundary octahedral tilts in an ABO_3

[1]Department of Chemistry, University of Liverpool, Crown Street, Liverpool L69 7ZD, UK. [2]Department of Physics, University of Liverpool, Liverpool L69 7ZE, UK.
*Present address: Department of Physics, West Virginia University, P.O. Box 6315, Morgantown, WV 26506, USA. †Corresponding author. E-mail: m.j.rosseinsky@liv.ac.uk (M.J.R.); j.b.claridge@liv.ac.uk (J.B.C.)

perovskite block coupled to translational symmetry breaking [from A-site order in 1:1 perovskite heterostructures or the shearing of adjacent even-numbered $(ABO_3)_n$ blocks by the interposed AO rock salt layer of the $(AO)(ABO_3)_n$ Ruddlesden-Popper (RP) structure (*12*)] can generate hybrid improper ferroelectricity (HIF) without requiring classical zone-center displacements. The polarization in a HIF arises from noncancellation of

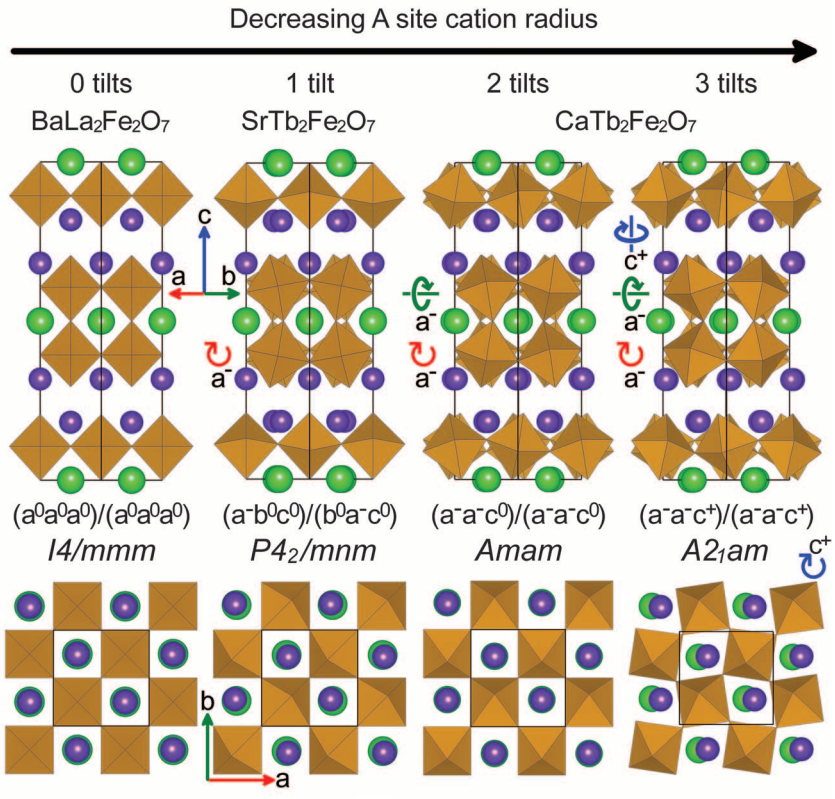


Fig. 1. Crystal chemical engineering of octahedral tilts in the *n* = 2 RP structure. Lowest-energy crystal structures and space group symmetries obtained by DFT calculations are shown for the previously reported compounds $BaLa_2Fe_2O_7$ and $SrTb_2Fe_2O_7$ and the hypothetical compound $CaTb_2Fe_2O_7$ where two candidate structures are shown. Reducing the A-site cation size from $BaLa_2Fe_2O_7$ to $SrTb_2Fe_2O_7$ introduces a single tilt. DFT calculations predict that reducing the A-site cation size further to form $CaTb_2Fe_2O_7$ would produce a polar structure in space group $A2_1am$ with three tilts, which is found to be lower in energy than the zero, one, or two tilt structures shown. Tilt axes are shown with arrows and are labeled according to Glazer notation (*22*), along with the space group symbol of the resulting structure. Alkaline earth atoms are shown in green, lanthanides in purple, and FeO_6 octahedra in brown.

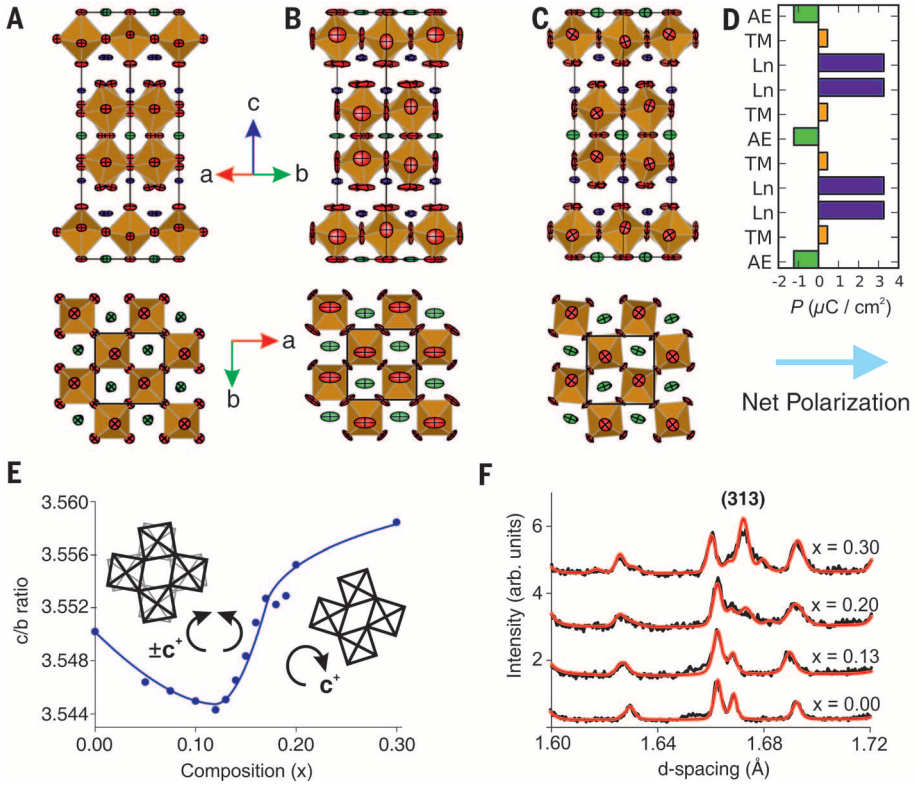


Fig. 2. Onset of an ordered c^+ tilt as a function of composition in the system $[1–x](Ca_ySr_{1–y})_{1.15}Tb_{1.85}Fe_2O_7–[x]Ca_3Ti_2O_7$, *y* = 0.563 and 0.60. (**A**) Refined crystal structure of $Sr_{1.1}Tb_{1.9}Fe_2O_7$ from NPD, which exhibits a single $(a^-b^0c^0)/(b^0a^-c^0)$ tilt in space group $P4_2/mnm$, projected along [110] and [001] with atoms plotted as 99% probability ellipsoids and FeO_6 plotted as solid octahedra. (**B**) Refined structure of *x* = 0.10, *y* = 0.60 from NPD, a nonpolar two-plus-disordered-third tilt system $(a^-a^-c^{<0>})/(a^-a^-c^{<0>})$ in space group *Amam*. The disordered c^+ tilt, indicated as $c^{<0>}$, can be described using either anisotropic displacement parameters highly elongated along the in-plane directions orthogonal to the tilt axis (used here for simple comparison with the other materials shown) or with the split site model shown in fig. S5A. (**C**) Refined structure of *x* = 0.20, *y* = 0.60, which adopts a polar three-tilt system $(a^-a^-c^+)/(a^-a^-c^+)$ in space group $A2_1am$. (**D**) Layer-by-layer polarization of $A2_1am$ *x* = 0.20, *y* = 0.60, calculated from the refined structure in (C), showing a net polarization along the *a* axis (AE, alkaline earth oxide; TM, transition metal oxide; Ln, lanthanide oxide). (**E**) *c*/*b* ratio plotted as a function of composition for series *y* = 0.60 in the range $0.0 \le x \le 0.30$ at RT, with schematic illustration of the ordered c^+ tilt onset shown (inset). (**F**) NPD data showing the emergence of the (313) nuclear Bragg reflection as polarity increases in the series *y* = 0.563. The intensity of this reflection cannot be fitted in *Amam* (fig. S8).

the antiferrodistortive displacements associated with the tilt at the interfaces between structural blocks. The structure is an improper ferroelectric, as the primary order parameter is the zone-boundary octahedral tilting. HIF is rare because the required structural features are difficult to attain; for example, via artificial cation order in thin-film heterostructures (*13*) or in appropriately tilted bulk $n = 2$ RP phases, of which there are only three known examples: the nonmagnetic phases $Ca_3Ti_2O_7$ (*14*), which has switchable polarization (*15*), and $Ca_3Ru_2O_7$ (*16*), as well as the low-temperature magnetically ordered phase $Ca_3Mn_2O_7$ (*17*).

The objective of this study is to combine both polarization and magnetization at RT in a bulk $n = 2$ RP oxide through control of composition to generate the required octahedral tilts in a system with a strongly magnetic B-site sublattice. For a magnetically ordered oxide insulator operating via superexchange, this requirement can be met by selection of Fe^{3+} as the B-site cation because of the good orbital energy match with oxide, enhancing the exchange constant J, and the large $S = 5/2$ spin of the d^5 cations in a mean-field approximation to the magnetic ordering temperature $T_{ord} \approx zJS(S + 1)/k_B$ (where z is the number of nearest neighbor cations and k_B is Boltzmann's constant). Fe^{3+}-O^{2-}-Fe^{3+} interactions are dominantly antiferromagnetic (AFM) when driven by conventional superexchange alone, but antisymmetric exchange between two Fe^{3+} cations without an inversion center between them can create canted spontaneous magnetization via the Dzyaloshinskii-Moriya mechanism, as in the weak ferromagnet α-Fe_2O_3 (*18*, *19*). Imposing the correct combination of octahedral tilts on an Fe^{3+}-based material could simultaneously generate spontaneous magnetization and polarization at RT.

The Fe^{3+} $n = 2$ RP phases $BaLa_2Fe_2O_7$ and $SrLa_2Fe_2O_7$ are untilted and adopt the nonpolar space group *I4/mmm* (Fig. 1). They order antiferromagnetically below 545 (*20*) and 535 K (*21*), respectively, demonstrating that Fe-based $n = 2$ RP oxides offer high-temperature magnetic ordering. Chemical substitutions are required to produce the three coupled octahedral tilts of the polar $(a^-a^-c^+)/(a^-a^-c^+)$ tilt system. Here a^-a^- signifies out-of-phase tilts of equal magnitude about orthogonal axes in the basal plane, and c^+ signifies an in-phase rotation of different magnitude about the stacking axis (*22*), with the brackets referring to the two $(ABO_3)_2$ blocks within the unit cell (Fig. 1). Introduction of a smaller lanthanide (Ln)

Table 1. Energies of the relaxed structures of $CaTb_2Fe_2O_7$ with polar and nonpolar tilt schemes. Energies are calculated by DFT and are quoted relative to the calculated energy in the polar $A2_1am$ structure.

Tilt scheme	Relative energy (eV/formula unit)	BVS (Fe)	BVS (Ca)	BVS (Tb)
$(a^-a^-c^+)/(a^-a^-c^+)$	0	2.80	1.80	3.02
$(a^-a^-c^0)/(a^-a^-c^0)$	0.17	2.99	1.58	2.94
$(a^-b^0c^0)/(b^0a^-c^0)$	0.10	2.99	1.47	2.99
$(a^0a^0a^0)/(a^0a^0a^0)$	0.95	3.34	1.16	2.71

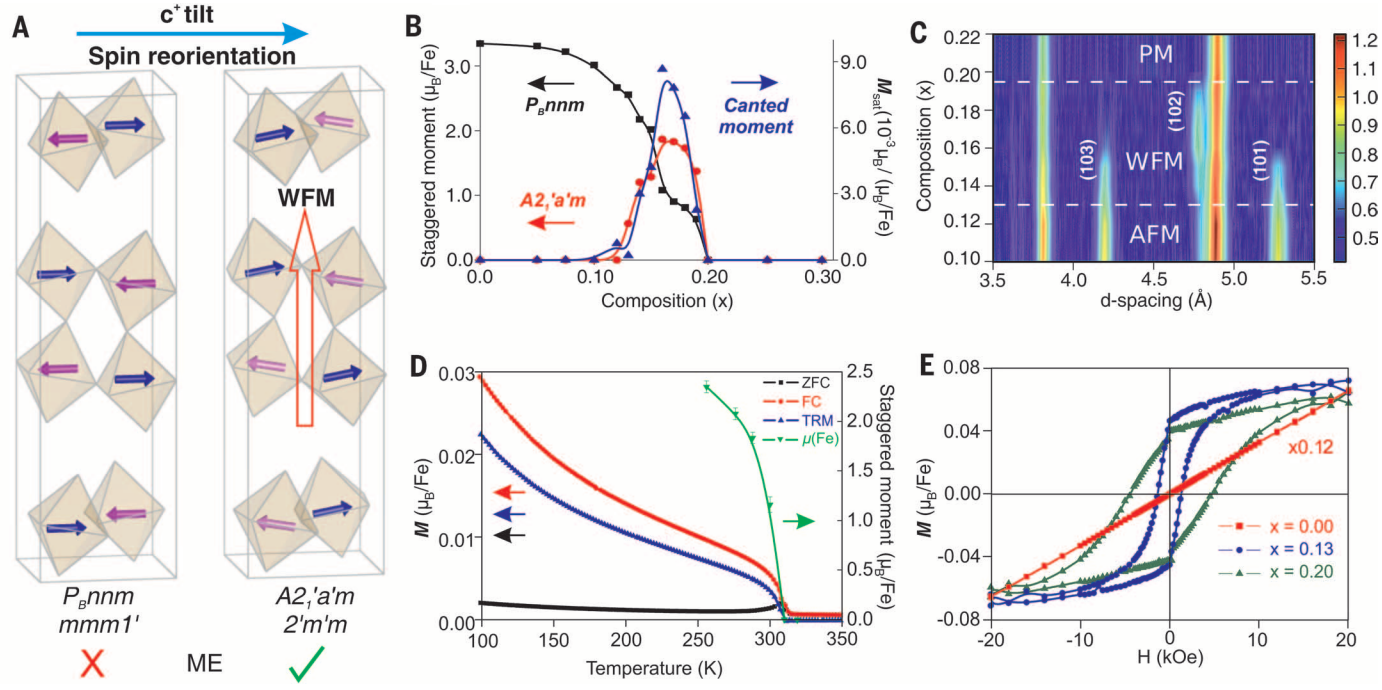


Fig. 3. Simultaneous magnetic spin-reorientation transition and onset of weak ferromagnetism at the transition to a polar crystal structure as a function of composition. (**A**) The P_Bnnm and $A2_1'a'm$ magnetic structures, with the magnetic Fe^{3+} sublattice illustrated as blue and pink arrows. The transition from the P_Bnnm to the $A2_1'a'm$ magnetic structure occurs with retention of the G-type ordering within the perovskite blocks but changes the ordering between the blocks (see also fig. S12) to allow spin canting and ME coupling in the polar materials. (**B**) Refined antiferromagnetically ordered (staggered) Fe moment (*25*) in each magnetic structure from NPD, and canted moment from $M(H)$ loops at 300 K, plotted versus composition at RT for series $0.0 \le x \le 0.30$, $y = 0.60$, showing the onset of the spin-canted $A2_1'a'm$ structure at $x = 0.13$. μ_B, Bohr magneton; M_{sat}, saturated weak ferromagnetic magnetization. (**C**) Contour plot constructed from 10 RT NPD patterns in the compositional range $0.10 \le x \le 0.20$, $y = 0.60$ [general materials diffractometer (GEM), $2\theta = 35°$ detector bank], showing the magnetic transitions from the AFM (P_Bnnm) structure to the WFM ($A2_1'a'm$) structure to paramagnetic (PM) with increasing x. The magnetic peaks are labeled with their *hkl* indices to distinguish them from the nuclear Bragg peaks persisting throughout the composition range. (**D**) $M(T)$ data and refined staggered Fe moment from NPD for the polar phase $x = 0.18$, $y = 0.60$, showing the onset of WFM at $T_N = 310$ K, as indicated by a sharp increase in thermoremanent magnetization (TRM), splitting of the field-cooled (FC) and zero-field cooled (ZFC) data, and the appearance of magnetic Bragg peaks from the $A2_1'a'm$ magnetic structure in the NPD patterns. (**E**) $M(H)$ loops at the ME measurement temperature (60 K) for samples of $x = 0.13$, 0.20 ($y = 0.563$), with linear contribution subtracted (fig. S25), and nonsubtracted data with an applied scale factor of 0.12 plotted for $x = 0.0$, $y = 0.563$.

cation induces rotation of the octahedral network to reduce the A-site coordination number; for instance, $SrTb_2Fe_2O_7$ was reported as a highly distorted $n = 2$ RP phase (*23*) that exhibits a Néel temperature (T_N) of 542 K (*24*), but reinvestigation of this compound by density functional theory (DFT) calculations predicts a conventional single octahedral rotation in the nonpolar $(a^-b^0c^0)/(b^0a^-c^0)$ tilt scheme (Fig. 1). The synthetically accessible phase was $Sr_{1.1}Tb_{1.9}Fe_2O_7$ (figs. S1 and S2) (*25*), and its refined crystal structure (Fig. 2A) adopts the tilt scheme predicted by DFT. The generation of a polar structure requires two further tilts to yield the $(a^-a^-c^+)/(a^-a^-c^+)$ tilt system, but this is not accessible with any Sr/Ln combination within the stability limits of the $n = 2$ RP structure (*25*).

Tilting in perovskites is driven by the need to compensate for underbonding caused by reduction of the A-site cation size, so we performed DFT calculations to evaluate the potential effect of substituting Ca^{2+} for Sr^{2+} in $SrTb_2Fe_2O_7$ to afford $CaTb_2Fe_2O_7$. The calculations show that Ca^{2+} is underbonded with less than three tilts [see bond valence sums (BVSs) in calculated structures (Table 1)], so introduction of Ca^{2+} should drive polarization by inducing a third (c^+) tilt. However, the calculations also indicate that it may be necessary to compensate the loss of bonding at Fe^{3+} to stabilize the polar $A2_1am$ structure (Fig. 1), which must be done without destroying the magnetism. The Tb^{3+} BVS is satisfied with one tilt but, unlike the Fe^{3+} BVS, it is not substantially affected by the key third (c^+) tilt.

We next explored the synthesis of three-tilt polar $n = 2$ RP phases. $CaTb_2Fe_2O_7$ itself could not be synthesized (because of phase separation to the $n = 1$ RP phase plus perovskite), but additional tilts were induced in the series $(Sr_{1-y}Ca_y)_{1.15}Tb_{1.85}Fe_2O_7$ for $0.55 \le y < 0.65$ (fig. S3). Neutron powder diffraction (NPD) measurements (figs. S6 and S7), which are highly sensitive to oxygen displacements, showed that these materials exhibited the two a^- tilts shown in Fig. 1 plus a disordered c^+ tilt (*26*), validating the synthetic strategy of introducing this in-phase tilt by Ca^{2+} substitution (Fig. 2B). The disordered nature of the c^+ tilt produced nonpolar *Amam* symmetry at RT. The *Amam* crystal structure was unchanged on cooling to 40 K. The magnetic ordering exhibited by this series was purely antiferromagnetic and persisted to high temperatures—for instance, the $y = 0.60$ member $(Ca_{0.60}Sr_{0.40})_{1.15}Tb_{1.85}Fe_2O_7$ has a T_N of 525 K (figs. S7 and S13).

The presence of the disordered third tilt suggests that the *Amam* structure in the $(Sr_{1-y}Ca_y)_{1.15}Tb_{1.85}Fe_2O_7$ series is near the target polar instability. Further substitution of Ca^{2+} into this series is not possible beyond $y = 0.65$ (fig. S3), so a different route was needed to order the c^+ tilt and access the polar structure. We coupled the further substitution of Ca^{2+} for Sr^{2+} with the replacement of Fe^{3+} ($r = 0.645$Å) with Ti^{4+} ($r = 0.605$Å) (*27*) to stabilize the structure by compensating for the underbonding at the B site that the third tilt brings (Table 1). We synthesized two series of solid solutions between purely AFM $[1-x](Ca_ySr_{1-y})_{1.15}Tb_{1.85}Fe_2O_7$ (with $y = 0.60$ or 0.563) and nonmagnetic $[x]Ca_3Ti_2O_7$ (which adopts the polar three-tilt structure), which formed stable $n = 2$ RP phases across the range $0.0 \le x \le 0.30$. Examination of the variation of the lattice parameters with x at RT showed a well-defined minimum and gradient change in the c/b ratio at $x = 0.13$ for both the $y = 0.60$ series (Fig. 2E) and the $y = 0.563$ series (fig. S22), suggesting that a compositionally driven structural phase transition occurred. Refinement of NPD data (fig. S4) shows that at and beyond this minimum, the materials with $x > 0.13$ adopt the targeted polar $A2_1am$ RT structure because of the onset of an ordered c^+ tilt, described by the X_2^+ distortion mode (Fig. 2E, inset, and fig. S11) and accompanied by a marked increase in the intensity of the (313) nuclear reflection (Fig. 2F). The refined

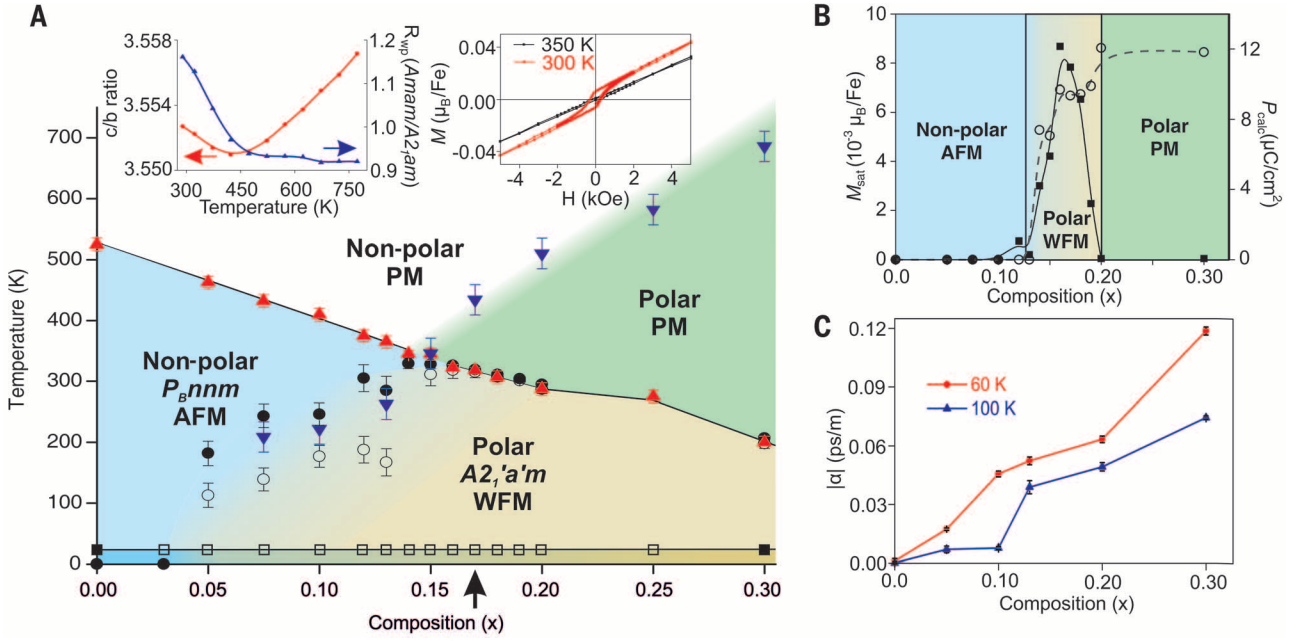


Fig. 4. Phase diagram of the series $0.0 \le x \le 0.30$, $y = 0.60$ and the occurrence of ME coupling in the polar and weak ferromagnetic region. (**A**) Structural-magnetic phase diagram showing dependence of crystal structure (polarity), magnetic structure, and magnetization on composition and temperature in the series $[1-x]$ $(Ca_ySr_{1-y})_{1.15}Tb_{1.85}Fe_2O_7$–$[x]Ca_3Ti_2O_7$ with $y = 0.60$. Red triangles, T_N from NPD (figs. S13 to S15); blue inverted triangles, T_{c+} from *c*/*b* ratio plots (fig. S16); black circles, T_{wfm} (onset WFM) from divergence of field-cooled and zero-field cooled $M(T)$ plots (figs. S17 to S19); open circles, T_{max} from maxima in $|\partial M(T)/\partial T|$ plots (figs. S17 to S19); black squares, Tb^{3+} magnetic ordering from NPD; open squares, Tb^{3+} magnetic ordering from $M(T)$ plots. Error bars denote the widths of the observed peaks (T_{c+}, T_{wfm}, and T_{max}) and the temperature resolution of the NPD patterns (T_N). Four distinct regions are marked by color blocks that correspond to PM/nonpolar (white), PM/polar (green), $A2_1'a'm$/WFM/polar (orange); and P_Bnnm/AFM/nonpolar (blue). Insets show the polar-nonpolar transition at 450 K by the temperature dependence of the lattice parameters (*c*/*b* ratio), the ratio of weighted-profile R-factor (R_{wp}) values for *Amam* and $A2_1am$ models, and the presence of WFM at 300 K by $M(H)$ loop for $x = 0.17$ (marked with a black arrow on the *x* axis). (**B**) Cross section of the phase diagram (A) at 300 K: Saturated magnetic moment per Fe from $M(H)$ loops (fig. S21) (black squares, with solid line as a guide to the eye) plotted with calculated polarization (P_{calc}) of the refined crystal structure from NPD (open circles, with dashed line as a guide to the eye) showing the simultaneous emergence of magnetization and polarization as *x* increases, with regions colored to correspond with those mapped in (A). (**C**) Linear ME susceptibility α versus composition at 60 and 100 K for the series $0.0 \le x \le 0.30$, $y = 0.563$, showing that the ME coupling increases with polarization.

polar RT structure of $x = 0.20$, $y = 0.60$ with the $(a^-a^-c^+)/(a^-a^-c^+)$ tilt system is shown in Fig. 2C, alongside the polarization arising from the three tilts (Fig. 2D), which is amplified by the observed A-site cation ordering because the smaller and more highly charged Tb^{3+} cations preferentially occupy the lower coordination number A sites in the rock salt layer (tables S1 and S2). The RT NPD refinement of $x = 0.17$, $y = 0.60$, which is polar and weakly ferromagnetic up to 315 K, is shown in fig. S10. For $0.05 \le x \le 0.13$, the ordering of the c^+ tilt occurs on cooling from RT, with the structural transition temperature T_{c+} signaled by the same minimum in c/b (figs. S16 and S20).

All of the materials ($0.0 \le x \le 0.30$) in both the $(Ca_{0.60}Sr_{0.40})_{1.15}Tb_{1.85}Fe_2O_7$ ($y = 0.60$)– and $(Ca_{0.563}Sr_{0.437})_{1.15}Tb_{1.85}Fe_2O_7$ ($y = 0.563$)–based series order magnetically. The magnetic structure was directly affected by the chemically induced polar c^+ tilt. Materials adopting the nonpolar *Amam* crystal structure display AFM order in the magnetic space group P_Bnnm, which has point symmetry *mmm1′* (Fig. 3A and fig. S12). This magnetic symmetry forbids weak ferromagnetism (*28*) confirmed by the observation of linear $M(H)$ (where H is magnetic field strength) isotherms for phases adopting this structure (Fig. 3E and fig. S21). The transition into the polar $A2_1am$ crystal structure driven by the c^+ tilt changed the magnetic structure, signaled by the appearance of the strong (102) magnetic Bragg reflection (Fig. 3C). The tilt drove reorientation of the spins by 180° between the two perovskite blocks in the unit cell, affording magnetic space group $A2_1'a'm$ (Fig. 3A and fig. S12). The point symmetry *2′m′m* then permitted weak ferromagnetic (WFM) canting of the Fe^{3+} moments in these polar materials (*28*), demonstrated by nonlinear, hysteretic $M(H)$ isotherms for phases adopting this magnetic structure (Fig. 3E and fig. S21). The onset of weak ferromagnetism in the dc magnetization measurements coincided with the appearance of the $A2_1'a'm$ magnetic structure (Fig. 3B for $y = 0.60$, and figs. S23 and S24 for $y = 0.563$). The Bragg scattering signaling long-range bulk magnetic order in the polar $A2_1am$ crystal structure was accompanied simultaneously by a sharp increase in $M(T)$ (where T is temperature) because of the spin canting (Fig. 3D). These observations both show that the observed magnetization was determined by the bulk magnetic structure of the $n = 2$ RP phase: The weak ferromagnetism was produced by canting of the moments in the polar structure due to the c^+ tilt. For compositions with $0.13 < x < 0.20$, $y = 0.60$, the two magnetic structures coexist at RT, with the canted magnetic structure becoming dominant when $x > 0.15$ (Fig. 3B).

We plot T_N, T_{c+}, and temperature dependence of the magnetic structures determined by neutron diffraction (figs. S13 to 16) together with the magnetization onset temperatures (T_{wfm} and T_{max}) from dc $M(T)$ scans (figs. S17 to 19) to produce the phase diagram in Fig. 4A, which demonstrates the interplay between c^+ tilt, magnetization, and magnetic structure. The structural and magnetic properties of the $y = 0.563$ series are consistent with those of the $y = 0.60$ series (figs. S9 and S22 to 24). T_N decreases with x from 525 K in $x = 0.0$, $y = 0.60$, as nonmagnetic Ti^{4+} substitutes for Fe^{3+}, whereas T_{c+} increases because of the higher Ca^{2+} content to 693 K for $x = 0.30$, $y = 0.60$. These opposing trends produce four regions in the phase diagram, with the nature of the magnetic order controlled by the tilt: Nonpolar AFM or polar WFM ground states are accessible, but there is no polar AFM state. The coexistence of weak ferromagnetism and polarization can persist to 330 K in the $y = 0.60$ series, with nonzero remanent magnetization above RT for the polar compositions between $0.13 < x < 0.20$ (Fig. 4B). The c^+ tilt thus produces both structurally detectable electrical polarization and spontaneous magnetization at RT for the compositions $0.13 < x < 0.20$, $y = 0.60$.

The parent composition $x = 0.0$, $y = 0.60$ is located in the nonpolar AFM region of Fig. 4A and does not show linear ME coupling (fig. S28), consistent with the magnetic and crystal symmetries determined above (*28*). The symmetry of the WFM polar structure region permits linear ME coupling [ME susceptibility $\alpha = \mu_0 M/E$ (where E is electric field strength) (*25*, *28*)], which was observed in pellets of both series in this region of the phase diagram (fig. S28 for $y = 0.60$ and figs. S29 and S30 for $y = 0.563$). ME experiments were conducted at 60 and 100 K, thus avoiding artifacts caused by leakage currents (fig. S27). A linear ME coupling was exhibited by all $x \neq 0$ compositions in the polar WFM region of the phase diagram (Fig. 4C). NPD refinements and $M(H)$ loops at the ME measurement temperatures show that these compositions all exhibited the same polar crystal structures and canted magnetic structures as found at RT (figs. S22, S23, and S25). The size of the ME coupling increased with the magnitude of the c^+ tilt, despite the complex composition-dependence of the saturated Fe moment: At 60 K the competition between increased canting and dilution of the Fe sublattice produces a maximum saturated moment at $x \approx 0.13$ (fig. S26), whereas the maximum α was observed at $x = 0.30$ where the c^+ tilt amplitude is greatest (Fig. 4C). This increase in α with x shows that the enhanced tilt (T_{c+} increased with x), which induced both polarization and magnetization, overcame the reduced strength of the magnetic exchange (T_N decreased with x) and controlled the ME coupling, in accordance with the x dependence of the structural and magnetic phase boundaries seen in Fig. 4A.

A single set of cooperative atomic motions (the c^+ tilt) was imposed by designed chemical substitution into the parent $n = 2$ Ruddlesden-Popper $A_3B_2O_7$ structure to produce RT polarization and magnetization simultaneously. As these two often antagonistic order parameters both arise from the chemically driven tilting of the oxygen network, ensuring that the resulting polarization is large enough to be structurally detectable, their coupling is observed in the ME response, although the magnetization arises from the Fe 3d electrons, whereas the polarization is produced by relative displacement of all of the constituent ions. Ferroelectric switching of the polarization will require higher electric fields than we access here, enabled by ceramic engineering to minimize loss (*29*). Using a strongly magnetic B-site sublattice along with the A-site composition to control the tilt structure, we synthesized materials that are both polar and weakly ferromagnetic at up to 330 K over a range of compositions, revealing the potential of crystal-chemically–driven symmetry control to generate multiple coupled functions.

REFERENCES AND NOTES

1. L.-D. Zhao *et al.*, *Nature* **508**, 373–377 (2014).
2. H. Sakai *et al.*, *Phys. Rev. Lett.* **107**, 137601 (2011).
3. N. A. Hill, *J. Phys. Chem. B* **104**, 6694–6709 (2000).
4. M. Bibes, A. Barthélémy, *Nat. Mater.* **7**, 425–426 (2008).
5. C. A. F. Vaz, *J. Phys. Condens. Matter* **24**, 333201 (2012).
6. D. M. Evans *et al.*, *Nat. Commun.* **4**, 1534 (2013).
7. J. van Suchtelen, *Philips Res. Rep.* **27**, 28–37 (1972).
8. N. J. Perks, R. D. Johnson, C. Martin, L. C. Chapon, P. G. Radaelli, *Nat. Commun.* **3**, 1277 (2012).
9. M. Ghita, M. Fornari, D. J. Singh, S. V. Halilov, *Phys. Rev. B* **72**, 054114 (2005).
10. N. A. Benedek, C. J. Fennie, *Phys. Rev. Lett.* **106**, 107204 (2011).
11. J. M. Rondinelli, C. J. Fennie, *Adv. Mater.* **24**, 1961–1968 (2012).
12. K. S. Aleksandrov, J. Bartolome, *J. Phys. Condens. Matter* **6**, 8219–8235 (1994).
13. J. Alaria *et al.*, *Chem. Sci.* **5**, 1599–1610 (2014).
14. M. M. Elcombe *et al.*, *Acta Crystallogr. B* **47**, 305–314 (1991).
15. Y. S. Oh, X. Luo, F.-T. Huang, Y. Wang, S.-W. Cheong, http://arxiv.org/abs/1411.6315.
16. Y. Yoshida *et al.*, *Phys. Rev. B* **72**, 054412 (2005).
17. M. V. Lobanov *et al.*, *J. Phys. Condens. Matter* **16**, 5339–5348 (2004).
18. I. E. Dzialoshinskii, *Sov. Phys. JETP* **5**, 1259–1272 (1957).
19. T. Moriya, *Phys. Rev.* **120**, 91–98 (1960).
20. D. Samaras, A. Collomb, *Solid State Commun.* **16**, 1279–1284 (1975).
21. N. N. M. Gurusinghe *et al.*, *Mater. Res. Bull.* **48**, 3537–3544 (2013).
22. A. M. Glazer, *Acta Crystallogr. B* **28**, 3384–3392 (1972).
23. D. Samaras, A. Collomb, J. C. Joubert, *J. Solid State Chem.* **7**, 337–348 (1973).
24. D. Hanžel, *J. Magn. Magn. Mater.* **5**, 243–246 (1977).
25. Materials and methods are available as supplementary materials on *Science* Online.
26. Crystallographic information files may be obtained from Fachinformationszentrum Karlsruhe, 76344 Eggenstein-Leopoldshafen, Germany (e-mail: crysdata@fiz-karlsruhe.de, www.fiz-karlsruhe.de/request_for_deposited_data.html) on quoting the appropriate Crystal Structure Depot number tabulated in table S4.
27. R. D. Shannon, C. T. Prewitt, *Acta Crystallogr. B* **25**, 925–946 (1969).
28. H. Schmid, *J. Phys. Condens. Matter* **20**, 434201 (2008).
29. I. M. Reaney *et al.*, *Appl. Phys. Lett.* **100**, 182902 (2012).

ACKNOWLEDGMENTS

This work is funded by the European Research Council (ERC grant agreement 227987 RLUCIM). Synchrotron x-ray diffraction work was carried out with the support of Diamond Light Source (UK). Time-of-flight neutron diffraction work was carried out with the support of ISIS Spallation Source (UK). We thank C. Tang, S. Thompson, J. Parker, and P. Adamson for assistance on beamline I11 (Diamond Light Source); W. Kockelmann and D. Keen for assistance on GEM (ISIS); A. Daoud-Aladine and K. Knight for assistance on the high-resolution powder diffractometer (ISIS); and R. Smith for assistance on POLARIS (ISIS). Computational resources were provided and funded by the N8 consortium and Engineering and Physical Sciences Research Council (EPSRC) (EP/K000225/1) and the UK's national high-performance computing service HECToR through the UK's High-Performance Computing Materials Chemistry Consortium funded by EPSRC (EP/L000202). M.J.R. is a Royal Society Research Professor.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/420/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S30
Tables S1 to S4
References (*30*–*44*)

3 October 2014; accepted 12 December 2014
10.1126/science.1262118

MATERIALS CHEMISTRY

Composition-matched molecular "solders" for semiconductors

Dmitriy S. Dolzhnikov,[1,2] Hao Zhang,[1,2] Jaeyoung Jang,[1,2] Jae Sung Son,[1,2] Matthew G. Panthani,[1,2] Tomohiro Shibata,[3,4] Soma Chattopadhyay,[3,4] Dmitri V. Talapin[1,2,5]*

We propose a general strategy to synthesize largely unexplored soluble chalcogenidometallates of cadmium, lead, and bismuth. These compounds can be used as "solders" for semiconductors widely used in photovoltaics and thermoelectrics. The addition of solder helped to bond crystal surfaces and link nano- or mesoscale particles together. For example, CdSe nanocrystals with $Na_2Cd_2Se_3$ solder was used as a soluble precursor for CdSe films with electron mobilities exceeding 300 square centimeters per volt-second. CdTe, PbTe, and Bi_2Te_3 powders were molded into various shapes in the presence of a small additive of composition-matched chalcogenidometallate or chalcogel, thus opening new design spaces for semiconductor technologies.

Two pieces of metal, such as copper wires, can be mechanically and electrically connected by using soldering, a process in which metallic items are joined together by introducing a filler metal (solder) with lower melting point (*1*). Soldering and related processes (such as brazing) are very useful in many areas, from microelectronics to plumbing. Unlike the case for metals, there are no established methods for joining semiconductor pieces under mild conditions without disrupting semiconducting properties at the joint. A "semiconductor solder" could affect numerous fields, including printable electronics, photovoltaics, and thermoelectrics, because it would add new techniques for device integration.

The electronic properties of semiconductor interfaces are much more sensitive to impurities and structural defects than are metals. Metal-metal contacts always show ohmic behavior, whereas semiconductor-semiconductor interfaces are more complex. Misalignment of the Fermi energy levels or trapping charge carriers at the interface creates a Schottky barrier (*2*). An ideal "semiconductor solder" should be constituted of a precursor that, upon mild heat treatment, forms a semiconducting material structurally and compositionally matched to the bonded parts. In this study, we designed chalcogenidometallate "solders" for technologically important II-VI, IV-VI, and V-VI semiconductors, including CdSe, CdTe, $Hg_xCd_{1-x}Te$, PbTe, $(Bi_xSb_{1-x})_2Te_3$, and Bi_2Te_3.

The use of chalcogenidometallates as soluble precursors for inorganic semiconductors was first introduced by Mitzi *et al.*, who used the ability of N_2H_4 to dissolve metal chalcogenides in the presence of elemental chalcogens (*3*), forming soluble hydrazinium chalcogenidometallates. These species cleanly decompose back into semiconducting metal chalcogenides upon heating. This simple approach works very well for the chalcogenides of Cu, Ga, Ge, In, Sb, Sn, and mixed phases such as $Cu(In_{1-x}Ga_x)Se_2$ and $Cu_2ZnSn(S,Se)_4$ (*4*–*6*) but showed no success for Cd-, Pb-, and Bi-chalcogenides (*7*, *8*), which are among the most widely used binary and ternary semiconductors. Here, we propose the origin of this limitation and a general solution, which substantially expands the list of solution-processed semiconductors.

In the above process, the formation of a soluble chalcogenidometallate begins with the reduction of an elemental chalcogen (Ch = S, Se, or Te) by N_2H_4 following Eq. 1 (*3*, *4*)

$$2n\mathrm{Ch} + 5\mathrm{N_2H_4} \rightarrow 4\mathrm{N_2H_5^+} + 2\mathrm{Ch}_n^{2-} + \mathrm{N_2} \quad (1)$$

The Ch_n^{2-} reacts with electron-deficient metal centers at the surface of a solid metal chalcogenide (M_xCh_y), generating soluble $M_xCh_{(y+m)}^{2m-}$ chalcogenidometallate ions:

$$2\mathrm{nM}_x\mathrm{Ch}_y + 2m\mathrm{Ch}_n^{2-} + 5(n-1)m\mathrm{N_2H_4} \rightarrow 2n\mathrm{M}_x\mathrm{Ch}_{(y+m)}^{2m-} + (n-1)m\mathrm{N_2} + 4(n-1)m\mathrm{N_2H_5^+} \quad (2)$$

The progression of reaction in Eq. 2 is determined by the balance between the lattice energy of the metal chalcogenide and the free energy of formation and solvation of the chalcogenidometallate complex. Reactivity of Ch_n^{2-} is highest at $n = 1$ and decreases as n increases. The reducing potential of N_2H_4 in the reaction in Eq. 1 is not sufficient to bring Se and Te into their most reactive Ch^{2-} state. X-ray absorption near-edge spectroscopy (XANES) measurements show that Eq. 1 generates soluble Se or Te species with an oxidation state near zero, which is equivalent to large n (fig. S1) (*9*). To increase the driving force for Eq. 2, we used A_2Se or A_2Te (A = Na, K, and Cs) instead of elemental chalcogens and found that for $n = 1$, Eq. 2 proceeds smoothly for a number of metal chalcogenides previously considered unreactive, including CdSe, CdTe, PbS, PbSe, PbTe, Bi_2S_3, Bi_2Se_3, Bi_2Te_3, $(Bi_xSb_{1-x})_2Te_3$, and some others, which are summarized in (*9*). These reactions could also be conveniently carried out in one pot—for example, by adding alkali metal hydride to the stoichiometric mixture of CdTe and Te in N_2H_4: CdTe + Te + 2NaH → Na_2CdTe_2 + gaseous products (*9*).

The structure of $K_2CdTe_2{\cdot}2N_2H_4$ crystallized from the reaction mixture is shown in Fig. 1A. It contains molecular chains built of edge-sharing, slightly distorted [$CdTe_4$] tetrahedrons with Te-Cd-Te angles of 99.4° and 119.1° (*9*). The bonds between Cd and Te are 2.81 and 2.83 Å in length, which are similar to that of zinc blende CdTe (2.81 Å). This structural motif with one-dimensional (1D) "molecular wires" has apparently not been reported for chalcogenidocadmates (*10*, *11*). In the crystal lattice, $[CdTe_2^{2-}]_\infty$ chains are separated with N_2H_4 molecules, which leads to the facile solubility of the compound in various solvents, up to 600 mg/mL in N_2H_4 (Fig. 1B). The solubility of ditellurometallates could be further tailored by means of cation exchange (*9*). For example, Na^+ or K^+ was exchanged for alkylammonium cations such as didodecyldimethyl ammonium (DDA^+) or tetraethyl ammonium (NEt_4^+), providing good solubility of DDA_2CdTe_2 in toluene and of $(NEt_4)_2CdTe_2$ in acetonitrile (CH_3CN) or N,N-dimethylfomamide (DMF), respectively (Fig. 1B). Exchange of K^+ with hydrazinium cations (*9*) could be used to facilitate thermal decomposition of the resulting hydrazinium salts.

X-ray diffraction and extended x-ray absorption fine structure (EXAFS) studies (figs S2 and S3 and tables S1 and S2) (*9*) of ditellurocadmates suggest that, in solution, $[CdTe_2^{2-}]_\infty$ chains can exist in the equilibrium forms outlined in Fig. 1C. In a strongly coordinating solvent such as N_2H_4, the equilibrium shifted toward the $[CdTe(\mu\text{-}Te)]_n^{2n-}$ structure, where each Cd atom has 3 Te neighbors and Te has on average 1.5 Cd atoms in the first coordination shell. On the other hand, weakly coordinating solvents such as CH_3CN shifted the equilibrium toward the structure with two bridging Te atoms per $[CdTe_2]^{2-}$ unit. These variations in coordination environment were reflected by fully reversible shifts of the absorption bands in mixtures of N_2H_4 and CH_3CN with various solvent ratios (fig. S4) (*9*).

Reacting Na_2CdTe_2 with a stoichiometric amount of $CdCl_2$ in N_2H_4 formed a white amorphous gel, further referred to as "CdTe-gel" (Fig. 1B), which is an important addition to the chalcogel family (*12*). The gel was easily separated from its NaCl by-product by means of centrifugation and then redispersed in *N*-methylformamide (NMF) or N_2H_4, forming a stable gel of $[CdTe_2]^{2-}$ polyions cross-linked with Cd^{2+}. Upon heating above 250°C, the CdTe-gel transformed into crystalline CdTe, as shown in fig. S5 (*9*).

$[CdTe_2^{2-}]_\infty$ ions, similar to other chalcogenidometallates (*13*), were used as capping ligands for colloidal nanocrystals (NCs). $[CdTe_2]^{2-}$-capped CdTe quantum dots showed bright band-edge photoluminescence (Fig. 1B and fig. S6) (*9*). The

[1]Department of Chemistry, University of Chicago, Chicago, IL 60637, USA. [2]James Franck Institute, University of Chicago, Chicago, IL 60637, USA. [3]Materials Research Collaborative Access Team (MRCAT) Sector 10 Advanced Photon Source, Argonne National Laboratory, Argonne, IL 60439, USA. [4]Department of Physics, Advanced Materials Group, Illinois Institute of Technology, Chicago, IL 60616, USA. [5]Center for Nanoscale Materials, Argonne National Laboratory, Argonne, IL 60439, USA.
*Corresponding author. E-mail: dvtalapin@uchicago.edu

presence of strong emission after the ligand exchange suggests that $[CdTe_2]^{2-}$ ions did not introduce fast recombination centers at the CdTe surface.

Following similar approaches, we synthesized and characterized a series of soluble Cd-, Pb-, and Bi chalcogenidometallates and chalcogels (*9*). These chalcogenidometallates could be dispersed in polar solvents, including N_2H_4, ethylenediamine, NMF, dimethyl sulfoxide (DMSO), formamide (FA), DMF, and N_2H_4/water mixtures (*9*). Such chemical versatility allowed for the design of compositionally matched "solders" for practical semiconductors. These solders were used in various capacities—for example, to consolidate nano- or mesoscopic grains or to bond single-crystal wafers. In one example, we applied $Na_2Cd_2Se_3$ solder to CdSe NCs: Colloidal 4.7-nm CdSe NCs capped with *n*-octadecylphosphonic acid (ODPA) were ligand-exchanged with $[Cd_2Se_3]^{2-}$ ions (*9*). The infrared (IR) absorption spectra confirmed a complete replacement of original ODPA ligands

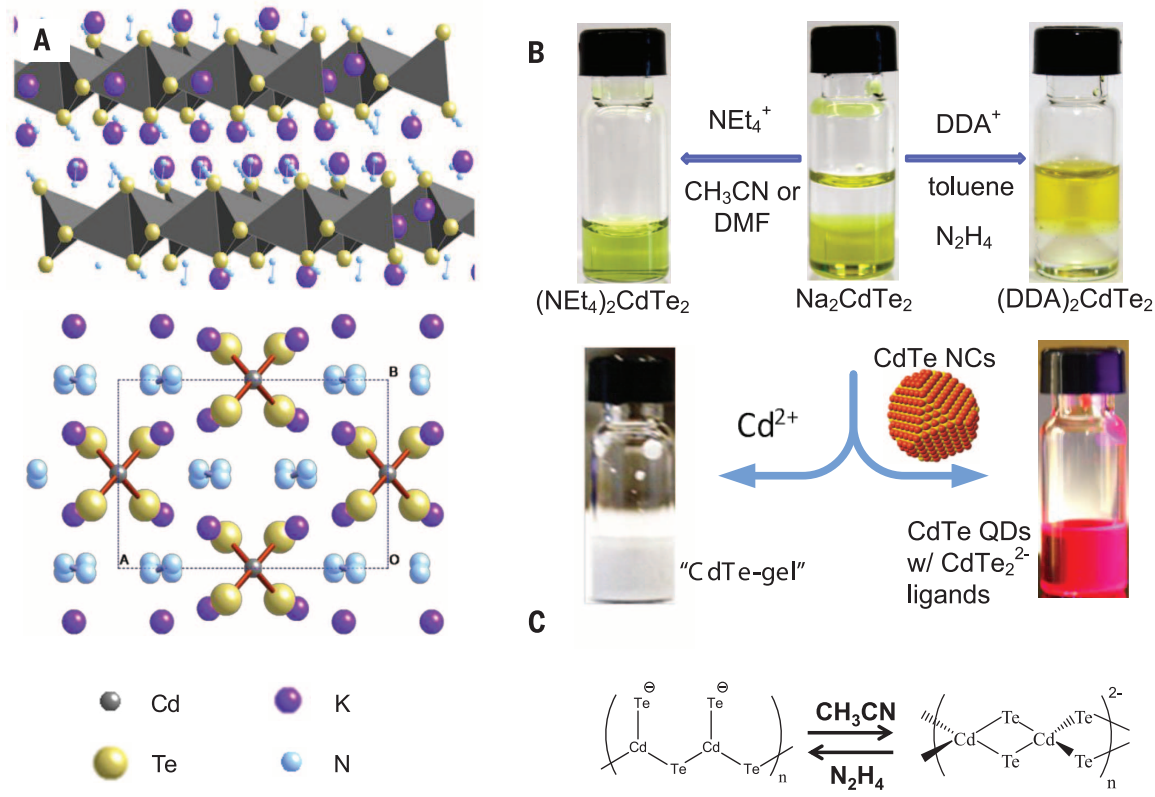
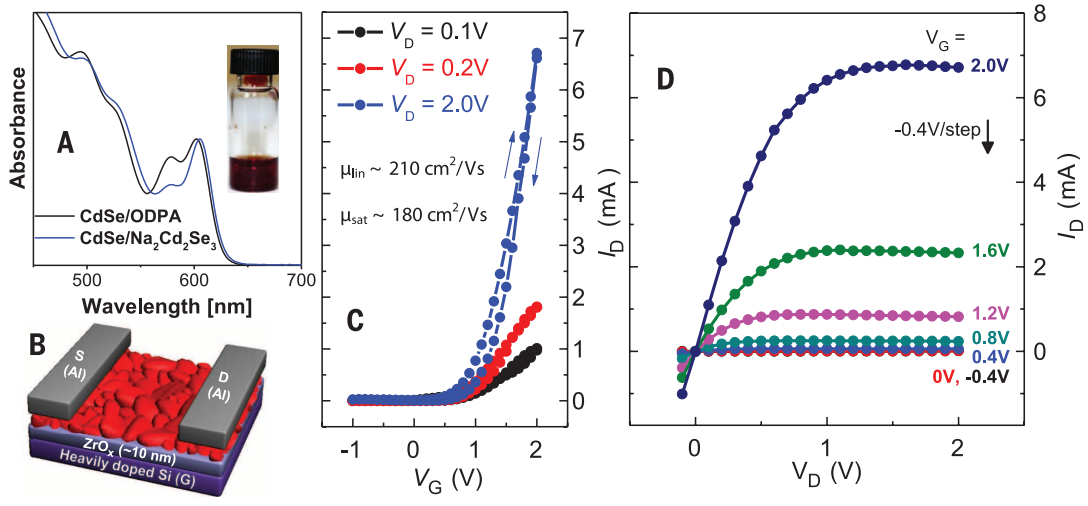


Fig. 1. Structure and properties of $CdTe_2^{2-}$ compounds. (**A**) Crystal structure of $K_2CdTe_2{\cdot}2N_2H_4$ determined by using single crystal x-ray diffraction. Hydrogen atoms are not shown for clarity. (**B**) Solubility of Na_2CdTe_2 in different solvents can be tuned by cation exchange from sodium (Na^+) to tetratethylammonium (NEt_4^+) or didodecyldimethylammonium (DDA^+). Polymeric $[CdTe_2^{2-}]_\infty$ ions can be cross-linked with Cd^{2+} ions, forming stable "CdTe-gel." $[CdTe_2^{2-}]_\infty$ ions can serve as ligands for luminescent CdTe quantum dots (QDs). (**C**) Depending on solvent polarity and coordinating ability, $[CdTe_2^{2-}]_\infty$ can exist in two equilibrium forms.

Fig. 2. Transport properties of CdSe NCs after ligand exchange. (**A**) Absorption spectra of 4.7-nm CdSe NCs capped with ODPA and $[Cd_2Se_3]^{2-}$ ligands. (Inset) Photograph of the solution of $[Cd_2Se_3]^{2-}$-capped CdSe NCs. (**B**) Scheme of a FET device with 10-nm ZrO_2 gate dielectric and Al source and drain electrodes. (**C**) Transfer characteristics and (**D**) output characteristics of an FET with a channel made of spin-coated 4.7 nm $[Cd_2Se_3]^{2-}$-capped CdSe NCs annealed at 250°C for 30 min. FET channel width and length are 1500 and 30 μm, respectively. The inset in (C) shows the linear and saturation regime electron mobility extracted from FET transfer characteristics (*9*).



with $[Cd_2Se_3]^{2-}$ ions (fig. S7) (*9*). A stable colloidal solution of $[Cd_2Se_3]^{2-}$-capped CdSe NCs was obtained, which preserved the original quantum dot excitonic features, with a small red shift of the first absorption peak (Fig. 2A). This shift was likely caused by the expansion of wave functions into the compositionally matched ligand shell. The solution of $[Cd_2Se_3]^{2-}$-capped CdSe NCs was spin-coated on heavily doped Si substrates with a ~10-nm-thick ZrO_x gate dielectric and annealed under nitrogen at various temperatures for 30 min. Mobility measurements were carried out in a standard field-effect transistor (FET) geometry (Fig. 2B) (*14*, *15*). Even annealing at only 250°C was sufficient to achieve excellent output and transfer characteristics (Fig. 2, C and D, and fig. S8) (*9*) with an electron mobility of 210 cm^2/V s, which exceeds the best reported values for any solution-processed inorganic semiconductor (table S3) (*9*). Annealing NC films at 300°C results in devices with FET mobility above 300 cm^2/V s (fig. S9) (*9*). In control experiments, we found that FETs with channels made of pure CdSe NCs, $Na_2Cd_2Se_3$, or $(N_2H_5)_2Cd_2Se_3$ showed mobilities of <4 cm^2/V s (figs. S10 to S12) (*9*).

In recent years, various organic and inorganic ligands (such as 1,4-phenylenediamine, $In_2Se_4^{2-}$, and SCN^-) have been used as "electronic linkers" for CdSe NCs (*16*–*20*). The highest reported FET mobilities of annealed CdSe NC samples were on the order of 35 cm^2/V s (table S3) (*9*), which is ~5% of the mobility of CdSe single crystals at 300 K (*21*). In our case, $[Cd_2Se_3]^{2-}$ ions worked as a compositionally matched solder that improved carrier transport up to almost half of the single-crystal value, which is surprising given that the Scherrer size of CdSe grains was about 15 and 22 nm for samples annealed at 250°C and 300°C, respectively (fig. S13) (*9*). This grain size was also confirmed with transmission electron microscopy (TEM) measurements of annealed films (fig. S14) (*9*). Various chemical treatments have been successfully used to passivate interfaces and control doping in polycrystalline semiconductors and NC solids (*22*, *23*). At the same time, it has proven difficult to achieve high mobility through sintering (*24*) or oriented attachment of NCs (*25*). Recent studies show that simply increasing CdSe grain size does not improve FET mobility (*26*). We therefore attribute this high mobility to exceptional electronic transparency of the grain boundaries, which did not introduce appreciable transport bottlenecks.

The molecular solders functioned not only for nanoscale but also for mesoscale grains. For example, PbTe, Bi_2Te_3, or CdTe films were obtained by soldering microparticles, obtained by ball-milling of ingots, by using Na_2PbTe_2, $(N_2H_5)_4Bi_2Te_5$, or CdTe-gel as solders, respectively (Fig. 3, B to D, and fig. S15) (*9*). In the first case, we applied the sodium salt because Na^+ is a commonly used p-type dopant for PbTe. A suspension of PbTe microparticles with a small amount [2 to 5 weight percent (wt %)] of Na_2PbTe_2 solder was drop-cast on a hydrophilized glass substrate and annealed at 500°C for 30 min, resulting in a continuous thin film with a uniform thickness of ~7 μm (*9*) with well-connected dense grains (Fig. 3B). The film exhibited a Hall mobility of 58 cm^2/V s, which is comparable with that (~100 cm^2/V s) of optimized Na-doped PbTe thermoelectric pellets made at a similar temperature by means of spark-plasma sintering or hot pressing (*27*, *28*). In control experiments, no continuous films were obtained from suspensions of PbTe and other microparticles without addition of solder. After annealing under identical conditions, the material remained powdery (Fig. 3A). The above examples show that chalcogenidometallates and chalcogels, mixed with nano- and microparticles, offer a viable path toward solution processing of thin-film semiconductors.

The solder effect was used to consolidate microparticles (such as PbTe, Bi_2Te_3, $Bi_{0.5}Sb_{1.5}Te_3$, and CdTe) into 3D blocks cast with a mold (Fig. 3E). The addition of a few weight percent of molecular or gel solder helped connect the microparticles and allowed the casting of solid pellets in graphite molds with no external pressure, similar to molding metal particles (*29*). Disk-, triangle-, square-, and bar-shaped blocks were obtained via annealing various combinations of nano- and microparticles (PbTe, Bi_2Te_3, $Bi_{0.5}Sb_{1.5}Te_3$, or CdTe) with corresponding compositionally matched molecular chalcogenidometallates or chalcogels (*9*). In control experiments without solders, the microparticles molded under the same conditions remained powdery (fig. S16) (*9*). Comparison of Fig. 3, A and B, shows that the addition of molecular solder facilitates sintering and grain growth compared with pure components. The microscopic origin of this effect is likely related to improved mass transport along and across the grain boundaries (*30*).

Chalcogenidometallate solders were also used for bonding metal chalcogenide crystals under modest pressures and temperatures. Shown in Fig. 3F is an example in which two polished (111) CdTe surfaces (with an area of 0.25 cm^2) were bonded at ~0.5 MPa and 500°C by using a minuscule amount (2 μL) of Na_2CdTe_2 solution in N_2H_4 (*9*). No physical bonding between CdTe crystals occurred in control experiments with similarly prepared crystal surfaces but no solder.

The above examples reveal the potential of bottom-up assembly of inorganic semiconductors with compositionally matched molecular solders. This approach enables the highest electron mobility achieved to date for solution-processed semiconductors, opening new opportunities for printable electronics and optoelectronics, and paves the way to future technological developments such as moldable or 3D-printable thermoelectric materials.

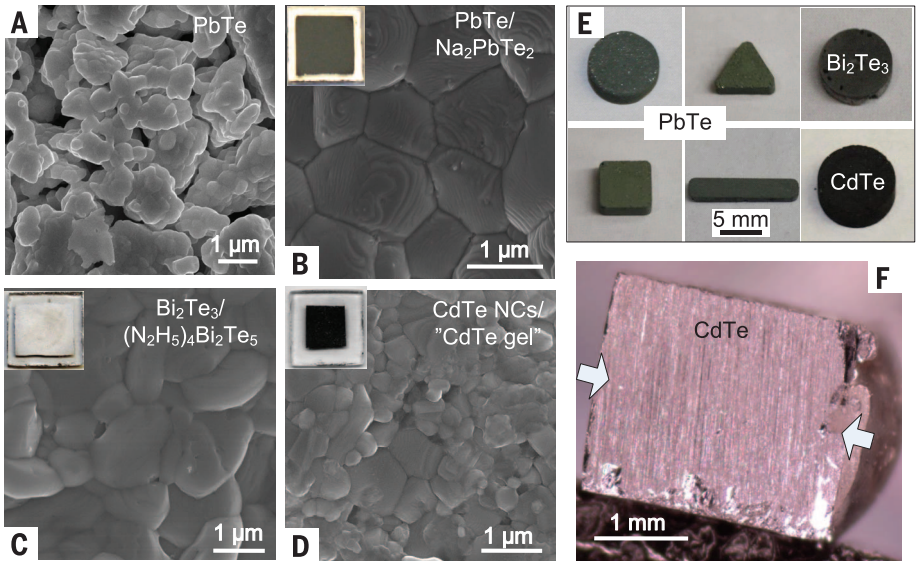


Fig. 3. Soldering various materials with novel chalcogenidometallates. Scanning electron microscopy images of (**A**) ball-milled PbTe microparticles annealed at 500°C for 30 min. (**B**) Film of similar PbTe microparticles annealed in the presence of 5 wt % Na_2PbTe_2. (**C**) Film made of Bi_2Te_3 microparticles annealed in the presence of 5 wt % $(N_2H_5)_4Bi_2Te_5$. (**D**) Film made of CdTe NCs annealed in the presence of 20 wt % "CdTe-gel." The films were annealed at 500°C, 400°C, and 450°C for 30 min, respectively. (Insets) Film optical micrographs. (**E**) 3D blocks of PbTe, Bi_2Te_3, and CdTe cast by using graphite molds of various shapes from microparticles mixed with corresponding "solders": Na_2PbTe_2 for PbTe, $Na_4Bi_2Te_5$ for Bi_2Te_3, and Na_2CdTe_2 for CdTe. (**F**) A photograph showing two CdTe crystals with mechanically polished (111) faces soldered with Na_2CdTe_2 at 500°C.

REFERENCES AND NOTES

1. *Encyclopædia Britannica Online*, s. v. "soldering" (2014); available at www.britannica.com/EBchecked/topic/553118/soldering.
2. S. M. Sze, *Semiconductor Devices: Physics and Technology* (John Wiley and Sons, New York, 2002).
3. D. B. Mitzi, L. L. Kosbar, C. E. Murray, M. Copel, A. Afzali, *Nature* **428**, 299–303 (2004).
4. D. B. Mitzi, *Adv. Mater.* **21**, 3141–3158 (2009).
5. D. B. Mitzi *et al.*, *Adv. Mater.* **20**, 3657–3662 (2008).
6. T. K. Todorov, K. B. Reuter, D. B. Mitzi, *Adv. Mater.* **22**, E156–E159 (2010).
7. R. Y. Wang *et al.*, *Chem. Mater.* **22**, 1943–1945 (2010).
8. D. H. Webber, R. L. Brutchey, *J. Am. Chem. Soc.* **135**, 15722–15725 (2013).

9. Materials and methods are available as supplementary materials on *Science* Online.
10. R. C. Burns, L. A. Devereux, P. Granger, G. J. Schrobilgen, *Inorg. Chem.* **24**, 2615–2624 (1985).
11. C.-W. Park, R. J. Salm, J. A. Ibers, *Can. J. Chem.* **73**, 1148–1156 (1995).
12. S. Bag, P. N. Trikalitis, P. J. Chupas, G. S. Armatas, M. G. Kanatzidis, *Science* **317**, 490–493 (2007).
13. M. V. Kovalenko, M. Scheele, D. V. Talapin, *Science* **324**, 1417–1420 (2009).
14. D. V. Talapin, C. B. Murray, *Science* **310**, 86–89 (2005).
15. D. S. Chung *et al*., *Nano Lett.* **12**, 1813–1820 (2012).
16. Y. Liu *et al*., *Nano Lett.* **10**, 1960–1969 (2010).
17. Y. Liu *et al*., *Nano Lett.* **13**, 1578–1587 (2013).
18. J. H. Choi *et al*., *Nano Lett.* **12**, 2631–2638 (2012).
19. J. S. Lee, M. V. Kovalenko, J. Huang, D. S. Chung, D. V. Talapin, *Nat. Nanotechnol.* **6**, 348–352 (2011).
20. D. Yu, C. Wang, P. Guyot-Sionnest, *Science* **300**, 1277–1280 (2003).
21. S. S. Devlin, *Physics and Chemistry of II-VI Compounds* (North Holland, Amsterdam, 1967).
22. B. M. Basol, S. S. Ou, O. M. Stafsudd, *J. Appl. Phys.* **58**, 3809–3813 (1985).
23. S. J. Oh *et al*., *Nano Lett.* **14**, 1559–1566 (2014).
24. B. A. Ridley, B. Nivi, J. M. Jacobson, *Science* **286**, 746–749 (1999).
25. W. J. Baumgardner, K. Whitham, T. Hanrath, *Nano Lett.* **13**, 3225–3231 (2013).
26. Z. M. Norman, N. C. Anderson, J. S. Owen, *ACS Nano* **8**, 7513–7521 (2014).
27. S. N. Girard *et al*., *J. Am. Chem. Soc.* **133**, 16588–16597 (2011).
28. K. Biswas *et al*., *Nature* **489**, 414–418 (2012).
29. R. Klajn *et al*., *Science* **316**, 261–264 (2007).
30. S.-J. L. Kang, *Sintering: Densification, Grain Growth and Microstructure* (Butterworth-Heinemann, Oxford, 2004).

ACKNOWLEDGMENTS

We thank S. Kwon and V. Zyryanov for help with EXAFS measurements and N. James for reading the manuscript. The work was supported by the II-VI Foundation, U.S. Department of Energy (DOE) SunShot program under award DE-EE0005312 and by NSF under award DMR-1310398. D.V.T. also thanks the Keck Foundation. This work used facilities supported by the NSF Materials Research Science and Engineering Center Program under awards DMR 08-20054 and DMR 14-20709. Use of the Center for Nanoscale Materials and Advanced Photon Source was supported by the DOE, Office of Science, Office of Basic Energy Sciences, under contract DE-AC02-06CH11357. MRCAT operations are supported by the DOE and MRCAT host institutions.

SUPPORTING ONLINE MATERIAL

www.sciencemag.org/content/347/6220/425/suppl/DC1
Materials and Methods
Figs. S1 to S16
Tables S1 to S3
References (*31*–*48*)

27 August 2014; accepted 8 December 2014
10.1126/science.1260501

DRUG RESISTANCE

K13-propeller mutations confer artemisinin resistance in *Plasmodium falciparum* clinical isolates

Judith Straimer,[1] Nina F. Gnädig,[1] Benoit Witkowski,[2]* Chanaki Amaratunga,[3]* Valentine Duru,[2]* Arba Pramundita Ramadani,[4,5]*† Mélanie Dacheux,[1] Nimol Khim,[2] Lei Zhang,[6] Stephen Lam,[6] Philip D. Gregory,[6] Fyodor D. Urnov,[6] Odile Mercereau-Puijalon,[7] Françoise Benoit-Vical,[4,5]‡ Rick M. Fairhurst,[3]‡ Didier Ménard,[2]‡ David A. Fidock[1,8]§

The emergence of artemisinin resistance in Southeast Asia imperils efforts to reduce the global malaria burden. We genetically modified the *Plasmodium falciparum K13* locus using zinc-finger nucleases and measured ring-stage survival rates after drug exposure in vitro; these rates correlate with parasite clearance half-lives in artemisinin-treated patients. With isolates from Cambodia, where resistance first emerged, survival rates decreased from 13 to 49% to 0.3 to 2.4% after the removal of K13 mutations. Conversely, survival rates in wild-type parasites increased from ≤0.6% to 2 to 29% after the insertion of K13 mutations. These mutations conferred elevated resistance to recent Cambodian isolates compared with that of reference lines, suggesting a contemporary contribution of additional genetic factors. Our data provide a conclusive rationale for worldwide K13-propeller sequencing to identify and eliminate artemisinin-resistant parasites.

The worldwide use of artemisinin (ART)–based combination therapies (ACTs) for the treatment of *Plasmodium falciparum* malaria is the foundation of renewed efforts to eradicate this leading cause of childhood mortality (*1*, *2*). The pharmacodynamic properties of clinically used ART derivatives [artesunate, artemether, and dihydroartemisinin (DHA)] can reduce the biomass of drug-sensitive parasites by four orders of magnitude every 48 hours (*3*), corresponding to a single cycle of asexual blood-stage *P. falciparum* development. The short half-life (typically <1 hour) of ART derivatives in plasma necessitates the use of longer-lasting partner drugs that can eliminate residual parasites once the ART component has dropped to subtherapeutic concentrations (*4*). The use of ACTs in expanded malaria control and elimination programs has yielded notable successes in recent years, contributing to an estimated 30% reduction in global mortality rates in the past decade (*5*).

These impressive gains, however, are now threatened by the emergence of ART resistance, first detected in western Cambodia and now observed in Thailand, Vietnam, and Myanmar (*6*, *7*). The severity of this situation is underscored by the fact that resistance to piperaquine, an ACT partner drug, is emerging in western Cambodia (*8*, *9*). No alternative, fully effective first-line therapy is currently available to replace ACTs, should ART fail globally. Clinically, ART resistance is defined as a long parasite clearance half-life (the time it takes for the peripheral blood parasite density to decrease by 50%) after treatment with ART monotherapy or an ACT (*6*, *10*, *11*). This metric correlates with the percentage of early "ring-stage" parasites (0 to 3 hours after invasion of human erythrocytes) that survive a pharmacologically relevant exposure to DHA (the active metabolite of all ARTs), as measured in the in vitro Ring-stage Survival Assay ($RSA_{0\text{-}3h}$) (*12*).

Recently, mutations in the propeller domain of the *K13* gene were identified as candidate molecular markers of ART resistance (*13*). This gene resides on chromosome 13 of the *P. falciparum* genome, near regions earlier associated with slow parasite clearance rates (*14*–*16*). K13 belongs to the kelch superfamily of proteins, whose propeller domain harbors multiple protein-protein interaction sites and mediates diverse cellular functions, including ubiquitin-regulated protein degradation and oxidative stress responses (*17*). The K13 M476I mutation was first observed in Tanzanian F32 parasites that were exposed in vitro to escalating concentrations of ART over 5 years, yielding the F32-ART line (*13*, *18*). [Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. In the mutants, other amino acids were substituted at certain locations; for example, M476I indicates that methionine at position 476 was replaced by isoleucine.] Subsequent genomic analysis of Cambodian isolates identified four prevalent K13-propeller mutations (Y493H, R539T, I543T, and C580Y) that were associated with elevated $RSA_{0\text{-}3h}$ survival rates in vitro and long parasite clearance half-lives (>5 hours) in patients (*13*, *19*).

[1]Department of Microbiology and Immunology, Columbia University College of Physicians and Surgeons, New York, NY, USA. [2]Malaria Molecular Epidemiology Unit, Institut Pasteur du Cambodge, Phnom Penh, Cambodia. [3]Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [4]Centre National de la Recherche Scientifique (CNRS), Laboratoire de Chimie de Coordination UPR8241, Toulouse, France. [5]Université de Toulouse, UPS, Institut National Polytechnique de Toulouse, Toulouse, France. [6]Sangamo BioSciences, Richmond, CA, USA. [7]Institut Pasteur, Parasite Molecular Immunology Unit, Paris, France. [8]Division of Infectious Diseases, Department of Medicine, Columbia University College of Physicians and Surgeons, New York, NY, USA.
*These authors contributed equally to this work. †Present address: Department of Pharmacology and Therapy, Faculty of Medicine, Gadjah Mada University, Yogyakarta, Indonesia. ‡These authors contributed equally to this work. **§To whom correspondence should be addressed. E-mail: df2260@columbia.edu**

9. Materials and methods are available as supplementary materials on *Science* Online.
10. R. C. Burns, L. A. Devereux, P. Granger, G. J. Schrobilgen, *Inorg. Chem.* **24**, 2615–2624 (1985).
11. C.-W. Park, R. J. Salm, J. A. Ibers, *Can. J. Chem.* **73**, 1148–1156 (1995).
12. S. Bag, P. N. Trikalitis, P. J. Chupas, G. S. Armatas, M. G. Kanatzidis, *Science* **317**, 490–493 (2007).
13. M. V. Kovalenko, M. Scheele, D. V. Talapin, *Science* **324**, 1417–1420 (2009).
14. D. V. Talapin, C. B. Murray, *Science* **310**, 86–89 (2005).
15. D. S. Chung *et al.*, *Nano Lett.* **12**, 1813–1820 (2012).
16. Y. Liu *et al.*, *Nano Lett.* **10**, 1960–1969 (2010).
17. Y. Liu *et al.*, *Nano Lett.* **13**, 1578–1587 (2013).
18. J. H. Choi *et al.*, *Nano Lett.* **12**, 2631–2638 (2012).
19. J. S. Lee, M. V. Kovalenko, J. Huang, D. S. Chung, D. V. Talapin, *Nat. Nanotechnol.* **6**, 348–352 (2011).
20. D. Yu, C. Wang, P. Guyot-Sionnest, *Science* **300**, 1277–1280 (2003).
21. S. S. Devlin, *Physics and Chemistry of II-VI Compounds* (North Holland, Amsterdam, 1967).
22. B. M. Basol, S. S. Ou, O. M. Stafsudd, *J. Appl. Phys.* **58**, 3809–3813 (1985).
23. S. J. Oh *et al.*, *Nano Lett.* **14**, 1559–1566 (2014).
24. B. A. Ridley, B. Nivi, J. M. Jacobson, *Science* **286**, 746–749 (1999).
25. W. J. Baumgardner, K. Whitham, T. Hanrath, *Nano Lett.* **13**, 3225–3231 (2013).
26. Z. M. Norman, N. C. Anderson, J. S. Owen, *ACS Nano* **8**, 7513–7521 (2014).
27. S. N. Girard *et al.*, *J. Am. Chem. Soc.* **133**, 16588–16597 (2011).
28. K. Biswas *et al.*, *Nature* **489**, 414–418 (2012).
29. R. Klajn *et al.*, *Science* **316**, 261–264 (2007).
30. S.-J. L. Kang, *Sintering: Densification, Grain Growth and Microstructure* (Butterworth-Heinemann, Oxford, 2004).

ACKNOWLEDGMENTS

We thank S. Kwon and V. Zyryanov for help with EXAFS measurements and N. James for reading the manuscript. The work was supported by the II-VI Foundation, U.S. Department of Energy (DOE) SunShot program under award DE-EE0005312 and by NSF under award DMR-1310398. D.V.T. also thanks the Keck Foundation. This work used facilities supported by the NSF Materials Research Science and Engineering Center Program under awards DMR 08-20054 and DMR 14-20709. Use of the Center for Nanoscale Materials and Advanced Photon Source was supported by the DOE, Office of Science, Office of Basic Energy Sciences, under contract DE-AC02-06CH11357. MRCAT operations are supported by the DOE and MRCAT host institutions.

SUPPORTING ONLINE MATERIAL

www.sciencemag.org/content/347/6220/425/suppl/DC1
Materials and Methods
Figs. S1 to S16
Tables S1 to S3
References (*31*–*48*)

27 August 2014; accepted 8 December 2014
10.1126/science.1260501

DRUG RESISTANCE

K13-propeller mutations confer artemisinin resistance in *Plasmodium falciparum* clinical isolates

Judith Straimer,[1] Nina F. Gnädig,[1] Benoit Witkowski,[2]* Chanaki Amaratunga,[3]* Valentine Duru,[2]* Arba Pramundita Ramadani,[4,5]*† Mélanie Dacheux,[1] Nimol Khim,[2] Lei Zhang,[6] Stephen Lam,[6] Philip D. Gregory,[6] Fyodor D. Urnov,[6] Odile Mercereau-Puijalon,[7] Françoise Benoit-Vical,[4,5]‡ Rick M. Fairhurst,[3]‡ Didier Ménard,[2]‡ David A. Fidock[1,8]§

The emergence of artemisinin resistance in Southeast Asia imperils efforts to reduce the global malaria burden. We genetically modified the *Plasmodium falciparum K13* locus using zinc-finger nucleases and measured ring-stage survival rates after drug exposure in vitro; these rates correlate with parasite clearance half-lives in artemisinin-treated patients. With isolates from Cambodia, where resistance first emerged, survival rates decreased from 13 to 49% to 0.3 to 2.4% after the removal of K13 mutations. Conversely, survival rates in wild-type parasites increased from ≤0.6% to 2 to 29% after the insertion of K13 mutations. These mutations conferred elevated resistance to recent Cambodian isolates compared with that of reference lines, suggesting a contemporary contribution of additional genetic factors. Our data provide a conclusive rationale for worldwide K13-propeller sequencing to identify and eliminate artemisinin-resistant parasites.

The worldwide use of artemisinin (ART)–based combination therapies (ACTs) for the treatment of *Plasmodium falciparum* malaria is the foundation of renewed efforts to eradicate this leading cause of childhood mortality (*1*, *2*). The pharmacodynamic properties of clinically used ART derivatives [artesunate, artemether, and dihydroartemisinin (DHA)] can reduce the biomass of drug-sensitive parasites by four orders of magnitude every 48 hours (*3*), corresponding to a single cycle of asexual blood-stage *P. falciparum* development. The short half-life (typically <1 hour) of ART derivatives in plasma necessitates the use of longer-lasting partner drugs that can eliminate residual parasites once the ART component has dropped to subtherapeutic concentrations (*4*). The use of ACTs in expanded malaria control and elimination programs has yielded notable successes in recent years, contributing to an estimated 30% reduction in global mortality rates in the past decade (*5*).

These impressive gains, however, are now threatened by the emergence of ART resistance, first detected in western Cambodia and now observed in Thailand, Vietnam, and Myanmar (*6*, *7*). The severity of this situation is underscored by the fact that resistance to piperaquine, an ACT partner drug, is emerging in western Cambodia (*8*, *9*). No alternative, fully effective first-line therapy is currently available to replace ACTs, should ART fail globally. Clinically, ART resistance is defined as a long parasite clearance half-life (the time it takes for the peripheral blood parasite density to decrease by 50%) after treatment with ART monotherapy or an ACT (*6*, *10*, *11*). This metric correlates with the percentage of early "ring-stage" parasites (0 to 3 hours after invasion of human erythrocytes) that survive a pharmacologically relevant exposure to DHA (the active metabolite of all ARTs), as measured in the in vitro Ring-stage Survival Assay ($RSA_{0\text{-}3h}$) (*12*).

Recently, mutations in the propeller domain of the *K13* gene were identified as candidate molecular markers of ART resistance (*13*). This gene resides on chromosome 13 of the *P. falciparum* genome, near regions earlier associated with slow parasite clearance rates (*14*–*16*). K13 belongs to the kelch superfamily of proteins, whose propeller domain harbors multiple protein-protein interaction sites and mediates diverse cellular functions, including ubiquitin-regulated protein degradation and oxidative stress responses (*17*). The K13 M476I mutation was first observed in Tanzanian F32 parasites that were exposed in vitro to escalating concentrations of ART over 5 years, yielding the F32-ART line (*13*, *18*). [Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. In the mutants, other amino acids were substituted at certain locations; for example, M476I indicates that methionine at position 476 was replaced by isoleucine.] Subsequent genomic analysis of Cambodian isolates identified four prevalent K13-propeller mutations (Y493H, R539T, I543T, and C580Y) that were associated with elevated $RSA_{0\text{-}3h}$ survival rates in vitro and long parasite clearance half-lives (>5 hours) in patients (*13*, *19*).

[1]Department of Microbiology and Immunology, Columbia University College of Physicians and Surgeons, New York, NY, USA. [2]Malaria Molecular Epidemiology Unit, Institut Pasteur du Cambodge, Phnom Penh, Cambodia. [3]Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [4]Centre National de la Recherche Scientifique (CNRS), Laboratoire de Chimie de Coordination UPR8241, Toulouse, France. [5]Université de Toulouse, UPS, Institut National Polytechnique de Toulouse, Toulouse, France. [6]Sangamo BioSciences, Richmond, CA, USA. [7]Institut Pasteur, Parasite Molecular Immunology Unit, Paris, France. [8]Division of Infectious Diseases, Department of Medicine, Columbia University College of Physicians and Surgeons, New York, NY, USA.
*These authors contributed equally to this work. †Present address: Department of Pharmacology and Therapy, Faculty of Medicine, Gadjah Mada University, Yogyakarta, Indonesia. ‡These authors contributed equally to this work. **§To whom correspondence should be addressed. E-mail: df2260@columbia.edu**

Determining whether K13-propeller mutations confer ART resistance in clinical isolates and assessing the contributions of individual polymorphisms in distinct genetic backgrounds is essential to defining the underlying molecular mechanisms.

We developed zinc-finger nucleases (ZFNs) (*20*) to enable targeted genetic engineering of *K13* in newly culture-adapted Cambodian isolates and older established reference lines of *P. falciparum* (tables S1 and S2). ZFNs were introduced into cultured intra-erythrocytic parasites via electroporation with plasmids containing *K13* donor templates. ZFNs triggered double-stranded breaks in the *K13* genomic target locus of this haploid organism, leading to DNA resection and repair events that captured mutations delivered by $pZFN^{K13}$-hdhfr plasmids (fig. S1). Donor plasmids contained additional synonymous mutations that preclude ZFN binding while preserving the K13-translated amino acid sequence across that same stretch of DNA base pairs. These silent ZFN binding-site mutations protected the donor sequence and prevented the edited recombinant locus from being recleaved by the nucleases. Plasmids contained either the wild-type *K13* allele or one of several mutations (present in the six-blade K13-propeller domain) found in ART-resistant Cambodian isolates or F32-ART. This strategy successfully introduced or removed mutations in a set of *P. falciparum* clinical isolates from Cambodia, the epicenter of emerging ART resistance, as well as reference laboratory lines from distinct geographic origins (Fig. 1 and table S3). Of note, $RSA_{0\text{-}3h}$ assays comparing parental and edited control parasites showed no difference if only the binding-site mutations were introduced into the K13-propeller domain, indicating that these synonymous mutations were phenotypically silent (Fig. 1 and fig. S2). Independent assays with the same parasite lines tested by our different groups yielded consistent survival rates between laboratories (fig. S3).

Using donor plasmids containing a wild-type *K13*-propeller sequence and silent binding-site mutations, we generated a series of clones in which individual K13 mutations were removed from ART-resistant Cambodian isolates. One of these isolates (Cam3.II) showed slow clearance after ART monotherapy (in vivo half-life 6.0 hours) (table S2). Parental $Cam3.I^{R539T}$ and $Cam3.II^{R539T}$ isolates harboring the R539T mutation showed 40 to 49% $RSA_{0\text{-}3h}$ survival, whereas edited $Cam3.I^{rev}$ and $Cam3.II^{rev}$ clones carrying the reverted wild-type allele showed only 0.3 to 0.7% survival (Fig. 2, A and B, and table S4). These highly significant differences in the survival rates of ring-stage parasites exposed to elevated DHA concentrations confirm the importance of R539T in mediating in vitro ART resistance in Cambodian isolates. Significant reductions in $RSA_{0\text{-}3h}$ survival rates were also observed upon removal of I543T (43% in $Cam5^{I543T}$ versus 0.3% in $Cam5^{rev}$) (Fig. 2C) and C580Y (13% in $Cam2^{C580Y}$ versus 2.4% in $Cam2^{rev}$) (Fig. 2D).

We also assessed the impact of introducing K13 mutations into a fast-clearing Cambodian isolate (CamWT; in vivo half-life 3.7 hours) (table S2), the $Cam3.II^{rev}$ clone, and three reference lines (V1/S, F32-TEM, and FCB). CamWT and $Cam3.II^{rev}$ parasites harboring wild-type *K13* alleles showed 0.6 to 0.7% $RSA_{0\text{-}3h}$ survival, whereas the corresponding C580Y-edited clones yielded 9 and 24% survival, respectively (Fig. 2, E and F). Introducing R539T into V1/S caused a similar increase in $RSA_{0\text{-}3h}$ survival (0.3 to 21%) (Fig. 2G and table S4). Editing F32-TEM to express M476I caused a moderate increase in $RSA_{0\text{-}3h}$ survival (<0.2% in F32-TEM to 1.7% in $F32\text{-}TEM^{M476I}$) (Fig. 2H). We also observed modest in vitro resistance in FCB parasites edited to express C580Y, with $RSA_{0\text{-}3h}$ survival increasing from 0.3% in the parental line to 1.9% in FCB^{C580Y} parasites (Fig. 2I). This result differs from a recent study of the use of Cas9 in *P. falciparum*, which reported a greater increase in $RSA_{0\text{-}3h}$ survival (11 to 15%) in two clones engineered to express K13 C580Y (*21*). That report used the drug-sensitive NF54 strain—which was isolated decades before ART use and the emergence of resistance (*22*)—and did not examine additional mutations or assess the impact of removing K13 mutations from ART-resistant clinical isolates.

In contrast to the substantial changes we observed in the $RSA_{0\text{-}3h}$, standard in vitro dose-response measurements by use of parental and *K13*-edited V1/S and Cam3.II parasites revealed no effect of R539T or C580Y on DHA or artesunate median inhibitory concentration (IC_{50}) values (fig. S4). This finding is consistent with earlier studies that showed no correlation between IC_{50} values and clinical ART resistance (*6*, *10*, *12*).

We subsequently investigated whether individual mutations confer different levels of ART resistance in the $RSA_{0\text{-}3h}$. In the Dd2 reference line, the introduction of M476I, R539T, or I543T mutations conferred considerably higher degrees of resistance than those of Y493H and C580Y (10 to 30% versus 2 to 4% survival, respectively) (Fig. 2J and table S4). These data corroborate the recent observation of higher levels of in vitro resistance in Cambodian isolates containing the R539T mutation as compared with Y493H or C580Y (*23*).

The relatively modest increase in survival of C580Y-expressing Dd2 parasites compared with

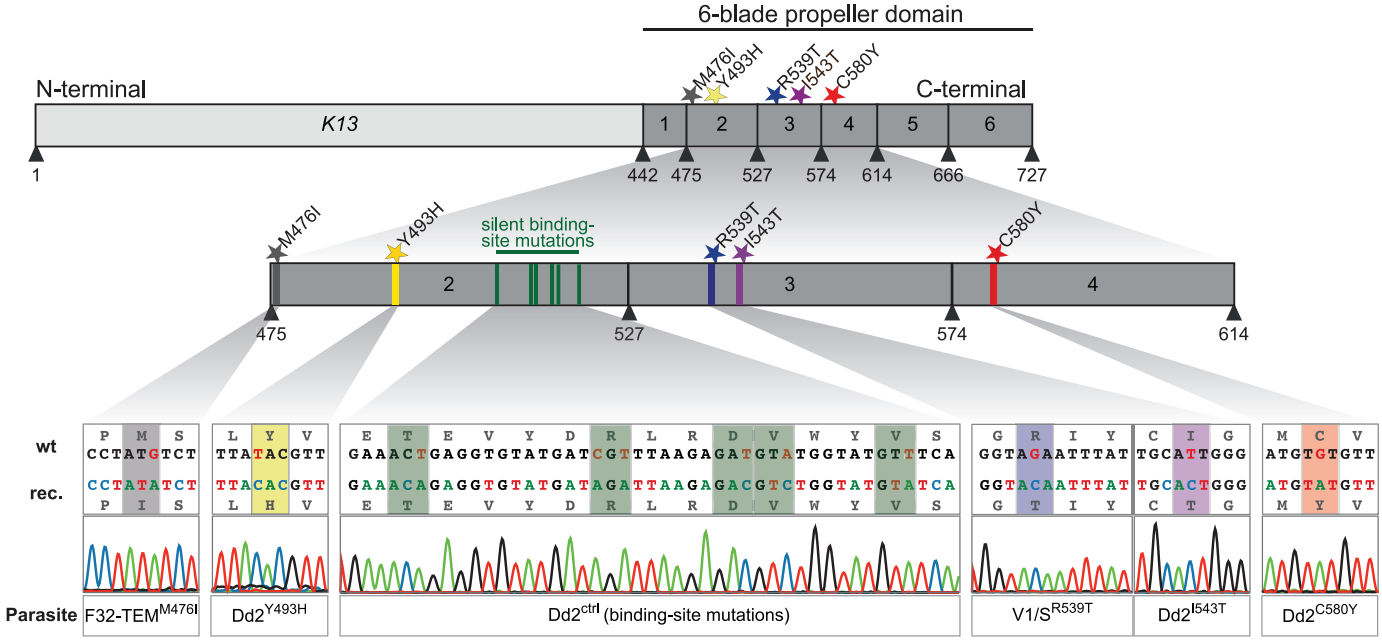


Fig. 1. Genetic modification of the K13-propeller domain. Location of K13-propeller mutations and sequencing results showing the insertion of individual mutations into recombinant parasites used in the $RSA_{0\text{-}3h}$. $Dd2^{ctrl}$ parasites contain only synonymous, phenotypically silent binding-site mutations and showed 0.7% survival rates, which is equivalent to those of parental Dd2 parasites (fig. S2).

R539T- and I543T-expressing clinical isolates and edited clones was quite unexpected, given that C580Y has rapidly become the predominant mutant allele in western Cambodia (*7*, *13*). We thus explored the impact of C580Y in different genetic backgrounds. Introducing C580Y conferred greater levels of resistance in three Cambodian isolates as compared with Dd2 and FCB parasites (Fig. 2K), suggesting a role for additional parasite factors in augmenting K13-mediated resistance in these contemporary field isolates. The disparity between relatively low in vitro resistance conferred by C580Y and its widespread dissemination in Cambodia might be explained by a lower fitness cost or increased transmission potential of C580Y-expressing parasites, or by the parasite genetic background.

Cambodian parasites are specifically characterized by sympatric subpopulations that show only limited genetic admixture and that generally harbor distinct K13 mutations (*16*). These findings suggest that K13 mutations might have arisen preferentially on backgrounds with favorable genetic factors. In this context, recent comprehensive analyses of K13 mutations across multiple sites in Southeast Asia have documented a series of additional mutations associated with slow clearance rates in Cambodia, Thailand, Myanmar, Laos, and Vietnam (*7*, *24*). K13 mutations have also been observed in African isolates (*7*, *25*, *26*), although none of these correspond to the most prevalent mutations in Cambodia, and ART or ACT treatments in African sites continue to show a high level of efficacy (*7*). A recent deep-sequencing study of the K13-propeller domain in more than 1110 *P. falciparum* infections collected from 14 sites across sub-Saharan Africa identified a large reservoir of naturally occurring K13-propeller variation, whose impact on artemisinin susceptibility is unknown and requires further investigation. These polymorphisms include one rare mutation previously observed in Cambodia (P553L) and several others (including A578S) close to known resistance-causing mutations in the propeller domain (*26*). Our gene-editing system can now be used to comprehensively dissect K13 polymorphisms across malaria-endemic regions and identify those that confer ring-stage ART resistance.

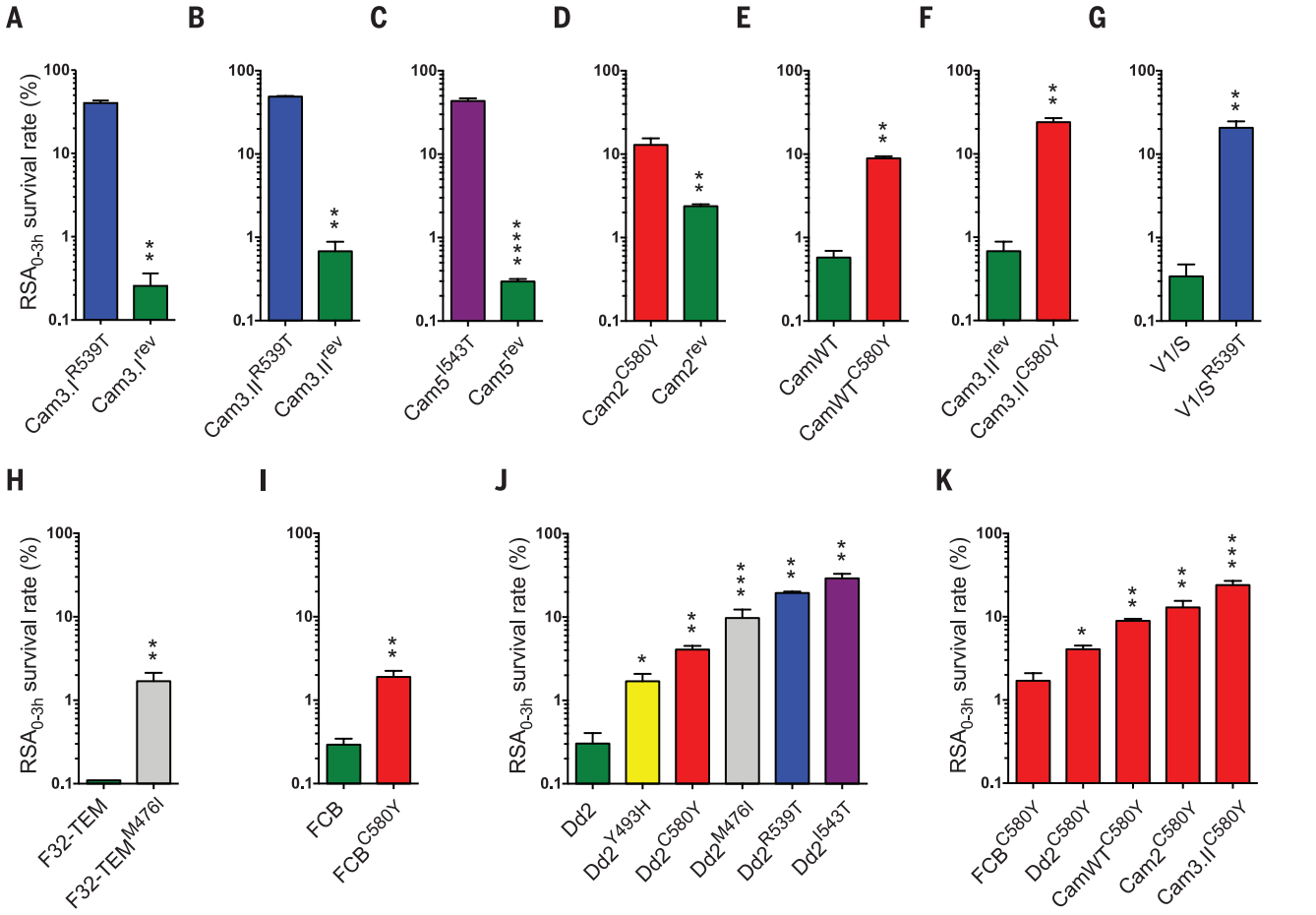


Fig. 2. K13-propeller mutations confer artemisinin resistance in clinical isolates and reference lines in vitro, as defined in the $RSA_{0\text{-}3h}$. Results show the percentage of early ring-stage parasites (0 to 3 hours after invasion of human erythrocytes) that survived a 6-hour pulse of 700 nM DHA (a pharmacologically relevant concentration of the active metabolite of ARTs), as measured by microscopy 66 hours later. Data show mean ± SEM percent survival compared with control dimethyl sulfoxide–treated parasites processed in parallel. (**A** to **D**) $RSA_{0\text{-}3h}$ survival for Cambodian isolates harboring native K13 mutations (shown in superscript) and ZFN-edited isogenic clones carrying wild-type *K13* alleles (superscript "rev"). (**E** to **I**) $RSA_{0\text{-}3h}$ survival for Cambodian isolates and reference lines harboring wild-type *K13* alleles and ZFN-edited isogenic clones carrying individual K13 mutations (shown in superscripts). (**J**) Impact of different K13 mutations on $RSA_{0\text{-}3h}$ survival in the Dd2 reference line, showing that I543T and R539T confer the highest levels of resistance. (**K**) Introduction of C580Y into multiple Cambodian clinical isolates and reference lines, showing that this mutation confers varying degrees of in vitro resistance depending on the parasite genetic background. The geographic origins and known drug-resistance genotypes of these isolates and lines are provided in table S2. Results were obtained from 3 or 4 independent assays performed in duplicate (values provided in table S4; F32-TEM showed <0.2% $RSA_{0\text{-}3h}$ survival). Two-sample *t* tests with unequal variances (performed with the STATA package) were used to assess for statistically significant differences between *K13*-edited clones and their comparator lines—the parental isolates listed on the left in (A) to (J) and the FCB^{C580Y} clone in (K) (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$). Statistical outputs (including calculations of the SE of the difference between the means of samples being compared and the *P* values) are listed in the supplementary materials.

Mode-of-action studies have shown that ARTs are active against all asexual blood stages of parasite development. In the more mature trophozoite stages, ARTs are activated after hemoglobin degradation and liberation of reactive heme whose iron moiety can cleave the endoperoxide linkage of these sesquiterpene lactone drugs (*27*). Activation generates free radicals that are thought to trigger oxidative stress and damage cellular macromolecules, including parasite membrane components, proteins, and neutral lipids (*28*, *29*). Recent evidence suggests that hemoglobin degradation begins early after merozoite invasion, potentially providing a source of ART activator in ring-stage parasites (*30*). Our $RSA_{0\text{-}3h}$ data support earlier evidence that reduced ring-stage susceptibility accounts for the clinical phenotype of slow parasite clearance after ART treatment (*12*, *31*). K13 mutations might achieve this by protecting parasites from the lethal effects of ART-induced oxidative damage, potentially via a cellular pathway similar to antioxidant transcriptional responses regulated by the mammalian ortholog Keap1 (*32*). Our set of K13-modified isogenic parasites with different levels of ART resistance on distinct genetic backgrounds now enables a search for K13-interacting partners and delivers tools to interrogate the underlying mechanism.

Our data demonstrate a central, causal role for K13-propeller mutations in conferring ART resistance in vitro and provide a molecular explanation for slow parasite clearance rates in patients (*6*, *7*, *10*). By exposing greater parasite biomasses to ACTs in vivo, K13-propeller mutations may promote the evolution of partner drug resistance (*8*, *9*) and higher-grade ART resistance. Our study thus offers a conclusive rationale for a global *K13* sequencing effort to track the spread of ART resistance and mitigate its impact on malaria treatment and control programs, particularly in hyperendemic regions in Africa.

REFERENCES AND NOTES

1. R. Feachem, O. Sabot, *Lancet* **371**, 1633–1635 (2008).
2. N. J. White *et al.*, *Lancet* **383**, 723–735 (2014).
3. N. J. White, *Parassitologia* **41**, 301–308 (1999).
4. R. T. Eastman, D. A. Fidock, *Nat. Rev. Microbiol.* **7**, 864–874 (2009).
5. World Health Organization, World Malaria Report: 2013. WHO Press: Geneva, Switzerland; available at www.who.int/malaria/publications/world_malaria_report_2013/en.
6. A. M. Dondorp *et al.*, *N. Engl. J. Med.* **361**, 455–467 (2009).
7. E. A. Ashley *et al.*, *N. Engl. J. Med.* **371**, 411–423 (2014).
8. D. L. Saunders *et al.*, *N. Engl. J. Med.* **371**, 484–485 (2014).
9. R. Leang *et al.*, *Antimicrob. Agents Chemother.* **57**, 818–826 (2013).
10. C. Amaratunga *et al.*, *Lancet Infect. Dis.* **12**, 851–858 (2012).
11. J. A. Flegg *et al.*, *Malar. J.* **12**, 411 (2013).
12. B. Witkowski *et al.*, *Lancet Infect. Dis.* **13**, 1043–1049 (2013).
13. F. Ariey *et al.*, *Nature* **505**, 50–55 (2014).
14. I. H. Cheeseman *et al.*, *Science* **336**, 79–82 (2012).
15. S. Takala-Harrison *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 240–245 (2013).
16. O. Miotto *et al.*, *Nat. Genet.* **45**, 648–655 (2013).
17. J. Adams, R. Kelso, L. Cooley, *Trends Cell Biol.* **10**, 17–24 (2000).
18. B. Witkowski *et al.*, *Antimicrob. Agents Chemother.* **54**, 1872–1877 (2010).
19. C. Amaratunga, B. Witkowski, N. Khim, D. Menard, R. M. Fairhurst, *Lancet Infect. Dis.* **14**, 449–450 (2014).
20. J. Straimer *et al.*, *Nat. Methods* **9**, 993–998 (2012).
21. M. Ghorbal *et al.*, *Nat. Biotechnol.* **32**, 819–821 (2014).
22. T. Ponnudurai, J. H. Meuwissen, A. D. Leeuwenberg, J. P. Verhave, A. H. Lensen, *Trans. R. Soc. Trop. Med. Hyg.* **76**, 242–250 (1982).
23. C. Amaratunga *et al.*, *Antimicrob. Agents Chemother.* **58**, 4935–4937 (2014).
24. S. Takala-Harrison *et al.*, *J. Infect. Dis.* 10.1093/infdis/jiu491 (2014).
25. M. D. Conrad *et al.*, *PLOS One* **9**, e105690 (2014).
26. S. M. Taylor *et al.*, *J. Infect. Dis.* 10.1093/infdis/jiu467 (2014).
27. N. Klonis, D. J. Creek, L. Tilley, *Curr. Opin. Microbiol.* **16**, 722–727 (2013).
28. C. L. Hartwig *et al.*, *Biochem. Pharmacol.* **77**, 322–336 (2009).
29. T. Antoine *et al.*, *J. Antimicrob. Chemother.* **69**, 1005–1016 (2014).
30. N. Klonis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11405–11410 (2011).
31. S. Saralamba *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 397–402 (2011).
32. Y. S. Keum, B. Y. Choi, *Molecules* **19**, 10074–10089 (2014).

ACKNOWLEDGMENTS

D.A.F. gratefully acknowledges funding from the NIH (R01 AI109023). This study was supported in part by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, NIH, the French "Agence Nationale de la Recherche" (ANR-13-BSV3-0018-01 and the Laboratoire d'Excellence IBEID), and the Institut Pasteur, Division International (ACIP A-10-2010). Parental and transgenic parasite lines have been deposited and are being made available through BEI Resources (www.mr4.org) with the following accession numbers: MRA-1240, Cam3.IR539T (also known as IPC 5202); MRA-1252, Cam3.Irev; MRA-1241, Cam5I543T (also known as IPC 4912); MRA-1253, Cam5rev; MRA-1236, Cam2C580Y (also known as IPC 3445); MRA-1254, Cam2rev; MRA-1250, CamWT; MRA-1251, CamWTC580Y; MRA-150, Dd2; MRA-1255, Dd2R539T. Parasite lines generated for this study will also be provided upon request from D.A.F. Requests for ZFNs should be directed to F.D.U. (FUrnov@sangamo.com); a materials transfer agreement is required. We extend our gratitude to F. Ariey (Institut Pasteur, Paris) for his important contribution to initiating this study, I. McKeague and O. Lieberman (Columbia University Medical Center) for their statistical and scientific input, and E. Rebar and the Production Group at Sangamo BioSciences for ZFN assembly and validation. L.Z., S.L., P.D.G., and F.D.U. declare that they are full-time employees of Sangamo, which designed, validated, and provided the ZFNs used in this study. B.W., O.M.-P., F.B.-V., and D.M., are co-inventors on the pending patents US61/904651 and US62/062439, and N.K. is a co-inventor on the pending patent US62/062439. Both patents are filed by Institut Pasteur. These patents cover the use of K13 mutations as a molecular marker of *P. falciparum* ART resistance. Sangamo holds patents on engineered DNA-binding proteins and the use thereof in targeted genome engineering and gene-specific regulation. All other authors declare no competing financial interests.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/428/suppl/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S5
Results of two-sample *t* tests with unequal variances
References (*33*–*36*)

5 September 2014; accepted 1 December 2014
Published online 11 December 2014;
10.1126/science.1260867

DRUG RESISTANCE

Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance

Sachel Mok,[1] Elizabeth A. Ashley,[2,3] Pedro E. Ferreira,[1] Lei Zhu,[1] Zhaoting Lin,[1] Tomas Yeo,[1] Kesinee Chotivanich,[4] Mallika Imwong,[5] Sasithon Pukrittayakamee,[4] Mehul Dhorda,[3,6,7] Chea Nguon,[8] Pharath Lim,[8,9] Chanaki Amaratunga,[9] Seila Suon,[8] Tran Tinh Hien,[10] Ye Htut,[11] M. Abul Faiz,[12] Marie A. Onyamboko,[13] Mayfong Mayxay,[14,15] Paul N. Newton,[2,3,14] Rupam Tripura,[2] Charles J. Woodrow,[2,3] Olivo Miotto,[2,16,17] Dominic P. Kwiatkowski,[16,17] François Nosten,[3,18] Nicholas P. J. Day,[2,3] Peter R. Preiser,[1] Nicholas J. White,[2,3] Arjen M. Dondorp,[2,3] Rick M. Fairhurst,[9] Zbynek Bozdech[1]*

Artemisinin resistance in *Plasmodium falciparum* threatens global efforts to control and eliminate malaria. Polymorphisms in the kelch domain–carrying protein K13 are associated with artemisinin resistance, but the underlying molecular mechanisms are unknown. We analyzed the in vivo transcriptomes of 1043 *P. falciparum* isolates from patients with acute malaria and found that artemisinin resistance is associated with increased expression of unfolded protein response (UPR) pathways involving the major PROSC and TRiC chaperone complexes. Artemisinin-resistant parasites also exhibit decelerated progression through the first part of the asexual intraerythrocytic development cycle. These findings suggest that artemisinin-resistant parasites remain in a state of decelerated development at the young ring stage, whereas their up-regulated UPR pathways mitigate protein damage caused by artemisinin. The expression profiles of UPR-related genes also associate with the geographical origin of parasite isolates, further suggesting their role in emerging artemisinin resistance in the Greater Mekong Subregion.

Artemisinin resistance in *Plasmodium falciparum* is spreading rapidly throughout Southeast Asia. Since it was first detected 7 years ago in Pailin, western Cambodia (*1*), artemisinin resistance has become prevalent in other Cambodian provinces (*2*, *3*), Thailand-Myanmar border areas (*4*, *5*), and southern Vietnam (*4*, *6*) and is emerging in southern Laos and central Myanmar (*4*). Artemisinin resistance threatens the

efficacy of all artemisinin-based combination therapies (ACTs) (*7*), as its spread from Southeast Asia to Africa would derail malaria control and elimination efforts worldwide. The parasite clearance half-life, derived from the log-linear decline in parasitemia over time, is now established as the best pharmacodynamics measure of *P. falciparum* sensitivity to artemisinins (*8*, *9*). Resistance is associated with a parasite clearance half-life of >5 hours (*4*) and is prevalent where standard 3-day courses of ACTs are now failing (*10*, *11*).

Artemisinin resistance is a heritable genetic trait (*12*) linked with three loci on chromosomes 10, 13, and 14 (*13*, *14*) and nonsynonymous single-nucleotide polymorphisms (SNPs) in the propeller domain of a *kelch* gene on chromosome 13 (*PF3D7_1343700*) (*3*). These "K13-propeller" polymorphisms are currently the best predictors of artemisinin resistance in mainland Southeast Asia, with the most common mutation, $Cys^{580} \rightarrow Tyr$, approaching genetic fixation in western Cambodia (*3*). A large multicenter clinical investigation by the Tracking Resistance to Artemisinins Collaboration (TRAC) has shown that the artemisinin resistance phenotype that is currently spreading through the Greater Mekong Subregion (GMS) is associated with K13 polymorphisms (*4*). Although K13-propeller mutations are highly predictive of resistance, very little is known about the molecular mechanisms that render *P. falciparum* insensitive to artemisinins. These drugs eliminate parasitemias more rapidly than other antimalarials because they accelerate the clearance of young ring-stage parasites by the spleen (*9*). Several lines of evidence indicate that resistance affects the early ring stage of the parasite's intraerythrocytic developmental cycle (*8*, *15*, *16*). Previously, we suggested that artemisinin resistance may be underlined by broad changes of the parasite's transcriptional program that alter its physiology (*17*).

Here, we carried out transcriptome analyses of 1043 clinical *P. falciparum* isolates to identify the transcriptionally underlined mechanisms that mediate artemisinin resistance. The parasite samples were obtained directly from the peripheral blood of patients with acute falciparum malaria enrolled into the TRAC study in 2011–2012 (*4*). This study was approved by the relevant local ethics committees and the Oxford Tropical Research Ethics Committee, and all adult patients or the parents of children gave written informed consent. The samples originated from 13 malaria-endemic regions of Southeast Asia and Africa. These include regions where artemisinin resistance is well established (Pursat and Pailin, Cambodia; Mae Sot, Thailand; and Binh Phuoc, Vietnam), emerging (Preah Vihear, Cambodia, and Shwe Kyin, Myanmar), or currently not detected [Ramu, Bangladesh,

[1]School of Biological Sciences, Nanyang Technological University, Singapore. [2]Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. [3]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [4]Department of Clinical Tropical Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. [5]Department of Molecular Tropical Medicine and Genetics, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. [6]WorldWide Antimalarial Resistance Network (WWARN), Asia Regional Centre, Mahidol University, Bangkok, Thailand. [7]WorldWide Antimalarial Resistance Network, University of Maryland School of Medicine, Baltimore, MD, USA. [8]National Center for Parasitology, Entomology and Malaria Control, Phnom Penh, Cambodia. [9]Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. [10]Oxford University Clinical Research Unit (OUCRU), Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam. [11]Department of Medical Research, Lower Myanmar, Yangon, Myanmar. [12]Malaria Research Group & Dev Care Foundation, Dhaka, Bangladesh. [13]Kinshasa School of Public Health, Kinshasa, Democratic Republic of the Congo. [14]Lao–Oxford–Mahosot Hospital–Wellcome Trust Research Unit, Mahosot Hospital, Vientiane, Lao PDR. [15]Faculty of Postgraduate Studies, University of Health Sciences, Vientiane, Lao PDR. [16]Medical Research Council (MRC) Centre for Genomics and Global Health, University of Oxford, Oxford, UK. [17]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. [18]Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand.
***Corresponding author. E-mail: zbozdech@ntu.edu.sg**

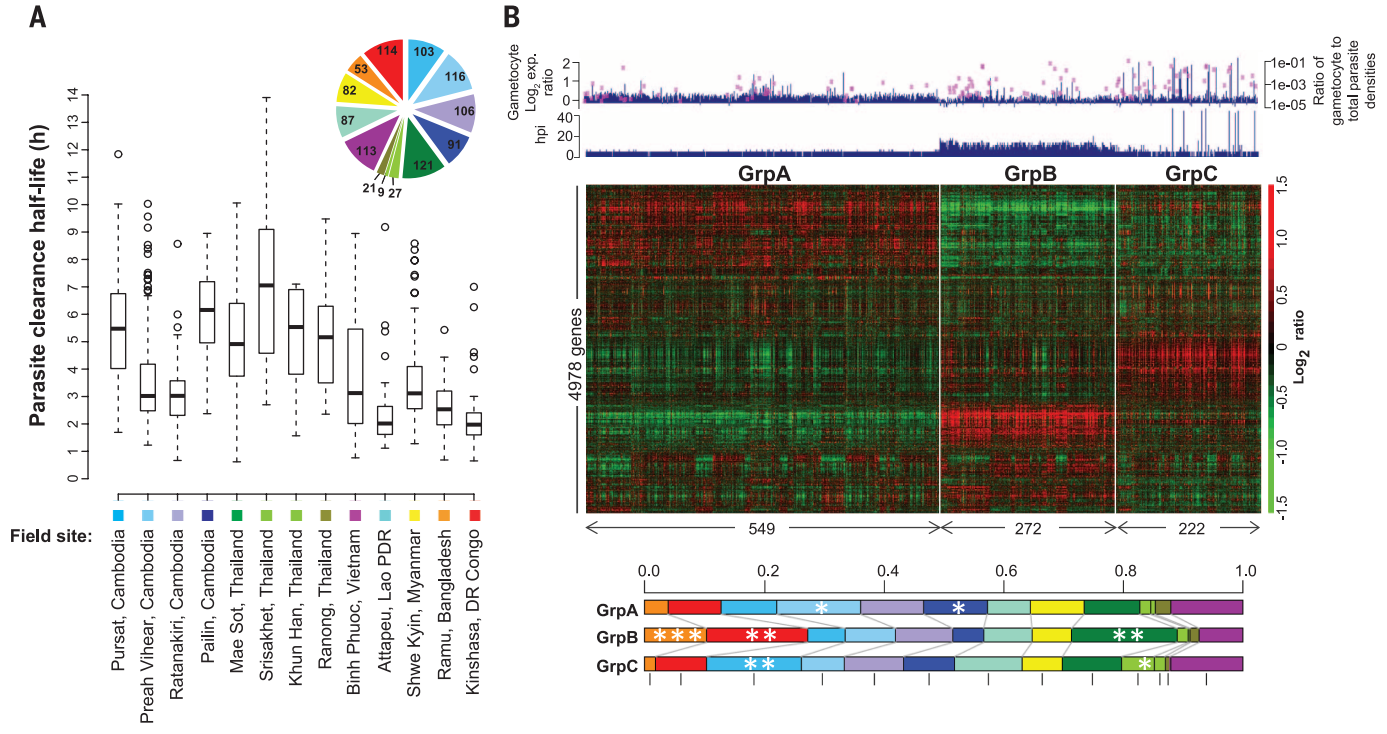


Fig. 1. Artemisinin response and transcription profiles of *P. falciparum* isolates from patients with acute malaria. (**A**) The pie chart and boxplots show the distribution of parasite isolates and parasite clearance half-lives, respectively, according to field site. (**B**) The heat map shows mean-centered relative expression levels (log_2 ratios) grouped by *k*-means clustering: GrpA (n = 549), GrpB (n = 272), and GrpC (n = 222). Corresponding gametocyte densities per microliter (pink squares) (*4*), relative expression levels of gametocyte-specific genes (blue bars), and parasite age (hpi) are shown above; distributions of geographical origins according to parasite group (colored bars) are shown below. *P < 0.05, **P < 0.01, ***P < 0.001.

and Kinshasa, Democratic Republic of Congo (DR Congo)] (Fig. 1A and table S1). By measuring the overall abundance of mRNA transcripts for 4978 of the ~5591 genes in the *P. falciparum* genome, we identified three types of transcription profiles among the parasite isolates (GrpA, GrpB, and GrpC) (Fig. 1B). Projecting the field isolate transcription profiles onto the in vitro reference (*17*) revealed that the three transcriptional groups are mainly defined by parasite developmental stage ($P < 2 \times 10^{-16}$) (fig. S1). This analysis showed that GrpA consisted mostly of early ring-stage parasites [8 to 10 hours post-invasion (hpi)], and GrpB consisted predominantly of middle and late ring-stage parasites (10 to 20 hpi). GrpC also consisted predominantly of early ring-stage parasites (8 to 10 hpi); however, in these isolates the density of gametocytemia was significantly higher than in the two other groups (fig. S1).

The three transcriptional patterns were significantly associated with several clinical and parasitological parameters (fig. S1). Artemisinin-resistant parasites were unevenly distributed among the three groups. GrpA was associated with a longer parasite clearance half-life (mean 4.16 hours) relative to GrpB (3.43 hours). However, GrpB contained a higher proportion of parasites with a half-life of <5 hours (fig. S1), largely due to its overrepresentation of parasites from areas of Bangladesh and DR Congo that are apparently free of artemisinin resistance (Fig. 1B). The transcriptional profiles of GrpA and GrpC parasites were found to be associated with parasite clearance times elapsed time until parasitemia is reduced by 50% (PC50) and 90% (PC90) of its initial admission value; $P < 2 \times 10^{-16}$] and the duration of the lag phase of the parasite clearance curve (*9*) ($P = 4 \times 10^{-15}$) (fig. S1). The lag phase represents the period immediately after administration of the first artemisinin dose when the initial parasite density remains constant before commencing a first-order decline. The extended lag phases in GrpA and GrpC patients may be caused by higher densities of early ring-stage parasites in the peripheral blood, resulting from recent schizont rupture. This is suggested by the fact that these patients presented to clinic with higher temperatures, likely caused by pyrogenic factors released from rupturing schizonts into the bloodstream (*18*).

GrpB patients were significantly less likely to present to clinic in the evening hours than GrpA and GrpC patients ($P = 8 \times 10^{-11}$) (fig. S1). This suggests that in areas where primary care is readily available, patients tend to seek medical care for fever shortly after synchronous schizont rupture anytime throughout the day, including the evening hours. Patients with high fever at clinical presentation may thus have a higher chance of developing a longer lag phase that prolongs the overall clearance time, regardless of their artemisinin resistance status.

Next, we focused on the uniformly young ring-stage parasites in GrpA ($n = 549$) and carried out linear regression between mRNA levels and parasite clearance half-life (Fig. 2A and fig. S2A). We found that 487 (9.6%) and 511 (10.1%) genes were

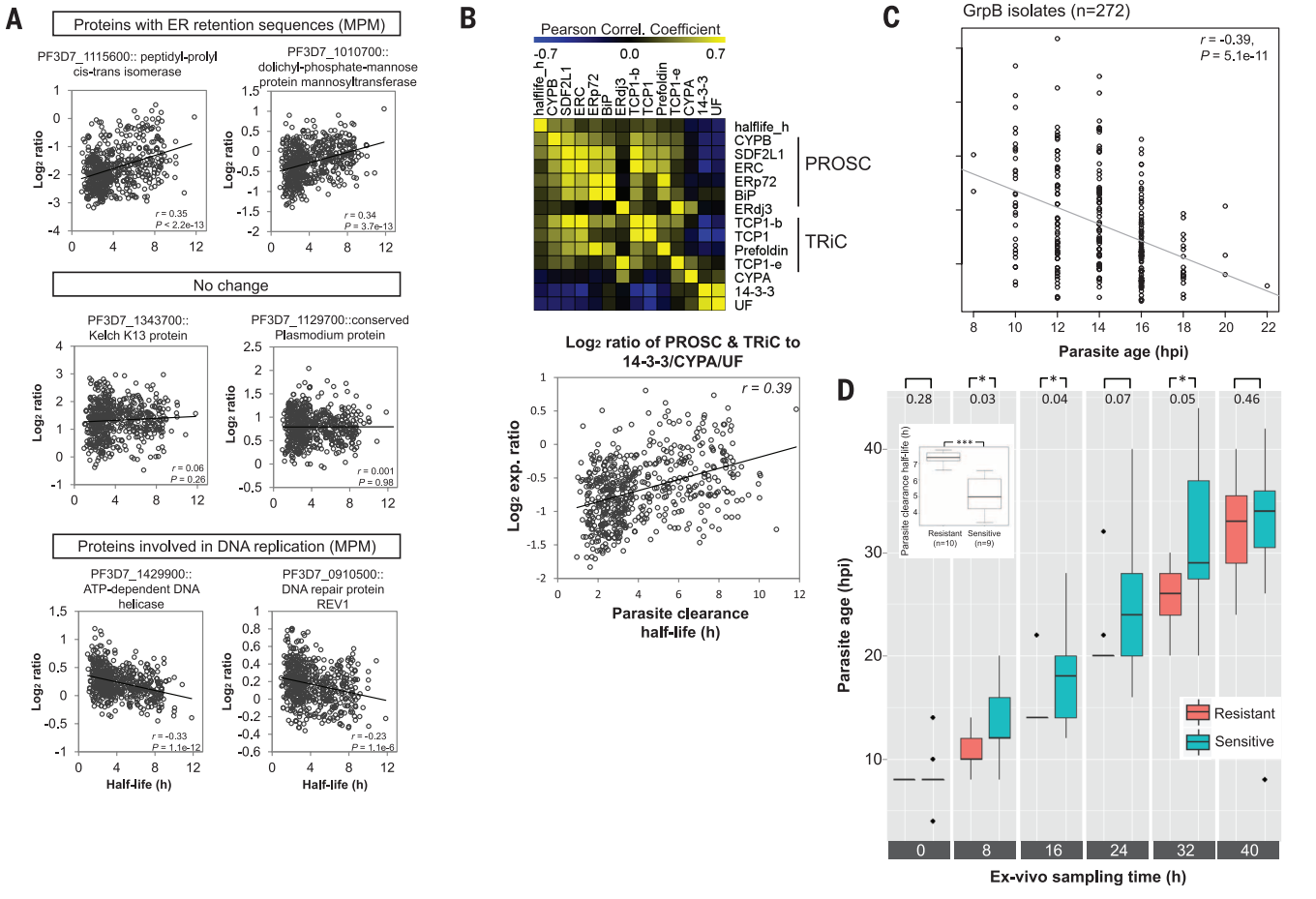


Fig. 2. Transcriptional features associated with artemisinin resistance. (**A**) Scatterplots depict examples of the linear regression analysis between mRNA levels and parasite clearance half-life for two genes with positive, negative, or no correlation. The middle left panel shows the correlation for the K13 artemisinin resistance marker. (**B**) The heat map shows the Pearson correlation coefficients of pairwise comparisons of the 13 differentially expressed genes within the chaperone network. The scatterplot shows the ratio of the average expression of *Plasmodium* PROSC and TRiC to that of the three down-regulated genes (14-3-3, Cyp19A, and UF). (**C**) Distribution of parasite age (hpi) and clearance half-life for 272 GrpB isolates. (**D**) Boxplots show the age (hpi) of parasites over 40 hours in ex vivo culture, stratified by their artemisinin resistance status (*$P < 0.05$).

significantly up-regulated and down-regulated, respectively, in association with the clearance half-life and thus artemisinin resistance [$P < 0.01$, false discovery rate (FDR) < 0.05] (fig. S2A and table S2). In particular, artemisinin resistance appears to be associated with up-regulation of genes involved in protein metabolism, such as endoplasmic reticulum retention sequences, unfolded protein binding, protein folding, protein export, posttranslational translocation, signal recognition particle (SRP), proteasome, and phagosome ($P < 0.05$, FDR < 0.25) (fig. S2A). A nearly identical set of functional pathways was found to be up-regulated in GrpA parasites with K13-propeller mutations (fig. S2B). There was no correlation between K13 mRNA levels and artemisinin resistance (Fig. 2A). Most of the up-regulated pathways are known to participate in the overall unfolded protein response (UPR) in other eukaryotic species. Hence, these results indicate that an up-regulated UPR may be a major mediator of artemisinin resistance in *P. falciparum* and is caused by single K13-propeller mutations. Similar functional representation of artemisinin resistance- and K13-linked up-regulated genes was found in GrpC, although the statistical significance was reduced because of greater gene expression variability caused by varying levels of gametocytes (fig. S2C).

Among the genes with the highest correlation with artemisinin resistance are those involved in protein folding and repair, including cyclophilin19B ($P < 2 \times 10^{-16}$), dolichyl-phosphate-mannose protein mannosyltransferase ($P = 3 \times 10^{-13}$), endoplasmic reticulum–resident calcium binding protein ($P = 0.004$), BiP/grp78 ($P = 0.004$), and a protein disulfide isomerase ($P = 10^{-6}$) (Fig. 2A and table S2). Consequently, we used previous experimental and bioinformatics data (*19*, *20*) to assemble a network of molecular chaperones consisting of two major protein complexes and other chaperones in the *P. falciparum* cell (fig. S3). These include subunits of two putative chaperonin complexes, *Plasmodium* reactive oxidative stress complex (PROSC) (table S3) and TCP-1 ring complex (TRiC), that participate in the UPR of other species (*21*, *22*). The mRNA levels for the PROSC and TRiC subunits correlated not only with each other but also with artemisinin resistance (Fig. 2B and fig. S3). Moreover, these mRNA levels inversely correlated with those of other chaperones such as Pf14-3-3, cyclophilin19A, and PF3D7_1024800 (protein of unknown function, UF), all of which are components of the chaperone network in *Plasmodium* (Fig. 2B). We also observed strong correlation

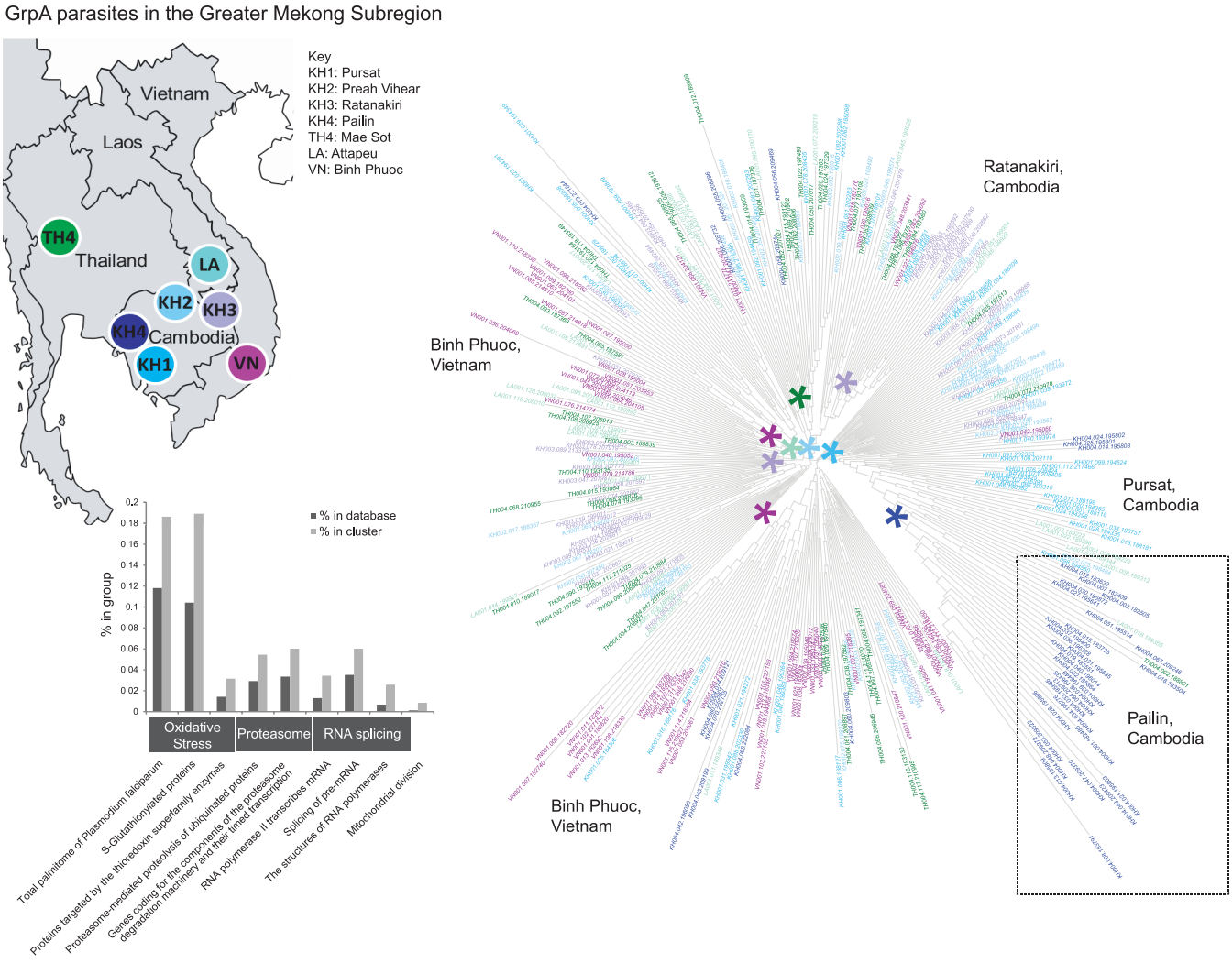


Fig. 3. Population transcriptomics analysis of *P. falciparum* isolates. Cladogram of the 348 GrpA isolates from the Greater Mekong Subregion, clustered by the Euclidean distance metric of 659 marker genes. The geographic origin of each isolate is indicated by the color key in the map. Stars indicate significant overrepresentation of isolates from a particular field site within the corresponding node ($P < 0.05$, hypergeometric test). Histograms of functional enrichment analysis show overrepresented and differentially expressed pathways in the demarcated Pailin clade ($P < 0.05$, hypergeometric test). Bars represent the proportion of Pailin clade-specific genes (light gray) relative to the total number of genes (dark gray).

between parasite clearance half-life and the relative expression ratio of PROSC and TRiC (average mRNA level) versus Pf14-3-3 (r = 0.39, P < 2.2 × 10^{-16}) (Fig. 2B). This correlation is significantly stronger (P = 0.01) than that between half-life and expression of any individual gene within or outside the chaperone network (e.g., maximum r = 0.35 for *cyclophilin19B*), indicating that artemisinin resistance is associated with the coordinated transcription of multiple chaperone partners.

In GrpA, artemisinin resistance and K13-propeller polymorphisms also associated with down-regulation of genes involved in DNA replication (Fig. 2A and fig. S2, A and B). Although DNA replication occurs later than 8 to 10 hpi (*23*), this down-regulation may represent a developmental stalling of resistant parasites in which transcription of the DNA replication genes is "halted." Accordingly, the age (hpi) of GrpB parasites correlated inversely with clearance half-life (r = −0.39, P < 0.001) (Fig. 2C). To analyze this relationship further, we investigated the ex vivo transcriptomes of 9 sensitive and 10 resistant parasites (median half-lives, 4.99 and 7.35 hours) from Pailin, sampled from patients at 0 hours and during ex vivo cultivation at 8, 16, 24, 32, and 40 hours (fig. S4). Estimating the parasite age at each time point, we observed that the resistant isolates progressed from rings to trophozoites to early schizonts at a slower rate than sensitive isolates (P < 0.05) (Fig. 2D). This deceleration in blood-stage development occurred during the first 32 hours. At 40 hours, presumably at the end of DNA replication, the resistant isolates reached the same stage of development as their sensitive counterparts, thus completing their intraerythrocytic life cycle at the same time.

Up-regulation of the UPR may enable resistant parasites to better withstand the deleterious effect of Fe^{2+}-activated artemisinins, which are believed to damage intracellular structures via direct alkylation and oxidation (*24*). An up-regulated UPR is likely to increase the capacity of parasites to quickly repair or degrade proteins (or other cellular components) that are damaged by brief artemisinin exposures in patients. The two chaperonin complexes, PROSC and TRiC, are likely to play a pivotal role in this process, as they do in the UPR of other eukaryotes (*21*, *22*). Up-regulation of the UPR is also consistent with the putative function of kelch proteins as negative regulators of signal transduction, triggering the UPR upon protein damage (*25*). In resistant parasites, mutant K13 proteins may lose their ability to negatively regulate this pathway, leading to its constitutive activation at baseline, increased activation upon artemisinin exposure, or both. Finally, inverse relationships between members of the *Plasmodium* chaperone network—especially that between cyclophilin19B and cyclophilin19A—may reflect specific adjustments in protein-folding activities in the *Plasmodium* cytoplasm. These proteins have been suggested to function as rate-limiting steps in protein folding in other eukaryotes, and their differential expression is also thought to reflect readjustments of the UPR (*26*).

Intriguingly, UPR-related genes, including those encoding PROSC and TRiC, show high variability among the GMS field sites. In investigating the geographical distribution of transcription profiles, we first noted two distinct clades of parasites based on Pearson correlation of the individual transcriptome in all three transcriptional groups (fig. S5). Whereas one clade contained mainly parasites from DR Congo and Bangladesh, the other was strongly enriched for parasites from the GMS. In the GMS clade for GrpA, we found 659 differentially expressed genes that define parasite populations in distinct areas. Among these, there was a high enrichment of biological processes associated with artemisinin resistance, including protein turnover and oxidative damage response (Fig. 3). The transcriptional variability of these genes is particularly evident in Pailin, where *P. falciparum* resistance to several classes of drugs (quinolines, antifolates, and artemisinins) has originated (*7*, *27*) and where the parasite population exhibits a distinct genetic structure (*28*). This suggests that the propensity of Pailin parasites to develop drug resistance may be maintained by differential expression of genes involved in response to oxidative or other types of stress. It seems reasonable to speculate that this transcriptional heterogeneity is driven by frequent exposure to oxidative stress due to the high prevalence of hemoglobinopathies in this region (*2*).

The decelerated progression of ring-stage development is also consistent with the artemisinin mode of action and putative mechanism of resistance. Ring-stage arrest has been associated with lower levels of endocytosis and hemoglobin digestion, which decreases free heme-mediated activation of artemisinin and enhances the parasite's tolerance to this drug (*29*). This phenotype may also relate to *P. falciparum* dormancy, which has been suggested to enable the parasite to withstand the drug pressure and resume growth in favorable conditions (*16*, *30*). Up-regulation of the UPR may be directly linked with decelerated parasite development as a mechanism to connect the repair of damaged proteins to cell cycle progression, as observed in other eukaryotes (*25*). Investigating these two phenomena in *P. falciparum* may improve our understanding of the molecular basis of artemisinin resistance and facilitate the development of new strategies to counter the threat it poses to global malaria control and elimination.

REFERENCES AND NOTES

1. A. M. Dondorp *et al.*, *N. Engl. J. Med.* **361**, 455–467 (2009).
2. C. Amaratunga *et al.*, *Lancet Infect. Dis.* **12**, 851–858 (2012).
3. F. Ariey *et al.*, *Nature* **505**, 50–55 (2014).
4. E. A. Ashley *et al.*, *N. Engl. J. Med.* **371**, 411–423 (2014).
5. A. P. Phyo *et al.*, *Lancet* **379**, 1960–1966 (2012).
6. T. T. Hien *et al.*, *Malar. J.* **11**, 355 (2012).
7. A. M. Dondorp *et al.*, *N. Engl. J. Med.* **365**, 1073–1075 (2011).
8. S. Saralamba *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 397–402 (2011).
9. N. J. White, *Malar. J.* **10**, 278 (2011).
10. R. Leang *et al.*, *Antimicrob. Agents Chemother.* **57**, 818–826 (2013).
11. D. L. Saunders, P. Vanachayangkul, C. Lon, *N. Engl. J. Med.* **371**, 484–485 (2014).
12. T. J. Anderson *et al.*, *J. Infect. Dis.* **201**, 1326–1330 (2010).
13. I. H. Cheeseman *et al.*, *Science* **336**, 79–82 (2012).
14. S. Takala-Harrison *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 240–245 (2013).
15. N. Klonis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5157–5162 (2013).
16. B. Witkowski *et al.*, *Lancet Infect. Dis.* **13**, 1043–1049 (2013).
17. S. Mok *et al.*, *BMC Genomics* **12**, 391 (2011).
18. M. S. Oakley, N. Gerald, T. F. McCutchan, L. Aravind, S. Kumar, *Trends Parasitol.* **27**, 442–449 (2011).
19. D. J. LaCount *et al.*, *Nature* **438**, 103–107 (2005).
20. S. R. Pavithra, R. Kumar, U. Tatu, *PLOS Comput. Biol.* **3**, e168 (2007).
21. L. Ellgaard, A. Helenius, *Nat. Rev. Mol. Cell Biol.* **4**, 181–191 (2003).
22. P. Tongaonkar, M. E. Selsted, *J. Biol. Chem.* **284**, 5602–5609 (2009).
23. Z. Bozdech *et al.*, *PLOS Biol.* **1**, e5 (2003).
24. S. R. Meshnick, T. E. Taylor, S. Kamchonwongpaisan, *Microbiol. Rev.* **60**, 301–315 (1996).
25. H. Digaleh, M. Kiaei, F. Khodagholi, *Cell. Mol. Life Sci.* **70**, 4681–4694 (2013).
26. F. U. Hartl, M. Hayer-Hartl, *Science* **295**, 1852–1858 (2002).
27. A. A. Khattak *et al.*, *Malar. J.* **12**, 310 (2013).
28. O. Miotto *et al.*, *Nat. Genet.* **45**, 648–655 (2013).
29. N. Klonis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11405–11410 (2011).
30. Q. Cheng, D. E. Kyle, M. L. Gatton, *Int. J. Parasitol. Drugs Drug Resist.* **2**, 249–255 (2012).

ACKNOWLEDGMENTS

Supported by Singapore National Medical Council grants NMRC/1265/2010 and NMRC/1292/2011 and by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases (NIAID). We are very grateful to the patients, doctors, nurses, and laboratory technicians who participated in the TRAC clinical studies. TRAC was funded by the UK Department for International Development (DFID) for the benefit of developing countries and coordinated by the Wellcome Trust Mahidol University Oxford Tropical Medicine Research Programme funded by the Wellcome Trust of Great Britain. Whole-genome sequencing and genotyping for TRAC were funded by the Wellcome Trust through core funding of the Wellcome Trust Sanger Institute (grant 098051). Additional support was given by the WorldWide Antimalarial Resistance Network, the Intramural Research Program of NIAID, and the Bill & Melinda Gates Foundation. The views expressed and information contained in this publication are not necessarily those of or endorsed by DFID or the other funders, which can accept no responsibility for such views or for any reliance placed on them. We thank W. Guo, K.-Y. Liong, and J. Sim for their assistance. The microarray data have been deposited in NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number GSE59099.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/431/suppl/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S3
References (*31*–*42*)

25 August 2014; accepted 1 December 2014
Published online 11 December 2014;
10.1126/science.1260403

CHRONIC INFECTION

Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds

M. Asghar,[1,2] D. Hasselquist,[1]*† B. Hansson,[1] P. Zehtindjiev,[3] H. Westerdahl,[1] S. Bensch[1]†

Recovery from infection is not always complete, and mild chronic infection may persist. Although the direct costs of such infections are apparently small, the potential for any long-term effects on Darwinian fitness is poorly understood. In a wild population of great reed warblers, we found that low-level chronic malaria infection reduced life span as well as the lifetime number and quality of offspring. These delayed fitness effects of malaria appear to be mediated by telomere degradation, a result supported by controlled infection experiments on birds in captivity. The results of this study imply that chronic infection may be causing a series of small adverse effects that accumulate and eventually impair phenotypic quality and Darwinian fitness.

The harmful consequences of infections are well understood in humans and wild and domesticated animals (*1*, *2*), and disease is known to be a key factor in sexual selection (*3*) and life-history evolution (*4*). In wild animals, the cost of disease is most apparent during acute infection, when the host shows sickness behaviors and becomes inefficient in avoiding predators, finding food, and defending territory (*5*). After surviving the acute phase, either the infection is cleared or it may become chronic, often without apparent clinical signs. To study the ultimate long-term effects of infection in natural populations, and the potential proximate mechanisms underlying such effects, detailed data are required from individuals measured repeatedly over their entire lives. Unfortunately, such data are scarce for animals.

We have investigated the ultimate (Darwinian fitness) costs of avian malaria caused by *Plasmodium* and *Haemoproteus* spp. in great reed warblers (*Acrocephalus arundinaceus*), using detailed life-history data collected during 25 years in a wild population breeding at Lake Kvismaren, Sweden (*6*–*8*). The great reed warbler breeds in Eurasia and winters in tropical Africa. Previous studies based on parasite mitochondrial cytochrome b genes have identified 17 different malaria parasite lineages in the great reed warbler study population, all transmitted in Africa (*9*, *10*). This host-parasite system is well suited for investigating long-term effects of infection, because avian malaria disease is characterized by a severe acute phase when individuals first contract the infection, after which most of the survivors retain a persistent, low-level chronic infection (*11*). In wild bird populations, the fitness costs to hosts of being infected are typically not detectable (*12*), probably because individuals with acute infections are difficult to find and study. Moreover, the short-term effects of chronic infection are often reported to be benign over the time scales that infected individuals have been studied (*10*, *12*). On blood samples from great reed warblers, we have used nested polymerase chain reaction (PCR) for DNA-based detection and identification of malaria parasites [*Haemoproteus* and *Plasmodium* (*13*, *14*)] and quantitative PCR (qPCR) for determining infection intensity (*9*, *14*, *15*). In our study population, uninfected birds and birds with low-level nonsymptomatic infections are equally common, and we have detailed knowledge of infection history and reproduction throughout the individuals' lives (*9*, *10*, *14*). Low-level chronic malaria appears to have no direct short-term costs to great reed warblers in our study population. There is no difference in activity patterns between infected and uninfected birds, either in song rate in males [mean ± SEM song rate (percent of time singing per 15 min): infected = 59.3 ± 1.47 (n = 14 birds), uninfected = 58.7 ± 0.83 (n = 12); $t_{1,24}$ = 0.32, P = 0.76; fig. S1A (*14*)] or in nestling feeding rate in females [mean ± SEM feedings per hour and nestling: infected = 2.07 ± 0.91 (n = 21), uninfected = 1.96 ± 0.92 (n = 28); $t_{1,47}$ = 0.41, P = 0.68; fig. S1B (*14*)].

We found that birds infected with malaria had significantly shorter life spans [mean ± SEM years: infected = 1.69 ± 0.15; uninfected = 2.53 ± 0.26; P = 0.006 (Fig. 1A)] and fewer fledglings during their lifetimes [mean ± SEM fledglings: infected = 4.03 ± 0.74; uninfected = 8.63 ± 1.52; P = 0.009 (Fig. 1B)] than uninfected birds. When controlling for life span [general linear model (GLM): $F_{1,75}$ = 78.17, P < 0.0001], the negative effect of malaria infection on lifetime fledging success was not significant [$F_{1,75}$ = 0.89, P = 0.35 (table S1)], which is consistent with a previous study in which we found no difference in annual reproductive success or survival to the next year between infected and uninfected great reed

[1]Department of Biology, Lund University, Ecology Building, 223 62 Lund, Sweden. [2]Infectious Disease Unit, Department of Medicine Solna, Karolinska Institute, 17176 Stockholm, Sweden. [3]Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, 2 Gagarin Street, 1113 Sofia, Bulgaria.
*Corresponding author. E-mail: dennis.hasselquist@biol.lu.se
†These authors contributed equally to this work.

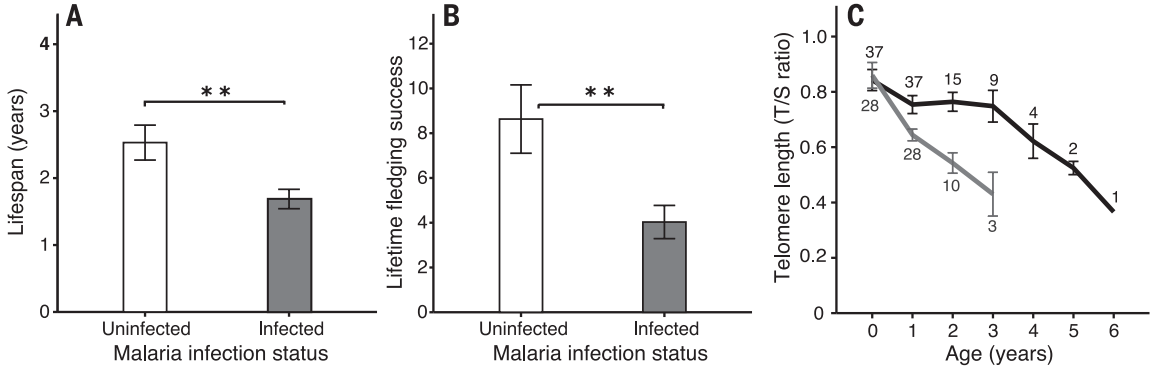


Fig. 1. Relationships between malaria infection status, life span, lifetime fledging success, and telomere length in wild great reed warblers. (**A**) Uninfected birds lived longer than birds with chronic malaria infection (t test: $t_{1,71.6}$ = 2.82, P = 0.006). (**B**) Uninfected birds produced more lifetime fledglings than chronically infected birds ($t_{1,67.5}$ = 2.71, P = 0.009). (**C**) Relationships between telomere length and age in uninfected (black line) and infected (gray line) great reed warblers (over life). Telomere length decreased with age (LME model: age, $F_{1,106}$ = 91.78, P < 0.0001) but at a steeper rate in infected birds (age × malaria status, $F_{1,106}$ = 27.09, P < 0.0001; table S2). Error bars represent mean ± SEM. n values are given for each age group (uninfected above and infected below error bars).

warblers (*10*). Taken together, these data show that a disease (chronic avian malaria) with seemingly negligible short-term costs over the long term induces physiological effects that shorten life span and reduce the number of lifetime offspring.

Recent research has linked telomere shortening to aging (*16*, *17*) and some degenerative diseases (*18*, *19*), and a few observational studies in humans (infected with, e.g., hepatitis C virus or HIV) and mice (infected with *Salmonella*) have shown that infected individuals have shorter telomeres (*20*, *21*). Telomeres are nucleoprotein structures that cap the ends of chromosomes to maintain chromosome integrity (*22*). Telomeres shorten at each cell division and if exposed to oxidative stress. When telomeres become too short, cell death occurs, which can accumulate in tissues to cause organ dysfunction and aging effects (*23*, *24*). Hence, telomere shortening is a possible mechanism for mediating long-term fitness costs of chronic or recurrent mild infections. We used qPCR (*14*, *25*, *26*) to quantify telomere length in blood samples of great reed warblers taken when 8 to 10 days old and then each year throughout their lives when breeding at our study site. We estimated telomere shortening as the between-year change in telomere length (*14*). These data allowed us to estimate early-life telomere length as well as the annual rate of telomere shortening. We found that telomere length in great reed warblers decreased with age ($P < 0.0001$) and more steeply so in infected than in uninfected birds [as shown by the significant interaction age × malaria infection status, $P < 0.0001$ (Fig. 1C and table S2]. Individuals infected with malaria experienced a significantly greater rate of telomere shortening than uninfected individuals [linear mixed-effect (LME) model: malaria infection status, $F_{1,61} = 24.2$, $P < 0.0001$ (table S3)]. The difference between infected and uninfected individuals was also present when we analyzed malaria infection status and telomere shortening over the birds' first year of life, for all parasite lineages combined ($P = 0.0001$) and for the three most common parasite lineages separately (Fig. 2A). The degree of telomere loss seemed to be related to the severity of the infection, because we found a marginally positive correlation between the rate of telomere loss over the birds' first year of life

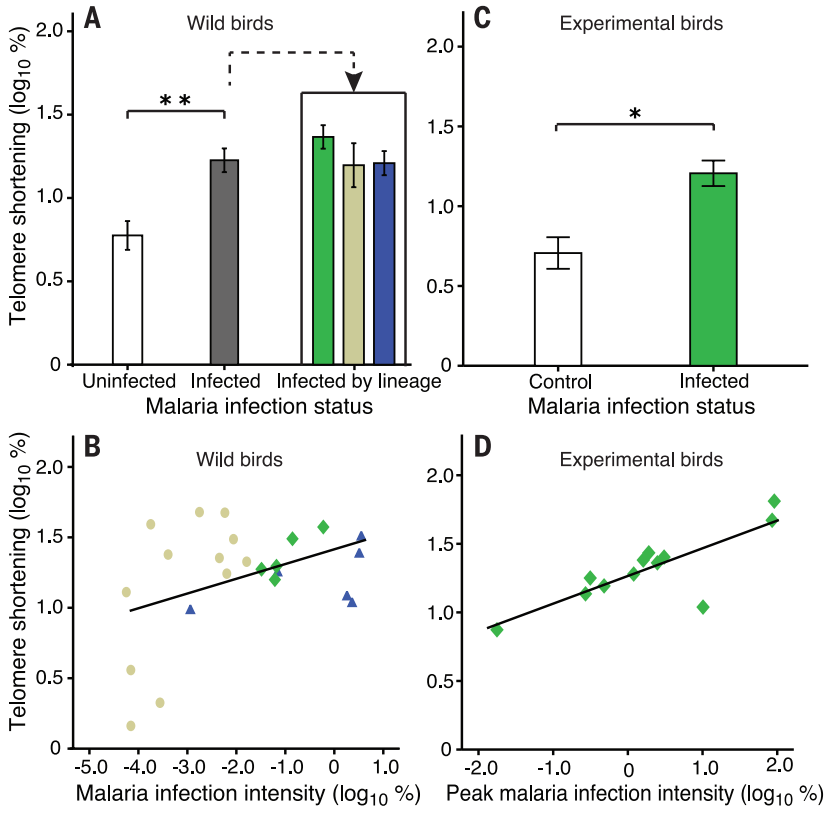


Fig. 2. Relationships between telomere shortening and malaria infection status in great reed warblers. (**A**) In wild birds, uninfected birds (white) experienced a lower degree of telomere shortening over their first year than infected birds (gray) for the common avian malaria lineages combined (*t* test: $t_{1,62.9} = 3.96$, $P = 0.0001$; $n_{uninfected} = 37$ birds, $n_{infected} = 28$), and separately for *P. ashfordi* (green; $t_{1,40} = 2.47$, $P = 0.018$), *P. relictum* (yellow; $t_{1,49} = 2.59$, $P = 0.013$), and marginally so for *Haemoproteus nucleocondensus* (blue; $t_{1,41} = 1.95$, $P = 0.058$). (**B**) In 1-year-old wild birds, there was a marginally positive correlation between malaria infection intensity and telomere shortening for the common malaria parasite lineages combined ($r = 0.41$, $P = 0.051$, $n = 23$), a relationship that was significant for *P. ashfordi* ($r = 0.89$, $P = 0.043$, $n = 5$) and *P. relictum* ($r = 0.60$, $P = 0.038$, $n = 12$) but not for *H. nucleocondensus* ($r = 0.52$, $P = 0.28$, $n = 6$). (**C**) In captive great reed warblers, there was a higher rate of telomere shortening over 9 to 10 weeks in birds experimentally infected with *P. ashfordi* than in uninfected controls (GLM: $F_{1,12} = 19.59$, $P = 0.0008$, $n_{infected} = 12$, $n_{control} = 4$), and (**D**) in the infected captive birds, there was a positive correlation between peak malaria infection intensity and telomere shortening ($r = 0.82$, $P = 0.001$, $n = 12$).

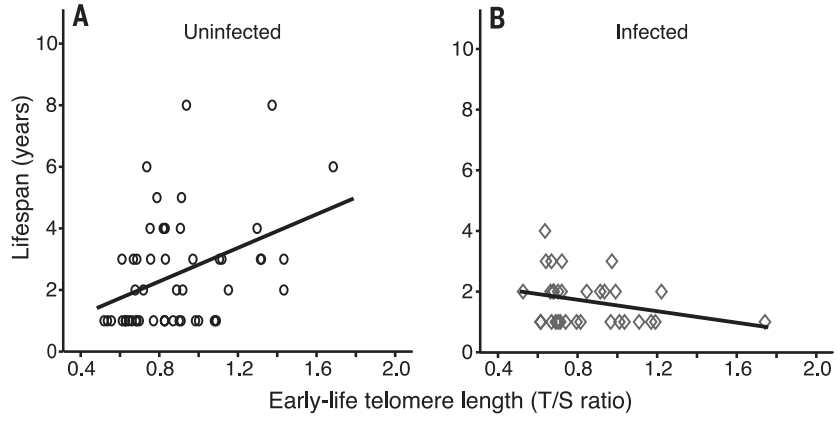


Fig. 3. Relationships between life span and early-life telomere length (TL; at 8 to 10 days of age) in uninfected and infected wild great reed warblers. (**A**) In uninfected birds (circles, black line), there was a significant positive relationship between early-life TL and life span ($r = 0.40$, $P = 0.005$, $n = 49$), whereas (**B**) in malaria-infected birds (diamonds, gray line), there was no such relationship ($r = -0.29$, $P = 0.11$, $n = 32$). The difference in slopes is supported by the significant interaction early-life TL × malaria status ($P = 0.003$).

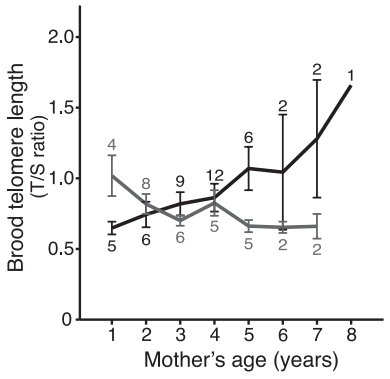


Fig. 4. Relationships between mother's age and the mean early-life telomere length of her offspring in relation to mother's malaria infection status in wild great reed warblers. Uninfected (black line) and malaria-infected (gray line) mothers differed in the relationship between the mother's age and the brood mean telomere length of her offspring. A LME model showed effects of mother's age × mother's malaria status ($F_{1,142}$ = 71.53, P < 0.0001), mother's age ($F_{1,142}$ = 6.96, P = 0.011), and brood year ($F_{1,142}$ = 26.76, P < 0.0001; table S6). Uninfected mothers (n = 20), as they became older, produced broods in which nestlings had longer mean telomeres (P < 0.0001), whereas in malaria-infected mothers (n = 19), this relationship was reversed (P = 0.0004). n values (broods) are given for each age group, and error bars represent mean ± SE of brood means (*14*). The statistical analysis was conducted using a LME model with the early-life telomere length of individual offspring as a dependent variable, the mother's identity as a random factor, and broods fitted into mother as a nested random factor.

and malaria infection intensity for the three parasite species combined (P = 0.051), a pattern that was significant in each of the two *Plasmodium* species (Fig. 2B).

In a combined analysis of malaria infection status and early-life telomere length on life span, we found a significant relationship between life span and malaria infection status (GLM: $F_{1,75}$ = 5.10, P = 0.027; table S4) and a significant positive relationship between life span and early-life telomere length (8- to 10-day-old nestlings; $F_{1,75}$ = 14.51, P = 0.0003). The interaction malaria infection status × early-life telomere length was also significant ($F_{1,75}$ = 9.15, P = 0.003). Thus, in birds remaining uninfected throughout their lives, there was a positive relationship between life span and early-life telomere length (r = 0.40, P = 0.005; Fig. 3A). In contrast, there was no such relationship in birds infected in their first year of life and onward (r = −0.29, P = 0.11; Fig. 3B).

Our study shows a positive relationship between early-life telomere length and life span in a species living under natural conditions. Such relationships have previously been found in comparisons across species and generally interpreted as reflecting the pace of life, where species with high energy turnover rates show fast telomere loss rates and short life spans (*16*, *27*). Within species, there is very little information on the relationship between early-life telomere length and life span apart from one study on captive birds (*16*).

To confirm the relationship between malaria infection and telomere shortening, we analyzed data from an infection experiment (*14, 28*) with *Plasmodium ashfordi*, one of the three common malaria parasites in great reed warblers. Experimentally infected great reed warblers experienced a higher rate of telomere loss than uninfected control birds over 9 to 10 weeks after inoculation (GLM: $F_{1,12}$ = 19.59, P = 0.0008; table S5 and Fig. 2C). In the experimentally infected birds, we also found a significant positive correlation between peak acute phase malaria infection intensity and rate of telomere loss (r = 0.82, P = 0.001; Fig. 2D).

Inspired by recent research suggesting that telomere length in offspring is influenced by parental effects (*25*, *29*), we investigated whether malaria infection in parent great reed warblers affected telomere length in their offspring. We found that malaria-infected mothers produced offspring with, on average, shorter telomeres in comparison to the offspring of uninfected mothers (t test: $t_{1,170.0}$ = 3.45, P = 0.0007), but no such effect was found in fathers ($t_{1,131.7}$ = −1.43, P = 0.16). Early-life telomere length in offspring was positively correlated with their mother's age when the eggs were laid (LME model: P = 0.011, Fig. 4), and there was also a significant interaction mother's age × mother malaria infection status (P < 0.0001; Fig. 4 and table S6). Hence, in uninfected mothers, there was a significant positive relationship between the mother's age (at the breeding event) and offspring early-life telomere length (LME model: $F_{1,75}$ = 29.42, P < 0.0001), whereas in malaria-infected mothers, this relationship was significantly negative ($F_{1,61}$ = 14.08, P = 0.0004; Fig. 4 and fig. S2). Thus, as uninfected mothers become older, they produce chicks with (over the years) successively longer early-life telomeres. In contrast, as infected mothers become older, they produce chicks with successively shorter early-life telomeres. These results suggest that telomere length in offspring, a proxy for phenotypic quality, can be influenced by environmental effects mediated by the mother (*29*, *30*), a finding that may have implications for our understanding of telomere dynamics and aging.

Our study combines longitudinal data from a wild population and experiments to show that individuals with chronic infections suffer long-term costs that act via telomere shortening to have consequences for survival, lifetime reproductive success, and offspring quality.

REFERENCES AND NOTES

1. T. M. Uyeki, *Clin. Infect. Dis.* **49**, 279–290 (2009).
2. S. I. Hay *et al.*, *PLOS Med.* **7**, e1000290 (2010).
3. W. D. Hamilton, M. Zuk, *Science* **218**, 384–387 (1982).
4. B. C. Sheldon, S. Verhulst, *Trends Ecol. Evol.* **11**, 317–321 (1996).
5. N. T. Ashley, J. C. Wingfield, in *Eco-Immunology*, R. M. Nelson, G. Demas, Eds. (Oxford Univ. Press, Oxford, 2012), pp. 45–91.
6. S. Bensch, *J. Anim. Ecol.* **65**, 283–296 (1996).
7. B. Hansson *et al.*, *Mol. Ecol.* **9**, 1529–1538 (2000).
8. D. Hasselquist, *Ecology* **79**, 2376–2390 (1998).
9. M. Asghar, D. Hasselquist, S. Bensch, *J. Avian Biol.* **42**, 530–537 (2011).
10. S. Bensch *et al.*, *J. Anim. Ecol.* **76**, 112–122 (2007).
11. M. Asghar *et al.*, *Parasitology* **139**, 1246–1252 (2012).
12. D. A. LaPointe, C. T. Atkinson, M. D. Samuel, *Ann. N.Y. Acad. Sci.* **1249**, 211–226 (2012).
13. J. Waldenström, S. Bensch, D. Hasselquist, O. Ostman, *J. Parasitol.* **90**, 191–194 (2004).
14. Materials and methods are available as supplementary materials on *Science* Online.
15. H. Westerdahl, M. Asghar, D. Hasselquist, S. Bensch, *Proc. Biol. Sci.* **279**, 577–584 (2012).
16. B. J. Heidinger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1743–1748 (2012).
17. R. M. Cawthon, K. R. Smith, E. O'Brien, A. Sivatchenko, R. A. Kerber, *Lancet* **361**, 393–395 (2003).
18. M. Jaskelioff *et al.*, *Nature* **469**, 102–106 (2011).
19. E. Sahin, R. A. Depinho, *Nature* **464**, 520–528 (2010).
20. P. Ilmonen *et al.*, *J. Evol. Biol.* **21**, 834–841 (2008).
21. C. M. Kong, X. W. Lee, X. Wang, *FEBS J.* **280**, 3180–3193 (2013).
22. T. de Lange, *Genes Dev.* **19**, 2100–2110 (2005).
23. J. H. J. Hoeijmakers, *N. Engl. J. Med.* **361**, 1475–1485 (2009).
24. E. Sahin *et al.*, *Nature* **470**, 359–365 (2011).
25. M. Asghar, S. Bensch, M. Tarka, B. Hansson, D. Hasselquist, *Proc. Biol. Sci.* **282**, 20142263 (2015).
26. F. Criscuolo *et al.*, *J. Avian Biol.* **40**, 342–347 (2009).
27. E. L. Barrett, T. A. Burke, M. Hammers, J. Komdeur, D. S. Richardson, *Mol. Ecol.* **22**, 249–259 (2013).
28. P. Zehtindjiev *et al.*, *Exp. Parasitol.* **119**, 99–110 (2008).
29. E. S. Epel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17312–17315 (2004).
30. D. Hasselquist, M. Tobler, J.-Å. Nilsson, in *Eco-Immunology*, R. M. Nelson, G. Demas, Eds. (Oxford Univ. Press, Oxford, 2012), pp. 165–224.

ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council (grants to D.H., B.H., H.W., and S.B.), the Higher Education Commission of Pakistan (to M.A.), and CAnMove (a Linnaeus research excellence environment financed by the Swedish Research Council and Lund University). The field work was conducted in collaboration with Kvismare Bird Observatory (report 174), and the experiment was conducted at Kalimok Field Station (Bulgarian Academy of Sciences, report 56). We thank C. Cornwallis, J.-Å. Nilsson, T. Alerstam, M. Strandh, and A. Marzal for manuscript comments and discussions. M.A., S.B., and D.H. designed the study. Field data and blood samples were collected by S.B., B.H., D.H., H.W., and many field assistants over 25 years. M.A. and S.B. developed the qPCR methods to quantify parasite intensity and telomere length. S.B. and P.Z. did the experimental infections. M.A. did the molecular analyses and analyzed the data together with S.B., D.H., and B.H. All authors interpreted the data. M.A., D.H., and S.B. wrote the manuscript, with extensive contributions from B.H., H.W., and P.Z. Data are available from the Dryad Digital Repository: http://doi.org/10.5061/dryad.d04h0. We declare no competing financial interests.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/436/suppl/DC1
Material and Methods
Figs. S1 to S2
Tables S1 to S7
References (*31–53*)

11 September 2014; accepted 16 December 2014
10.1126/science.1261121

PROTEASOMES

A molecular census of 26S proteasomes in intact neurons

Shoh Asano,* Yoshiyuki Fukuda,* Florian Beck, Antje Aufderheide, Friedrich Förster, Radostin Danev, Wolfgang Baumeister†

The 26S proteasome is a key player in eukaryotic protein quality control and in the regulation of numerous cellular processes. Here, we describe quantitative in situ structural studies of this highly dynamic molecular machine in intact hippocampal neurons. We used electron cryotomography with the Volta phase plate, which allowed high fidelity and nanometer precision localization of 26S proteasomes. We undertook a molecular census of single- and double-capped proteasomes and assessed the conformational states of individual complexes. Under the conditions of the experiment—that is, in the absence of proteotoxic stress—only 20% of the 26S proteasomes were engaged in substrate processing. The remainder was in the substrate-accepting ground state. These findings suggest that in the absence of stress, the capacity of the proteasome system is not fully used.

The ubiquitin proteasome system (UPS) plays a key role in many cellular processes, including cell cycle control and proteostasis (*1*). The 26*S* proteasome executes the processive degradation of proteins marked for destruction by the attachment of polyubiquitin chains (*2*, *3*). Malfunctions of the UPS result in the accumulation of damaged or misfolded proteins and are implicated in a number of diseases (*4*). The 26*S* proteasome is involved in many fundamental processes in neurons (*5*), including the regulation of synaptic strength and of the presynaptic (*6*) as well as postsynaptic proteome (*7*). However, the exact molecular mechanisms underlying these phenomena remain to be elucidated, and quantitative structural studies performed in situ should contribute to a deeper understanding.

Cellular electron cryotomography combines the power of high-resolution three-dimensional (3D) imaging with the best possible structural preservation (*8*). This is particularly important for fragile structures such as the 26*S* proteasome. Electron cryotomography performed with the recently developed "Volta" phase plate (*9*) allows the identification and localization of molecular structures in the cellular environment with high fidelity.

The 26*S* proteasomes could readily be identified by visual inspection (Fig. 1A, insets). Nevertheless, for a systematic and comprehensive statistical analysis of the 26*S* proteasomes, an automated template-based search was performed (*10*). To enable detection of both single-capped and double-capped 26*S* proteasomes, a double-capped 26*S* particle [Electron Microscopy Data Bank (EMDB), 2165] (*11*) with one regulatory particle (RP) removed computationally was low-pass filtered to a resolution of 40 Å and used as a template for an initial search of cytosolic volumes. Particles identified by cross-correlation coefficients (CCC) above a generous cut-off value (Fig. 1B) showed an isotropic orientation distribution (fig. S1) and were validated by visual inspection. Both single- and double-capped particles yielded almost identical CCCs (mean CCC ~ 0.28 for both populations) (Fig. 1B), indicating that the template-based search did not bias the statistics toward one or the other assembly state. In addition, we performed a search with a 20*S*-Cdc48 template (*12*); we were unable to distinguish a class populated by such a complex.

The CCC values had an approximately Gaussian distribution (Fig. 1B), suggesting that almost all (>3σ) 26*S* proteasome particles present in the tomograms were detected. In contrast, we could not identify 26*S* proteasome particles reliably in cryotomograms acquired without the Volta phase plate (*9*). To quantify the signal quality of particles acquired with the phase plate, we compared the single-capped 26*S* proteasomes to a high-resolution map obtained by cryoelectron microscopy (cryo-EM) single-particle analysis [Protein Data Bank (PDB), 4cr2] (*13*) by Fourier cross-resolution (FCR). The resolution of the majority of the subtomograms, each depicting an individual single-capped 26*S* proteasome particle, ranged from 35 to 50 Å (Fig. 1C).

Previous analyses of isolated proteasomes suggest that double-capped proteasomes, single-capped proteasomes, uncapped proteasomes, and different assembly intermediates all coexist in the cell (*14*, *15*). However, the relative abundance of the different assemblies is unclear because cell lysis and the subsequent purification steps may cause partial disassembly. To analyze the structural heterogeneity of our in situ particle ensemble, we subjected it to statistical analysis by subtomogram classification and averaging. First, we coherently aligned and averaged the subtomograms without using any external starting model to avoid template or reference bias (*16*) (figs. S2 and S3). The resulting subtomogram average showed a single-capped 26*S* proteasome. However, the global average, as well as the corresponding variance map, exhibited considerable heterogeneity at the uncapped end of the core particle (CP) (Fig. 2A). Both the faint density at the CP end and the variance maps suggested the coexistence of single- and double-capped 26*S* proteasomes in our data set, consistent with in vitro observations (*15*, *17*). We used a variance-based 3D focused classification (*18*) approach to divide the particle set into two structurally homogeneous classes, as well as a background bin comprising hypervariable subtomograms (fig. S4). The background bin comprised 20*S* core particles, (dis-)assembly intermediates and, possibly, a few remaining false positives. The particles assigned to the background bin were discarded after each classification step because the various molecular species contained in this bin were too few to yield meaningful averages. The main difference between the two homogeneous classes was the presence or absence of the second RP (Fig. 2B). Thus, the first class depicts an average of double-capped 26*S* proteasomes, whereas the second represents the single-capped population. Double-capped 26*S* proteasomes were far less abundant (425 particles, 27%) than the single-capped complexes (1156 particles, 73%), very similar to the ratio observed in native gel densitometry of rat cortex (*19*). On average, 23 single- or double-capped 26*S* proteasomes were found per tomogram, which translates into a mean cytosolic concentration of ~190 nM, a value close to the 26*S* proteasome subunit concentration previously reported for *Saccharomyces cerevisiae* (140 to 200 nM) (*20*).

To reveal differences in the two RPs of double-capped 26*S* proteasomes, all proteasome particles were subjected to in silico cutting between the two β rings of the 20*S* core particle (Fig. 2C). The two halves of double-capped particles were then treated as separate subtomograms. Because of the larger effective number of subtomograms, the resolution of the subtomogram average depicting a single-capped 26*S* proteasome improved to 31 Å. The corresponding variance map indicated a high variability for distinct RP subunits, such as Rpn6 and Rpn1 (Fig. 2C). Moreover, the variance was high in a region surrounded by Rpn1, Rpn2, Rpn10, and Rpn11, where there is no density in single-particle reconstructions of purified 26*S* proteasomes (*11*, *13*, *21*, *22*). We then separated the truncated particles into two major classes, class 1 (at 27 Å resolution) and class 2 (31 Å), which were found in both single- and double-capped 26*S* proteasomes (Fig. 2D). The size of the two classes was unbalanced: Class 1 accounted for 80% of all RPs (1367 particles) and class 2 for only 20% (339 particles). The difference map between the two classes (Fig. 2D) was in good agreement with the variance map, which was most intense for Rpn1, Rpn6, and the area enclosed by Rpn1, Rpn2, Rpn10, and Rpn11 (Fig. 2D).

A very large cryo-EM data set of isolated *S. cerevisiae* 26*S* proteasomes has recently revealed coexisting conformations of purified 26*S* proteasomes referred to as s1, s2, and s3 (*13*). Substrates

Department of Molecular Structural Biology, Max-Planck Institute of Biochemistry, 82152 Martinsried, Germany.
*These authors contributed equally to this work. †Corresponding author. E-mail: baumeist@biochem.mpg.de

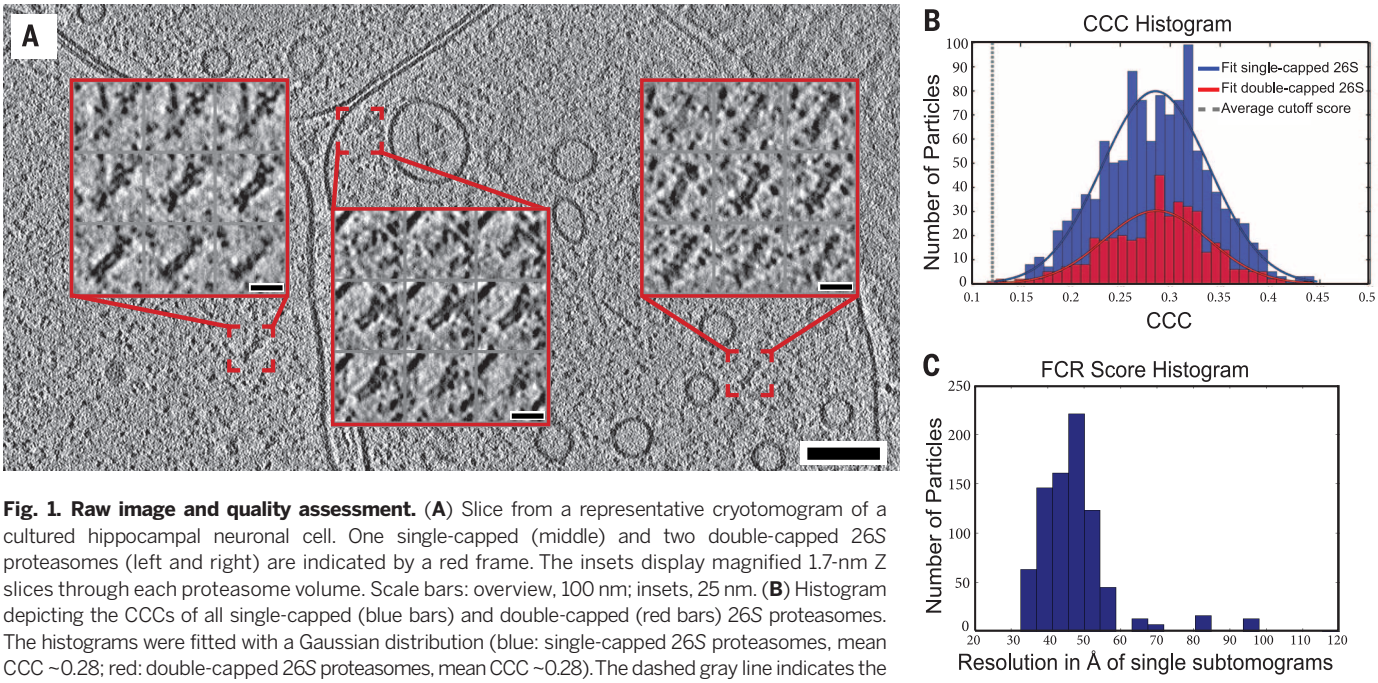
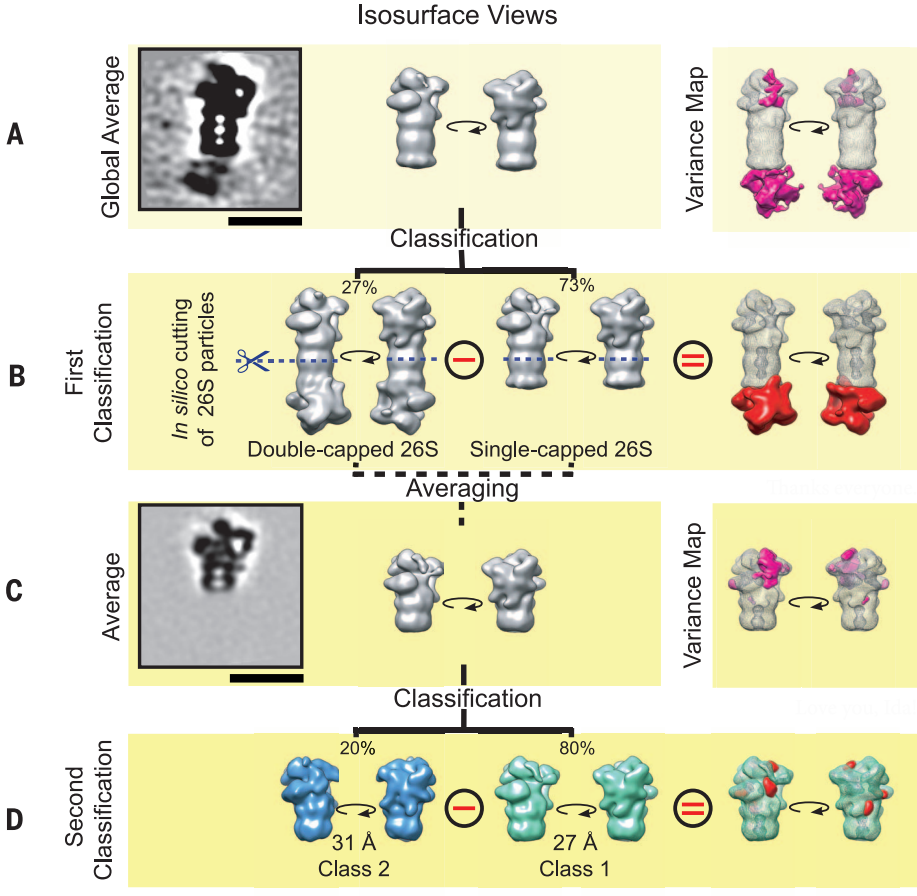


Fig. 1. Raw image and quality assessment. (**A**) Slice from a representative cryotomogram of a cultured hippocampal neuronal cell. One single-capped (middle) and two double-capped 26*S* proteasomes (left and right) are indicated by a red frame. The insets display magnified 1.7-nm Z slices through each proteasome volume. Scale bars: overview, 100 nm; insets, 25 nm. (**B**) Histogram depicting the CCCs of all single-capped (blue bars) and double-capped (red bars) 26*S* proteasomes. The histograms were fitted with a Gaussian distribution (blue: single-capped 26*S* proteasomes, mean CCC ~0.28; red: double-capped 26*S* proteasomes, mean CCC ~0.28). The dashed gray line indicates the average threshold (~0.12), below which template matching results have been discarded. (**C**) Histogram showing the FCR of each single-capped 26*S* proteasome [blue in (B)] and the atomic model of the *S. cerevisiae* single-capped 26*S* proteasome (PDB, 4cr2). The mean cross resolution is at 47 Å.

Fig. 2. Subtomogram averaging and classification workflow. (**A**) Global average from all particles displayed as center slice (left) and isosurface view (middle), as well as isosurface representation of the variance map (right; variance in violet). (**B**) First classification round separating single-capped from double-capped 26*S* proteasomes. The difference map between the two classes (right; difference shown in red) shows the additional RP. The in silico cutting was performed along the blue dashed line between the two 20*S* β rings. (**C**) Resulting average displayed as slice (left) and isosurface (middle), as well as corresponding variance map (right). (**D**) Second classification of the cut particles into two classes. Class 2 (left, light blue isosurface) and Class 1 (middle, light green isosurface) mainly differ at Rpn1, Rpn6, and the substrate entry location (right, red isosurface). Scale bars for the slices, 25 nm.



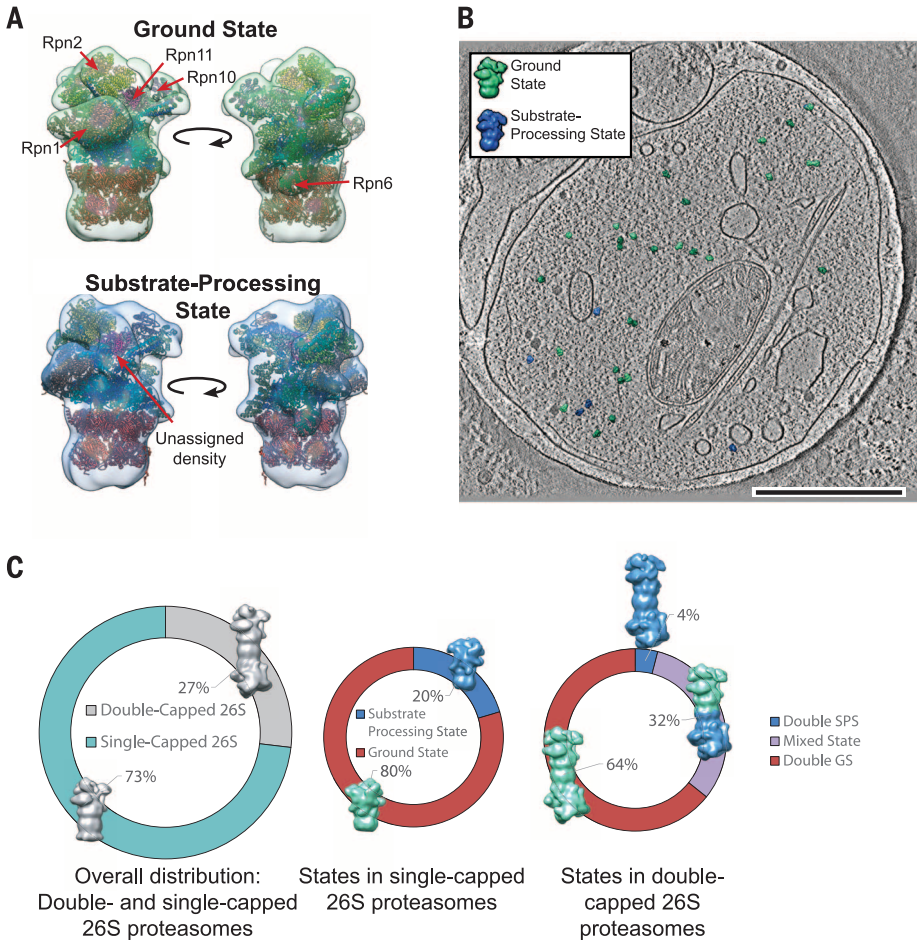


Fig. 3. Rigid body fitting of atomic subunit models into EM densities. (**A**) Fitted atomic models of the *S. cerevisiae* 26*S* proteasome subunits in the GS and SPS EM densities (Fig. 2D). The subunits Rpn9/5/6/7/3/12 are colored in different shades of green, Rpn8/Rpn11 in light/dark magenta, Rpn10 and Rpn13 in purple, Rpn1 in brown, Rpn2 in yellow, the AAA-ATPase hexamer in blue, and the CP in red (*13*). Red arrows indicate selected Rpns and the unassigned density in the SPS. (**B**) Single-capped and double-capped 26*S* proteasomes displayed as green (GS) and blue (SPS) isosurfaces overlaid on a slice of a representative tomogram. Scale bar, 500 nm. (**C**) Chart showing the overall distribution of single-capped and double-capped 26*S* proteasomes (left) and of the different states within single-capped (middle) and double-capped (right) 26*S* proteasomes.

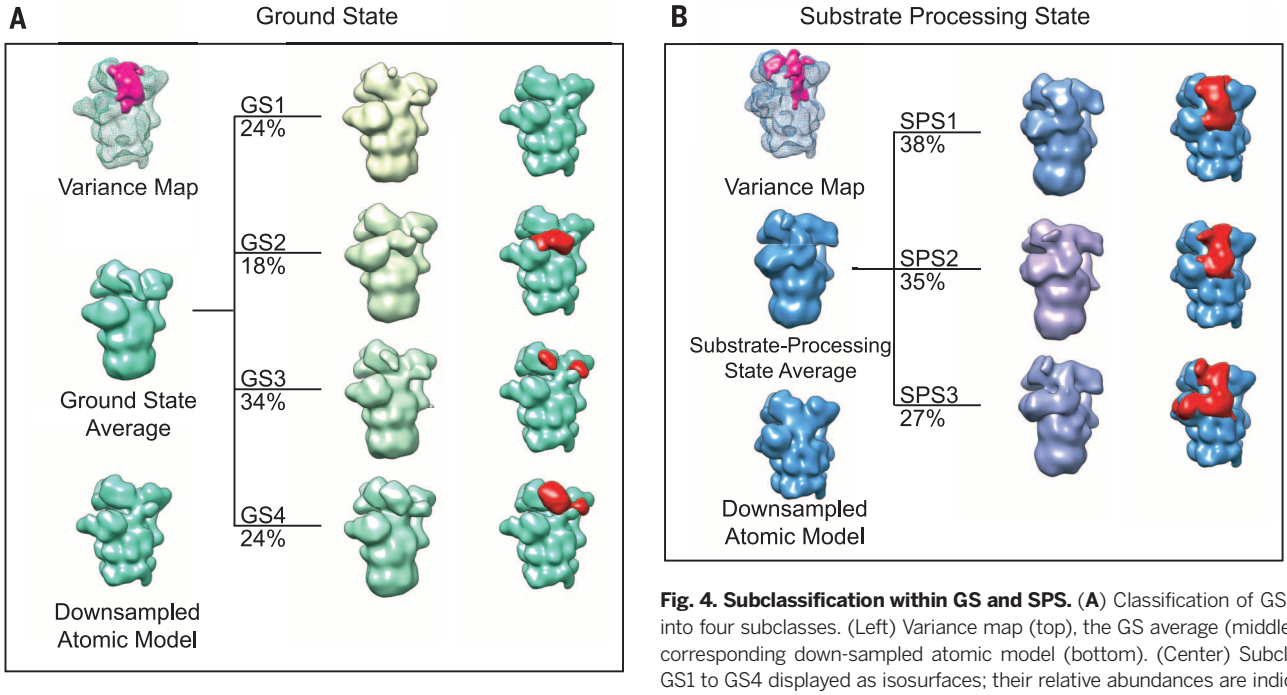


Fig. 4. Subclassification within GS and SPS. (**A**) Classification of GS class into four subclasses. (Left) Variance map (top), the GS average (middle) and corresponding down-sampled atomic model (bottom). (Center) Subclasses GS1 to GS4 displayed as isosurfaces; their relative abundances are indicated. (Right) Differences between GS1 to GS4 and the down-sampled atomic model (red). (**B**) Classification of SPS class into three subclasses, SPS1 to SPS3, displayed as in (A).

most likely primarily bind to the most abundant low-energy state s1; s2 mediates their tighter binding ("commitment") and activates the Rpn11 deubiquitylating module; and s3 finally enables translocation of substrates into the CP (*13*). To compare the mammalian in situ classes to the *S. cerevisiae* in vitro conformations, atomic models of the *S. cerevisiae* 26*S* proteasome subunits were rigidly fitted into the densities of class 1 and class 2 (Fig. 3A and fig. S5). Class 1 was most similar to s1 (substrate accepting) [root mean square deviation (RMSD) ~ 8.9 Å], whereas class 2 most closely resembled s3 (translocating) (RMSD ~ 7.4 Å) (fig. S5). In vitro, the 26*S* proteasome in complex with a polyubiquitylated substrate, GFP-I27-Ub4, adopts a conformation that is essentially identical to s3 and hence similar to class 2 (*23*). Thus, class 2 represents proteasomes that are engaged in substrate processing. Accordingly, we refer to the major conformational states represented in class 1 and class 2 as "ground state" (GS) and "substrate-processing state" (SPS), respectively. Only 20% of 26*S* proteasomes in the analyzed neurons were in a SPS. We generated a 3D atlas of the 26*S* proteasome within a cell with their individual orientations and conformation states (Fig. 3B).

Next, we analyzed whether the conformations of the two RPs in each double-capped 26*S* proteasomes correlated with one another. This analysis matched the observed distribution with the values predicted for uncorrelated states: GS-GS, 64% (measured) versus 63% (predicted); GS-SPS, 32% versus 33%; and SPS-SPS, 4% versus 4% (Fig. 3C), similar to previous single-particle cryo-EM studies (*13*, *23*), favoring the notion that both RPs can act independently.

The GS density (Fig. 4A) showed a high variance in the vicinity of the Rpn1 and Rpn13 subunits. The corresponding subtomograms were classified into four classes to deconvolute the underlying major structural differences (fig. S6). The first class (GS1, 24%) essentially did not differ from the density simulated with the fitted atomic model of a complex lacking the ubiquitin receptor Rpn13 (*24*) (Fig. 4A). Rpn13 was also absent in class GS2 (18%) and was found only in ~58% of all proteasome particles, in agreement with a dynamic binding of Rpn13 (*25*). Compared with GS1, GS2 exhibits a prominent additional density associated with Rpn1, which functions as a hub for a number of proteasome-interacting proteins (PIPs) (*2*). The remaining two ground-state classes, GS3 (34%) and GS4 (24%), showed density adjacent to Rpn2 that colocalizes with Rpn13 in *S. cerevisiae* proteasomes (*11*, *13*, *21*, *22*). This density was less prominent in GS3 but substantially larger in GS4, most likely due to the presence of PIPs.

The subtomograms contributing to the SPS were separated into three classes (Fig. 4B). The differences of all three classes, when compared to the atomic model, are located in a region above the "mouth" of the proteasomal ATPase Associated with diverse cellular Activities (AAA)–adenosine triphosphatase (AAA-ATPase). In single-particle studies, substrate was localized where we observed the additional masses in situ (*23*). However, the additional densities in each of the three in situ classes correspond to ~150 kD to 250 kD, which is substantially larger than a typical proteasomal substrate or the one used in the in vitro studies (*23*). In fact, one would anticipate a high degree of structural variability of the proteasome-associated densities if they corresponded to substrates only, given their heterogeneous nature and different stages of processing. Most likely, the proteasome-associated densities in the SPS classes depict mostly substrate-processing cofactors of the 26*S* proteasomes, such as deubiquitylating enzymes and E3 ubiquitin ligases, which are frequently found in association with 26*S* proteasomes (*2*).

Advances in technology, such as direct detectors and the contrast-enhancing Volta phase plate open up opportunities for structural studies in situ and visual proteomics (*26*). Fragile and highly dynamic macromolecular complexes can be studied in their functional and unperturbed cellular environments, providing quantitative information about their states of assembly and conformation. A challenge for the future is to correlate this information in a systematic manner with topographic information about the cellular environments they inhabit.

REFERENCES AND NOTES

1. A. Hershko, A. Ciechanover, A. Varshavsky, *Nat. Med.* **6**, 1073–1081 (2000).
2. D. Finley, *Annu. Rev. Biochem.* **78**, 477–513 (2009).
3. D. Voges, P. Zwickl, W. Baumeister, *Annu. Rev. Biochem.* **68**, 1015–1068 (1999).
4. E. Leroy *et al.*, *Nature* **395**, 451–452 (1998).
5. H. C. Tai, E. M. Schuman, *Nat. Rev. Neurosci.* **9**, 826–838 (2008).
6. K. Willeumier, S. M. Pulst, F. E. Schweizer, *J. Neurosci.* **26**, 11333–11341 (2006).
7. M. D. Ehlers, *Nat. Neurosci.* **6**, 231–242 (2003).
8. V. Lucić, F. Förster, W. Baumeister, *Annu. Rev. Biochem.* **74**, 833–865 (2005).
9. R. Danev, B. Buijsse, M. Khoshouei, J. M. Plitzko, W. Baumeister, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15635–15640 (2014).
10. J. Bohm *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14245–14250 (2000).
11. F. Beck *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14870–14875 (2012).
12. D. Barthelme, J. Z. Chen, J. Grabenstatter, T. A. Baker, R. T. Sauer, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E1687–E1694 (2014).
13. P. Unverdorben *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 5544–5549 (2014).
14. T. Yoshimura *et al.*, *J. Struct. Biol.* **111**, 200–211 (1993).
15. J. M. Peters, Z. Cejka, J. R. Harris, J. A. Kleinschmidt, W. Baumeister, *J. Mol. Biol.* **234**, 932–937 (1993).
16. Y. Chen, S. Pfeffer, T. Hrabe, J. M. Schuller, F. Förster, *J. Struct. Biol.* **182**, 235–245 (2013).
17. S. Bohn *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20992–20997 (2010).
18. Y. Chen, S. Pfeffer, J. J. Fernández, C. O. S. Sorzano, F. Förster, *Structure* **22**, 1528–1537 (2014).
19. H. C. Tai, H. Besche, A. L. Goldberg, E. M. Schuman, *Front. Mol. Neurosci.* **3**, 12 (2010).
20. C. G. Pack *et al.*, *Nat. Commun.* **5**, 3396 (2014).
21. G. C. Lander *et al.*, *Nature* **482**, 186–191 (2012).
22. K. Lasker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1380–1387 (2012).
23. M. E. Matyskiela, G. C. Lander, A. Martin, *Nat. Struct. Mol. Biol.* **20**, 781–788 (2013).
24. K. Husnjak *et al.*, *Nature* **453**, 481–488 (2008).
25. Z. Wang *et al.*, *Mol. Cell. Proteomics* **7**, 1639–1650 (2008).
26. C. V. Robinson, A. Sali, W. Baumeister, *Nature* **450**, 973–982 (2007).

ACKNOWLEDGMENTS

We thank Y. Chen for technical assistance and troubleshooting and P. Unverdorben for support in modeling and discussions. The research leading to these results has received funding from the European Commission under FP7 GA no. ERC-2012-SyG_318987-ToPAG. The work was additionally supported by the Deutsche Forschungsgemeinschaft Excellence Cluster CIPSM and SFB 1035 (both to W.B.) and FO 716/3-1 (to F.F.). Data availability: The cryo-EM maps for the GS and SPS were deposited into the EMDB with the accession codes EMD-2830 and EMD-2831, respectively. FEI Company has submitted a patent for the Volta phase plate, which is currently pending. Information on materials and methods is available on *Science* Online

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6220/439/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
References (*27*–*41*)

12 September 2014; accepted 15 December 2014
10.1126/science.1261197

Introducing

bio-techne®

HIGHEST QUALITY PRODUCT PORTFOLIO

TRUSTED GLOBAL BRANDS

R&D systems

NOVUS BIOLOGICALS™

TOCRIS

bio-techne®

LEARN MORE bio-techne.com/launch

Building Innovation Opportunities

bio-techne.com
info@bio-techne.com
techsupport@bio-techne.com

North America
TEL 800 343 7475

Europe • Middle East • Africa
TEL +44 (0)1235 529449

China
info.cn@bio-techne.com
TEL +86 (21) 52380373

Rest of World
bio-techne.com/find-us/distributors
TEL +1 612 379 2956

Bio-Techne is a trading name for R&D Systems

Lamont-Doherty Earth Observatory
COLUMBIA UNIVERSITY | EARTH INSTITUTE

www.ldeo.columbia.edu/vetlesen-prize

THE VETLESEN PRIZE 2015

Achievement in the Earth Sciences

The Lamont-Doherty Earth Observatory and The G. Unger Vetlesen Foundation
are pleased to congratulate
the 2015 Vetlesen Prize Laureate

STEPHEN SPARKS

for his ground-breaking work
in the field of volcanology.

Professor Sparks is the Chaning Wills Professor of Geology in the Department of Earth Sciences at the University of Bristol.

The Vetlesen Prize honors scientific achievement that results in a clearer understanding of the Earth, its history, or its relations to the universe.

RNA/Protein Detection Assay

The PrimeFlow RNA Assay is the first and only flow cytometry assay capable of simultaneous detection of RNA and protein within millions of cells at single-cell resolution. With the novel PrimeFlow assay, researchers can now incorporate the simultaneous analysis of RNA transcripts and proteins to elevate their understanding of single-cell dynamics. PrimeFlow RNA Assay enables high throughput detection of RNA and protein expression; therefore, it can be used to characterize, mechanistically and phenotypically, co-expression of RNAs with the functional proteins at the single-cell level. This assay, with a user-friendly protocol that has many similarities with standard antibody-staining procedures and data acquisition of flow cytometry, is an invaluable tool for immunology labs performing translational research. The major advantage of this technology is the ability to detect, with high sensitivity, mRNAs for which flow cytometry antibodies against the corresponding proteins perform poorly or are not available.

eBioscience
For info: 888-999-1371
www.ebioscience.com

Gene Engineering System

The Dharmacon Edit-R Gene Engineering System is a novel platform for research which creates permanent and heritable gene disruption (knockouts) in cells utilizing the CRISPR-Cas9 gene engineering technology. Harnessing the natural defense mechanism of *S. Pyogenes*, the Edit-R platform comprises the three components required for gene editing in mammalian cells: (1) A plasmid encoding Cas9 nuclease, (2) a chemically synthesized trans-activating RNA (tracRNA), and (3) a custom CRISPR RNA (crRNA) designed to target the site of interest. The Edit-R platform greatly simplifies and accelerates the experimental requirements and workflow compared to existing approaches.
Through easy co-transfection of synthetic tracrRNA and crRNAs using DharmaFECT Duo transfection reagent, researchers can program the Cas9 nuclease to target their genes of interest without the need to clone and sequence individual guide RNAs into expression vectors. The simplified workflow for this powerful reverse genetics method can generate functional knockouts in one to two weeks versus the current one-month timescale.

GE Life Sciences
For info: 800-235-9880
dharmacon.gelifesciences.com

Multiblock PCR Thermal Cycler

The new Mastercycler nexus X2 is ideal for researchers looking to carry out two polymerase chain reactions (PCR) simultaneously, without any compromise on the number of samples. The instrument comprises of two asymmetric blocks, consisting of 64 and 32 wells, which can be programmed and run completely independently, enabling two separate PCR protocols to be run in parallel. With reduced noise emission (<40 dB), low power consumption and a small, well-designed footprint, the Mastercycler nexus X2 is perfectly suited for use in busy academic laboratories, or any institution with multiple users. In comparison to other dual block cyclers, the Mastercycler nexus X2 provides an elegant solution for users wishing to run procedures using a large number of samples, without taking up a large amount of bench space. As the latest addition to the popular Mastercycler range, the Mastercycler nexus X2 continues Eppendorf's legacy of exceptional design and ease of use combined with efficiency and accuracy.

Eppendorf
For info: 800-645-3050
www.eppendorfna.com

Multiplex Amplicon Panels

An early access program is now available for Accel-NGS Amplicon Panels for use with Illumina Next Generation Sequencing (NGS) systems. The assays are compatible with FFPE and circulating cell free DNA from 10 ng of genomic DNA. The underlying technology enables hundreds of primer pairs to be amplified in a single tube. The panels can be used to target either contiguous coverage of a single gene, multiple loci throughout the genome, or a combination of both. The initial product is a comprehensive TP53 panel, a clinically important tumor suppressor gene. This panel covers all coding exons of TP53 in a single tube format. Another panel under development based on this technology targets over 200 clinically relevant oncology mutations. The multiplex amplicon panels achieve high-performance metrics, such as percent on-target and coverage uniformity. This new technology has a rapid workflow allowing identification of important mutations in a cost-effective manner.

Swift Biosciences
For info: 734-330-2568
www.swiftbiosci.com

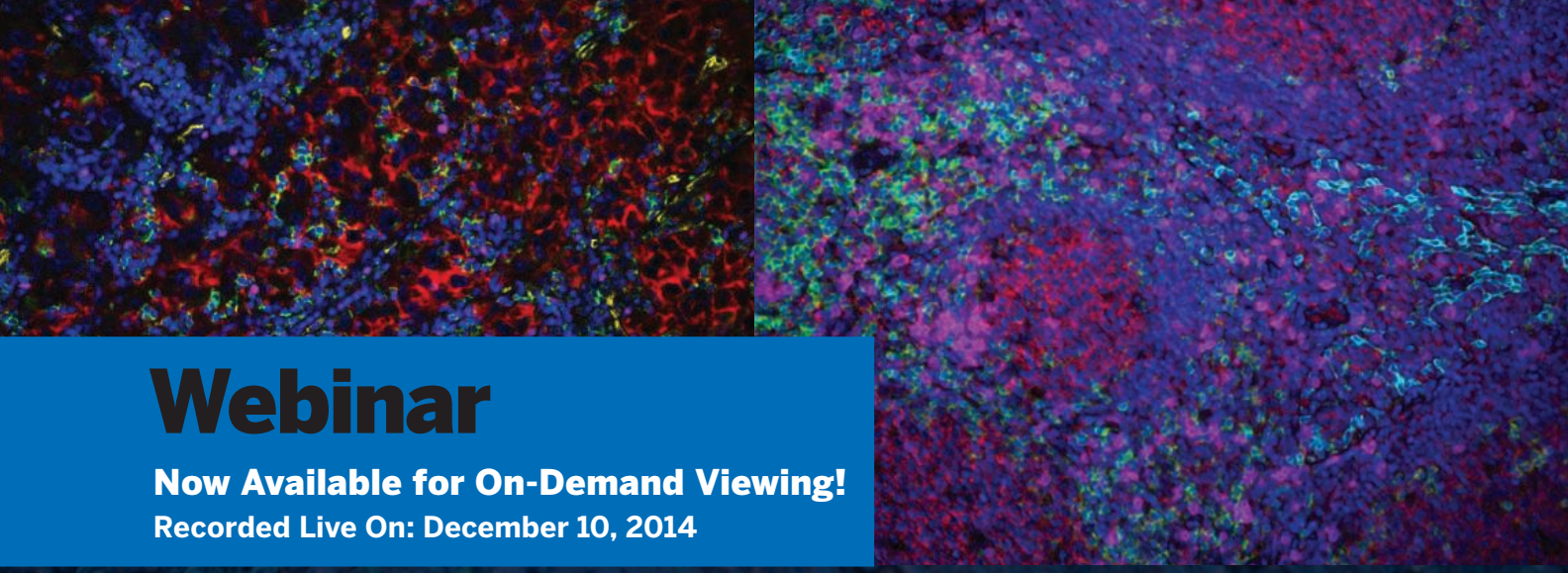
Gene Capture Probes

xGen Predesigned Gene Capture Pools and Plates are ideal for creating customized target capture panels for enrichment of 10 or more genes or for enhancing the performance of existing panels. Predesigned Gene Capture Pools and Plates are ideal for scientists who need to reduce the cost of their next generation sequencing (NGS) experiments without sacrificing quality and specificity. xGen Predesigned Gene Capture Pools are available for the coding regions of any human RefSeq gene and are delivered premixed in tubes or in individual plate wells for selective mixing. NGS target capture technologies have previously been available mostly as either fixed, off-the-shelf probe pools, or panels that do not always meet experimental needs or expensive custom panels. With xGen Predesigned Gene Capture Pools and Plates, researchers and core labs can now generate relevant, custom gene panels in-house. This customization allows smaller starting volumes per panel, resulting in less waste.

Integrated DNA Technologies
For info: 800-328-2661
www.idtdna.com

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.



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

www.cellsignal.com/science125

CST Guide: Pathways and Protocols. From the inception of the antibody as a research tool in the 1890s, to the most up-to-date research, applications, and tools, this is your complete resource for cellular research.

14PADCORPNONE0235END_00

Cell Signaling TECHNOLOGY®

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

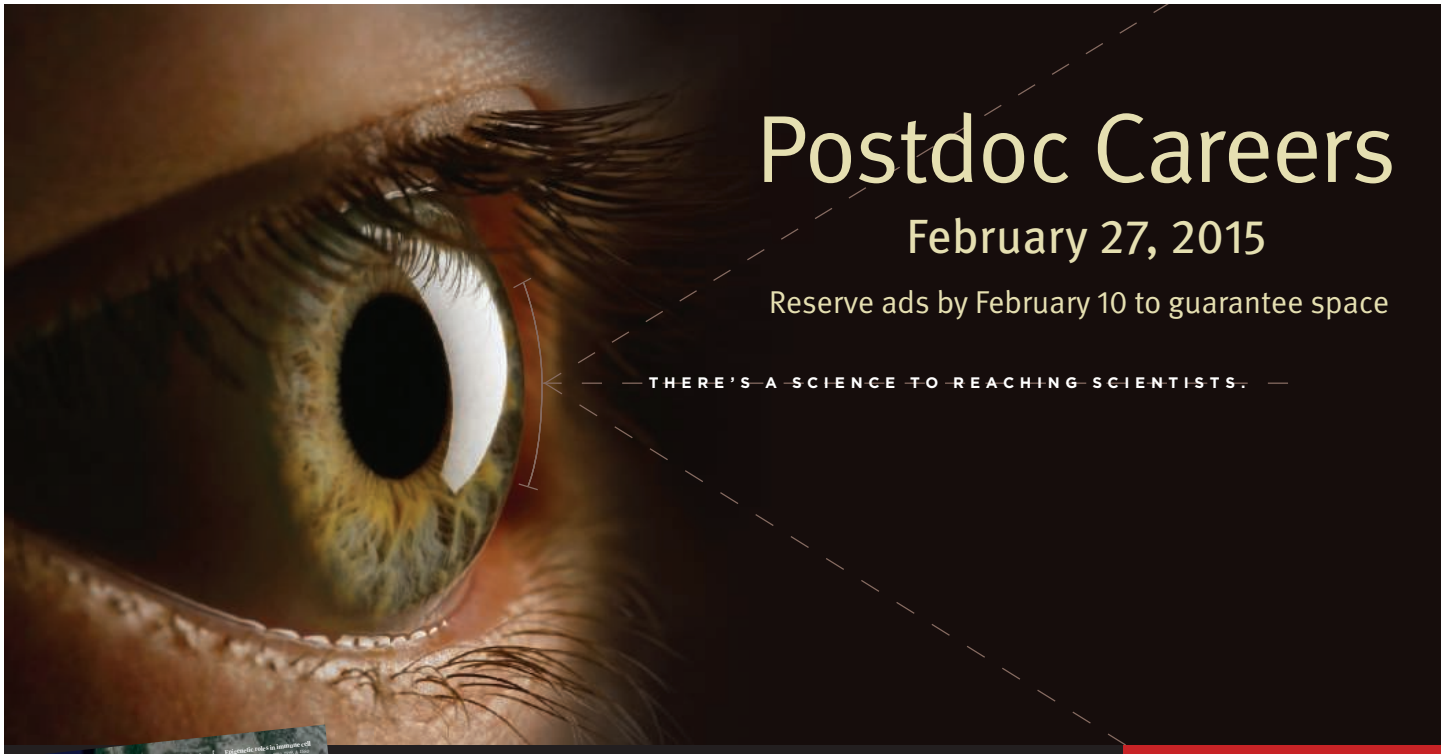
ScienceCareers

Cernet

"《科学》职业" 已经与Cernet/赛尔互联开展合作。中国大陆的高校可以直接联系Cernet/赛尔互联进行国际人才招聘。

请访问
Sciencecareers.org/CER
点得联系信息。

招募学术精英，《科学》是您的不二之选 Science



For recruitment in science, there's only one Science

Besides leading their own research team, academics are increasingly seeking multidisciplinary collaborations in which team leadership skills are important. This feature explores the various ways leadership can benefit early career scientists' careers and available training resources and programs to build such skills.

Why you should advertise in this issue of *Science*:

Reach: Your job ad is seen by 570,400 readers around the globe from varied backgrounds and it sits on special bannered pages promoting postdoc positions. 60% of our weekly readers work in academia and 67% are Ph.D.s. *Science* connects you with more scientists in academia than any other publication.

Results: If you are looking for your next postdoc, *Science* offers a simple formula: relevant content that spotlights your ad + a large qualified audience= hiring success.

Special bonus distribution to 50,000 scientists

Reserve space by February 10, 2015.

To book your ad, contact:

advertise@sciencecareers.org

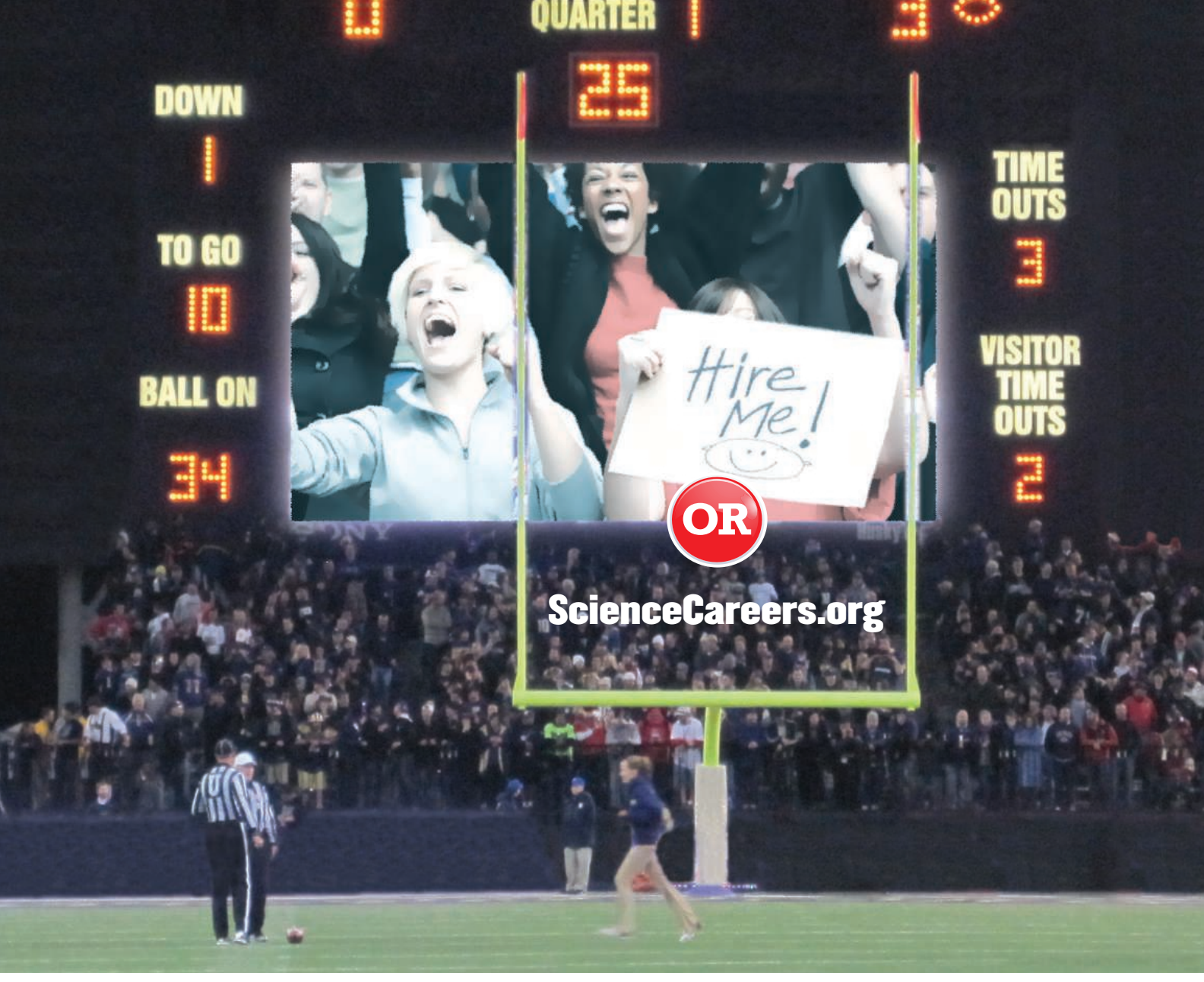
THE AMERICAS
202 326 6582

EUROPE/ROW
+44 (0) 1223 326500

JAPAN
+81 3 3219 5777

CHINA / KOREA / SINGAPORE / TAIWAN
+86 186 0082 9345

Produced by the *Science*/AAAS Custom Publishing Office.



Learn more and keep your job search out of the cheap seats.

- 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**.

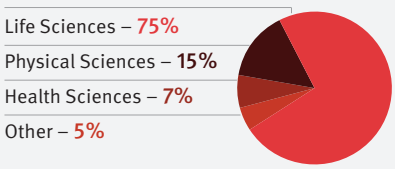
ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

ONLINE CAREER FAIR

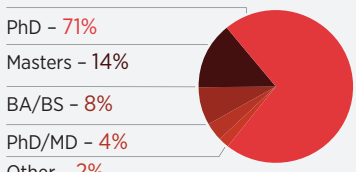
March 4, 2015 | 10:00 AM – 4 PM EST

Register now for this exciting virtual career fair and engage, screen, and recruit hundreds of targeted candidates.

ATTENDEE DEMOGRAPHICS



Field Experience



Degree Type

Over 3,000 pre-registrants; over 900 live participants

HOW THE EVENT WILL WORK

Employers receive a fully customized "booth" tailored to meet their recruiting needs. This landing page can include open positions, company information, testimonials and branding videos.

During the live event, candidates browse your booth and then choose to chat one-on-one. The conversations are timed to allow companies to meet as many candidates as possible.

BENEFITS OF EXHIBITING

- Connect with candidates from the comfort of your own desk
- Showcase your employer brand online to your target audience
- No travel or accommodation costs
- Get access to resumes to build your talent community
- Candidate information is displayed during the chat
- Powerful follow up tools to move top candidates through the hiring process
- Screen candidates with pre-qualifying questions to connect with the most relevant candidates

Book your booth today!

For more information, please visit:
ScienceCareers.org/onlinecareerfairemployers

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS