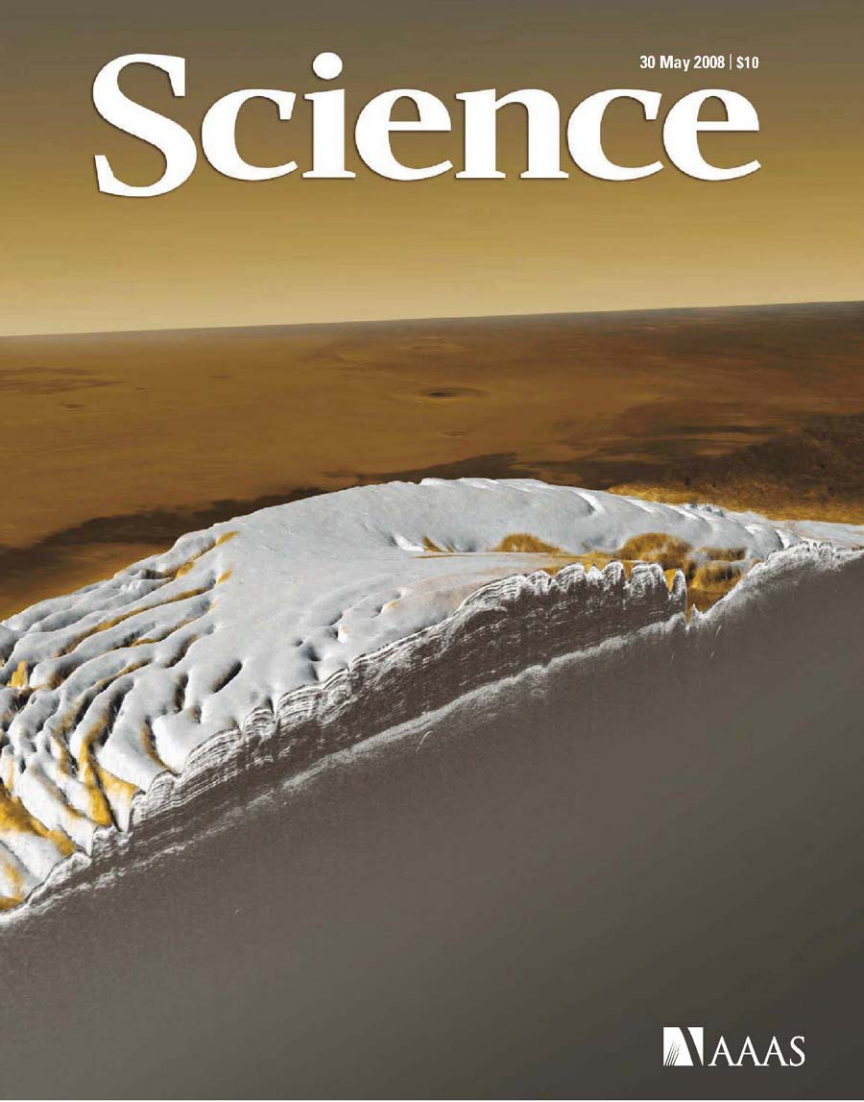
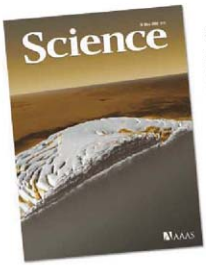
30 May 2008 | $10

Science



AAAS



COVER

Synthetic image of the north polar cap of Mars, about 1000 kilometers across, with cutaway revealing interior structure as imaged by the Shallow Radar (SHARAD) on the Mars Reconnaissance Orbiter mission. Vertical exaggeration is about 60:1. In this false-color rendition, bright lines in the cross section are reflections from internal layer boundaries within the icy cap; their depth distribution provides clues to recent climate history. See page 1182.

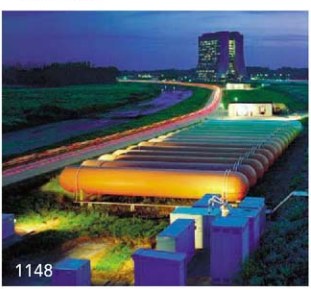
Image: NASA/JPL-Caltech/Agenzia Spaziale Italiana/ Università di Roma/Southwest Research Institute; additional data credits: Mars Orbiter Laser Altimeter Science Team, Mars Orbiter Camera (NASA/JPL/ Malin Space Science Systems); digital graphics: JPL/Eric De Jong/Mike Stetson/Jason Craig

DEPARTMENTS

EDITORIAL

NEWS OF THE WEEK

NEWS FOCUS



LETTERS

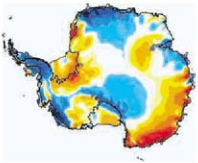
BOOKS *ET AL.*


1163

EDUCATION FORUM

PERSPECTIVES

CONTENTS continued >>



SCIENCE EXPRESS

www.sciencexpress.org

CLIMATE CHANGE

Elevation Changes in Antarctica Mainly Determined by Accumulation Variability
M. M. Helsen et al.
Satellite data since 1995, corrected for the conversion of snow to denser ice, imply that ice sheets grew slowly in the East Antarctic but decreased around the Amundson Sea.
10.1126/science.1153894

PERSPECTIVE: A Matter of Firn
K. M. Cuffey 10.1126/science.1158683

BIOCHEMISTRY

Structural Basis of Trans-Inhibition in a Molybdate/Tungstate ABC Transporter
S. Gerber, M. Comellas-Bigler, B. A. Goetz, K. P. Locher
A class of membrane transporters is subject to product inhibition: The imported substrate binds to a regulatory domain that sterically inhibits further adenosine triphosphate hydrolysis.
10.1126/science.1156213

GENETICS

Paleo-Eskimo mtDNA Genome Reveals Matrilineal Discontinuity in Greenland
M. T. P. Gilbert et al.
Ancient human DNA sequences from Greenland suggest that the earliest inhabitants of the far north were from a lineage distinct from extant Native Americans and Eskimos.
>> News story p. 1146
10.1126/science.1159750

PHYSICS

Probing Cold Dense Nuclear Matter
R. Subedi et al.
Electron-beam experiments reveal that some neutrons within ^{12}C nuclei tend to form close, dynamical pairs with protons but that pairs of the same particle type are rare.
10.1126/science.1156675

TECHNICAL COMMENT ABSTRACTS

OCEAN SCIENCE

Comment on "Mixed-Layer Deepening During Heinrich Events: A Multi-Planktonic Foraminiferal $\delta^{18}O$ Approach" 1161
C. Hillaire-Marcel and A. de Vernal
full text at www.sciencemag.org/cgi/content/full/320/5880/1161a

Response to Comment on "Mixed-Layer Deepening During Heinrich Events: A Multi-Planktonic Foraminiferal $\delta^{18}O$ Approach"
H. Rashid and E. A. Boyle
full text at www.sciencemag.org/cgi/content/full/320/5880/1161b

REVIEW

PSYCHOLOGY

How We See Ourselves and How We See Others 1177
E. Pronin

BREVIA

ECOLOGY

Ecosystem Feedbacks and Nitrogen Fixation in Boreal Forests 1181
T. H. DeLuca et al.
As burned areas in boreal forests recover, cyanobacteria in moss carpets fix nitrogen as it is lost from dead leaves and litter, ensuring a constant nitrogen supply.

RESEARCH ARTICLES

PLANETARY SCIENCE

Mars North Polar Deposits: Stratigraphy, Age, and Geodynamical Response 1182
R. J. Phillips et al.
Radar mapping shows that Mars' thick north polar ice cap contains four dust-rich layers recording variation in the planet's orbit and only slightly depresses the underlying crust. *>> Perspective p. 1171*

MOLECULAR BIOLOGY

Widespread Translational Inhibition by Plant miRNAs and siRNAs 1185
P. Brodersen et al.
Plant microRNAs and small interfering RNAs, thought to inhibit gene expression by cleavage of their RNA targets, also interfere with the translation of these RNAs into protein.

PSYCHOLOGY

Predicting Human Brain Activity Associated with the Meanings of Nouns 1191
T. M. Mitchell et al.
A model trained to associate specific nouns with resultant images of brain activity can predict what activity pattern will occur when a participant is shown a picture of a new noun.

CONTENTS continued >>

REPORTS

1168
& 1220

ADVANCING SCIENCE, SERVING SOCIETY

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. Copyright © 2008 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic institutional membership and subscription (51 issues): $944 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $770; 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 #125439322. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid (includes surface postage; bulk rates on request. **Authorization to photocopy** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $20.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for Science is 0036-8075. Science is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

CONTENTS continued >>

Best friends forever.

SCIENCE**NOW**

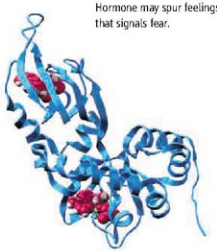
www.sciencenow.org

HIGHLIGHTS FROM OUR DAILY NEWS COVERAGE

Rover II, Going Once, Going Twice…
California company plans to auction off pet-cloning service; bids start at $100,000.

Tracing Humanity's Path
Modeling study suggests modern humans peopled the world in nine phases.

The Neurochemistry of Forgiving and Forgetting
Hormone may spur feelings of trust by suppressing brain region that signals fear.



Inhibiting PKA with Rp-cAMPS.

SCIENCE **SIGNALING**

www.sciencesignaling.org

THE SIGNAL TRANSDUCTION KNOWLEDGE ENVIRONMENT

REVIEW: Pharmacological PKA Inhibition—All May Not Be What It Seems
A. J. Murray
Widely used inhibitors of cAMP-dependent protein kinase (PKA) have nonspecific side effects.

PERSPECTIVE: Steering in Quadruplet—The Complex Signaling Pathways Directing Chemotaxis
E. C. Rericha and C. A. Parent
Multiple pathways act in concert to direct a cell's migration.

Help wanted, inquire within.

SCIENCE **CAREERS**

www.sciencecareers.org/career_development

FREE CAREER RESOURCES FOR SCIENTISTS

Educated Woman, Postdoc Edition, Chapter 16: Still Perky After All These Years
M. P. DeWhyse
As she prepares to walk away from the bench, Micella is at peace.

A Rare Opportunity to Take Flight
E. Pain
The European Space Agency has launched its first recruitment for astronauts since 1992.

Grad School Campus Visits
J. Borchardt
Campus visits can help you make a better choice about graduate schools.

From the Archives: The Perfect Postdoc
J. Austin
If you're looking for a postdoc position, first take a hard look at your prospects and plans.

SCIENCE**PODCAST**

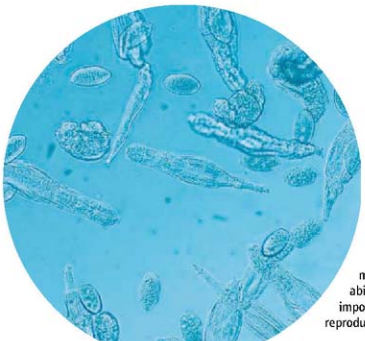
www.sciencemag.org/about/podcast.dtl

FREE WEEKLY SHOW

Download the 30 May *Science* Podcast to hear about kin selection and the evolution of eusociality, origins of the Cassiopeia A supernova remnant, Fermilab budget cuts, and more.

CREDITS: (SCIENCENOW) BIGARTS INTERNATIONAL INC.; (SCIENCE CAREERS) EUROPEAN SPACE AGENCY; (SCIENCE SIGNALING) C. BICKEL

Separate individual or institutional subscriptions to these products may be required for full-text access.



<< Rotifer Gene Scavenging

Rotifers of the Class *Bdelloidea* are small freshwater invertebrates known for their unusual ability to withstand desiccation at any life stage and for having evolved for millions of years apparently without sexual reproduction. **Gladyshev *et al.*** (p. 1210) find that *Bdelloidea* have incorporated large numbers of genes from diverse foreign sources, including bacteria, fungi, and plants into their genomes. These foreign genes have accumulated mainly in the telomeric regions at the ends of chromosomes and apparently retain their functional integrity. This ability to incorporate horizontally transferred genes may have been important in the shaping of the *Bdelloid* genome, which lacks sexual reproduction as a source of variation.

Noun Recognition Software

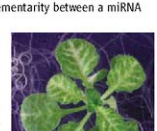
Discovering how knowledge is represented within the brain seems an impossibly complicated task, given how even simple nouns, such as celery, might be expected to activate neurons in various areas, such as those involved in eating and tasting. **Mitchell *et al.*** (p. 1191) describe a computational model that is based upon a theory that represents nouns as 25-dimensional vectors, where each dimension is a sensory-motor feature such as eating. The vectors are computed from the co-occurrences of words within a data set containing 10^{12} words. The model is then trained with neural activity data from subjects presented with 60 pictures (for instance, a piece of celery). Previously, a model trained on data from one subject was successful at identifying the noun represented by the pattern of neural activity in another subject. The current model can identify a noun that has not been included in the training set and predict the pattern of activity when a subject is shown a picture of a new noun.

Misguided Good Intentions?

Large volcanic eruptions that inject massive amounts of sulfate aerosols into the stratosphere, like that of Mount Pinatubo in 1991, cause significant climate cooling. This observation has led to the suggestion that artificial seeding of the stratosphere with sulfur could counteract the warming effects of CO_2 and other anthropogenic greenhouse gases. While such a strategy is certainly possible, what unintended consequences might arise? **Tilmes *et al.*** (p. 1201, published online 24 April; see the Perspective by **Robock**) explore one potential consequence: the destruction of stratospheric ozone at high latitudes. The injection of sulfur into the stratosphere would be likely to cause a dramatic increase in the extent of ozone loss in the Arctic, and delay the recovery of the Antarctic ozone hole by 30 to 70 years.

Micro Managing Translational Repression

MicroRNAs (miRNAs) are small noncoding RNAs that generally repress gene expression. The degree of complementarity between a miRNA and its RNA target is thought to determine the mode of repression. In plants, most characterized miRNAs are highly complementary to their targets, resulting in target cleavage (slicing), whereas those in animals have reduced complementarity and instead downregulate messenger RNA (mRNA) translation. **Brodersen *et al.*** (p. 1185, published online 15 May) conducted a screen in *Arabidopsis* for factors involved in miRNA-mediated gene regulation and found that, in addition to directing cleavage, highly complementary miRNAs also repressed translation. Furthermore, small interfering RNAs, which are precisely complementary to their mRNA targets and generally thought to repress gene expression by slicing, also have a translational repression component.



Star Light, Star Bright

Cassiopea A, which is the brightest radio emitter in the sky, is the remnant of a supernova explosion that occurred about 300 years ago. Due to the lack of contemporaneous recorded observations, little information is available about what kind of supernova it was. **Krause *et al.*** (p. 1195; see the Perspective by **Fabian**) now report their observations of an optical echo of the original explosion taken with the Spitzer Space Telescope, three centuries after the direct light from the supernova passed Earth. The spectral signature of the echo suggests that Cassiopea A resulted from the collapse of a red supergiant star.

Probing Martian Rocks

The Mars rovers and orbiters have shown that many of the oldest rocks on Mars contain a variety of sulfates and other minerals that probably formed due to the evaporation of water. **Tosca *et al.*** (p. 1204) present a thermodynamic model to calculate the properties of the desiccating water and explain the sequences of minerals observed. Water in equilibrium with the observed or inferred minerals must have been extremely saline, more so than any known terrestrial fluid, and also have had a low pH. These conditions exceed the known limits of life of Earth and, given that many waters originated in equilibrium with a basaltic crust, waters across Mars may have evolved to be inhospitable.

Cytosolic Iron Chaperone

Iron is an essential nutrient for virtually every organism because it is required for the activity of

Continued on page 1131

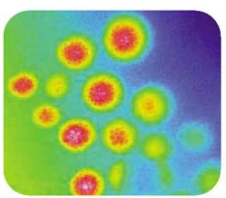
CREDITS (TOP TO BOTTOM): EUGENE GLADYSHEV; BRODERSEN ET AL.

Continued from page 1129

iron-dependent enzymes, which control a wide range of essential metabolic processes. Reduced iron is very redox active and can catalyze the formation of damaging reactive oxygen species within cells. For this reason, cells are thought to maintain "free" cytosolic iron at an exceedingly low level. Nevertheless, intracellular iron must be delivered to sites of utilization and storage, and this delivery is thought to be accomplished by iron chaperones, which would be predicted to bind iron and deliver it to other target proteins, such as enzymes or transporters. **Shi *et al.*** (p. 1207) now report the discovery of a cytosolic iron chaperone in human cells, Poly r(C)-Binding Protein 1 (PCBP1), which delivers iron to ferritin.

Counting Counts

Concepts of number and of space seem so fundamental to human experience as to be deeply embedded in cognitive structures. In general, young, preschool children intuitively map quantities onto a line where, in most cases, small numbers are placed to the left and large numbers to the right. The scale used in this mapping is logarithmic and not linear, meaning that 10 is placed in the middle of a line spanning 1 to 100. With the advent of schooling, along with exposure to cultural instantiations of number words such as rulers, the mapping shifts to a linear representation. From the latest in a series of visits to an Amazonian indigenous people, the Mundurucu, who lack formal number systems, **Dehaene *et al.*** (p. 1217) have obtained evidence that in these people the logarithmic mapping seen in youngsters persists in adults. Thus the log-to-linear shift depends upon culture—the existence of integer words separated by unit increments (for instance, twenty-three and twenty-four) and/or education in linear arithmetical operations such as addition and subtraction.



Lovely Bubbly

Trapped air bubbles influence the taste, smell, and feel of foods and cosmetics, but keeping air bubbles stable within a liquid for long periods of time can be tricky. When the bubbles get to be micrometers in size, they rapidly expand at the expense of smaller bubbles. **Dressaire *et al.*** (p. 1198) show that using a mixture of mono- and di-ester surfactants can stabilize bubbles for more than a year. The surfactants assemble on the surface of the bubbles in a hexagonal pattern, and it is this packing that leads to the long-term stability.

Right Time, Right Place

Regulatory T cells are highly immunosuppressive lymphocytes that help the body avoid autoimmunity and overzealous immune reactions. However, their existence also presents a dilemma for the immune system, because they might inadvertently shut down useful pathogen-specific immune responses. **Lund *et al.*** (p. 1220, published online 24 April; see the Perspective by **Kassiotis and O'Garra**) present evidence that suggests that regulatory T cells can in fact optimize immune responses during the early stages of infection. Using a mouse model of herpes simplex virus infection, depleting regulatory T cells delayed the arrival of the relevant immune cells at the site of infection. At the same time, inflammatory chemokines became elevated in the lymph nodes. Thus, under normal circumstances, regulatory T cells may minimize the expression of these soluble factors in the lymph nodes in order to redirect immune cells for a timely arrival at the site of infection.

Transcription Factor's Double Life Exposed

Mutations in the transcription factor MeCP2 (transcriptional repressor methyl-CpG binding protein 2) cause a broad range of neurobehavioral abnormalities, including autism, mild learning disabilities, and mental retardation. MeCP2 has been widely believed to regulate a small number of target genes in the brain by repressing their expression. By applying microarray technology to mouse models that either lack or overexpress MeCP2, **Chahrour *et al.*** (p. 1224; see the Perspective by **Cohen *et al.***) now find that this transcription factor regulates more than 2000 genes in the hypothalamus alone and that MeCP2 in fact appears to activate the expression of about 85% of these genes. The discovery that MeCP2 regulates such a large number of genes suggests that therapeutic strategies for MeCP2-related disorders should focus on restoring neuronal function rather than correcting the function of individual gene targets.

CREDIT: COURTESY OF E. DRESSAIRE, R. BEE, D. C. NEELI, A. LIPS, AND H. A. STONE

David Hamburg is the DeWitt Wallace Distinguished Scholar at Weill Cornell Medical College in New York City. He is working in the field of prevention of mass violence.

Education to Protect Humanity

AFTER MILLENNIA OF MASS EXTERMINATIONS, GENOCIDES SUCH AS THOSE IN RWANDA, Bosnia, and Darfur continue to plague the world. Given today's ready access to lethal weapons and technology-assisted incitements to hatred, the plague is poised to spread. How can we reverse this potential for malignant growth?

For decades, biological and behavioral sciences have explored factors in human conflict. But explicit focus on the prevention of mass violence, especially genocide, has lagged behind until recently. A finding of practical significance is that genocide-prone behavior can be foreseen years in advance, and this period of time could be used for prevention by applying pertinent knowledge and skills, mostly of recent origin.

Studies by international collaborators, mainly social scientists and policy analysts in universities, the United Nations, and several commissions, clarify what various organizations and institutions can do to build "pillars of prevention" that can greatly reduce the risk of genocide and other mass violence. These pillars include preventive diplomacy (such as efforts by former United Nations Secretary-General Kofi Annan to resolve conflict in Kenya) and long-term measures, especially democratic governance, equitable socioeconomic development, and constraints on weaponry.*

The pillar of education is vital in this context. Educating societies about hatred and violence is an old story in human history. But there is a better way. For example, decades of research on intergroup contact show how age-appropriate education can help people to live together harmoniously, even across previously adversarial barriers. Such intergroup contact is most effective when imbedded in shared, mutually rewarding activities that are supported by relevant authorities with a mutual aid ethic and seen as a source of benefits for all—so-called superordinate goals. Such goals may enhance intergroup relations within a school or community but also help substantially in preventing mass violence. There is no more vivid example than the recognition by the United States and the Soviet Union that avoiding nuclear war was a superordinate goal that could only be achieved by cooperation. Could global problems of food, water, health, and climate be made superordinate goals for unfriendly groups or states to tackle cooperatively?

Just as lifelong learning in mathematics, science, and technology is essential for the success of a modern economy, so too the teaching of pro-social behavior across the life-span can help to prevent immense destruction. This involves explicit information and hands-on experience with conflict resolution, violence prevention, mutual accommodation between groups, and conditions conducive to peaceful living. Evaluations of programs that emphasize pro-social behavior have identified methods that effectively teach children in these areas. Both science education and peace education require periodic updating and reinforcement on a long-term basis, and curriculum reforms cannot afford to neglect the latter as a vital component of modern education.

Schools, universities, academies, and international organizations can work together to develop and disseminate curricular material to overcome ethnocentrism, prejudice, and predisposition to violent pseudo-solutions. Good examples of such collaboration are the InterAcademy Panel and InterAcademy Council. These multinational organizations of science academies were created to mobilize the scientific community on a worldwide basis to address global needs and opportunities by advising national governments and international bodies, including the United Nations and the World Bank. It is vital to understand throughout the world the nature of human diversity and the fact of our common humanity.

These efforts are highly interdisciplinary and international, linking research with policy. They require international centers of knowledge and skill in prevention to identify warning signals and prepare appropriate responses well before a genocide begins. Such centers are now getting under way in the United Nations and the European Union, and they must be maintained.

– David Hamburg

10.1126/science.1160201

*D. A. Hamburg, *Preventing Genocide: Practical Steps Toward Early Detection and Effective Action* (Paradigm Publishers, Boulder, CO, 2008).

CHEMISTRY

Upside of a Double Negative

Chemical bonds tend to form most easily by the attack of an ion on an oppositely polarized center, or by the neutral interactions of radical centers bearing unpaired electrons. Bisai *et al.* took the somewhat counterintuitive approach of preparing two negatively charged sites and then adding an oxidant to draw them together. The deprotonation of both a carbon and a nitrogen center came at a late stage in the synthesis of the polycyclic organic natural product lyconadin A. When more traditional approaches toward linking C and N to close the ring had failed, the authors relied on the likelihood that a lithium cation would stabilize a dianion through simultaneous coordination to both centers. Treatment with iodine then induced C-N bond formation in 80% yield, affording an impressive overall product yield of 10% over the 18 steps of the synthesis. — JSY

J. Am. Chem. Soc. **130**, 10.1021/ja8028069 (2008).

NEUROSCIENCE

Out With the Old, In With the New

Might this adage, which some pundits have claimed as the basis for the vernal electoral calamities that have befallen the Labour Party in the United Kingdom, apply equally forcefully to the turnover of neurons in the brain? Adar *et al.* have performed a painstaking histological and immunofluorescence accounting of the survival likelihoods of newly born neurons in the brain of the zebra finch, a songbird that serves as an animal model for studying innate and learned influences on vocal communication. They focused on the nidopallium caudale (NC) region because it participates in auditory processing and is activated by social stimuli (other songbirds in this notably social species). By varying the complexity of the social environment, they found that the youngest cells—which had recently migrated from the site of their birth and were still becoming integrated, quite literally, as they established syn-aptic connections with existing NC neurons—were more likely to have survived if the bird had been exposed to a large group of male and female birds; conversely, in birds housed with only one other individual, the survival of older (though still relatively young) cells was enhanced. One interpretation of these data is that an increase in demand—in the form of an upturn in auditory/social inputs needing to be processed—acts as a selective pressure favoring the survival of new recruits. — GJC

J. Neurosci. **28**, 5394 (2008).

CLIMATE SCIENCE

At a Loss

Sea-ice coverage in the Arctic plummeted in the summer of 2007 to levels never before observed, surprising even experts who had witnessed the decades-long decline and predicted that the ice pack would continue to shrink at an increasingly rapid rate. Why did so much ice disappear? Zhang *et al.* conducted a retrospective modeling study of the evolution of Arctic sea-ice coverage and found that preconditioning, anomalous winds, and ice-albedo feedback were responsible for most of the retreat. Years of warming climate there preconditioned the ice for disappearance by thinning it significantly, pushing it ever closer to the point of complete melting, while stronger than normal winds pushed unusually large amounts out of the Arctic basin. The ice thinning and exposure of open water that these processes caused left the remaining ice even more susceptible than normal to loss due to heating of the upper ocean, increasing the intensity of the positive ice-albedo feedback and accelerating the rate of ice loss. Once summer had passed and temperatures had dropped low enough for ice to begin to regrow, 10% more ice than usual had vanished, 70% of it due to melting and 30% due to ice advection. The large ice loss, coupled with prevailing climate trends, suggests that Arctic sea ice has become particularly vulnerable to anomalous atmospheric forcing. — HJS

Geophys. Res. Lett. 10.1029/2008GL034005 (2008).

BIOMEDICINE

The School of Hard Knocks

The identification of a disease-causing gene mutation in humans is typically followed by a flurry of research aimed at elucidating the normal function of the gene and how disruption of that function produces the specific pathological features of the disease. These projects often rely on the phenotypic characterization of mice in which the murine ortholog of the gene has been inactivated (knock-out mice) or in which the specific disease-causing mutation has been introduced into the murine germline (knock-in mice).

Although such models are informative, a recent analysis serves as a reminder that mice are

CREDITS: (TOP TO BOTTOM) KEVIN SCHAFER/ALAMY; JEFF MORGAN POLITICS AND GOVERNMENT/ALAMY

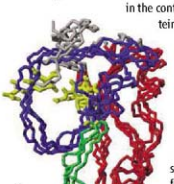
not men, especially when it comes to protein quality-control systems. By examining the Mouse Genome Informatics database and the literature, Liao and Zhang identified 120 genes known to be essential for human survival and found that 22% of these are nonessential in mice. Interestingly, nearly half of these 27 genes encode proteins localized to vacuoles, which are membrane-bound compartments that help remove cellular waste such as misfolded proteins. In independent studies, Kobuke *et al.* and Bartoli *et al.* found that a missense mutation responsible for a specific type of muscular dystrophy in humans (an R77C substitution in alpha-sarcoglycan) caused no disease phenotype when introduced into mice. In this case also, the cross-species difference was tentatively traced to the quality-control systems that recognize and process defective proteins. — PAK

Proc. Natl. Acad. Sci. U.S.A. **105**, 6987 (2008); *Hum. Mol. Genet.* **17**, 1201; 1214 (2008).

BIOCHEMISTRY

Close Enough

The classical view that protein function is associated with a well-defined three dimensional fold has been eroded, slightly, by the discoveries that intrinsically disordered proteins do exist and that disorder might play an important role in protein interactions. Catalysis would seem more structurally demanding. Although it is increasingly recognized that dynamics contributes to enzyme activity, most would have assumed that this occurs in the context of a folded protein; nevertheless, catalytic activity has been observed in an engineered enzyme with molten globule properties, where activity is apparently coupled to substrate-induced folding. Bemporad *et al.* show that the partially folded *Sulfolobus solfataricus* acylphosphatase is active and that this does not derive from a global substrate-induced folding. Molecular dynamics simulations revealed that an ensemble of partly folded molecules is characterized by substantial structural flexibility of the catalytic region; the remainder of the protein forms a scaffold that restricts the conformational space of the flexible region so that in a large fraction of the ensemble, residues important for catalysis remain close together. The authors suggest that scaffold regions in proteins might allow functional regions to mutate without compromising overall stability, thus facilitating the evolution of new activities. It remains to be seen if this is a rare case or whether catalysis in the absence of well-defined folded structure is a more common but underappreciated enzyme property. — VV



An ensemble of partly folded enzymes.

CREDIT: BEMPORAD *ET AL.*, *EMBO J.* **27**, 10.1038/EMBOJ.2008.82 (2008)

EMBO J. **27**, 10.1038/emboj.2008.82 (2008).

CHEMISTRY

Free to Switch

In many of the reported examples of switching molecular conductivity from low to high states, conformational changes have been induced with voltage pulses from a scanning tunneling microscope (STM) tip. Such changes are highly localized and involve serial processing, whereas optically induced changes could allow for faster parallel processing. However, the quenching of excited states near metal surfaces, as well as steric constraints, often limit this approach. Kumar *et al.* show that azobenzene-functionalized thiol derivatives, chemisorbed within domains of decanethiol monolayers on gold, undergo reversible trans-cis photoswitching upon irradiation with ultraviolet and visible light, respectively. The exposed azobenzene group exhibits a large change in apparent height (1.4 Å) upon switching from the much more conductive trans conformation to the cis state. — PDS

Nano Lett. **8**, 10.1021/nl080323+ (2008).

CELL BIOLOGY

I Need My Space

Protein chaperones not only promote protein folding but also help to prevent protein misfolding. The GroEL-GroES chaperonin system operates as a nanomachine in the folding of proteins in *Escherichia coli* and contains an internal compartment thought to protect substrate proteins as they fold. Tang *et al.* wanted to ascertain whether this folding compartment was important in intact cells when they folded newly synthesized proteins. To do this, they generated mutant forms of the chaperonine with a smaller folding compartment or whose compartment characteristics had been altered to be less amenable to the folding reaction. The viability of cells expressing the altered chaperonins was reduced when the size of the folding compartment was reduced or when the folding cavity was rendered less friendly. Despite this, for the protein substrate green fluorescent protein, the rate of folding actually increased with a slightly smaller folding cage, probably because of steric effects. Thus, even though mutant chaperonins are able to promote the folding of proteins that do not enter the folding chamber in vitro, the chamber itself is indeed critical for cell viability. — SMH

EMBO J. **27**, 10.1038/emboj.2008.77 (2008).

www.sciencemag.org

Science

1200 New York Avenue, NW
Washington, DC 20005
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6581, FAX 202-371-9227
Bateman House, 82-88 Hills Road
Cambridge, UK CB2 1LQ
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

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 please call 202-326-6755 for any questions or information

REPRINTS: Author Inquiries 800-635-7181
Commercial Inquiries 803-359-4578

PERMISSIONS 202-326-7074, FAX 202-682-0816

MEMBER BENEFITS AAAS/Barnes&Noble.com bookstore www.aaas.org/bn; AAAS Online Store http://www.apisource.com/aaas/ code MKB6; AAAS Travels: Betchart Expeditions 800-252-4910; Apple Store www.apple/eppstore/aaas; Bank of America MasterCard 1-800-833-6262 priority code FAA3YU; Cold Spring Harbor Laboratory Press Publications www.cshlpress.com/affiliates/aaas.htm; GEICO Auto Insurance www.geico.com/landingpage/go51.htm?logo=17624; Hertz 800-654-2200 CDP#343457; Office Depot https://bsd.officedepot.com/portalLogin.do; Seabury & Smith Life Insurance 800-424-9883; Subaru VIP Program 202-326-6417; VIP Moving Services http://www.vipmayflower.com/domestic/index.html; Other Benefits: AAAS Member Services 202-326-6417 or www.aaasmember.org.

science_editors@aaas.org	(for general editorial queries)
science_letters@aaas.org	(for queries about letters)
science_reviews@aaas.org	(for returning manuscript reviews)
science_bookrevs@aaas.org	(for book review queries)

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

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science and innovation throughout the world for the benefit of all people. The goals of the association are to: foster communication among scientists, engineers and the public; enhance international cooperation in science and its applications; promote the responsible conduct and use of science and technology; foster education in science and technology for everyone; enhance the science and technology workforce and infrastructure; increase public understanding and appreciation of science and technology; and strengthen support for the science and technology enterprise.

INFORMATION FOR AUTHORS

See pages 634 and 635 of the 1 February 2008 issue or access www.sciencemag.org/about/authors

EDITOR-IN-CHIEF **Bruce Alberts**
EXECUTIVE EDITOR **Monica M. Bradford**
DEPUTY EDITORS **R. Brooks Hanson, Barbara R. Jasny, Katrina L. Kelner**
NEWS EDITOR **Colin Norman**

EDITORIAL SUPERVISORY SENIOR EDITOR Phillip D. Szuromi; SENIOR EDITOR/PERSPECTIVES Lisa D. Chong; SENIOR EDITORS Gilbert J. Chin, Pamela J. Hines, Paula A. Kiberstis (Boston), Marc S. Lavine (Toronto), Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Valda Vinson; ASSOCIATE EDITORS Jake S. Yeston, Laura M. Zahn; ONLINE EDITOR Stewart Wills; ASSOCIATE ONLINE EDITORS Robert Frederick, Tara S. Marathe; WEB CONTENT DEVELOPER Martyn Green; BOOK REVIEW EDITOR Sherman J. Suter; ASSOCIATE LETTERS EDITOR Jennifer Sills; EDITORIAL MANAGER Cara Tate; SENIOR COPY EDITORS Jeffrey E. Cook, Cynthia Howe, Harry Jach, Barbara P. Ordway, Trista Wagoner; COPY EDITORS Chris Filiatreau, Lauren Kmec, Peter Mooreside; EDITORIAL COORDINATORS Carolyn Kyle, Beverly Shields; PUBLICATIONS ASSISTANTS Ramatoulaye Diop, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Scott Miller, Jerry Richardson, Jennifer A. Seibert, Brian White, Anita Wynn; EDITORIAL ASSISTANTS Carlos L. Durham, Emily Guise, Patricia M. Moore; EXECUTIVE ASSISTANT Sylvia S. Kihara; ADMINISTRATIVE SUPPORT Maryrose Madrid

NEWS DEPUTY NEWS EDITORS Robert Coontz, Eliot Marshall, Jeffrey Mervis, Leslie Roberts; CONTRIBUTING EDITORS Elizabeth Culotta, Polly Shulman; NEWS WRITERS Yudhijit Bhattacharjee, Adrian Cho, Jennifer Couzin, David Grimm, Constance Holden, Jocelyn Kaiser, Richard A. Kerr, Eli Kintisch, Andrew Lawler (New England), Greg Miller, Elizabeth Pennisi, Robert F. Service (Pacific NW), Erik Stokstad; INTERN Elsa Youngsteadt; CONTRIBUTING CORRESPONDENTS Jon Cohen (San Diego, CA), Daniel Ferber, Ann Gibbons, Mitch Leslie, Charles C. Mann, Virginia Morell, Evelyn Strauss, Gary Taubes; COPY EDITORS Rachel Curran, Linda B. Felaco, Melvin Gatling; ADMINISTRATIVE SUPPORT Scherraine Mack, Fannie Groom; BUREAUS New England: 207-549-7755, San Diego, CA: 760-942-3252, FAX 760-942-4979, Pacific Northwest: 503-963-1940

PRODUCTION DIRECTOR James Landry; SENIOR MANAGER Wendy K. Shank; ASSISTANT MANAGER Rebecca Doshi; SENIOR SPECIALIST Chris Redwood; SPECIALIST Steve Forrester; PREFLIGHT DIRECTOR David M. Tompkins; MANAGER Marcus Spiegler; SPECIALIST Jessie Mudjitaba

ART DIRECTOR Kelly Buckheit Krause; ASSOCIATE ART DIRECTOR Aaron Morales; ILLUSTRATORS Chris Bickel, Katharine Sutliff; SENIOR ART ASSOCIATES Holly Bishop, Laura Creveling, Preston Huey, Nayomi Kevitiyagala; ASSOCIATE Jessica Newfield; PHOTO EDITOR Leslie Blizard

SCIENCE INTERNATIONAL

EUROPE (science@science-int.co.uk) EDITORIAL: INTERNATIONAL MANAGING EDITOR Andrew M. Sugden; SENIOR EDITOR/PERSPECTIVES Julia Fahrenkamp-Uppenbrink; SENIOR EDITORS Caroline Ash, Stella M. Hurtley, Ian S. Osborne, Stephen J. Simpson, Peter Stern; EDITORIAL SUPPORT Deborah Dennison, Rachel Roberts, Alice Whaley; ADMINISTRATIVE SUPPORT John Cannell, Janet Clements, Louise Smith; NEWS: EUROPE NEWS EDITOR John Travis; DEPUTY NEWS EDITOR Daniel Clery; CONTRIBUTING CORRESPONDENTS Michael Balter (Paris), John Bohannon (Vienna), Martin Enserink (Amsterdam and Paris), Gretchen Vogel (Berlin); INTERN Lauren Cahoon

ASIA Japan Office: Asca Corporation, Eiko Ishioka, Fusako Tamura, 1-8-13, Hirano-cho, Chuo-ku, Osaka-shi, Osaka, 541-0046 Japan; +81 (0) 6 6202 6272, FAX +81 (0) 6 6202 6271; asca@os.gulf.or.jp; ASIA NEWS EDITOR Richard Stone (Beijing: rstone@aaas.org); CONTRIBUTING CORRESPONDENTS Dennis Normile (Japan: +81 (0) 3 3391 0630, FAX 81 (0) 3 5936 3531; dnormile@gol.com); Hao Xin (China: + 86 (0) 10 6307 4439 or 6307 3676, FAX +86 (0) 10 6307 4358; cindyhao@gmail.com); Pallava Bagla (South Asia: +91 (0) 11 2271 2896; pbagla@vsnl.com)

AFRICA Robert Koenig (contributing correspondent, rob.koenig@gmail.com)

EXECUTIVE PUBLISHER **Alan I. Leshner**
PUBLISHER **Beth Rosner**

FULFILLMENT SYSTEMS AND OPERATIONS (membership@aaas.org); DIRECTOR Waylon Butler; CUSTOMER SERVICE SUPERVISOR Pat Butler; SPECIALISTS Laurie Baker, Latoya Casteel, LaVonda Crawford, Vicki Linton; DATA ENTRY SUPERVISOR Cynthia Johnson; SPECIALIST Tarrika Hill

BUSINESS OPERATIONS AND ADMINISTRATION DIRECTOR Deborah Rivera-Wienhold; ASSISTANT DIRECTOR, BUSINESS OPERATIONS Randy Yi; SENIOR FINANCIAL ANALYSTS Michael LoBue, Jessica Tierney; FINANCIAL ANALYSTS Benjamin Aronin, Priti Pamnani; RIGHTS AND PERMISSIONS: ADMINISTRATOR Emilie David; ASSOCIATE Elizabeth Sandler; MARKETING DIRECTOR John Meyers; MARKETING MANAGERS Allison Pritchard, Darryl Walter; MARKETING ASSOCIATES Aimee Aponte, Alison Chandler, Mary Ellen Crowley, Marcia Leach, Julianne Wielga, Wendy Wise; INTERNATIONAL MARKETING MANAGER Wendy Sturley; MARKETING EXECUTIVE Jennifer Reeves; MARKETING/MEMBER SERVICES EXECUTIVE Linda Rusk; SITE LICENSE SALES DIRECTOR Tom Ryan; SALES MANAGER Russ Edra; SALES AND CUSTOMER SERVICE Mehan Dossani, Iquo Edim, Kiki Forsythe, Catherine Holland, Phillip Smith, Philip Tsolakidis; ELECTRONIC MEDIA: MANAGER Lizabeth Harman; PROJECT MANAGER Trista Snyder; ASSISTANT MANAGER Lisa Stanford; SENIOR PRODUCTION SPECIALISTS Christopher Coleman, Walter Jones; PRODUCTION SPECIALISTS Nichele Johnston, Kimberly Oster

ADVERTISING DIRECTOR WORLDWIDE AD SALES Bill Moran

PRODUCT (science_advertising@aaas.org); MIDWEST Rick Bongiovanni: 330-405-7080, FAX 330-405-7081; WEST COAST/W. CANADA Teola Young: 650-964-2266; EAST COAST/E. CANADA Christopher Breslin: 443-512-0330, FAX 443-512-0331; UK/EUROPE/ASIA Michelle Field: +44 (0) 1223 326524, FAX +44 (0) 1223 326532; JAPAN Mashy Yoshikawa: +81 (0) 33235 5961, FAX +81 (0) 33235 5852; SENIOR TRAFFIC ASSOCIATE Delandra Simms

COMMERCIAL EDITOR Sean Sanders: 202-326-6430

CLASSIFIED (advertise@sciencecareers.org); US: RECRUITMENT SALES MANAGER Ian King: 202-326-6528, FAX 202-289-6742; INSIDE SALES MANAGER: MIDWEST/CANADA Daryl Anderson: 202-326-6543; INSIDE SALES REPRESENTATIVE Karen Foote: 202-326-6740; KEY ACCOUNT MANAGER Joribah Able; NORTHEAST Alexis Fleming: 202-326-6578; SOUTHEAST Tina Burks: 202-326-6577; WEST Nicholas Hintibidze: 202-326-6533; SALES COORDINATORS Erika Foard, Rohan Edmonson, Shirley Young; INTERNATIONAL: SALES MANAGER Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532; SALES Marium Hudda, Alex Palmer, Alessandra Sorgente; SALES ASSISTANT Louise Moore; JAPAN Mashy Yoshikawa: +81 (0)3 3235 5961, FAX +81 (0)3 3235 5852; ADVERTISING PRODUCTION OPERATIONS MANAGER Deborah Tompkins; SENIOR PRODUCTION SPECIALISTS Robert Buck, Amy Hardcastle; SENIOR TRAFFIC ASSOCIATE Christine Hall; PUBLICATIONS ASSISTANT Mary Lagnaoui

GRIM OUTLOOK FOR CHINESE APES

Gibbons used to roam throughout southern China. Now only a handful remain in the southwest, and another species has bitten the dust. A survey of the Nangunhe Nature Reserve in China's tropical Yunnan Province has revealed no trace of the white-handed gibbon (*Hylobates lar, above*), whose piercing whoops were last heard in 1992.

A team from the Gibbon Conservation Alliance, based at the University of Zürich and the Kunming Institute of Zoology, surveyed areas where there had been reports of the species in the past 2 decades. In a report to the alliance last week, the team estimated that the last of the white gibbons disappeared a decade ago, victims of hunting and habitat destruction.

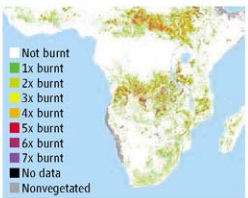
CREDITS: (TOP TO BOTTOM) COURTESY OF THOMAS GEISSMANN; COURTESY OF NANCY STEVENS, UNIVERSITY OF LEICESTER; ELIEZER STEIN-ISRAEL ANTIQUITIES AUTHORITY

Dino Dents

Very few dinosaur fossils have been found on the Arabian Peninsula. Now scientists have published the first report on dino footprints in the area—from a 150-million-year-old rock formation, formerly a coastal mud flat, in Yemen. This track of 15 56-centimeter-wide prints was made by a bipedal, plant-eating ornithopod, according to Anne Schulp of the Maastricht Museum of Natural History in the Netherlands and colleagues from the United States and Yemen. The scientists, whose paper appears in the 21 May issue of *PLoS ONE*, say the site also has footprints of a herd of 11 sauropods.

Scorched Earth

This map shows the frequency and extent of fires over 7 years in sub-Saharan Africa. Data came from Europe's SPOT satellite, which registers how burning of vegetation changes the reflectance of Earth's surface. The use of satellite data on fires "has really taken off" in global models on climate, vegetation, and atmospheric pollution, says geographer Kevin Tansey of the University of Leicester, U.K. "Ninety-nine percent of all fires on the planet are lit by humans" for clearing land, burning crop stubble, scaring up game, and introducing ash for fertilizer, Tansey says. Africa has the most burning of all, mostly in grasslands, which are often set ablaze every year. Around the world, up to 4.5 million square kilometers of vegetation—an area larger than India—burn every year, the scientists report in the current issue of *Geophysical Research Letters*.



Crusaders' Bug

Paleopathologists have nailed the cause of dysentery that devastated Crusaders invading the Holy Land in the 12th and 13th centuries.

"I had read the descriptions of dysentery in Crusader armies recorded in medieval chronicles but could not tell which organisms were responsible," says Piers Mitchell of Imperial College London. Detecting dysentery-causing parasites by microscope from archaeological samples is difficult because the cysts are tiny and degrade in soil. So the team used an assay (ELISA) that uses antibodies specific to proteins produced by the parasites.

Samples were taken from two locations in Israel: a cesspool used by the citizens of Acre and the Hospital of St. John, whose latrines were used by knights, soldiers, and pilgrims. The researchers unveiled traces of two dysentery-causing parasites, *Entamoeba histolytica* and *Giardia duodenalis*, at the St. John Hospital latrine. No parasites were found in samples from the cesspool, suggesting that locals didn't suffer from the problem, the researchers write in July's *Journal of Archaeological Science*.

"We have lots of literary data on health problems during the Crusades, but until now, diagnostic abilities were not all that accurate," says anthropologist Joe Zias of the Hebrew University of Jerusalem. "I'm thinking of using [the technique] at Qumran," location of the Dead Sea Scrolls.


Latrine at St. John Hospital.

SPRING IN HIS STEP? Biomechanist Hugh Herr was intrigued by a study showing that the prosthetic legs used by sprinter Oscar Pistorius, who last week won a fight to compete for a spot in this summer's Olympic games, used 25% less energy to run at the same speed as legged athletes. "I thought, wow, it would be amazing if this were true," says the Massachusetts Institute of Technology researcher. "It is my goal as a technologist to design prosthetics that outperform human limbs." So he agreed to help Pistorius's lawyers review the study's claims as part of their plea to an international sports body.

The study, by kinesiologist Gert-Peter Brüggemann of the German Sport University in Cologne, showed that Pistorius, a South African who races in J-shaped Cheetah Flex-Foot prostheses, used less oxygen than able-bodied athletes in a 400-meter sprint. But such a short run is mainly powered by anaerobic metabolism. Herr and others redid the energy calculations based on new measurements of the athlete's oxygen consumption during the aerobic portion of a longer run and found it to be within the normal range for able-bodied athletes. Also, although Pistorius's prostheses were more efficient than human ankles at storing and releasing energy, the researchers successfully argued that this was not sufficient evidence of an advantage.

Herr hopes to do a study of a number of amputees who use the Flex-Foot.

NONPROFIT WORLD

BIGGER CANVAS. Francis Eberle has been named the new executive director of the National Science Teachers Association (NSTA). Eberle comes to the job after 11 years as head of the Maine Mathematics and Science Alliance, where he helped the state get $25 million in new funds for math and science education in the state's schools.

In his new role, Eberle plans to focus on efforts to scale up successful science-education projects in various places across the country. "There's excellence in pockets," he says. He hopes NSTA will help the U.S. Congress see how to transfer the ideas behind these local successes to a national scale. "You have to think about the system, not specific projects," he says. But NSTA can't transform education alone: Eberle wants his organization to promote a sense of responsibility among all scientists to speak highly of science education and to help prepare science teachers for their jobs.

Eberle succeeds Gerry Wheeler, who is stepping down this summer after a 13-year stint.

THEY SAID IT

"You have essentially become a figurehead."

—Representative Henry Waxman (D–CA), chair of the House Oversight and Government Reform Committee, to U.S. Environmental Protection Agency Administrator Stephen Johnson at a 20 May hearing on input from the White House in recent EPA decisions to regulate carbon dioxide emissions and determine a health standard for smog.

Got a tip for this page? E-mail people@aaas.org

CREDITS: (TOP TO BOTTOM) ALESSANDRO DI MEO/EPA/CORBIS; (INSET) WEBB CHAPPELL/NSTA

Pioneers >>

A VIRTUAL SOCIETY. The fact that millions of people are already connected through Facebook didn't stop Ijad Madisch from dreaming up a networking site tailored to life scientists and social scientists. So last week, Researchgate.net debuted with more than 1000 users.

Madisch, 27, an M.D. and Ph.D. virologist, says he and some friends came up with the idea when he moved from his native Germany for a research traineeship at Harvard Medical School in Boston. He found that collaborating with colleagues in Germany was awkward because e-mail wasn't an efficient way to share updated protocols and drafts of papers. Meanwhile, sites such as Facebook and LinkedIn were taking off. "I thought, 'This adapted to the requirements of researchers would be of big benefit to every researcher in the world,' " says Madisch, who is now working full-time as ResearchGATE Corp.'s CEO.

Like Facebook, Researchgate.net lets users post their profiles and link up with contacts. But they can also include information such as publications and research skills. Tools for collaborating, such as virtual conferencing, are coming soon. Madisch's team admits that the big challenge will be making the site profitable—mainly with tailored ads and job postings—without driving away advertising-averse scientists.

Sichuan quake levels science city 1145

Eskimos' surprising roots 1146

PHYSICS

Energy Department Pulls Plug on Overbudget Fusion Experiment

Fusion may someday yield cheap power, but a troubled experimental reactor has proved too pricey for the U.S. Department of Energy (DOE). This week, DOE terminated the National Compact Stellarator Experiment (NCSX) at the Princeton Plasma Physics Laboratory (PPPL) in New Jersey. The not-yet-completed reactor would have been one of four large "magnetic confinement" reactors in the United States.

Donald Rej, a plasma physicist at PPPL and NCSX project manager since February, says he was disappointed to hear the news. "My colleagues have put a good fraction of their careers into this," Rej says. "It's a technological tour de force."

Like a supernova, the stellarator's budget had exploded. In 2001, it was proposed as a $58 million project to be completed in 5 years. It won approval in 2005 with a "baseline" budget of $102 million and a completion date of 2009. But in April, a DOE review showed that the cost had ballooned to $170 million and the machine could not be completed until 2013 at the earliest. The review suggested that even those estimates might not hold. "We were unable at this point to rebaseline, to formally say that we knew what it would take to finish," says Raymond Fonck, associate director for fusion energy sciences in DOE's Office of Science.

Had it been completed, the NCSX might have served as the prototype for the next great fusion experiment to come after ITER, the $12 billion machine that will be built in Cadarache, France. In a magnetic confinement reactor, scientists heat an ionized gas, or plasma, of light nuclei to 100 million degrees while trapping and squeezing it with magnetic fields. ITER and the three remaining machines in the United States are tokamaks, reactors built around a tubelike magnetic coil curled into a doughnut shape, or torus. The coil produces magnetic fields that go straight around in the horizontal torus. But to confine the plasma, the field must be modified so that it spirals around the doughnut like the stripes on a candy cane. To make that happen, the plasma must flow around the tokamak to produce an electrical current. The flow is generated by applying pulses of magnetic field or by other, more complicated means.

A stellarator, by contrast, uses bizarrely shaped coils to generate a spiraling field from scratch. Those coils are immensely complicated. NCSX researchers were able to fabricate them, but they needed more time and money to assemble the parts into a whole machine. The problems were far trickier than sticking tab A into slot B and required overcoming engineering challenges such as how to weld the massive asymmetrical pieces without deforming them, Rej says.

Surreally hard. The reactor's coils look like something Salvador Dalí might have sculpted. Joining them proved hugely difficult and expensive.

CREDIT: ELLE STARKMAN/PPPL

The cancellation of the NCSX strikes a body blow to the United States's domestic fusion program, says Stewart Prager, a plasma physicist at the University of Wisconsin, Madison, and chair of DOE's Fusion Energy Sciences Advisory Committee. "The loss of important scientific knowledge is very large," Prager says. "NCSX would have tested a fascinating physics concept and advanced understanding of a very promising fusion configuration."

The cut intensifies the uncertainty already facing plasma physicists. Over the past decade, DOE's budget for fusion research has stagnated at $300 million, and since the United States rejoined the ITER collaboration in 2003, researchers have fretted that money for smaller experiments at home might be siphoned off to pay for the nation's commitment overseas. (This year, however, the U.S. Congress zeroed out a scheduled $149 million contribution to ITER and bumped up the budget for running domestic facilities to $93.5 million, $6 million more than DOE had requested.)

Fonck says he would prefer to keep the $19.6 million requested for NCSX next year within the domestic fusion program. That money could be used simply to run the other U.S. fusion experiments longer. For example, PPPL's tokamak, the National Spherical Torus Experiment (NSTX), ran for only 13 of a possible 25 weeks in 2007 and will

Fermilab's looming black hole
1148

Clues to Antarctica's climate
1152

run for 15 weeks this year. "The plan for next year is down to 9 weeks," says A. J. Stewart Smith, dean of research at Princeton University, which runs PPPL. "This was necessary to accommodate NCSX."

It's too early to know whether the cancellation of NCSX will lead to layoffs at PPPL, Fonck says. The lab has a staff of 420 and a budget of $77 million this year.

Some physicists argue that, in spite of the cost overruns, DOE should have stuck with the NCSX. "Given the energy problem we have, it doesn't make any sense to shut down projects like this," says Miklos Porkolab of the Massachusetts Institute of Technology in Cambridge. He notes that China and South Korea, both members of ITER, completed their own large fusion experiments in 2006 and 2007, respectively. The United States hasn't completed a new machine since PPPL finished the NSTX in 1999. **–ADRIAN CHO**

PARKINSON'S DISEASE

Streamlined Clinical Trials, From a Home Computer

A Parkinson's research and treatment center and the genetic testing company 23andMe, both in California's Silicon Valley, are experimenting with an unusual new approach to clinical trials: have participants assess themselves from their home computers potentially using everything from videos of tremors to a mouse that senses motor abilities. If it works—still a big if—the strategy could greatly reduce the need for doctors' visits and make trial participation vastly cheaper and possible from anywhere in the world.

Researchers have long considered how the Web might enhance clinical trials, in particular its ability to aggregate data from the thousands of people needed for studies tracking genetic and environmental factors linked to common diseases. But the new alliance comes with many uncertainties. Will there be a bias in who completes the assessment? Can the approach capture subtle changes in disease symptoms? Will participants understand what they're signing up for, and how will the data they supply be used?

"Can you really do valid clinical research via the Web?" asks Katie Hood, chief executive officer of the Michael J. Fox Foundation for Parkinson's Research in New York City, which is funding the project with a $625,000, 2-year grant. "That's what this will show."

The Parkinson's Institute and Clinical Center in Sunnyvale, California, will recruit 150 people, half with Parkinson's disease and half without, who recently participated in a traditional study examining occupational risk factors behind Parkinson's; through that study, researchers will have amassed a wealth of data on them. 23andMe will help design or adapt computer technologies and questionnaires for the 150 participants to see whether they accurately capture their health and health history. If they do, the tools could be expanded to much larger populations and to other research groups. Volunteers will also submit saliva samples to 23andMe, and those samples, like all supplied to the company, will be sequenced for more than 580,000 single-base variations. Eventually, their genetic information will be matched with information they provide online and, say the study leaders, stored securely.

Getting creative. Could Michael J. Fox and others with Parkinson's assess symptoms themselves online?

"We want to build this engine basically to be able to power genomewide association studies," says Linda Avey, a co-founder of 23andMe. Such enormous studies compare the genomes of those with a particular disease to the genomes of those without. To accomplish this, 23andMe needs to collect large online cohorts of people with different ailments, including Parkinson's disease. Eventually, 23andMe might sell information on participants, with their consent, to pharmaceutical companies looking to recruit for clinical trials. Avey notes that the company is considering similar strategies in other diseases, although she declined to say which ones.

"We're all trying to come up with new clinical approaches that can help us get large cohorts," says Kenneth Marek, a neurologist at the Institute for Neurodegenerative Disorders in New Haven, Connecticut. Marek is experimenting with another in-home assessment for Parkinson's disease, independent of the 23andMe venture: a sniff test that includes 40 different odorants. One of the first symptoms of the disease is a diminished sense of smell. Marek and his colleagues have begun recruiting 10,000 people whose family members have Parkinson's to see whether the disease can be detected in its earliest stages.

One concern, says Marek, is that relying on volunteers to run through such tests out of sight of clinicians may lead to bias. People who struggle with the tasks could "get frustrated and may be less willing to complete this at home than they would be in a clinic," he says. "We have some sense that that might be happening," but more information is needed. A question for 23andMe is whether participants in a disease-focused study will be ready for all the other genetic information the company offers when it surveys their genomes. Will someone in a Parkinson's trial want to know whether they—and by extension, their family members—are at increased risk for heart attacks or prostate cancer? The project leader, Parkinson's Institute scientific director and CEO J. William Langston, expects that the institutional review board considering the ethics of the venture, which could start later this year, will scrutinize it closely.

As that process gets under way, Langston and 23andMe are bringing together technology companies to explore how far computers can be pushed to assess Parkinson's disease from a person's home. The Michael J. Fox Foundation, meanwhile, is gearing up to spend as much as $1 million more on other Web-based assessment tools: It put out a call for applications in March. **–JENNIFER COUZIN**

CREDIT: AP

WENCHUAN EARTHQUAKE

Damaged University Mourns Its Dead—and Plans Fast Recovery

MIANYANG, CHINA—At 2:28 p.m. on 12 May, Gao Kun was relaxing in his fifth-floor dorm room at Southwest University of Science and Technology (SWUST) when the building began to shake violently. The 23-year-old computer science major scrambled down the stairwell. Outside, people were screaming. As the ground convulsed, some students jumped from dorm windows. Gao, ducking chunks of falling masonry, helped a young woman who had jumped and broken her leg to limp away from the building. He ran back and dragged a more severely hurt jumper to safety. Minutes later, several professors arrived to aid the injured and soothe the terrified students.

Mianyang, a sprawling science city with a population of 5.2 million, was one of the worst hit by the rupture of the Longmenshan Fault, 50 kilometers to the west. As *Science* went to press, the number of local dead had climbed above 20,000, and the overall official toll for the Wenchuan earthquake stood at 67,183. In addition to SWUST, Mianyang has several institutes, including the China Air Dynamics Research and Development Center, and a high-technology R&D park anchored by the Sichuan Changhong Electronics Group. Mianyang may be best known, however, for China's main nuclear weapons design laboratory, the Chinese Academy of Engineering Physics. Although damage assessments at "China's Los Alamos" and nuclear weapons–fabrication facilities northwest of Mianyang are not publicly available, there have been no reports of radiation leaks. In the meantime, a lake formed by a landslide in the mountains west of the city is rising (*Science*, 23 May, p. 996), prompting the evacuation of more than 70,000 Mianyang residents as *Science* went to press.

North Mianyang, where SWUST is located, suffered heavier damage than other parts of the city. The campus is mourning the loss of three students—one who jumped from a dorm window, one hit by a falling brick, and one who disappeared and is presumed killed on a field trip in devastated Beichuan County. Another 99 students and staff were injured, 13 severely. "We are deeply shocked by the loss of life," says SWUST President Xiao Zhengxue, a professor of rock mechanics. The quake, he says, severely damaged scientific equipment and a quarter of the buildings; losses could tally $72 million.

A week after the earthquake, SWUST was picking up the pieces. Administrators were working out of temporary offices in a medical clinic and other lightly damaged single-story buildings. At least two-thirds of the university's 24,000 students have gone home. "Officials urged us to leave if we could," says Gao, who was planning to return to Jiangsu Province. Others pitched tents on campus.

When the earthquake struck, only a few weeks were left in the spring term at the technical university, which specializes in engineering, computer science, and agriculture and is jointly managed by the nuclear weapons lab. Officials hope that students who have returned home will complete the term's work online. "One of our strengths is distance education," says Xiao. Striving for normality, professors and students remaining on campus have been holding classes outdoors. Psychologists are counseling traumatized students.

Badly shaken. Students camp outside a damaged SWUST building.

School staff have organized students to assist relief workers aiding the thousands of injured in Mianyang and refugees who lost homes in devastated areas along the fault. Some good has come out of the tragedy, Xiao says: "Professors and students have banded together. They are closer than they ever were before." Reconstruction will begin as soon as possible, he says. "We are determined to restore the campus by September," when the next term starts. **–RICHARD STONE**

With reporting by Chen Xi.

CREDIT: R. STONE/SCIENCE

Push for Antimatter Search

A controversy over whether to put an experiment to detect antimatter aboard the space station has reached a high boil. The NASA authorization bill approved by the House Science and Technology committee's space subcommittee last week proposes $150 million for the launch of the Alpha Magnetic Spectrometer aboard the space shuttle. NASA last year knocked the experiment out of the shuttle queue, arguing that its first priority is to finish construction of the station. But several House lawmakers say that NASA's failure to carry out the international project would jeopardize the reputation of the United States as a reliable science partner. The bill is expected to be considered by the full House this summer.

–ANDREW LAWLER

Dream Teams to Tackle Cancer

The London-based charity Cancer Research UK (CRUK) has begun appointing scientist "dream teams" to collaborate with pharmaceutical companies on emerging cancer therapies. The handpicked teams will include up to five experts from different fields and will receive $1 million over 2 years, after which CRUK hopes industrial partners will take the therapies to market.

CRUK announced this week that its first team, led by Nicol Keith of the University of Glasgow, will focus on cell senescence, the mechanism by which aging cells stop dividing. Although the group includes only U.K. researchers, CRUK says future teams will enlist scientists from around the world and tackle topics such as cancer stem cells and chromatin modification. The goal, says CRUK's Simon Youlton, is to fill the gap between "what's coming out of academic research and what's being pursued by the pharmaceutical companies."

–LAUREN CAHOON

NIST Funds Bricks and Mortar

The National Institute of Standards and Technology (NIST) has announced a competition to give out $29 million for the construction of science facilities at two or three academic or nonprofit institutions. Federal agencies generally don't like giving universities grants for bricks and mortar, but Congress inserted the provision in a spending bill last year.

Such funding has ordinarily come only in the form of earmarks. But Tobin Smith of the Association of American Universities says the competition is a good idea in view of tight state budgets. If the idea catches on, he says, it could help replace some science earmarks with grant competitions. **–ELI KINTISCH**

BIOSECURITY

Fate of Plum Island Animal Lab Still Unclear

The decision over what to do about the U.S. government's antiquated Plum Island Animal Disease Center off Long Island, New York, seems little closer than it was almost 2 years ago when plans to replace the facility were first announced.

At a hearing by a House Energy and Commerce subcommittee last week, officials unveiled a Government Accountability Office (GAO) report sharply critical of possible plans to move the lab, which does research on highly infectious foot-and-mouth disease, to a mainland location. Committee chair John Dingell (D–MI) wasn't having any part of assurances that modern biosafety technology renders the need to conduct such research on an island obsolete. "I would note that history is littered with the smoking wreckage of the impregnable, the indestructible, and the unsinkable," he declared.

Plum Island, established in the 1950s for the study of diseases of livestock, is largely devoted to finding a vaccine for foot-and-mouth disease, the virus that led to the culling of more than 6 million animals in the United Kingdom in 2001. The Department of Homeland Security (DHS), which has administered the lab since 2002, wants to build a new one, the National Bio and Agro-Defense Facility (NBAF), to address emerging diseases as well as bioterror threats (*Science*, 2 September 2005, p. 1475).

Last July, DHS narrowed the list of possible locations for NBAF to five sites in the South and West in addition to Plum Island. Arguments advanced in favor of moving such a lab to the mainland range from the high cost of Long Island housing for employees to the desirability of being situated near other research facilities.

At the hearing, GAO official Nancy Kingsbury testified that DHS is in too much of a hurry to justify a mainland location and is relying on a 2002 U.S. Department of Agriculture (USDA) study that says safe research would be "technically feasible" at such a location. But, she said, DHS "has neither conducted nor commissioned any study to determine" whether work on foot-and-mouth disease can actually be done safely in light of inevitable human errors. The GAO report argues that islands offer "an extra layer of protection" in case of accidents and points out that Denmark and Germany both recently chose island locations for similar research labs.

USDA and DHS officials insisted that they're taking this decision seriously and haven't made up their minds yet where the lab should go. Another possibility is to keep foot-and-mouth research at Plum Island and build a new lab elsewhere. Plum Island operates at the next-to-highest biosafety level, or BSL-3. To take on other diseases—such as those caused by Nipah and Hendra viruses, which, unlike foot-and-mouth, also affect humans—would require BSL-4 facilities, which New Yorkers don't want in their neighborhood.

Island of safety? Experts debate whether an island location is necessary to give the lab an extra measure of biosecurity.

Jay Cohen, DHS undersecretary for science and technology, said a draft environmental impact statement on all the options will be available next month. A new facility was initially expected to cost about $445 million and open in 2011. Now cost estimates are rising steeply, with no new lab expected before 2015.

–CONSTANCE HOLDEN

GENETICS

Ancient DNA From Frozen Hair May Untangle Eskimo Roots

Humans began to brave the frozen northern reaches of Alaska, Canada, and Greenland about 4500 years ago, according to archaeological evidence. Researchers have long pondered the family history of these so-called Paleo-Eskimos, who were skillful hunters. Did they descend from the same Asian peoples who 10 millennia earlier crossed the Bering Strait and headed south, giving rise to Native Americans? And were they the ancestors of modern Eskimos? Now DNA recovered from an ancient clump of hair suggests that the answer to both questions is no.

This week, researchers from Europe and Greenland report online in *Science* (www.sciencemag.org/cgi/content/abstract/1159750) the sequencing of mitochondrial DNA from a male Paleo-Eskimo who lived in western Greenland roughly 4000 years ago—the first near-complete ancient mtDNA genome ever published. The sequence is distinct from that of both Native Americans and modern Eskimos but closely resembles that of small populations living today in the Bering Sea area, implying that the earliest Eskimos derived from an independent wave of migration from this region that left no living descendants. "The methodology appears to be excellent and their conclusions are believable," says Michael Crawford, a biological anthropologist at the University of Kansas, Lawrence.

For decades, archaeologists have attempted to trace the peopling of the far north. The earliest Paleo-Eskimos show up all across the Arctic region about 4500 years ago. But by about 1000 years ago, the Paleo-Eskimos were replaced by new migrants called the Neo-Eskimos, which researchers have concluded are the ancestors of modern Eskimo groups such as the Inuit.

Eske Willerslev, who specializes in ancient DNA at the University of Copenhagen in Denmark, learned recently that Bjarne Grønnow, an archaeologist at the National Museum of Denmark, had found human hair during his excavations 20 years ago at a Paleo-Eskimo site on the west coast

CREDIT: AP

EPIDEMIOLOGY

Children's Study Needs Pilot Testing, Panel Finds

An expert panel last week urged that several changes be made in a controversial $2.7 billion health study of 100,000 U.S. children before the first pregnant mothers are enrolled in September. The National Institute of Child Health and Human Development says it is responding to some of the 24 suggestions, but others would be too expensive.

Mandated by the U.S. Congress in 2000, the National Children's Study (NCS) will follow children from before birth to age 21 and explore how environmental factors, including pollution and television, influence disease and normal development. The study "has had a long and difficult gestation," notes the report from the National Research Council and the Institute of Medicine. Congress appropriated $69 million for this project in 2007 and $111 million this year, although the White House and National Institutes of Health Director Elias Zerhouni opposed it (*Science*, 9 February 2007, p. 751). Last year, the child health institute produced a 700-page research plan describing the study's 28 broad hypotheses. It then requested a review.

The 12-member review panel, chaired by Samuel Preston, a sociologist at the University of Pennsylvania, praises NCS's large size and its plan to visit randomly chosen households at 105 locations. The easy way would have been to find participants through health care providers, but a random sample means the results will be nationally representative and will provide "valuable" data.

But the panel also points to "important weaknesses and shortcomings." For one, it faults the absence of a built-in pilot stage for recruiting patients, collecting data, and managing the "huge databases." The panel suggests a 6- to 12-month delay between enrollment at the vanguard centers later this year and at the next wave of centers. Extra time is needed because "it is such a complicated study with so many variables," says Preston.

National Children's Study Timeline

2007	Reviews and approvals begin
2008	Seven vanguard centers begin enrollment
2010	Begin full study at other centers
2011	Initial study results
2016	Full data set for pregnancy outcomes

The reviewers also saw a problem in follow-up: After the babies turn 1, they will be evaluated in person once every few years. In the meantime, researchers will rely on telephone interviews with mothers to find out about sicknesses and diagnoses. The report urges that NCS conduct more frequent home or clinic visits and gather medical records, school records, and other documents that could help confirm illnesses and problems such as child abuse.

NCS study director Peter Scheidt says that "we agree absolutely" on the need for a pilot phase and proposed one a few years ago, but it wasn't funded. The vanguard centers will test the protocol, he says, but he admits that NCS will have less than ideal flexibility. Still, NCS is delaying full enrollment by 6 months to January 2010 to allow for revisions. Scheidt says NCS will follow several more recommendations, such as reconvening a working group on health disparities.

But adding more in-person visits or collecting paper medical records, Scheidt says, would simply be too expensive. Pediatrician Philip Landrigan of the Mount Sinai School of Medicine in New York City, an early proponent of NCS and a leader of a vanguard center, says his team would have to hire two full-time nurses to visit 100 medical practices. He thinks interviews will suffice: "If it's a significant illness, like diabetes or a surgical procedure, parents remember," he says.

Preston notes that the panel was not asked to weigh in on whether the study should go forward or is worth the cost. "That's beyond our level of expertise," he says, "but it's not an unimportant issue." The study's annual cost rises to $192 million next year.

–JOCELYN KAISER

CREDIT: A. T. P. GILBERT ET AL.; SOURCE: NCS/NICHD

Ancient heirloom. The tresses of a prehistoric Eskimo were found in the Greenland permafrost.

of Greenland. The hair was "a huge clump," says Willerslev, and its DNA was well-preserved by the permafrost.

The mtDNA genome derived from the hair bore a relatively rare genetic marker called D2a1, which is absent in modern Native Americans. To check possible links between the Paleo-Eskimo sample and Neo-Eskimos, Willerslev's team also sequenced the mtDNA genomes from 14 Greenlandic Inuits; none had the D2a1 marker. Nor does the marker show up in the handful of partial ancient DNA sequences from Neo-Eskimo skeletons that have been excavated and analyzed by others.

The genetic marker is also absent from Europeans, one reason Willerslev is confident that the mtDNA his team analyzed is not from the Danish-Inuit team that found the hair. Beth Shapiro, an expert on ancient DNA at Pennsylvania State University in State College, agrees that the mtDNA sequence is "not likely" to be the result of the contamination problems that often plague ancient DNA studies.

The D2a1 marker found in the Paleo-Eskimo mtDNA is closely related to a marker called D2a1a, which is found in present-day inhabitants of the Bering Sea area such as the Aleuts and a group of Siberian Eskimos. Scientists agree that this suggests that Paleo-Eskimos originated from this area. Still, notes Agnar Helgason, a biological anthropologist at deCODE Genetics in Reykjavik, Iceland, "this interpretation rests on a single sequence." Grønnow and his colleagues found two smaller clumps of hair during the 1980s excavations, and Willerslev's team plans to repeat their ancient DNA analysis. "I think there is a good chance of getting one more individual," Willerslev says.

–MICHAEL BALTER

Twilight? The denizens of Fermilab's iconic Wilson Hall worry for the lab's future.

But the United States's position in particle physics has been slipping, and this year the decline has snowballed into a crisis. In the past 3 months, U.S. researchers have shuttered colliders at Cornell University and the Stanford Linear Accelerator Center in Menlo Park, California. Only the 25-year-old Tevatron remains, and it will shut off in 2010. Fermilab's smaller experiments will end at about the same time. In this country, the cupboard is bare, and physicists have only unapproved plans with which to restock it.

The immediate cause of the turmoil at Fermilab is the last-minute budget Congress passed in December. It trimmed Fermilab's budget to $320 million this year from $342 million in 2007, $52 million less than requested by the U.S. Department of Energy (DOE), which owns the lab. Congress zeroed out $36 million for a proposed neutrino experiment called NOvA; slashed $14 million for work on the proposed multibillion-dollar International Linear Collider (ILC), which American physicists hope someday to build at the lab; and clipped $18 million for research on superconducting accelerator technology. "You took all the things that the lab was working toward for a future facility and you lopped them off," says Fermilab Director Piermaria Oddone. "When you do that, you're pointing the laboratory straight for the rocks."

The roots of the problem reach further back. Knowing that the LHC would eclipse the Tevatron, many U.S. physicists and some DOE officials have pushed to start building the ILC at the lab as early as 2016. In their haste, they scrapped smaller projects that otherwise might have protected Fermilab from cost cutters looking for vulnerable research and development expenditures. Or so others say. "It's pretty clear that there was a plan at DOE to clear the decks for the ILC, and a number of us saw that this was incredibly risky," says Sheldon Stone of Syracuse University in New York, who calls the current jam "predictable."

"What is matter made of? How does the universe work? We will always have these questions. The only question is, is [the U.S.] going to be involved in getting the answers."

—FLORENCIA CANELLI, FERMILAB

SOURCE: FERMILAB; CREDIT: FERMILAB

THE ONE AND ONLY

Major Experiments Built at Fermilab Since 1999			
MINOS	Studies neutrinos	$170 million	Completed in 2005
Major Experiments Canceled at Fermilab Since 1999			
Proton Driver	Various studies	~$1 billion	Canceled 2006
BTeV	To study B mesons	$200 million	Canceled 2005
CKM	To study K mesons	$100 million	Cancelled 2003
KAMI	To study K mesons	>$50 million	Cancelled 2001

Fermilab is not giving up. To secure their future through the next decade, researchers have proposed a relatively modest billion-dollar proton accelerator, dubbed Project X, to feed neutrino studies and other smaller scale experiments. The stakes are high: If all the accelerators are overseas, U.S. particle physics may simply die, researchers say. "Five years down the road, Congress may look at it and say, 'If there's nothing here, why are we funding this at all?' " says Robert Harr of Wayne State University in Detroit, Michigan. The odds may be long: DOE has completed just one major project at Fermilab in 9 years.

The ILC: A gamble that didn't pay

On 19 February 2006, William Foster penned an open letter to his colleagues. Foster had worked at Fermilab for 22 years, the last five as head of a project that, he thought, would provide the lab with a decade of research to do. Dubbed the Proton Driver, the billion-dollar linear accelerator would have pumped out protons that would crash into targets to generate neutrinos and particles called muons, K mesons, and D mesons for experiments.

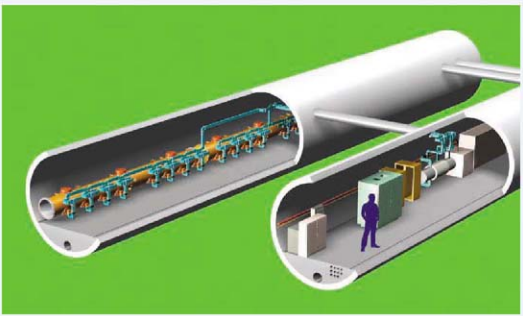
But weeks earlier, officials from DOE's Office of Science decided that they would not put the project up for the first of five "critical decision" reviews that any DOE project must pass as it wends its way from idea to facility. The reason for DOE's refusal, Foster wrote, was that the Proton Driver, the concept for which had been proposed by others as early as 1994, would interfere with efforts to get the ILC built as quickly as possible. "This position apparently applies not only to the Proton Driver, but to any intermediate-scale projects which might provide any alternate or interim future for U.S. [high-energy physics] at a cost significantly less than the [approximately] $10 billion estimated cost of the ILC," Foster wrote. "I fear that this approach is likely to end very badly. …"

Foster quit the project, the lab, and the field. "The Proton Driver represented my last best effort for Fermilab," Foster says. "And when it was clear that it wouldn't go through, I wanted to try something else where I could be more successful." That something else was politics: In March, Foster won a seat in the U.S. House of Representatives (*Science*, 14 March, p. 1470).

Numerous factors put Fermilab in its riches-to-rags predicament. For more than a decade, the DOE's high-energy physics budget failed to keep pace with inflation; this year it fell 8.5% to $688 million, from $752 million in 2007. In 2001, after a major upgrade, the Tevatron performed so poorly that the lab had to throw all it had at the problem (*Science*, 8 February 2002, p. 942). Plainly put, in the past 2 decades, particle physics has produced few of the major discoveries that thrill the public and secure generous funding.

However, many Fermilab researchers argue that the lab is in trouble because, following the lead of advisory panels stacked with ILC supporters, DOE officials have sacrificed modest experiments that the department can afford for a chance at a dream machine that may never come. "We ended up in an ILC-or-bust mode," says Stephen Holmes, Fermilab's associate director for accelerators.

In the past decade, physicists at Fermilab have proposed several hundred-million-dollar experiments that would have searched for new physics by studying the decays of familiar particles in great detail (see table, above). For example, the Tevatron feeds two particle detectors, named CDF and D0, that are searching for the Higgs boson and other new particles. The proposed BTeV detector would have studied well-known particles called B mesons, which the Tevatron produces in spades. The $200 million effort was



Only longer. An artist's conception of the International Linear Collider, which many physicists say is the future of the field.

Whither the International Linear Collider?

Efforts to develop the International Linear Collider (ILC), a 40-kilometer-long, straight-shot particle smasher, have taken some thumps in the past 16 months. But like a seasoned pugilist, the ILC has rolled with the blows, project leaders say. "We've been slowed down," says Barry Barish, a physicist at the California Institute of Technology in Pasadena, who leads the ILC Global Design Effort (GDE). Still, he says, "in terms of the threat of it being turned off, I don't think there's much chance of that."

Physicists generally agree that the ILC or something like it represents the future of particle physics. This summer, the European lab, CERN, near Geneva, Switzerland, will turn on the Large Hadron Collider (LHC), which could cough up a slew of new particles and perhaps reveal new dimensions of space. The LHC will fire protons into protons, each of which is a knot of particles called quarks and gluons, so it will produce extremely messy collisions. The ILC would collide indivisible electrons and positrons and produce cleaner collisions, which should allow researchers to study in detail the new particles glimpsed by the LHC.

The ILC's troubles began in February 2007, after the GDE released a cost estimate for the machine (*Science*, 9 February 2007, p. 746). It set the "value" of the ILC at $6.7 billion, not including contingency or inflation during planning and construction. Adding those factors meant that, if the United States hosted the ILC and paid for half of it, its share would total $7.5 billion. Two weeks later, Raymond Orbach, undersecretary for science at the U.S. Department of Energy (DOE), warned that it could take until 2025 or later to get the go-ahead for the machine (*Science*, 2 March 2007, p. 1203).

Then in December, the U.K.'s Science and Technology Facilities Council announced that Britain was pulling out of the project entirely, saying it could "not see a practicable path towards the realization of this facility" (*Science*, 21 December 2007, p. 1851). Two weeks later, a quarter of the way into the 2008 fiscal year, Congress cut funding for ILC research and development from a requested $60 million to $15 million, stopping work in the United States for the year (*Science*, 11 January, p. 142).

Still, physicists in Europe and Asia continue to soldier on and make progress, Barish says. And Spain and India have recently joined the effort. What has really suffered, Barish says, are the chances that the machine will be built in the United States, at Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois. "The most likely place that a machine like this will be built is CERN," Barish says. "It's hard to see a scenario that would bring it to Fermilab at this point."

The United States's prospects for hosting the machine suffered not so much because of the cuts to ILC development but more because of cuts to the U.S. contribution to the international fusion experiment, ITER, which will be built in Cadarache, France. Congress zeroed out the $149 million that the United States was supposed to contribute this year, leaving the six other ITER partners in the lurch. "At the moment, the U.S. is not a reliable partner for long-term projects, with the obvious consequence that few people think the U.S. is a probable candidate" for hosting the ILC, says Albrecht Wagner, head of the German Electron Synchrotron Laboratory (DESY) in Hamburg and chair of the International Committee for Future Accelerators.

The United States is not giving up on hopes for landing the ILC, however. "There was never—never—a suggestion in my comments or my actions that we were somehow moving away from the ILC," Orbach told *Science* in an interview in January. DOE has requested $35 million for ILC research and development in 2009. However, observers say Congress is likely to continue with the present budget until a new president takes office in January 2009, at which point the United States's role in the ILC will lie in the hands of the next Administration. **–A.C**

canceled in 2005, just as physicists expected the go-ahead for construction. Since 1999, only the MINOS neutrino experiment, which shoots a beam of the elusive particles to a detector in the Soudan Mine in Minnesota, has made it to completion.

Some physicists say it's unfair to blame the ILC for the demise of smaller experiments. In tight budgets, those projects simply weren't worth the costs, they say. For example, BTeV would have required running the Tevatron into the middle of the next decade at a cost of $40 million per year. "I don't think that anyone can say that BTeV was canceled because of the ILC," says Barry Barish, a physicist at the California Institute of Technology in Pasadena and leader of the ILC Global Design Effort.

But Raymond Orbach, DOE undersecretary for science, who declined to be interviewed for this article, has indicated in the past that DOE was foregoing smaller projects in favor of the ILC. "There is a fear, and the fear is well-grounded, that we may be sacrificing a lot and that [the ILC] may not come to pass," Orbach told *Science* in a June 2006 interview. "But if we don't take the risk, then we won't have the ILC on shore. … I want it here, and I want the United States to maintain its leadership in this area, and it's the only way I know how to do it."

Orbach struck a more cautious tone 8 months later. On 8 February 2007, physicists working on the ILC design released a cost estimate that indicated that if the machine were built in the United States, it would cost upward of $10 billion, of which the nation's share would be roughly $7.5 billion (*Science*, 9 February 2007, p. 746). Two weeks later, Orbach told researchers on DOE's High Energy Physics Advisory Panel that the ILC probably could not be built until the middle of the 2020s and asked for smaller projects that the United States could pursue in the meantime (*Science*, 2 March 2007, p. 1203).

CREDIT: FERMILAB

Orbach's warning suddenly presented Fermilab physicists with a gap of 15 years or more without an accelerator project. And with nothing beyond the planning stage, many physicists say, the lab became an easy target for congressional budget cutters who had to shear $22 billion from the 2008 budget to avoid a veto by President George W. Bush.

Project X: Too little, too late?

Now Fermilab researchers have come up with a plan to restore their future. To some measure, lab leaders hope to take up where they left off before this year's crisis. Congress did not cancel the NOvA neutrino experiment, notes Oddone, and physicists hope to resume work on it. Similarly, DOE has requested $35 million for ILC work in 2009. But most say that, if it is to survive, Fermilab needs a new accelerator project, and getting one may be difficult because the United States's particle physics community has painted itself into a corner, says Joel Butler, a 28-year veteran of Fermilab. "The things we can do are deemed not grandiose enough, and the things that are grandiose enough we can't afford," he says.

Fermilab hopes to solve that paradox with Project X, the conceptual son of the Proton Driver. Similar to the Proton Driver, Project X would consist of a superconducting linear accelerator measuring 700 meters long. Like Proton Driver, it would produce intense beams of protons that could be used to generate neutrinos and other familiar particles. But unlike the Proton Driver, the guts of Project X—the "cavities" through which particles surf on electromagnetic waves—would be more like those in the ILC, says Young-Kee Kim, deputy director at the lab. "The technology is aligned, so any progress we make with Project X will help us with our efforts to host the ILC," she says.

Fermilab hopes to have Project X up and running by 2016, but securing it is not a slam dunk. Many physicists question whether the menu of experiments it would support—precision studies of muons, K mesons, and neutrinos—is hearty enough to justify the expense and sustain the lab. "Thus far, I haven't seen anybody stand up and make the case that the physics that can be done with Project X is as important as building Project X" for the sake of the accelerator program, says Peter Cooper, a physicist at Fermilab. Kim says that everyone to whom she's presented the science case seems convinced.

Project X also has competition. Japanese physicists will fire up their own proton source, the Japan Proton Accelerator Research Complex (J-PARC), this year. It will pursue much of the same physics as Project X. At least in its first incarnation, J-PARC will produce a beam only half as intense as Project X's. But researchers already plan to upgrade the facility, and Japanese scientists will enjoy a head start of at least 8 years over their Fermilab rivals.

If Project X is going to help, then lab and DOE officials will have to hustle it along. The Tevatron will shut down in just over 2 years, and the accelerators that feed it and the current MINOS neutrino experiment probably won't run much longer. And if the lab goes too long without a working accelerator, it will likely lose the people it needs to build a new one, says Fermilab accelerator boss Holmes. "If you're not operating an accelerator, then you're not going to be able to design and construct a future facility," he says.

Fermilab is hoping to get DOE's preliminary okay in 2009 and start construction in 2012. If nothing is in the works by the time the Tevatron shuts down, then Fermilab will likely cut another 10% to 15% of its staff, Oddone says. Fermilab's Butler warns that, having shelved the Proton Driver once, physicists and DOE may have a tough time selling Congress on a similar project and the experiments it can do. "It's going to be difficult to say that the things that we said weren't that important are now the most important things," Butler says.

"The things we can do are deemed not grandiose enough, and the things that are grandiose enough we can't afford."

—JOEL BUTLER, FERMILAB

The no longer unthinkable

If Project X does not come to fruition, Fermilab won't vanish as soon as the Tevatron shuts down. It will still be the national headquarters for the 630 physicists from the United States who are working on CMS, one of four gargantuan particle detectors at CERN that will be fed by the LHC. The lab is also broadening its mission into astrophysics and cosmology. For example, Fermilab is one of 25 institutions in the Sloan Digital Sky Survey, which since 1998 has used a 2.5-meter telescope on Apache Point, New Mexico, to map 1/5 of the sky. Fermilab also leads the proposed Dark Energy Survey, which would use the 4-meter Blanco Telescope at Cerro Tololo in Chile to probe the bizarre dark energy that is accelerating the expansion of the universe.

But such efforts cannot sustain the lab at its present size. What's more, if Fermilab has no operating accelerator to anchor it, these other activities could be moved to other institutions, researchers worry. "The worst [possible outcome] is that we get shut down," says Fermilab physicist Stephen Pordes, "and the question is, do we get shut down quickly or slowly."

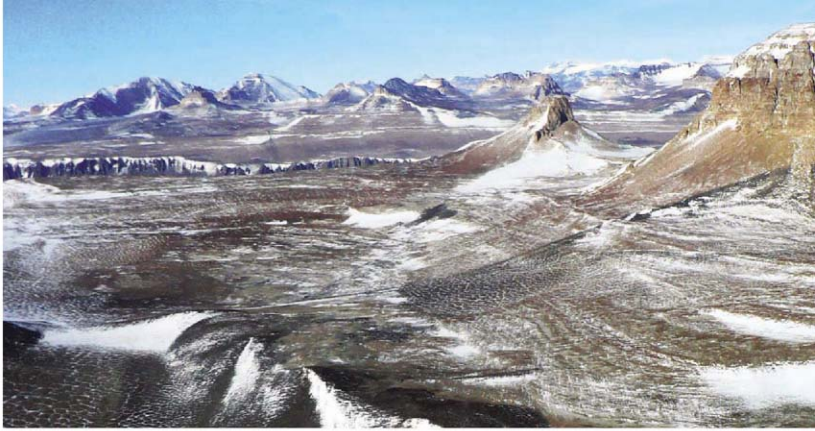
Still, there are rays of hope for the lab, and Florencia Canelli is one of them. One of Fermilab's best and brightest, the 35-year-old Argentinean holds a Wilson Fellowship—the equivalent of a tenure-track professorship at a university—and has been working at the lab since 1997, when she was a grad student at the University of Rochester, New York. She and her husband, a postdoc from Ohio State University in Columbus, both work on the CDF particle detector, which is fed by the Tevatron. They have fielded offers of dual professorships from three different universities. But Canelli has just decided to stay, taking a joint position with the lab and the University of Chicago, and her husband is taking a staff position at the lab. The two want to devote themselves full-time to exploiting the Tevatron data and gearing up for the LHC, she says.

Their decision marks a small victory not only for Fermilab but also for the U.S. program. Canelli has Italian citizenship, and her husband holds a passport from the United Kingdom; in principle, they could take off for Europe. But the United States offers opportunities that Europe does not, Canelli says, such as the chance for anyone with talent to climb to the top. "It's a good thing about the U.S.," she says. "I haven't heard of a lot of non-Italians getting a position in Italy or a lot of non-French people getting positions in France."

Canelli says she remains optimistic that the United States won't drop out of the most fundamental physics. "What is matter made of? How does the universe work?" she says. "We will always have these questions. The only question is, is this country going to be involved in getting the answers." Right now, the answer is a definite maybe.

–ADRIAN CHO

CREDIT: FERMILAB



ANTARCTICA

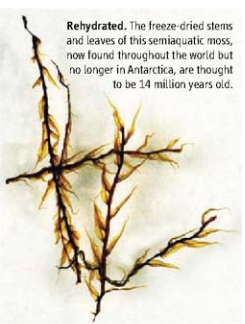
Freeze-Dried Findings Support a Tale of Two Ancient Climates

A surprising cache of ancient plant material adds evidence for divergent climate histories of the East and West Antarctic ice sheets over the past 14 million years

MCMURDO STATION, ANTARCTICA—Setting out on foot from camp on a clear, cold November day, three graduate students picked their way around the boulder-strewn flank of Mount Boreas, a naked butte in the Olympus Range of the Dry Valleys in Antarctica. The students had wandered 20 kilometers in search of glacial deposits or volcanic ash that would help them reconstruct the geologic history of this region 120 km from McMurdo Station. As they climbed over a pile of boulders, they spotted a low drift of white powder concealed in a depression.

At first, it looked like volcanic ash. But when they dropped to their knees and dug in with their fingers, they discovered papery layers of brown and white that resembled the stacked pages of a buried dictionary. Along the edges, inch-long fibers fluttered in the wind. The soft organic material was out of place in this sterile, stone-paved landscape. Once the trio had returned to camp, their microscopes revealed the brown and white layers to be desiccated slabs of an ancient lakebed that had been pried up by powerful winds. The white powder on the surface turned out to be thick deposits of diatoms. And the fibers were tiny brown stems and shriveled leaves.

The 2000 discovery, only now coming to light through a series of recent meeting presentations, represents the last gasp of a tundra ecosystem before it plunged 14 million years ago into a winter from which it has never emerged. It's one of several findings that point to an alpinelike ecosystem millions of years ago that has remained in a freeze-dried state ever since. "For 14 million years, presumably, they've been close to the surface," says paleoecologist Allan Ashworth of North Dakota State University (NDSU) in Fargo, who is coordinating analysis of the specimens. "If the climate had gotten warmer and wetter, microbes would have mined these deposits as carbon sources."



Rehydrated. The freeze-dried stems and leaves of this semiaquatic moss, now found throughout the world but no longer in Antarctica, are thought to be 14 million years old.

Juxtaposed against these findings are results from two recent cores drilled into the seabed of McMurdo Sound, 120 kilometers from the Olympus Range. The analysis of the ANDRILL (ANtarctic geological DRILLing) cores, taken in the fall of 2006 and 2007, indicate huge fluctuations in temperature over the same period in that general region.

These findings appear to be contradictory at first glance, but in fact they buttress an evolving view among scientists that the two major features of the continent, the western and eastern ice sheets, have experienced vastly different climate histories. Data from the Dry Valleys reveals an East Antarctic Ice Sheet that is high, dry, cold, and stable, at least in its central area. And the ANDRILL cores suggest a more volatile West Antarctic Ice Sheet that is subject to the changing temperatures of the sea in which it wades. "It reaffirms the fragility of the West Antarctic Ice Sheet [WAIS] and the stability of the central part of the East Antarctic Ice Sheet," says Peter Barrett, a sedimentologist at the Victoria University of Wellington (VUW) in New Zealand, who advised the ANDRILL project.

Those results also have major implications for what lies in store for the two ice

CREDITS: (TOP TO BOTTOM) ADAM LEWIS; ALLAN ASHWORTH

A hot ice story. Paleoclimate records from the ANDRILL cores, the Olympus Range, and the West Antarctic Ice Sheet reveal divergent climate histories for East and West Antarctica. The picture shows Mount Boreas in the Olympus Range of Dry Valleys, near one of the sites of the freeze-dried lake deposits.

sheets, which together cover more than 95% of the frozen land mass. Their fate represents a key element in any scenario about the long-term effects of global warming.

In fact, there's so much at stake, and the fossil findings were so unexpected, that David Marchant, a geomorphologist at Boston University (BU), initially suspected a practical joke when his students asked him to look through the microscope. "I thought someone had snuck in grass from New Zealand," he says. "It looked so fresh. As soon as I knew it was a real find, I thought this was among the biggest discoveries in the Dry Valleys, and maybe all of continental Antarctica, in the last decade."

Freeze-dried samples

Adam Lewis, then pursuing his Ph.D. at BU under Marchant, had already spent several years studying glacial deposits in the Olympus Range. That work had given him a nodding acquaintance with dating them and allowed him to draw a general picture of past climates in the region. So although the withered roughage spotted by him and fellow grad students Jane Willenbring and Brett VandenHeuvel was novel, he immediately sensed its significance. "We knew it could be used to reconstruct temperature," says Lewis. "We knew that we would be able to say, 15 million years ago, how warm it was in summer."

It was some time before the graduate students could indulge their curiosity, however.

"We bagged it up, wrote down our notes, and made a little rock pile," says Lewis, now on the faculty at NDSU. "Then we had to get back to work on glacial geology." The group—Willenbring is now a postdoc at the University of Hanover, Germany, and VandenHeuvel is an environmental lawyer in Oregon—kept the find secret to prevent the site from being excavated before its stratigraphy was properly dated.

But Lewis returned 2 years later, and the story is now appearing in a series of recent presentations and upcoming publications. Lewis and Ashworth also star in a new documentary, *Ice People*, that describes their work in the Dry Valleys.

The shriveled fibers Lewis and his colleagues stumbled across were 14-million-year-old aquatic mosses whose leafy stems once undulated in shallow, lazy currents. When dunked in water, which Ashworth immediately tried with some of them, they reinflated and unfurled. "These things look like freeze-dried museum samples," says Ashworth. The bodies of pea-sized freshwater seed shrimp, or ostracods, have also turned up with their soft mouth parts intact. A second lakebed, also in the Dry Valleys, offered up pollen, leaves, and twigs of southern beech—the twigs still composed of pliable, burnable wood. And lake sediments have coughed up megaspores of aquatic quillwort plants and fragments of insects, including a species of weevil.

These fossils paint a picture of an alpine lake dammed behind glacial moraines, surrounded by tundra and weather-beaten southern beech shrubs. The setting is similar to modern, above-tree-line communities in New Zealand or Patagonia. The wide range of diatom species show that the lake didn't dry up seasonally, says Alex Wolfe of the University of Alberta in Edmonton, Canada, who examined the diatoms: "It was a proper lake ecosystem."

This spongy lawn of alpine tundra suggests an average summer temperature well above freezing, according to early results presented last summer by Marchant, Lewis, and Ashworth at the 10th International Symposium on Antarctic Earth Sciences in Santa Barbara, California. (More complete climate reconstructions are in press.) That's at least 20°C warmer than current conditions, in which only a few minuscule nematodes, springtails, and mites eke out an existence below the surface.

> *"So as well as seeing variability, we're also getting this sense of extended periods when the West Antarctic Ice Sheet was very small, if not gone altogether."*
>
> —TIMOTHY NAISH, VICTORIA UNIVERSITY OF WELLINGTON

Isotope analysis of volcanic ash sprinkled over one lakebed dates the tundra at 14.1 million years. Last November, Marchant and Lewis published geomorphic surveys suggesting that the Olympus Range shifted within a few hundred thousand years from warm, seasonally melting glaciers that fed lakes to dry-based glaciers that sublimated rather than melted. The dry glaciers have advanced and evaporated since then, but cosmogenic nuclide dating, which estimates the times that rocks on the surface have been exposed to cosmic rays, suggests that this landscape—perhaps the oldest on Earth—hasn't tasted liquid water in the last 14 million years. "We'll move a boulder, and we'll think to ourselves, 'Our ancestors were swinging in the trees, and that boulder was already there,'" says Lewis.

Cork in the bottle

While the portion of Antarctica east of the Transantarctic Mountains stayed locked in the freezer for the last 14 million years, West Antarctica, which constitutes about a fifth of the continent, was having a much more on-again, off-again relationship with its ice. The two 1200-meter cores from the ANDRILL project provide a 20-million-year climate record. The 2006 core was of particular interest because it came from seabed beneath the northwest corner of the Ross Ice Shelf, a Spain-size slab of floating ice that feeds out of the West Antarctic Ice Sheet.

Glaciologists regard the Ross Ice Shelf as a barometer of the ice sheet's health. It buttresses the flow of major glaciers flowing out of WAIS, providing what some call "the cork in the bottle." Sea levels could rise by 5 meters were WAIS to melt. Satellite studies indicate that it is currently losing as much as 150 cubic kilometers of ice per year, and many observers consider it at risk for collapse because it slides on a bed that sits below sea level, where rising and warming seas could destabilize it. (All other ice sheets that survived the last deglaciation sit atop land.) The disappearance of the Ross Ice Shelf would lead to an acceleration of WAIS's glaciers and maybe major ice loss.

Alternating layers of diatoms and glacial debris from the ANDRILL cores revealed 60 cycles of advance and retreat of the Ross Ice Shelf over 14 million years, according to results reported last December. "The variation surprised us," says Timothy Naish, a paleoclimatologist at VUW, who co-led the 2006 ANDRILL effort and presented in April at the assembly of the European Geosciences Union in Vienna, Austria. Most striking of all was a 90-meter layer of green, diatom-rich sediment that revealed 200,000 years of biologically productive, ice-free sea starting 4.2 million years ago. "So as well as seeing variability," says Naish, "we're also getting this sense of extended periods when the West Antarctic Ice Sheet was very small, if not gone altogether."

For Lewis, an ice-free McMurdo Sound teeming with life at various points in the last 14 million years is consistent with having the Olympus Range maintained in a freeze-dried limbo if one factors in the powerful winds screaming off the polar plateau. "There has to be a fairly big ice sheet perched right behind those mountains [in East Antarctica]," he says, "to blow all that cold air down the slope."

"I've got this gut feeling that [more of] these deposits are going to be turning up."

—ALLAN ASHWORTH, NORTH DAKOTA STATE UNIVERSITY

That the Ross Ice Shelf underwent major collapses rather than minor fluctuations is supported by glacial drop stones found in the 2006 ANDRILL core. The stones come from 300 kilometers south, where Byrd Glacier pours through the Transantarctic Mountains. Rocks scooped up by Byrd were delivered to McMurdo Sound because the Ross Ice Shelf bent the glacier's flow. But in sections of core showing open sea, rocks come from local sources. To Naish, it means that the Ross Ice Shelf was absent, or least too small to bend Byrd Glacier toward McMurdo.

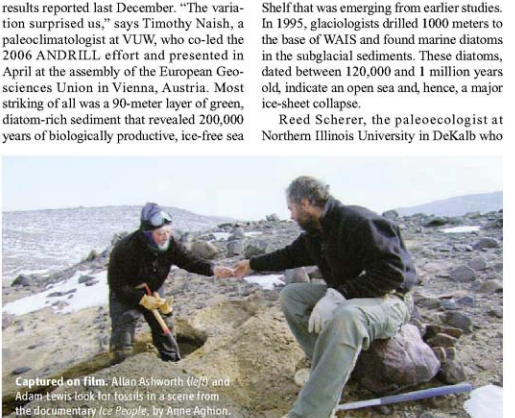
The ANDRILL cores support a volatile view of West Antarctica and the Ross Ice Shelf that was emerging from earlier studies. In 1995, glaciologists drilled 1000 meters to the base of WAIS and found marine diatoms in the subglacial sediments. These diatoms, dated between 120,000 and 1 million years old, indicate an open sea and, hence, a major ice-sheet collapse.

Reed Scherer, the paleoecologist at Northern Illinois University in DeKalb who dated the diatoms, sees corroboration for WAIS collapse in other records. Beaches, corals, and water lines suspended in sea cliffs high above current water levels reveal major sea level spikes on at least three occasions during the time when Scherer's diatoms might have grown: at 125,000 years, 400,000 years, and 1.07 million years ago. The million-year-old event corresponds with an episode of open sea in the 2006 ANDRILL core. It also corresponds to a spike in $^{16}O/^{18}O$ ratios in marine cores worldwide, a sign that ice sheets had injected fresh water into the oceans, because ice sheets preferentially incorporate water molecules containing ^{16}O over ^{18}O.

One potential payoff from knowing Antarctica's history is a better understanding of how its ice will respond to global warming. The 90-meter segment of diatom-rich core that Naish helped drill in 2006 points to an open Ross Sea 4 million years ago, a time when marine-core records from other parts of the world suggest that CO_2 levels hovered around 400 parts per million, with global temperatures 3° to 4°C warmer than today. Naish notes that those conditions correspond to optimistic predictions in last year's Intergovernmental Panel on Climate Change report of CO_2 levels in 2100. "If we can control our carbon emissions, we might be able to keep warming to 3°C above present and stabilize CO_2 at 450 parts per million," he says. "But even at those levels, we don't have a West Antarctic Ice Sheet," implying that WAIS would disappear in the coming centuries.

Marchant and Lewis are already looking for fossils in other parts of the Dry Valleys. Ashworth advocates broadening the search to provide a more nuanced view of how Antarctica's two massive ice sheets have changed over time. For example, the East Antarctic Ice Sheet as a whole seems to have been stable over time. But some areas around its fringes sit on marine beds like WAIS does, and people would love to know more about the history of those areas. "All the way around the edge of Antarctica, you could have pockets of these fossils preserved," says Ashworth. "I've got this gut feeling that [more of] these deposits are going to be turning up." If he's right, what was once a rare discovery may someday become a standard tool for understanding the history of Antarctica's ice sheets and climate. **–DOUGLAS FOX**

Douglas Fox is a science journalist based in northern California.



Captured on film. Allan Ashworth (*left*) and Adam Lewis look for fossils in a scene from the documentary *Ice People*, by Anne Aghion.

CREDIT: © DRY VALLEYS PRODUCTIONS

Science fiction? Hydrogen sulfide might someday help people survive trauma and illness in a hibernation-like state.

CREDIT: JOHN LAMB/GETTY IMAGES

MEDICINE

Nothing Rotten About Hydrogen Sulfide's Medical Promise

Despite its toxicity and famously bad odor, hydrogen sulfide's ability to lower metabolism and create a hibernation-like state has scientists wondering whether the gas can help soldiers and other people withstand injuries or surgeries

It's foul-smelling, corrosive, flammable, and deadly. It's the bane of oil fields, sewage treatment plants, and farms because at concentrations workers sometimes encounter, a single breath of it can kill. Hydrogen sulfide (H_2S), the rotten egg gas, is not something you would think to pump into sick or injured people.

But that's exactly what some scientists plan to do, reflecting the gas's improving reputation over the past 2 decades. Long known for its distinctive smell and toxicity—it starves our cells by disabling an enzyme necessary for extracting energy from food—the molecule has proven to be an influential physiological signal, with effects on everything from blood flow to hormone secretion. Eager to capitalize on these newfound capabilities, scientists are trying to exploit hydrogen sulfide to tame the side effects of common painkillers, for example, and curb heart attack damage. After announcing last week that injecting low doses of hydrogen sulfide into healthy people produced no dangerous side effects, one company plans later this year to start testing the molecule as a treatment for several conditions, possibly including restricted blood flow to the liver and lung injuries. "We are right at the beginning of an expanding field that could have enormous clinical implications," says cardiovascular physiologist David Lefer of Albert Einstein College of Medicine in New York City.

In the past few years, hydrogen sulfide research has also veered into science-fiction territory, as investigators have found that nonlethal doses of the molecule can send small animals into a hibernation-like state. Whether this unexpected effect can be reproduced in large animals, or people, remains a matter of debate, but that hasn't stopped some scientists—and the U.S. military—from investigating whether the gas could allow patients to better survive severe injuries or traumas, such as a stroke, by placing them in a form of suspended animation.

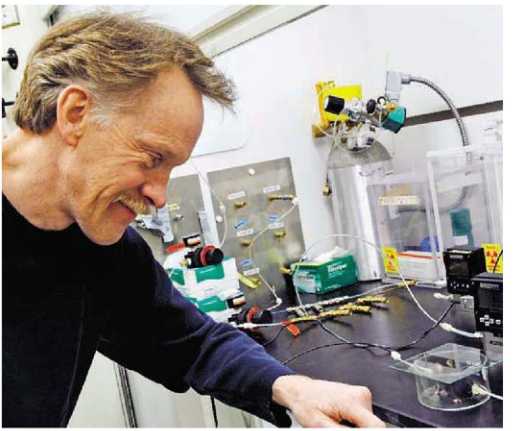
Not just sewer gas anymore

Toxicologists know plenty about the downside of hydrogen sulfide. Even at 10 parts per million (ppm), the exposure limit set by the U.S. National Institute for Occupational Safety and Health, the gas can irritate the eyes. If you breathe 500 ppm, you can die within half an hour, and 1000 ppm knocks you out instantly and kills within a few minutes.

So researchers were surprised to discover that the human body naturally makes the potentially lethal molecule, although at much lower concentrations. This ability represents a legacy from some of the earliest microbes, says physiologist Rui Wang of Lakehead University in Thunder Bay, Canada. These organisms—like a few living today—relied on sulfur, not oxygen, to obtain energy from metabolism. Although oxygen took over this job for the most part, organisms still put hydrogen sulfide to use for numerous other functions.

Hydrogen sulfide isn't the only noxious gas to reveal a good side. Nitric oxide was the first so-called gasotransmitter that scientists identified. Nobel Prize–winning work that started in the mid-1980s demonstrated that nitric oxide relaxes blood vessels, quells inflammation, nudges the hypothalamus to release hormones, and even transmits signals between the brain's neurons. Another killer gas, carbon monoxide, later joined the gasotransmitter family (*Science*, 21 November 2003, p. 1320). Despite these precedents, hydrogen sulfide researchers say they still get grief for their work, particularly from toxicologists. "A week doesn't go by when I'm not answering questions like, 'Isn't this a poison?'" says pharmacologist John Wallace of the University of Calgary in Canada.

Interest in hydrogen sulfide's possible benefits bubbled up about 20 years ago, when three papers described surprisingly high concentrations in brain samples from animals and humans. What was the molecule doing in healthy tissue, researchers wondered. It might help us to learn, replied neuroscientist Hideo Kimura and postdoc Kazuho Abe, then at the Salk Institute for Biological Studies in San Diego, California. In 1996, the pair found that in slices of rodent brain tissue, hydrogen sulfide spurs long-term potentiation, an increase in synapse sensitivity that can promote learning and memory. Unlike nitric oxide, hydrogen sulfide doesn't appear to transmit messages between neurons, says Kimura, who's now at the Institute of Neuroscience in Tokyo. Rather, it acts as a neuromodulator that



Reanimator. Mark Roth of the Fred Hutchinson Cancer Research Center prepares to give a mouse a dose of hydrogen sulfide gas.

adjusts the responsiveness of neural circuits.

The molecule's résumé includes other responsibilities, subsequent studies have shown. Many of its effects stem from the power to open the membrane channels that allow potassium to leave the cell, says Wang. For example, in 2001, his group showed that by prodding these channels, hydrogen sulfide relaxes smooth muscle cells in the walls of blood vessels, suggesting that the molecule regulates blood pressure. Researchers are still trying to resolve the sometimes contradictory reports on hydrogen sulfide's effects—for instance, studies conflict about whether it is pro- or anti-inflammatory—and nail down all of its functions. But physiologist David Kraus of the University of Alabama, Birmingham, predicts that hydrogen sulfide "will rival the prominence of nitric oxide."

One apparent function of the molecule—soothing cells under stress—has attracted medical interest. Kimura and colleagues revealed 4 years ago that hydrogen sulfide can shield cultured neurons from oxidative damage, which often occurs after a stroke. Instead of combating oxidants directly, the molecule spurs cells to bump up the levels of glutathione, neurons' natural antioxidant. Researchers still need to test whether hydrogen sulfide curbs brain damage from a stroke in animals, Kimura says.

However, administration of hydrogen sulfide does seem to limit the damage from a heart attack, as Lefer and colleagues revealed last year. To simulate a clogged artery, the researchers temporarily tied off one of the vessels that delivers blood to the left ventricle in mice. An injection of hydrogen sulfide directly into the heart cut the amount of scarring and inflammation that resulted once blood flow resumed. Protection of mitochondria, the cell's energy-generating organelles, might explain that finding. In the untreated mice, the organelles' capacity to use oxygen plummeted. The mitochondria were swollen, and their complex interior structure appeared scrambled. This damage was absent in treated mice. Lefer and colleagues are already assessing hydrogen sulfide for other conditions, including heart failure, and plan to test sulfide-releasing pills that are now under development.

Pills that promote the creation of hydrogen sulfide might also protect the gut. Cells of the gastrointestinal (GI) tract naturally make the molecule, possibly to regulate blood flow or shield intestinal linings. In rodent studies, Wallace and his colleagues have found that intravenous administration of hydrogen sulfide fends off side effects, such as GI bleeding and ulcers, frequently caused by nonsteroidal anti-inflammatory drugs (NSAIDs), the drug class that includes aspirin and ibuprofen. Seeking easier-to-stomach NSAIDs, the researchers designed versions that emit small amounts of hydrogen sulfide. In tests on rats, the modification almost eliminated intestinal injury from the NSAID diclofenac, the team reported last year. These new NSAIDs "don't cause any GI damage at all and are as potent as the parent drugs," Wallace says. He predicts that the drugs will reach clinical trials within 18 months. And a sulfide-delivering version of mesalamine, a treatment for inflammatory bowel disease that causes similar GI anguish, could be ready for testing later this year, he says.

A big sleep

While researchers have been gradually unearthing the physiological roles of hydrogen sulfide, Mark Roth's heart-slowing experiments with the gas grabbed headlines 3 years ago. A physiologist at the Fred Hutchinson Cancer Research Center in Seattle, Washington, Roth was studying how animals drastically reduce their metabolism, as some do during hibernation. His lab found that small animals such as nematodes and zebrafish embryos could endure oxygen concentrations supposedly below the minimal level for survival. The animals remained barely alive—the embryonic fish's hearts often stopped beating—but revived when researchers cranked up the oxygen levels. The team then started looking for ways to induce the same effect by preventing animals from using available oxygen. Carbon monoxide worked for nematodes, but they thought it would be too risky for humans.

Heart healthy. After a simulated heart attack, the scarred heart of a control mouse (*left*) contrasts with the ruddy tissue of a mouse treated with hydrogen sulfide (*right*).

CREDITS: (TOP TO BOTTOM) FRED HUTCHINSON CANCER RESEARCH CENTER; ELROD *ET AL.*, *PNAS*, 2 SEPTEMBER 2007, VOL. 104, NO. 36, P. 15560

Roth says he decided to try hydrogen sulfide instead after seeing a TV documentary that mentioned its dangers to cavers.

When Roth's team exposed mice to 80 ppm of the gas, which toxicology studies had established was safe for rodents, the animals passed out. Their core body temperature plunged more than 20°C, their oxygen consumption fell, and their carbon dioxide output—an indicator of metabolic rate—tumbled. Once the researchers shut off the gas and provided heat, the mice were up and gnawing within an hour. They passed a battery of behavioral tests, indicating that they incurred no brain damage during their down time. In a follow-up study published last year, Roth's team showed that hydrogen sulfide could enable mice to survive low oxygen concentrations that are otherwise lethal to rodents.

After being gassed, the mice slip into an altered state that differs from the unconsciousness of sleep, hibernation, and anesthesia. Their eyes are closed, but unlike anesthetized patients, they are not paralyzed and respond to pain, says Warren Zapol, chief of anesthesia at Massachusetts General Hospital in Boston, who has studied the effects of hydrogen sulfide on the rodents. "You couldn't take their appendix out." he jokes. Pinch a tail, and they try to wriggle away, Zapol says. The animals aren't hibernating, either, as that involves much more gradual changes in everything from fat deposition to protein synthesis.

In an April paper in *Anesthesiology*, Zapol and colleagues provided the most detailed view yet of the cardiovascular changes induced by hydrogen sulfide. After breathing 80 ppm of gas, the animals' heart rate declined more than 50%, to about 250 beats per minute, the team reported. Typically, when a mouse's heart rate plunges, the organs can begin to run short of oxygen, Zapol notes. But the rodents' blood pressure remained steady, suggesting that the delivery of oxygen throughout the body didn't falter, he says.

hydrogen sulfide appears to serve as a master metabolic regulator, Roth says. He notes that 18th century British chemist Joseph Priestly likened humans to burning candles, consuming oxygen to keep the flame going. "We have stumbled across the mechanism by which human beings are regulating the degree to which they burn their candle," Roth says.

Scent of survival

Roth now wants to lower the flame on people's candles. He envisions temporarily turning down metabolism and oxygen demand with a dose of hydrogen sulfide, buying time for patients who have suffered heart attacks, strokes, or wounds that produce drastic blood loss. In a study that *The Journal of Trauma* will publish in July, Roth and colleagues show that 66% of rats injected with a hydrogen sulfide solution can survive the loss of 60% of their blood, versus 14% of rats given a control solution. He's now testing for the same benefits in larger animals and has received a grant from the U.S. Defense Advanced Research Projects Agency to design injectable hydrogen sulfide kits that troops could carry in the field. Wounded soldiers could be "deanimated," suggests Roth, until they can be evacuated to a hospital.

Roth acknowledges that using hydrogen sulfide in this way faces some public relations hurdles. People expect doctors and paramedics to resuscitate their loved ones, not put them into a near-death state, he says: "The notion that you are better off deanimated than animated is not something people want to think about."

The bigger problem, say some critics, is that hydrogen sulfide won't elicit the same effect in people as it does in mice. Two studies on large animals bolster their skepticism. Last fall, for instance, a French team reported no metabolic decline in sheep that breathed 60 ppm of hydrogen sulfide gas.

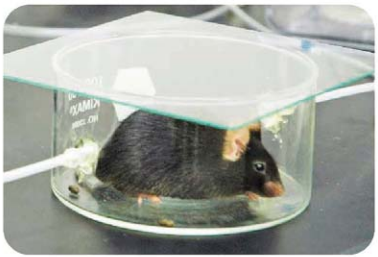
Recent work on pigs by cardiologist Andrew Redington of the Hospital for Sick Children in Toronto, Canada, and colleagues suggests that nonlethal concentrations of the gas could exacerbate patients' problems. In their experiment, piglets that had been anesthetized much like a surgical patient breathed gradually increasing concentrations of hydrogen sulfide gas, from 20 ppm to 80 ppm. Instead of slowing the metabolism of the piglets, hydrogen sulfide revved it up, the team reported in January. Heart rate, blood pressure, and cardiac output were all higher in animals that breathed the gas than they were in controls. Many of the target patients for hydrogen sulfide treatment have weakened hearts, and forcing the organ to work harder could kill them, says Redington. His take is that in large animals, hydrogen sulfide "has no effect at best but possibly a detrimental effect." Resuscitation researcher Samuel Tisherman of the University of Pittsburgh Medical Center in Pennsylvania isn't as harsh but concedes that "some of the promise seems to be slipping away."

Roth counters that both large-animal studies share a crucial flaw: The concentration of hydrogen sulfide was too low. Although 80 ppm will knock out a mouse, a pig or sheep requires a bigger hit, he claims. He and his colleagues are conducting their own tests on large animals, but he says they're not ready to discuss the results.

Hydrogen sulfide doesn't need to produce dramatic effects to provide benefits, Roth notes. Ikaria, the company that he co-founded 3 years ago and for which Lefer serves as a consultant, has just announced the results of the first safety trial of an injectable form of hydrogen sulfide. Last week, the company revealed that 36 volunteers who received doses well below the level that causes unconsciousness in people showed no ill effects. Csaba Szabo, Ikaria's chief scientific officer, says that before the end of the year, the company plans to launch phase II trials. What types of patients will receive the drug isn't certain, but they could include people who have suffered heart attacks or who are undergoing operations such as a heart or lung bypass.

If suspended animation isn't futuristic enough for you, Roth published a paper last year hinting at another dramatic effect: increased longevity. His lab found that exposure to the gas stretches the life span of nematodes by up to 70%. Whether or not other organisms react the same way, it's clear that hydrogen sulfide has come a long way from just being a killer.

–MITCH LESLIE



Gassing up. As this mouse breathes hydrogen sulfide, its heart rate, metabolic rate, and body temperature will plummet.

CREDIT: FRED HUTCHINSON CANCER RESEARCH CENTER

LETTERS

edited by Jennifer Sills

Joshua Lederberg's Interest in Ignorance

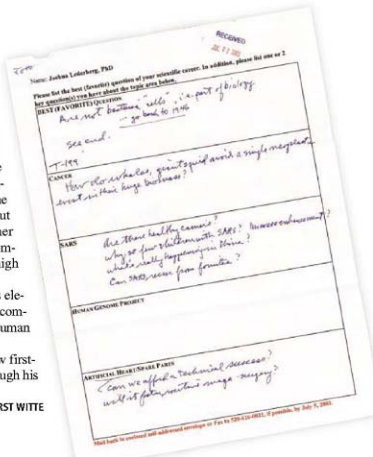
IN OCTOBER 2000, TO MY SURPRISE AND DELIGHT, I RECEIVED A REQUEST FROM Joshua Lederberg to "be kind enough to favor me with a copy" of an article I had written on "Ignorance in infectious diseases" (*1*) and to "please tell me more about the Ignorance Agenda." Thus began my own "continuing conversation" with Joshua Lederberg, the subject of the Retrospective by S. S. Morse in the 7 March issue (p. 1351). We exchanged letters, reflections, and references about how to use ignorance—unanswered questions and unquestioned answers—rather than knowledge as the terrain for learning and discovery. Along the way, he completed the handwritten homework assignment that I assigned to our summer high school student researchers.

Note that Dr. Lederberg's favorite question ("Are not bacteria 'cells'...?"), in its elegant paradigm-shattering simplicity, would probably not score well in higher-order complexity according to Bloom's taxonomy. And curiously, he left the space under Human Genome Project blank!

I never had the opportunity to meet Dr. Lederberg in person, but I came to know firsthand the grace, humility, uncanny intellect, and childlike wonder that endured through his latter years.

MARLYS HEARST WITTE

Department of Surgery, University of Arizona College of Medicine, Tucson, AZ 85724–5200, USA.



Reference

1. M. H. Witte, C. L. Witte, *Lymphology* **33**, 95 (2000).

Women in Science: A Top-Down Approach

APRIL 28 MARKED A DECADE SINCE THE European Union Women and Science Conference in Brussels, which generated several initiatives that have been serving EU women in science well. One initiative (*1*) set a goal to seek at least 40% representation (on average) for women in Marie Curie scholarships, advisory groups, and assessment/ monitoring panels throughout the EU Fifth Framework Programme; when the EU initiative began in 1999, representative participation of women was 23% in evaluation panels, 23% in monitoring panels, 29% in advisory groups, and 21% in program committees (*2*). The EU Commission's report, "Women and science: Mobilizing women to enrich European research," was passed into legislation by the European Parliament on 3 February 2000 (*3*). Subsequent EU reports on the representation of women in these positions chronicle progress toward the 40% goal. By 2001, the representation of women had increased to 27% in EU evaluation panels, 35% in monitoring panels, 28% in advisory groups, and 22% in program committees (*2*).

In the United States, there are few comparable data, but in 2001, we determined the percentage of underrepresented groups among American Chemical Society (ACS) publication editors (13.6% women, 94.3% white) and their advisory boards (13.3% women, 88.6% white) (*4*). Now in 2008, we have found (*5*) that representation of women in these ACS positions has increased (editors are 14.5% women; advisory boards are 17.5% women), but not to the EU levels given above.

The EU legislation also had an effect on Marie Curie scholarships. For example, during 1994 to 1998, the success rate of female students was 83% that of male students; in 1999, it rose to 92% of male students (*6*). The EU 40% legislation has been described as part of a "top-down" approach, with the goal to increase the participation of women at all levels in science, technology, engineering, and mathematics (STEM). In contrast, most U.S. programs and efforts seem focused on increasing underrepresented groups at the lowest rather than highest levels of the academic pipeline.

Perhaps we should add the EU "top-down" approach to our repertoire of existing programs in the United States. Although the representation of women among award recipients cannot be legislated, their representation among committee members selecting award recipients can be. The representation of minorities among committee members could be addressed simultaneously.

An additional benefit of such a program is that young women and minorities, whom we

want to attract to STEM, would see more of their role models in leadership positions. It is illogical to expect members of a group to be inspired into science and engineering if senior members of that group are seen not to share equally in the fruits of the profession.

DONNA J. NELSON* AND CHRISTOPHER N. BRAMMER

Department of Chemistry, University of Oklahoma, Norman, OK 73019, USA.

*To whom correspondence should be addressed. E-mail: djnelson@ou.edu

References

1. European Commission, "Women and science: Mobilizing women to enrich European research," 17 February 1999 (COM(1999)76 final).
2. European Commission, "Women and science: Excellence and innovation: gender equality in science," 11 March 2005 (EUR21784).
3. European Parliament, "Resolution on the communication from the Commission entitled: Women and science: Mobilising women to enrich European research," 3 February 2000 (P5_TA(2000)0056).
4. D. J. Nelson, *AWIS Magazine* **30**, 33 (summer 2001); http://cheminfo.chem.ou.edu/faculty/djn/diversity/Pubs/AWIS_2001_Summer_p33-39.pdf.
5. D. J. Nelson, C. Brammer, "Representation of women among 2000 ACS editors and editorial advisory boards" (Diversity in Science Association, 2000); http://cheminfo.chem.ou.edu/faculty/djn/diversity/top50.html.
6. European Commission European Technology Assessment Network (ETAN), "Science policies in the European Union: Promoting excellence through mainstreaming gender equality" (Report from the ETAN Expert Working Group on Women and Science, 2000), Tables 4.3 and 4.4.

Biological Basis of the Third-Cousin Crush

IN THEIR REPORT, "AN ASSOCIATION BETWEEN the kinship and fertility of human couples" (8 February, p. 813), A. Helgason and colleagues provided solid evidence for an association between kinship and fertility in human couples. They demonstrated that greatest reproductive success was obtained by couples who are third or fourth cousins, not among unrelated couples as a less-informed observer might expect, and concluded that this association is likely to have a hitherto-unknown biological basis. The likely biological mechanism for this phenomenon has, however, already been identified. The major histocompatibility complex (MHC) has been shown to mediate mate choice (*1*–*3*) in most vertebrates studied. Although MHC genes are best known for their role in immune recognition and transplantation success (*4*), early studies demonstrated that they contribute to the well-known phenomenon of inbreeding avoidance (i.e., potential mates with a high number of matching MHC alleles are avoided) (*2*, *3*). Later, these results were demonstrated in humans when MHC-dependent mating preference was suggested in isolated religious enclaves, where couples were much less likely to share MHC haplotypes than expected by chance (*5*, *6*). We know now that mate choice is governed not by the quest of maximum outbreeding, but by the pursuit of a balance between the costs of inbreeding and outbreeding. Bateson (*7*) originally postulated that mate selection was governed by weighting the allure of familiar (kin) with that of novel (unrelated),

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

thus creating an optimal level of outbreeding. Bateson (*8*) later experimentally validated this hypothesis by demonstrating that birds investigate relatives on the level of first or second cousins more than either closer kin or unrelated birds. The similarities between Bateson's hypothesis and results and those of Helgason and colleagues are striking. More recently, Jacob and colleagues (*9*) elegantly demonstrated that human mate choices are influenced by paternally, but not maternally, inherited MHC alleles communicated by our body odors. Helgason and colleagues' finding thus provides both a genealogical demonstration of the curvilinear effect one would expect to see from the optimal-level outbreeding hypothesis and the first ecological test of the hypothesis by measurement of reproductive outcome rather than preference.

JOHAN N. LUNDSTRÖM,[1,2]*
CHARLES J. WYSOCKI,[1,3] MATS J. OLSSON,[4]
GEORGE PRETI,[1] KUNIO YAMAZAKI[1]

[1]Monell Chemical Senses Center, Philadelphia, PA 19104, USA. [2]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA. [3]Department of Animal Biology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [4]Section for Psychology, Department of Clinical Neuroscience, Karolinska Institute, 171 77, Stockholm, Sweden.

*To whom correspondence should be addressed. E-mail: jlundstrom@monell.org

References

1. M. Milinski *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4414 (2005).
2. W. K. Potts, C. J. Manning, E. K. Wakeland, *Nature* **352**, 619 (1991).
3. K. Yamazaki *et al.*, *J. Exp. Med.* **144**, 1324 (1976).
4. J. Klein, *Natural History of the Major Histocompatibility Complex* (Wiley, New York, 1986).
5. C. Ober, *Hum. Reprod. Update* **5**, 103 (1999).
6. C. Ober *et al.*, *Am. J. Hum. Genet.* **61**, 497 (1997).
7. P. Bateson, *Nature* **273**, 659 (1978).
8. P. Bateson, *Nature* **295**, 236 (1982).
9. S. Jacob, M. K. McClintock, B. Zelano, C. Ober, *Nat. Genet.* **30**, 175 (2002).

TECHNICAL COMMENT ABSTRACTS

COMMENT ON "Mixed-Layer Deepening During Heinrich Events: A Multi-Planktonic Foraminiferal $\delta^{18}O$ Approach"

Claude Hillaire-Marcel and Anne de Vernal

Rashid and Boyle (Reports, 19 October 2007, p. 439) analyzed oxygen isotopes in planktonic foraminera from marine sediments and concluded that Heinrich events (massive iceberg discharges into the North Atlantic Ocean) caused upper water masses to deepen. We question the robustness of this interpretation and argue that a strongly stratified mixed layer characterized by dense sea-ice cover and production of oxygen-18–depleted brines likely prevailed during such events.

Full text at www.sciencemag.org/cgi/content/full/320/5880/1161a

RESPONSE TO COMMENT ON "Mixed-Layer Deepening During Heinrich Events: A Multi-Planktonic Foraminiferal $\delta^{18}O$ Approach"

Harunur Rashid and Edward A. Boyle

Hillaire-Marcel and de Vernal question our hypothesis of the "homogenization of the upper water masses" of the North Atlantic during Heinrich events 0, 1, and 4. Here, we respond to their specific points and synthesize data from the southern Labrador Sea cores, the basic tenet of Hillaire-Marcel and de Vernal's arguments.

Full text at www.sciencemag.org/cgi/content/full/320/5880/1161b

SCHOLARLY COMMUNICATION

Sharing Data, Constructing Science

Karla L. Hahn

We have now traveled a short distance down the path into what is widely recognized as a paradigm shift in how new knowledge is created and exchanged. Twenty-first century research requires the ubiquity of digital information. We expect the networked environment to house the scientific literature and foster collaboration between researchers on disparate continents. Now primary data are moving into the electronic realm as well, with shared digital data stores from GenBank to the Sloan Digital Sky Survey feeding entire research specialties.

While science prospers under the new order, there is much contention over the practices that will determine the ethos of the digital age. What will be the new norms for "sharing" and "access"? The movement of documents and data into a digital realm has placed great strains on the structures supporting the exchange of research knowledge—social structures as well as technologies. Major battles are being fought over such issues as the appropriate intellectual property regime for a digital age. Publishing systems that formerly maximized the exchange of scientific research are becoming, instead, barriers to research distribution. Both issues underlie the recently implemented National Institutes of Health Public Access Policy that requires that articles resulting from NIH funding be deposited in PubMed Central and become available to readers within a year of publication. Such questions extend beyond research reports and publications to touch research data as well.

Christine Borgman's *Scholarship in the Digital Age: Information, Infrastructure, and the Internet* looks across the emerging landscape, identifying key choices and exploring the perspectives of different disciplines. Borgman (a professor in information studies at the University of California, Los Angeles) has been both a keen observer and an active participant in recent shifts to new digital modes of information exchange, keeping an eye as well on the shifting cultures of the scholarly tribes that engage in the research enterprise. She has received several major grants from the National Science Foundation for digital library infrastructure and data management projects and also been active in information policy development. Her previous book, *From Gutenberg to the Global Information Infrastructure* (*1*), also looked at digital library development, information infrastructure, and information policy issues, but comparing the two books merely reveals how much has changed since the first was published in 2000—both in the external environment, including policy issues, and in understanding the social environment of networked science.

Scholarship in the Digital Age
Information, Infrastructure, and the Internet
by Christine L. Borgman
MIT Press, Cambridge, MA, 2007. 360 pp. $35, £22.95. ISBN 9780262026192.

A striking feature of *Scholarship in the Digital Age* is its synthetic approach, integrating policy documents and published scholarship. Borgman lays out the trajectory of change in information exchange with the advent of the digital age, but her main concern is developing "a model of cyberinfrastructure in which scientific databases and digital libraries form an 'information and content layer' above the middleware layer that provides the services and underlying core network technologies." A dichotomy of research data and authored works structures her analysis of this content layer.

Drawing on the findings of decades of academic study as well as policy analysis, Borgman characterizes researchers' changing practices for creating and using scholarly works. For scientists with a general interest in this area, the book provides a readable, up-to-date, and well-integrated overview.

In contrast to the body of knowledge on the culture and practices of scholarly communication, no such foundational research exists for researchers' use of data. Consequently, Borgman's analysis of the challenges posed by the burgeoning possibilities for building collections of digital research data is her most original and insightful contribution. She demonstrates that, although behaviors regarding publications and data are tightly interconnected, they tend not to be analogous. Even the incentives and disincentives for sharing documents are notably different from those for sharing data.

Yet data and documents cannot be treated separately. The problems of connecting the two predate the digital era, but in the dichotomous content layer, the need to create and manage connections between data and documents assumes new prominence despite the inherent challenges. In theory, data can be embedded or co-published with documents, but this is minimally useful. Documents and data rarely exist in a one-to-one relationship. Rather, some documents draw on multiple data sources, and many data sources support multiple documents. These relations negate simplistic solutions such as storing data with individual documents on disparate publisher repositories; they require investment in broadly interactive, network-based solutions.

Borgman also devotes substantial attention to researchers' social and cultural practices. Using generalized representations of the sciences, the social sciences, and the humanities, she compares and contrasts communication, research, and data traditions. The sciences serve as a touchstone contextualizing scarcer commentary on the social sciences and the humanities. Inevitably, this approach suffers from overgeneralization at various points, but her analysis of the incentives and disincentives for the scholarly sharing needed to build the content layer is illuminating.

CREDIT: GETTY IMAGES

Borgman persuasively argues that creating incentives to increase the shared body of content should be an integral part of the investment in developing infrastructure for e-research. This content layer does not simply archive the products of research—it is a key resource for research. Claims for access rights and calls for sharing, she believes, need to pay greater attention to the perspectives of the researchers who create both documents and data.

"Open science" has always been essential

The reviewer is at the Association of Research Libraries, 21 Dupont Circle, Washington, DC 20036, USA. E-mail: karla@arl.org

to the success of the research enterprise, with U.S. policy notable, as Borgman observes, for its enactment of this principle. But openness in the digital age has a more expansive meaning, and the changes needed to honor traditions of open science will be difficult—making some researcher resistance to new norms of openness predictable. Borgman's analysis successfully highlights key barriers to the development of new norms for access and sharing. However, some of the pessimism she expresses for policy strategies has already proven premature, as new policies from the National Institutes of Health and Harvard University's Faculty of Arts and Sciences create mandates for public access to research that truly change expectations and incentives.

Reference

1. C. L. Borgman, *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World* (MIT Press, Cambridge, MA, 2000).

10.1126/science.1155430

EXHIBITIONS: NATURAL HISTORY ART

Paper Museums

Elizabeth Regina can boast of possessing 600 drawings by Leonardo da Vinci (1452–1519), but she is shamefully modest when it comes to showing them off. Currently, you can see 18 selected by David Attenborough from the Royal collections on exhibition in *Amazing Rare Things: The Art of Natural History in the Age of Discovery* at the Queen's Gallery, Buckingham Palace .

What immediately strikes one about this set of Leonardo's drawings is how modern they seem, yet how old they look. It was clearly intensely important to Leonardo to work out the inner anatomy on which the outer structures hung. Sketches of the flaring nostrils of noble horses and exquisite tiny cats writhing across the paper combine in his depictions of St. George strenuously vanquishing mythical but anatomically convincing dragons. Several of his botanical illustrations were preparation for the foreground details of a now-lost painting of Leda and the swan. Leonardo was striving to depict a landscape as fecund as Leda standing among the eggshells of her twins' births. The drawings are so accurate that they provide the earliest record of the introduction of the grass known as Job's tears (*Coix lachryma-jobi*) into Europe from China. It is interesting to see this plant in the context of fertile Leda: it is used in China as an herbal medicine for reproductive control (*1*) and was possibly collected and brought back to Europe for this purpose.

By the 16th century, Italy was a hub for explorers who brought marvels from around the world. An adjoining room offers a dazzling display of paintings from the "Museum of Paper" amassed by Cassiano dal Pozzo (1588–1657). Horticulturalists in particular cultivated paper museums as records of their ephemeral subjects. Several intriguing examples on show originated from the private florilegium of the English gentleman gardener Alexander Marshal (1620–1682) (*2*). There are pictures containing puzzling juxtapositions of objects among the flowers: A jay with a broken leg lies dead in one. In others, Marshal's pet dogs and birds romp or snooze in miniature among giant plants. In the corner of another sheet, a speckled frog waits damply for a caterpillar to come closer. Marshal was also fascinated by insects, to the point of illustrating the tiny details of insect damage on his plants.

Maria Sibylla Merian (1647–1717) shared with Marshal an expertise in making durable colors for painting and a fascination with the metamorphoses of insects and amphibians. Driven by her passion for the natural world, she flouted convention by leaving home at the age of 52 and traveling to Surinam to paint. Nevertheless—like Mark Catesby (1682–1749), some of whose pictures from North America are included in the exhibition—she tended not to challenge anatomical preconceptions. Merian used a distinctive curvaceous style, even when drawing from life. Despite her stylistic foibles, Merian's exuberantly colorful illustrations are as amazing now as they must have seemed in the 17th century: bird-eating spiders pouncing, snouted lantern bugs with widespread wings, and snakes wrestling caimans.

Mark Catesby's great hogfish [*Lachnolaimus maximus*] (c. 1725).

Detail from page in Marshal's florilegium (17th century).

Although the paintings are beautiful and the Leonardo drawings make you gasp, *Amazing Rare Things* is a very conventional exhibition. More could have been done to place the pictures in a scientific context, and ultimately, gazing at the catalog in one's own private palace is almost as good as viewing the originals. **–Caroline Ash**

Amazing Rare Things
The Art of Natural History in the Age of Discovery

Royal Collection Curators

The Queen's Gallery, Buckingham Palace, London. Through 28 September 2008. www.royalcollection.org.uk/microsites/amazingrarethings/

Amazing Rare Things
The Art of Natural History in the Age of Discovery

by David Attenborough, Susan Owens, Martin Clayton, and Rea Alexandratos

Royal Collection, London, 2007. 223 pp. £18.95. ISBN 9781902163468. Paper, £9.95. ISBN 9781902163994. Yale University Press, New Haven, CT, 2007. 223 pp. $37.50. ISBN 9780300125474.

References and Notes

1. S.-M. Hsia, *Exp. Biol. Med.* **232**, 1181 (2007).
2. Marshal's florilegium has just been published as *Mr. Marshal's Flower Book* (Royal Collection, London, 2008).

10.1126/science.1159434

DIVERSITY

Culture, Gender, and Math

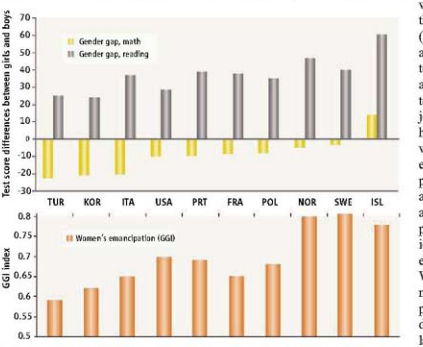
Analysis of PISA results suggests that the gender gap in math scores disappears in countries with a more gender-equal culture.

Luigi Guiso,[1]* Ferdinando Monte,[2]* Paola Sapienza,[3]*† Luigi Zingales[4]*

The existence (*1*), degree (*2*), and origin (*3*, *4*) of a gender gap (difference between girls' and boys' scores) in mathematics are highly debated. Biologically based explanations for the gap rely on evidence that men perform better in spatial tests, whereas women do better in verbal recall ones (*1*, *5*, *6*). However, the performance differences are small, and their link with math test performance is tenuous (*7*). By contrast, social conditioning and gender-biased environments can have very large effects on test performance (*8*).

To assess the relative importance of biological and cultural explanations, we studied gender differences in test performance across countries (*9*). Cultural inequalities range widely across countries (*10*), whereas results from cognitive tests do not (*6*). We used data from the 2003 Programme for International Student Assessment (PISA) that reports on 276,165 15-year-old students from 40 countries who took identical tests in mathematics and reading (*11*, *12*). The tests were designed by the Organisation for Economic Co-operation and Development (OECD) to be free of cultural biases. They are sufficiently challenging that only 0.6% of the U.S. students tested perform at the 99th percentile of the world distribution.

Girls' math scores average 10.5 lower than those of boys (2% less than the mean average score for boys), but the results vary by country (see chart, above): in Turkey, –22.6, whereas, in Iceland, 14.5. A similar variation exists in the proportion of girls over boys who score above 95%, or 99% of the country-level distribution (fig. S2A).



Math and reading gender gaps. In more gender-equal cultures, the math gender gap disappears and the reading gender gap becomes larger. **(Top)** Gender gaps in mathematics (yellow) and reading (gray) are calculated as the difference between the average girls' score and the average boys' score. A subset of countries is shown here (see SOM for complete data set and calculations). In many countries, on average, girls perform more poorly than boys in mathematics. In all countries, girls perform better than boys in reading. The gender gap in mathematics and reading correlates with country measures of gender status within the culture, one of which measures is the GGI (**bottom**). Larger values of GGI point to a better average position of women in society. Besides USA, the countries are abbreviated as their first three letters, except for PRT, Portugal, and ISL, Iceland.

The gender gap is reversed in reading. On average, girls have reading scores that are 32.7 higher than those of boys (6.6% higher than the mean average score for boys), in Turkey, 25.1 higher and in Iceland, 61.0 higher (see chart). The effect is even stronger in the right tail of the distribution. In spite of the difference in levels, the gender gap in reading exhibits a variation across countries similar to the gender gap in math. Where girls enjoy the strongest advantage in reading with respect to boys, they exhibit the smallest disadvantage (sometime even an advantage) in math. [The correlation between the average gender gaps in mathematics and reading across countries is 0.59 (fig. S4)].

To explore the cultural inputs to these results, we classified countries according to several measures of gender equality. (i) The World Economic Forum's Gender Gap Index (GGI) (*10*) reflects economic and political opportunities, education, and well-being for women (see chart). (ii) From the World Values Surveys (WVSs) (*13*), we constructed an index of cultural attitudes toward women based on the average level of disagreement to such statements as: "When jobs are scarce, men should have more right to a job than women." (iii) The rate of female economic activity reflects the percentage of women age 15 and older who supply, or are available to supply, labor for the production of goods and services. (iv) The political empowerment index computed by the World Economic Forum (*8*) measures women's political participation, which is less dependent on math skills than labor force participation. These four measures are highly correlated (table S2).

We find a positive correlation between gender equality and gender gap in mathematics (fig. S5). If Turkey, a low gender-equality country (GGI = 0.59), were characterized by the degree of gender equality manifested in Sweden (GGI = 0.81), our statistical model suggests that the mean score performance in mathematics of girls relative to boys would increase by 23 points, which would eliminate the Turkish gender gap in math (see table, p. 1165). In more gender-equal countries, such as Norway and Sweden, the math gender gap disappears. Similar results are obtained when we use the other indicators of women's roles in society. These results are true not only at the mean level, but also in the tail of the distribution (table S3). In Iceland, the ratio of girls to boys who score above the 99th percentile of the country distribution in math scores is 1.17.

There are many unobserved reasons why countries may differ in a way that affects the

[1]European University Institute, Villa San Paolo, Via della Piazzuola 43, 50133 Florence, Italy. [2]Economics Department, University of Chicago,1126 East 59th Street, Chicago, IL 60637, USA. [3]Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. [4]Graduate School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA.

*These authors contributed equally to this work. †To whom correspondence should be addressed. E-mail: Paola-Sapienza@northwestern.edu

Differences in Test Scores Correlated with Indicators of Gender Equality ♀♂

	LHS: Gender difference in math				LHS: Gender difference in reading			
Women's emancipation (GGI)	105.49± 26.92**				83.56± 30.43**			
Avg. WVS indicators		13.21± 7.06				16.39± 8.46		
Female economic activity rate			0.45± 0.14**				0.34± 0.15*	
Women's political empowerment				29.10± 10.05**				24.35± 10.86*
Log GDP per capita, 2003	-6.56± 2.40**	1.09± 2.26	-3.12± 1.93	-4.95± 2.52	-2.23± 2.71	0.52± 2.71	-0.56± 2.15	-1.06± 2.73
Constant	-19.62± 20.01	-57.16± 23.27*	-2.75± 17.72	32.43± 23.72	-3.02± 22.62	-16.09± 27.90	21.49± 19.80	39.03± 25.63
Observations (no.)	37	32	39	36	37	32	39	36
R^2	0.32	0.15	0.23	0.21	0.20	0.14	0.12	0.15

Culture affects the gap. More gender-equal cultures are associated with reducing the negative gap in math and further enlarging the positive gap in reading in favor of women. Test scores are positively correlated with indicators of gender equality in society (GGI, WVSs, see text). Economic conditions are accounted for by per capita Gross Domestic Product (GDP). The correlation persists among high achievers on both tests (table S3). See SOM for details of statistical analysis. The constant is where the regression line intercepts the *y* axis, representing the amount the dependent *y* (gender gap) will be when all the independent variables are set to 0. LHS, left-hand side variable in the least-squares regression analysis. *$P < 0.05$; **$P < 0.01$.

math gender gap. Without appropriate controls, we run the risk of capturing a spurious correlation between the unobserved factors and our measures of gender equality. We reran our regression at the student level, inserting a dummy variable for each country, to control for unobserved heterogeneity (table S4). The interaction between gender and GGI index remains statistically significant at the 1% confidence level in a two-tailed *t* test, which suggests that the correlation between gender equality and girls' math scores is not driven by unobserved heterogeneity. This interaction between gender gap and GGI remains significant even when we insert an interaction between gender and log of GDP per capita, which suggests that the improvement in math scores is not just related to economic development, but to the improvement of the role of women in society.

To investigate whether the disappearance of the math gender gap in some countries translates into an overall improvement of girls or is simply limited to mathematics scores, we correlated reading performance differences with measures of women's equality (see table, above). In countries where women are more emancipated, girls' comparative advantage in reading widens. Comparing Turkey (GGI = 0.59) and Sweden (GGI = 0.81), we see an increase in the mean score performance of girls relative to boys in reading by 18 points, which almost doubles Turkey's reading gap in favor of girls.

To verify that these results are not driven by biological differences across countries, we analyzed whether they persist in populations that have a similar or identical evolutionary history. To assess history, we used a genetic distance measure (*14*–*17*) based on the frequency of each allele across DNA polymorphisms.

According to this measure, there are 13 European countries with genetic distance equal to zero and 26 European countries with genetic distance less than 100 (table S5). When we restrict the regression of the table (above) to either one of these two groups, our findings are substantially unchanged (table S6).

These results suggest that the gender gap in math, although it historically favors boys, disappears in more gender-equal societies. The same cannot be said for how boys score in mathematics compared with how boys score in readings. Boys' scores are always higher in mathematics than in reading, and although the difference between boys' math and boys' reading scores varies across countries, it is not correlated with the GGI index or with any of the other three measures of gender equality (table S7A). Hence, in countries with a higher GGI index, girls close the gender gap by becoming better in both math and reading, not by closing the math gap alone. The gender gap in reading, which favors girls and is apparent in all countries, thus expands in more gender-equal societies. Similarly, although the gender gaps in all math subfields decrease in societies with more gender equality, the difference between the gender gap in geometry (where the boys' advantage relative to the girls' is the biggest) and arithmetic (where the boys' advantage relative to the girls' is the smallest) does not (table S7B).

This evidence suggests that intra-gender performance differences in reading versus mathematics and in arithmetic versus geometry are not eliminated in a more gender-equal culture. By contrast, girls' underperformance in math relative to boys is eliminated in more gender-equal cultures. In more gender-equal societies, girls perform as well as boys in mathematics and much better than them in reading. These findings shed some light on recent trends in girls' educational achievements in the United States, where the math gender gap has been closing over time (*2*).

References and Notes

1. L. V. Hedges, A. Nowell, *Science* **269**, 41 (1995).
2. C. Goldin, L. F. Katz, I. Kuziemko, *J. Econ. Perspect.* **20**, 133 (2006).
3. C. P. Benbow, J. C. Stanley, *Science* **210**, 1262 (1980).
4. "The science of gender and science: Pinker vs. Spelke, a debate," *Edge*, 10 May 2005, no. 160; www.edge.org/documents/archive/edge160.html#d.
5. S. Baron-Cohen, *The Essential Difference: Men, Women, and the Extreme Male Brain* (Allen Lane, London, 2003).
6. D. Kimura, *Sex and Cognition* (MIT Press, Cambridge MA, 1999).
7. E. S. Spelke, *Am. Psychol.* **60**, 950 (2005).
8. D. Halpern, J. Wai, A. Saw, in *Gender Differences in Mathematics*, A. M. Gallagher and J. C. Kaufman, Eds. (Cambridge Univ. Press, New York, 2005), pp. 48–72.
9. Materials and methods are available as supporting material on *Science* Online.
10. R. Hausmann, L. D. Tyson, S. Zahidi, *The Global Gender Gap Report* (World Economic Forum, Geneva, Switzerland, 2006).
11. OECD, Programme for International Student Assessment (PISA), 2nd Assessment (OECD, Paris, 2003).
12. PISA includes originally 41 countries; we drop Liechtenstein because it contains only 165 observations, which makes problematic any calculation of the tail of the distribution. All other countries have at least 639 observations.
13. R. Inglehart *et al.*, "World Values Surveys and European Values Surveys, 1981–1984, 1990–1993, and 1995–1997" [Computer files; Interuniversity Consortium for Political and Social Research (ICPSR) version] (Institute for Social Research, Ann Arbor, MI, 2000), distributed by ICPSR.
14. This measure was originally computed at the population level by (*15*).
15. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, 1996).
16. This measure has been mapped on modern countries by (*17*).
17. E. Spolaore, R. Wacziarg, "The diffusion of development," Centre for Economic Policy Research Discussion Paper 5630 (CEPR, London, 2006).

18. We thank S. Bafiga; C Hoxby; seminar participants at Northwestern University, University of Chicago, and the National Bureau of Economic Research (NBER); and two anonymous referees for their helpful comments. The Initiative on Global Markets at the University of Chicago provided financial support.

10.1126/science.1154094

Supporting Online Material
www.sciencemag.org/cgi/content/full/320/5880/1164/DC1

ATMOSPHERIC SCIENCE

Whither Geoengineering?

Costs, benefits, and harms associated with geoengineering must be assessed before it is used to mitigate climate change.

Alan Robock

According to the Intergovernmental Panel on Climate Change (IPCC) (*1*), global warming will soon have severe consequences for our planet. The IPCC also estimates (*2*) that mitigation would only cost ~0.1% of the global gross national product per year for the next 30 years, a price far smaller than the damage that would occur. As a potential route to mitigation, the old idea of "geoengineering" has gotten much attention in the last 2 years (*3*, *4*). On page 1201 of this issue, Tilmes *et al.* (*5*) quantify the effects of one geoengineering approach—the introduction of additional aerosols into Earth's stratosphere, akin to a volcanic eruption—on high-latitude stratospheric ozone concentrations.

A polar stratospheric cloud over McMurdo, Antarctica, on 24 August 2004. These clouds cause ozone depletion every spring because of anthropogenic chlorine in the stratosphere. The ozone hole is expected to disappear by the middle of this century, but with geoengineering, the Antarctic ozone hole would continue to form for another 30 to 70 years.

Geoengineering involves trying to reduce the amount of sunlight reaching Earth's surface to compensate for the additional long-wave infrared radiation from greenhouse gases, thereby reducing or reversing global warming (*6*). Even if it works, there are problems with this approach (*7*). If perceived to be a possible remedy for global warming, it would reduce societal pressure to reduce greenhouse gas emissions. It could reduce overall precipitation, particularly Asian and African summer monsoon rainfall, threatening the food supply of billions. It would allow continued ocean acidification, because some of the carbon dioxide humans put into the atmosphere continues to accumulate in the ocean. Weather modification could be used as a weapon (*8*), thus violating the 1977 U.N. Convention on the Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques. There would be rapid warming if geoengineering stopped suddenly. If geoengineering worked, whose hand would be on the thermostat? How could the world agree on an optimal climate?

Department of Environmental Sciences, Rutgers University, New Brunswick, NJ 08901, USA. E-mail: robock@envsci.rutgers.edu

Nevertheless, for some schemes, the benefits may outweigh the problems, especially if used on a temporary basis. To date, only some schemes have been investigated in detail. Furthermore, proponents of geoengineering, especially the fossil fuel industry, will continue to push for its use.

Sunshades in orbit around Earth (*9*) or cloud seeding to brighten them (*10*) have been proposed, but most geoengineering ideas focus on emulating explosive volcanic eruptions by injecting SO_2 or H_2S into the stratosphere, producing a sulfuric acid cloud to scatter solar radiation back to space and cool the planet. Deciding whether this is a good idea or not requires detailed analysis of the costs, benefits, and harm to the planet that such a strategy would entail, and comparison to the same metrics for mitigation and sequestration. Given the need for rapid mitigation, these ideas need rapid and thorough investigation.

It has been suggested (*3*, *4*) that the cooling of the global climate for a couple years after large volcanic eruptions—like the 1991 Mount Pinatubo eruption—serves as an innocuous model for what humans could do by creating a permanent stratospheric aerosol layer. However, volcanic eruptions actually serve as a warning about geoengineering: They produce drought (*11*), hazy skies, much less direct solar radiation for use as solar power, and ozone depletion (*12*).

We now have an ozone hole over Antarctica every spring because the polar stratospheric clouds that form there (see the figure) serve as surfaces for heterogeneous chemistry that releases chlorine, which then catalytically destroys ozone. Polar stratospheric clouds only form when the temperature falls below ~195 K, but additional sulfate aerosols provided by geoengineering or volcanic eruptions alter these temperature restrictions and provide more surface area for the chemistry, allowing more chlorine to be activated and more ozone to be destroyed.

Advocates of geoengineering suggest that this ozone problem would not be important, because the stratospheric concentration of chlorine is slowly decreasing as a result of global environmental agreements (*13*). However, Tilmes *et al.* show that even with the projected chlorine declines, ozone depletion (and increased ultraviolet flux) would be prolonged for decades by geoengineering of the stratospheric sulfate layer. In their model, the effects would occur every spring in the Southern Hemisphere and in most springs in the warmer Northern Hemisphere. The presence of sulfate aerosols would raise the temperature needed for chlorine activation over 200 K, expanding both vertically and horizontally the regions of polar ozone depletion.

A U.S. Department of Energy white paper (*14*) in October 2001 recommended a $13 million/year national geoengineering research effort, but the paper was never released. According to the paper, "any effort to deliberately moderate or ameliorate threats that may arise or become more likely as a result of climate change should be undertaken only in extraordinary circumstances.... In view of the risk of significant consequences to society and the environment from either inaction or

poorly understood actions, research should be initiated now to examine possible options to moderate adverse climate threats; to ensure that these options are effective, affordable, reversible and sustainable."

It is not too late to make up for lost time, but further delay must be avoided. A research program, more generously funded than that proposed in 2001, supported by the U.S. federal government with international cooperation, will allow us to compare the efficacy, costs, and consequences of the various options of responding to global warming—mitigation, sequestration, geoengineering, or doing nothing—so that an informed public can agree on the best courses of action.

References and Notes

1. IPCC, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon *et al.*, Eds. (Cambridge Univ. Press, Cambridge, UK, and New York, NY, 2007).
2. IPCC, *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, L. A. Meyer, Eds. (Cambridge Univ. Press, Cambridge, UK, 2007).
3. P. J. Crutzen, *Climatic Change* **77**, 211 (2006).
4. T. M. L. Wigley, *Science* **314**, 452 (2006).
5. S. Tilmes, R. Müller, R. Salawitch, *Science* **320**, 1201 (2008).
6. I use "geoengineering" to refer to schemes designed to reduce solar radiation input to the climate system; I exclude the broader meaning that includes sequestration of atmospheric carbon dioxide, for example, by iron fertilization of the oceans [an idea that has been shown to be premature (15)], afforestation, and reforestation.
7. A. Robock, *Bull. Atomic Scientists* **64(2)**, 14 (2008).
8. J. R. Fleming, *Wilson Q.* **2007**, 46 (spring 2007).
9. R. Angel, *Proc. Nat. Acad. Sci. U.S.A.* **103**, 17184 (2006).
10. K. Bower, T. Choularton, J. Latham, J. Sahraei, S. Salter, *Atm. Res.* **82**, 328 (2006).
11. K. Trenberth, A. Dai, *Geophys. Res. Lett.* **34**, L15702, 10.1029/2007GL030524 (2007).
12. S. Solomon, *Rev. Geophys.* **37**, 275 (1999).
13. L. Lane, K. Caldeira, R. Chatfield, S. Langhoff, Eds., *Workshop Report on Managing Solar Radiation, NASA/CP–2007-214558* (NASA, Ames Research Center, Moffett Field, CA, 2007).
14. E. Khan *et al.*, *Response Options to Limit Rapid or Severe Climate Change* (Department of Energy, Washington, DC, 2001).
15. K. O. Buesseler *et al.*, *Science* **319**, 162 (2008).

16. I thank R. Salawitch, S. Tilmes, G. Stenchikov, and A. Marquardt for valuable suggestions. Supported by NSF grant ATM-0730452.

10.1126/science.1159280

ASTRONOMY

A Blast from the Past

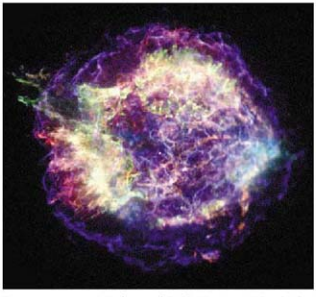
Echoes of light, reflections from nearby gas and dust clouds, can be used to reconstruct past astronomical events.

Andrew C. Fabian

Have you ever wanted to view an event that happened many years ago? Most of the light from that event is still traveling through space and can, in principle, be reflected back to us to reconstruct the event. This is, of course, completely impractical for events that occur on a human scale, but when a star explodes as a supernova, so much light is emitted that it may be possible to see a delayed reflection from surrounding dust clouds. On page 1195 of this issue, Krause *et al.* (*1*) report their observations of a light echo for the outburst of Cassiopeia A (Cas A), which is the most recent nearby supernova known to have occurred in our Galaxy.

The remnant of Cas A was first discovered in 1947 and identified optically in 1950. From its observed expansion, it can be deduced that the explosion itself would have occurred around 1680, as viewed from Earth. A recent x-ray image of the remnant is shown in the figure.

More recently, infrared images made with the Spitzer Space Telescope revealed moving light echoes around Cas A 4 years ago (*2*). These echoes were monitored last year with the Calar Alto optical telescope in Spain, and a spectrum of a particularly bright patch was taken by the Subaru telescope in Hawaii. The echo spectrum clearly shows light from the supernova. When a star of 10 to 20 solar masses explodes, an energy equivalent to about 1% of the mass of the Sun is turned into kinetic energy of the stellar envelope, which then expands into space at velocities of 10,000 km/s or more. The spectrum shows emission and absorption lines Doppler-broadened by such large velocities. The presence of hydrogen lines in the spectrum places it in the category of a type II supernova, which results from collapse of the core of a massive star when it runs out of fuel, as was long suspected from the properties of the still-expanding remnant. The spectrum is remarkably similar to that of supernova 1993J (SN 1993J), a type IIb supernova seen (in 1993) in the nearby galaxy M81.



Supernova remnant. An image of the Cas A remnant taken by the Chandra X-ray Observatory (CXC).

CREDIT: NASA/CXC/MIT/UNIV. OF MASSACHUSETTS AMHERST/M. D. STAGE ET AL.

Light echoes also have recently been seen from SN 1993J (*3*), and from other supernovae in our satellite galaxy, the Large Magellanic Cloud (*4*), including the famous SN 1987A (*5*), which is the only supernova to have been seen with the naked eye since the invention of the telescope more than 400 years ago. Van den Bergh (*6*) in 1966 had tried to look for an echo around Cas A. However, we now know that it was much too faint to be seen with the photographic plates available at that time.

The light echo spectrum from Cas A is notable primarily because Cas A is a type IIb supernova and its remnant has been so well studied due to its proximity and youth. We can assume (*7*) that Cas A was a red giant before it exploded, and that it probably had a binary companion at some stage. The progenitor of SN 1993J was predicted to have been a member of a binary, and a massive star consistent with a companion remains at the site (*8*). There is no such stellar companion remaining at the position of Cas A, so it possibly spiraled into the progenitor some time before the explosion. A faint non-variable pointlike x-ray source has been found (*9*) close to the center of the remnant and is probably a neutron star.

Institute of Astronomy, Cambridge University, Madingley Road, Cambridge CB3 0HA, UK. E-mail: acf@ast.cam.ac.uk

A nagging question is why no one saw the initial explosion of Cas A here on Earth. It would have appeared brighter than the full Moon at the reflecting dust clouds 300 light-years from the explosion. At 10,000 light-years away on Earth, it would have been just brighter than Venus except for the attenuation by a factor of about 30 produced by further dust clouds along our line of sight. This diminished brightness places it in about the third magnitude, or like some of the fainter stars familiar to the naked eye. A star was mapped by the astronomer John Flamsteed in the right place in 1680, but there has been suspicion that this was an error. It seems more plausible now that he did see the explosion of Cas A, but did not register anything special about it.

The dust that is needed to make reflections possible also dims them. The interstellar medium must be crisscrossed with light echoes of dramatic past events from novae and supernovae. Any bright time-varying source can accomplish this, and it need not be optical radiation: X-ray echoes also have been seen from gamma-ray bursts (*10*) and our Galactic center (*11*), showing that it was likely active just 100 years ago. Following the observations by Krause *et al.* of the echo from Cas A, one can now speculate on how the past history of events in our Galaxy, and others, can in the future be reconstructed. At some level, the signals from these events must become confused with each other, but it does offer some fascinating possibilities. We now have the ultimate action replay. The explosion of Cas A, for example, can be seen again and again, even from different angles, over the next few hundred years.

References

1. O. Krause *et al.*, *Science* **320**, 1195 (2008).
2. O. Krause *et al.*, *Science* **308**, 1604 (2005).
3. J.-T. Liu, J. N. Bregman, P. Seitzer, *Astrophys. J.* **582**, 919 (2003).
4. A. Rest *et al.*, http://arxiv.org/abs/0801.4762 (2008).
5. N. B. Suntzeff *et al.*, *Nature* **334**, 135 (1988).
6. S. van den Bergh, *Publ. Astron. Soc. Pac.* **78**, 74 (1966).
7. P. Young *et al.*, *Astrophys. J.* **640**, 891 (2006).
8. J. R. Maund *et al.*, *Nature* **427**, 129 (2004).
9. G. G. Pavlov *et al.*, *Astrophys. J.* **531**, L53 (2000).
10. S. Vaughan *et al.*, *Astrophys. J.* **639**, 323 (2006).
11. K. Koyama *et al.*, *Publ. Astron. Soc. Jpn.* **60**, 201 (2008).

10.1126/science.1158538

IMMUNOLOGY

Immunity Benefits from a Little Suppression

Regulatory T cells thought to limit the extent of an inflammatory response to viral infection can actually optimize the response by redirecting it to sites of infection.

George Kassiotis and Anne O'Garra

Successful resolution of an infection requires a highly coordinated immune response. Dendritic cells of the immune system survey tissues for signs of microbial infection. When an invader is found, innate immune cells are alerted and local dendritic cells capture microbial antigens at the site of infection and transport them to the draining lymph nodes. There, dendritic cells stimulate the proliferation and activation of effector T cells that specifically recognize the foreign antigen and fight the infection. However, regulatory T cells—a subset whose primary task is to prevent catastrophic immune responses to self antigens—can pose a considerable obstacle to initiating immune responses to infection. On page 1220 of this issue, Lund *et al.* (*1*) report their surprising finding that induction of the early antiviral immune response at a site of infection is actually promoted by regulatory T cells.

The existence of "suppressor" T cells that halt immune responses was suggested soon after the discovery of T cells, but the lack of robust experimental evidence reversed early enthusiasm and by the mid-1990s, the term had gained notoriety. Interest in immune-suppressive T cells was renewed by the description of a distinct regulatory T cell population (*2*). This was eventually followed by the landmark discovery of the transcription factor Foxp3, which is indispensable for regulatory T cell development and function (*3*–*5*). Mutations in *Foxp3* result in fatal multiorgan inflammatory disease in both humans and mice (*6*–*8*), due to failure to generate regulatory T cells. This highlights their essential role in preventing lethal autoimmunity.

However, the involvement of regulatory T cells in the immune response to infection has been much less clear. Suppression of an excessive immune response is desirable, as it limits the pathological consequences of infection. But suppression of the early response to infection is harmful, as it prevents pathogen control. Lund *et al.* have discovered an unexpected benefit of immune suppression by regulatory T cells during viral infection.

Lund *et al.* show in mice that regulatory T cells increase in number and activation during vaginal infection with herpes simplex virus. Moreover, accumulation, activation, and proliferation of these T cells, both at the site of infection and in the draining lymph nodes, are similar to those of effector T cells. This supports recent observations that the expansion and contraction kinetics of Treg and effector T cells are synchronized, such that their ratio remains relatively constant (*9*). Thus, although regulatory T cells were long considered unresponsive with respect to proliferation, they not only can suppress immune responses but also proliferate extensively in response to immunization or infection.

These observations raise the question of what regulatory T cells actually respond to. Similar to effector T cells, regulatory T cells could respond to virus-derived antigens. However, regulatory T cells and "conventional" T cells differ in their T cell antigen receptor repertoire and diversity (*10*, *11*), so whether regulatory T cells are equipped to respond to foreign antigens requires further investigation. Coincident expansion of regulatory and effector T cells could also be explained by responsiveness of the former to cytokines (interleukin-2) produced by effector T cells (*12*).

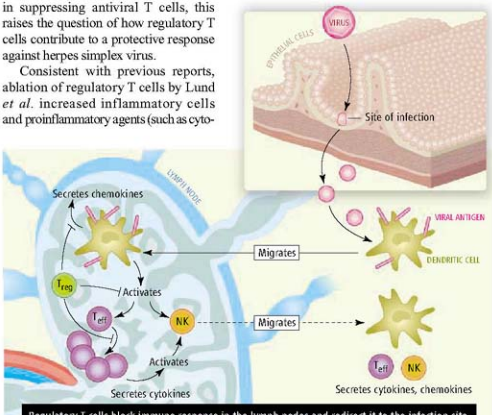
Although the precise signals that activate regulatory T cells during infection remain unclear, Lund *et al.* demonstrate their requirement in early protection against herpes simplex virus. The authors took advantage of the universal expression of Foxp3 in regulatory T cells to selectively kill them. Injection of diphtheria toxin in mice expressing toxin receptor (driven by Foxp3) resulted in complete elimination of regulatory T cells expressing the transcription factor ($Foxp3^+$). Surprisingly, this rendered mice more susceptible to replication of herpes simplex virus in vaginal epithelial cells. The infection eventually

Division of Immunoregulation, MRC National Institute for Medical Research, The Ridgeway, London NW7 1AA, UK. E-mail: aogarra@nimr.mrc.ac.uk; gkassio@nimr.mrc.ac.uk

spread to the central nervous system, resulting in paralysis and death. Lund *et al.* also show that increased susceptibility of mice to viral infection is due exclusively to ablation of $Foxp3^+$ regulatory T cells and not to inadvertent ablation of cells belonging to any other lineage, an important control to be considered in future studies with specific ablation of $Foxp3^+$ cells. Given their described role in suppressing antiviral T cells, this raises the question of how regulatory T cells contribute to a protective response against herpes simplex virus.

Consistent with previous reports, ablation of regulatory T cells by Lund *et al.* increased inflammatory cells and proinflammatory agents (such as cytokines and chemokines) in lymph nodes. The authors show that ablation of regulatory T cells triggers the release of chemokines (CCL2, CXCL9, and CXCL10) from dendritic cells, natural killer cells (a type of innate immune cell), and stromal cells in the lymph node. These chemokines further attract circulating dendritic cells, natural killer cells, and T cells to the lymph node, thus amplifying the immune response. The end result is elevated production of type I and type II interferons—cytokines that by block viral replication within host cells and boost the immune response—in the lymph node.

Although this response is a powerful weapon against herpes simplex virus, it is happening in the wrong place. Despite, or perhaps because of, the inflammatory response in lymph nodes, ablation of regulatory T cells profoundly reduces or delays recruitment of inflammatory dendritic cells, natural killer cells, and T cells to the site of infection. Failure to attract inflammatory cells to the vaginal epithelium is consistent with a lack of local induction of chemokine production (most importantly CCL5, which plays an active role in recruiting immune cells to the site of infection) and the subsequent absence of antiviral interferon production at the infection site. This provides the virus with enough time to replicate locally and spread, thus accelerating the infection.



Guiding the early immune response. Immune cell recruitment either to the site of infection or to the draining lymph node differs depending on regulatory T cell function. The presence of regulatory T cells in the lymph node suppresses proinflammatory cytokine and chemokine production and prevents constitutive influx and activation of effector T cells (T_{eff}), natural killer (NK) cells, and dendritic cells. This allows migration of immune cells out of the lymph node and into the site of virus infection, where an effective response is then initiated. In the absence of regulatory T cells, immune cells are trapped in the inflamed lymph node, resulting in their reduced migration into the infection site, which allows virus replication and spread (not shown).

Given the complexity of cellular and molecular changes caused by the absence of regulatory T cells in the lymph nodes during viral infection, it is important to define the targets of regulatory T cell–mediated suppression. Persistent ablation of these T cells in adult mice results in lymphoproliferative disease with expansion of lymphocytes and myeloid cells in the lymphoid organs (*13*). Dendritic cell expansion is thought to arise as a result of excessive cytokine production by dysregulated effector T cells. However, Lund *et al.* show that in contrast to lymphoproliferation, elevated chemokine production by lymph node dendritic cells and stromal cells is a direct result of a deficiency in regulatory T cells. This reveals a surprising role for these T cells in controlling the chemotactic environment that determines the migration of immune cells between organized lymphoid tissues and the site of infection where they need to act. It also raises the question of what the chemokine levels are in mice without regulatory T cells.

By suppressing the inflammatory response in lymph nodes, regulatory T cells direct immune cells to the site of infection, thereby promoting the antiviral response where it is most needed. When interpreting the results of Lund *et al.*, it is important to keep in mind that regulatory T cells do not simply suppress the lymph node reaction to infection. They suppress the constitutive activation of T cells, natural killer cells, and dendritic cells by signals that are yet to be defined. Indeed, changes in proinflammatory chemokine and cytokine production and accumulation of immune cells caused by ablation of regulatory T cells in the lymph node occur independently of the infection. Thus, the inflammatory lymph node environment generated by complete loss of regulatory T cells disrupts the orchestration of an early antiviral response at the site of infection. It will be important, however, to establish whether or not more physiological fluctuations in regulatory T cell number and function—such as those observed in the genetically diverse human population, during infection or with aging of the immune system—affect the antiviral response in the same way.

References

1. J. M. Lund, L. Hsing, T. T. Pham, A. Y. Rudensky, *Science* **320**, 1220 (2008); published online 24 April 2008 (10.1126/science.1155209).
2. S. Sakaguchi *et al.*, *J. Immunol.* **155**, 1151 (1995).
3. S. Hori, T. Nomura, S. Sakaguchi, *Science* **299**, 1057 (2003); published online 9 January 2003 (10.1126/science.1079490).
4. J. D. Fontenot, M. A. Gavin, A. Y. Rudensky, *Nat. Immunol.* **4**, 330 (2003).
5. R. Khattri, T. Cox, S. A. Yasayko, F. Ramsdell, *Nat. Immunol.* **4**, 337 (2003).
6. M. E. Brunkow *et al.*, *Nat. Genet.* **27**, 68 (2001).
7. R. S. Wildin *et al.*, *Nat. Genet.* **27**, 18 (2001).
8. C. L. Bennett *et al.*, *Nat. Genet.* **27**, 20 (2001).
9. D. Haribhai *et al.*, *J. Immunol.* **178**, 2961 (2007).
10. C. S. Hsieh *et al.*, *Immunity* **21**, 267 (2004).
11. R. Pacholczyk, H. Ignatowicz, P. Kraj, L. Ignatowicz, *Immunity* **25**, 249 (2006).
12. A. R. M. Almeida, B. Zaragoza, A. A. Freitas, *J. Immunol.* **177**, 192 (2006).
13. J. M. Kim, J. P. Rasmussen, A. Y. Rudensky, *Nat. Immunol.* **8**, 191 (2007).

10.1126/science.1159090

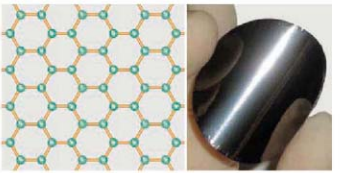
MATERIALS SCIENCE

Graphene-Based Materials

Advances in synthesizing graphene offer opportunities for making novel materials for nanoelectronics and many other applications.

Dan Li[1] and Richard B. Kaner[2]

Graphenes—monolayers of carbon atoms arranged in a honeycomb network (see the figure, left panel)—are prevalent in many carbon-based materials. For example, graphite can be considered as stacks of graphene layers. Although the weak interactions that hold the graphene sheets together in graphite allow them to slide readily over each other, the numerous weak bonds make it difficult to separate the sheets. When physicists recently managed to isolate individual graphene sheets from mechanically cleaved graphite, they discovered unusual electronic properties arising from confinement of electrons in two dimensions (*1*). Researchers are now making rapid progress toward creating processable graphene.



Simple structure, great potential. (**Left**) In graphene, carbon atoms (green dots) are bonded together through sp^2 hybridization (orange lines). (**Right**) Shiny and flexible graphene paper is formed by controlled restacking of graphene sheets (*8*).

Apart from its electronic properties, graphene displays several other unusual attributes. Graphene is a giant aromatic macromolecule that conducts both electricity and heat well in two dimensions. The theoretical specific surface area of individual graphene sheets (*2*) is more than double that of the finely divided activated carbon used in water purification. Their mechanical strength is comparable to that of carbon nanotubes—which can be considered graphene with a twist. Furthermore, recent studies (*2*–*9*) suggest that the production cost of graphene sheets in large quantities could be much lower than that of carbon nanotubes. Graphene sheets are thus attractive as atomically thin yet robust components for nanoelectronic and nanoelectromechanical devices (*1*, *10*) and as nanoscale building blocks for new materials.

To create useful graphene-based materials, graphene sheets must be available in large quantities. Approaches for producing graphene sheets include mechanical cleavage of graphite (*1*), epitaxial growth (*11*), bottom-up organic synthesis (*12*), and chemical exfoliation of graphite (*2*–*9*). Given that graphite is relatively inexpensive and available in large quantities, the top-down method—splitting graphite into individual sheets mechanically or chemically—has received the most attention with respect to large-scale production of graphene.

High-yield direct exfoliation of graphite, however, has so far remained a challenge. Many researchers are now focusing on derivatives of graphite, especially graphite oxide (*2*–*9*). The latter is hydrophilic and has a larger interlayer distance than graphite; it can readily exfoliate into individual graphene oxide sheets in water and forms stable dispersions after ultrasonication. Subsequent deoxygenation via chemical reduction can restore electrically insulating graphene oxide to conductive graphene (*3*–*5*). Rapid thermal expansion can also cause graphite oxide to delaminate, but many of the resultant sheets are crumpled and wrinkled (*2*).

In all synthetic routes, keeping the graphene sheets individually separated is the most important and challenging part. Bulk graphene sheets—if left unprotected—will spontaneously agglomerate and even restack to form graphite. Chemical functionalization or the use of dispersants is generally needed to prevent agglomeration (*4*–*7*). The graphite oxide synthetic route is especially attractive for stabilizing individual sheets in solution. The oxygen-containing groups that exist in graphite oxide (carboxyl and hydroxyl) provide reactive sites for chemical modification using known carbon surface chemistry. Chemical attachment of appropriate organic groups onto graphite oxide surfaces not only leads to physical separation of the resultant graphene sheets, but also makes it possible to directly form stable graphene dispersions during the synthetic process (*5*, *7*), something that has been difficult to accomplish with carbon nanotubes.

Chemical conversion from graphene oxide leaves residual oxygen-containing groups and makes the resulting graphene surfaces negatively charged when dispersed in water. By controlling the colloid chemistry, graphene sheets can form stable aqueous colloids via electrostatic repulsion without the need for foreign polymeric or surfactant dispersants (*8*). By alternately immersing a substrate into a negatively charged graphene colloid and a positively charged polycation solution, graphene sheets can be integrated with other functional materials at a molecular/nanometer scale to create multifunctional graphene-based composites (*8*).

Successful dispersion of graphene enables the use of low-cost solution processing techniques to fabricate various potentially useful graphene-based materials. For example, graphene films can be made by drop-casting, spraying, electrostatic adsorption, filtration, and dip- or spin-coating of graphene or precursor dispersions (*6*–*9*, *13*–*17*). Chemically functionalized graphene can be readily mixed with other polymers in solution, producing a new class of electrically conductive nanocomposites at relatively low cost (*5*).

Dispersion of graphene in solution facilitates assembly of the atomically thin sheets into highly ordered macroscopic architectures that are otherwise difficult to make. Ruoff and co-workers have shown that, under a directional flow induced by vacuum filtration, graphene oxide sheets in solution can interlock in a near-parallel manner, yielding a paper-like material that is remarkably stiff and strong—much stronger than its nanotube counterpart (*13*). Analogously, we have shown that filtration of chemically reduced graphene colloids readily yields ultrastrong, electrically conductive, thermally stable graphene paper (see the figure, right panel) (*8*).

Graphene sheets can be further modified by engineering their shape, size, and chemical structure. For example, Dai and co-workers have used controlled ultrasonication of thermally expanded graphite to produce graphene

PHOTO CREDIT: ALLAN DU AND DAN LI/UNIVERSITY OF WOLLONGONG

[1]ARC Centre of Excellence for Electromaterials Science, Intelligent Polymer Research Institute, University of Wollongong, Wollongong, NSW 2522, Australia. E-mail: danli@uow.edu.au [2]Department of Chemistry and Biochemistry, Department of Materials Science and Engineering, and California NanoSystems Institute, University of California, Los Angeles, CA 90095, USA. E-mail: kaner@chem.ucla.edu

nanoribbons less than 10 nm wide. Transistors based on these nanoribbons exhibit high on/off ratios (*18*).

Facile synthesis and solution-processing of chemically derived graphene facilitates not only the fabrication of graphene-based nanoelectronic devices (*14*, *15*, *18*), but also the exploration of its use in many other areas—for example, as transparent electrodes for solar cells (*16*). Nonetheless, research toward the application of graphene-based materials has just begun. Many challenges and opportunities remain. For example, to enable applications in batteries and supercapacitors, in separation technologies, and as supports for catalysts, the hierarchical structure of graphene assemblies must be controlled to make the surface of individual sheets maximally accessible. Additionally, graphene synthesis through chemical conversion from graphite introduces a considerable amount of defects, reducing electrical conductivity. New strategies are needed to produce more conducting yet processable graphene.

References

1. A. K. Geim, K. S. Novoselov, *Nat. Mater.* **6**, 183 (2007).
2. M. J. Allen, V. C. Tung, R. B. Kaner, *Chem. Rev.* **110**, 132 (2010).
3. S. Stankovich *et al.*, *Nature* **442**, 282 (2006).
4. S. Stankovich *et al.*, *Carbon* **45**, 1558 (2007).
5. S. Stankovich *et al.*, *J. Mater. Chem.* **16**, 155 (2006).
6. S. Niyogi *et al.*, *J. Am. Chem. Soc.* **128**, 7720 (2006).
7. Y. Xu, H. Bai, G. Lu, C. Li, G. Shi, *J. Am. Chem. Soc.* **130**, 5856 (2008).
8. D. Li, M. B. Mueller, S. Gilje, R. B. Kaner, G. G. Wallace, *Nat. Nanotechnol.* **3**, 101 (2008).
9. G. Eda, G. Fanchini, M. Chhowalla, *Nat. Nanotechnol.* **3**, 270 (2008).
10. J. S. Bunch *et al.*, *Science* **315**, 490 (2007).
11. C. Berger *et al.*, *Science* **312**, 1191 (2006); published online 12 April 2006 (10.1126/science.1125925).
12. J. S. Wu, W. Pisula, K. Mullen, *Chem. Rev.* **107**, 718 (2007).
13. D. A. Dikin *et al.*, *Nature* **448**, 457 (2007).
14. S. Gilje, S. Han, M. Wang, K. L. Wang, R. B. Kaner, *Nano Lett.* **7**, 3394 (2007).
15. C. Gomez-Navarro *et al.*, *Nano Lett.* **7**, 3499 (2007).
16. X. Wang, L. J. Zhi, K. Mullen, *Nano Lett.* **8**, 323 (2008).
17. H. A. Becerril *et al.*, *ACS Nano* **2**, 463 (2008).
18. X. Li, X. Wang, L. Zhang, S. Lee, H. Dai, *Science* **319**, 1229 (2008); published online 23 January 2008 (10.1126/science.1150878).

10.1126/science.1158180

PLANETARY SCIENCE

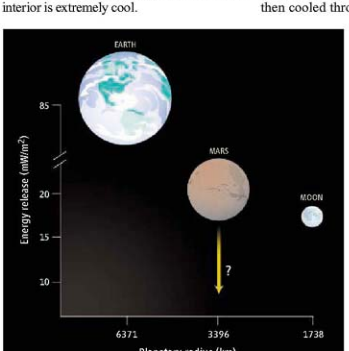
Is Mars Geodynamically Dead?

The observation that Mars' northern polar cap barely deforms its crust implies that its planetary interior is colder than expected.

Matthias Grott

The temperature of a planet's interior is one of the key factors controlling the amount of geodynamic activity at its surface. Planetary scientists can estimate interior temperatures if they can assess the deformability of the planetary surface, which is greater when interior temperatures are higher (warm rocks will more easily deform than cold ones). Mass loading by glaciers and mountains deflects the surface, and this deformation can be used to estimate the elasticity of the surface layer when the load was emplaced. On Mars, the polar caps are a geologically recent load. On page 1182 of this issue, Phillips *et al.* (*1*) report that the amount of deflection caused by the polar caps leads to unexpected constraints on the current thermal state of the planet.

To investigate the ice-hidden planetary surface, the authors used radar-sounding data of the northern polar cap obtained by SHARAD, the shallow-radar instrument onboard the Mars Reconnaissance Orbiter, which has orbited Mars since 2006. The radar-wave travel times can be converted into distances by assuming that the polar cap consists of water ice and dust. The vertical deformation of the surface was extremely small—less than 100 m—and requires that the surface be very stiff to support these loads. The elastic lithosphere thickness, a measure of the surface's resistance to flexure, needs to be greater than 300 km at the poles today. This value is almost twice as great as previously estimated from other measures of deformation and theoretical considerations (*2*, *3*). Because the stiffness of the rocks is connected to temperature, Phillips *et al.*'s findings imply that the martian interior is extremely cool.



A colder Mars. The energy release at a planetary surface is a measure for the temperatures in the planetary interior. Large planets, which have a lower surface-to-volume ratio, should be warmer than small ones if they contain equal concentrations of heat-producing elements. The average heat flow of Earth is 87 mW/m², and the lunar heat flow determined by the Apollo experiments is 18 mW/m² (*9*). Mars was expected to release 20 mW/m², but the value inferred from Phillips *et al.*'s findings would be much lower.

This result is surprising. First, the temperatures in the interior of terrestrial planets should be proportional to their radius if they started with the same amount and distribution of radioactive, heat-producing elements and then cooled through surface losses (see the figure). In this case, Mars would be expected to plot between Earth and the Moon. However, the new estimates imply that the martian heat flow, a measure for the temperatures in the planetary interior, is below even that of the Moon, even though Mars is about twice the diameter. Either there is a large degree of compositional heterogeneity among the terrestrial planets, or heat transport is very inefficient inside Mars.

Second, theoretical considerations of the planet's thermal evolution have unanimously concluded that Mars should have retained more internal heat than is consistent with the observed stiff lithosphere (*4*, *5*). This discrepancy could be resolved if the amount of radioactive elements in Mars (potassium, thorium, and uranium) were smaller

Institute of Planetary Research, German Aerospace Center (DLR), Rutherfordstraße 2, 12489 Berlin, Germany. E-mail: matthias.grott@dlr.de

than previously assumed by 20 to 30%. However, this conclusion is difficult to reconcile with the composition of the so-called SNC meteorites believed to have originated from Mars (*6*). They hold heat-producing elements in concentrations that would result in heating rates and interior temperatures much too high to be consistent with the new data.

A third problem with a cold martian interior is that low temperatures limit the partial melting of rocks in the upper mantle, which reduces the amount of magma that can be erupted onto the surface. However, Mars shows recent volcanic activity (*7*) that, although sparse, requires some degree of partial melting in the recent geologic history of the planet. Whether this is feasible if interior temperatures are as low as required by the new observations seems questionable. Also, mechanisms that lead to surface eruptions of magmas derived from partial-melt zones much deeper than 300 km are hard to imagine. For such deep reservoirs, intrusive volcanism—the injection of magma into the crust rather than onto the surface—appears to be the more plausible explanation.

These problems could be resolved if subsurface temperatures on Mars were highly spatially variable; the estimate obtained by Phillips *et al.* would then not be representative for the bulk of the planet. However, data from the gamma-ray spectrometer onboard the Mars Odyssey spacecraft indicate that lateral variability of heat-producing elements on the surface is low (*8*), and the absence of plate tectonics on Mars casts doubt on this hypothesis.

To solve these puzzles, direct evidence is needed, which requires drilling and in situ temperature logging. Apart from Earth, such measurements have been performed successfully only during the Apollo 15 and 17 missions to the Moon (*9*). The next step forward will be in situ geophysical experiments, which are scheduled for launch with the European Space Agency's ExoMars mission in 2013. In particular, the seismometer and heat flow probe, which will conduct Apollo-like experiments, will provide ground truth for the degree of Mars' geodynamic activity. Until then, the interesting data reported by Phillips *et al.* will cause some controversy and stimulate activity as the planetary science community tries to fit their results into our current view of Mars.

References

1. R. J. Phillips *et al.*, *Science* **320**, 1182 (2008); published online 15 May 2008 (10.1126/science.1157546).
2. M. Grott, D. Breuer, *Icarus* **193**, 503 (2008).
3. M. A. Wieczorek, *Icarus*, 10.1016/j.icarus.2007.10.026 (2008).
4. S. Schumacher, D. Breuer, *J. Geophys. Res.* **111**, E02006 (2006).
5. S. A. Hauck, R. J. Phillips, *J. Geophys. Res.* **107**, 10.1029/2001JE001801 (2002).
6. H. Wänke, G. Dreibus, *Philos. Trans. R. Soc. London Ser. A* **349**, 285 (1994).
7. G. Neukum *et al.*, *Nature* **432**, 971 (2004).
8. G. J. Taylor *et al.*, *J. Geophys. Res.* **111**, E03S10 (2006).
9. M. Langseth *et al.*, *Proc. Lunar Sci. Conf.* **7**, 3143 (1976).

10.1126/science.1159365

MEDICINE

Activating a Repressor

Sonia Cohen, Zhaolan Zhou, Michael E. Greenberg

A factor implicated in autism spectrum disorders can both suppress and activate the expression of genes involved in neurodevelopment.

Autism spectrum disorders are characterized by abnormal social behavior, deficits in communication, and repetitive behavior. A diverse array of genetic mutations have been linked to autism (*1*), suggesting that what these behaviorally defined disorders share is not a single genetic lesion, but a disruption of biological processes important in cognitive development. As a result, attempts to understand the etiology of autism spectrum disorders have focused on the neurobiology of related syndromes caused by single-gene mutations that are more easily modeled in the laboratory. One example is mutation of the gene encoding methyl-CpG binding protein 2 (MeCP2), which gives rise to a spectrum of related but distinct postnatal neurodevelopmental disorders including Rett syndrome and autism spectrum disorders (*2*). Initially identified on the basis of its ability to bind methylated DNA (*3*), MeCP2 was thought to repress transcription. Mutations disrupting MeCP2 function were thus expected to increase gene expression, perturb neuronal function, and give rise to behavioral disorders such as Rett syndrome. However, the search for genes repressed by MeCP2 has so far identified only a few convincing targets. On page 1224 in this issue, Chahrour *et al.* (*4*) demonstrate that MeCP2 may play a more complex role, coordinating either transcriptional repression or activation, depending on the molecular context.

To investigate the effect of *MeCP2* gene dosage on transcriptional regulation, Chahrour *et al.* analyzed gene expression in the hypothalamus of *MeCP2*-deficient, wild-type, and *MeCP2*-overexpressing mice. This approach revealed small but important changes in the expression of thousands of genes, many of which are affected in opposite ways with deletion or overexpression of *MeCP2*. Although these data are consistent with previous studies indicating that many genes may be subtly misregulated in the brains of *MeCP2*-deficient animals (*5*), the ability of Chahrour *et al.* to correlate changes in the expression of particular genes with the level of MeCP2 expression convincingly links these transcriptional changes to the presence or absence of MeCP2. Surprisingly, most of the genes identified in these experiments showed increased expression in the hypothalamus of animals overexpressing MeCP2, and decreased expression in the absence of MeCP2. The study reveals a large number of new genes whose expression depends on MeCP2, and suggests a role for the factor in activating transcription.

A fundamental step toward understanding how mutations in *MeCP2* disrupt neuronal development and function is to identify the mechanisms by which MeCP2 regulates transcription. One possibility is that MeCP2 functions as a classical transcription factor that binds directly to regulatory sequences within the promoters of target genes. Repression or, as suggested by Chahrour *et al.*, activation of the targeted genes would depend on the interaction of MeCP2 with molecular complexes that promote or inhibit transcription (see the figure). Alternatively, or perhaps additionally, MeCP2 may play a structural role in the higher-order organization of chromatin architecture (*6*, *7*). Transcriptional repression and activation of individual genes would still be controlled by classical sequence-specific transcription factors, but MeCP2 might influence the chromatin landscape, thus dictating whether regions of genomic DNA are silenced or expressed. Chahrour *et al.* show that MeCP2 is indeed enriched at the promoters of

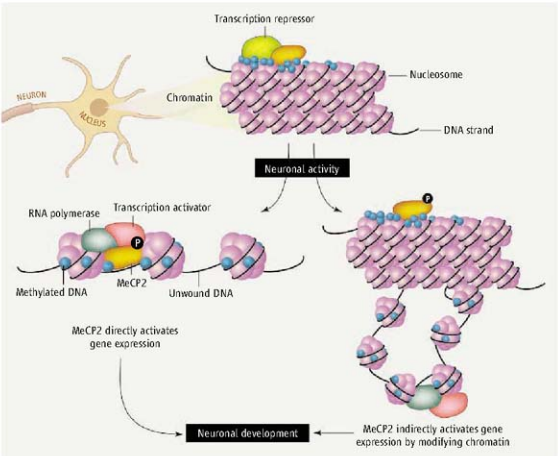
F. M. Kirby Neurobiology Center, Children's Hospital Boston and Departments of Neurology and Neurobiology, Harvard Medical School, 300 Longwood Avenue, Boston, MA 02115, USA. E-mail: michael.greenberg@childrens.harvard.edu

several genes that are either activated or repressed by MeCP2 in the hypothalamus. These data are consistent with the possibility that MeCP2 directly regulates the expression of at least a subset of the genes identified. However, because MeCP2 is highly expressed in the nucleus and likely binds widely within the genome, further approaches are needed to address whether binding of MeCP2 to specific promoters determines the expression of the targeted genes.

The identification of thousands of genes that are either activated or repressed by MeCP2 raises the intriguing problem of how the transcriptional outcome is determined for a particular MeCP2 target gene. Chahrour *et al.* approach this question in two ways. First, using mass spectrometry, they identify the transcriptional activator cAMP response element–binding protein (CREB) on chromatin that is associated with MeCP2. Using chromatin immunoprecipitation, the authors verify that CREB is present at the promoter of a gene activated by MeCP2 in the hypothalamus, and is absent from the promoter of a gene repressed by MeCP2. These data suggest that interactions with regulatory complexes present at the promoters of genes bound by MeCP2 may determine whether MeCP2 activates or represses transcription. The authors also find that the promoters of genes activated by MeCP2 are enriched for CpG islands (short regions of DNA enriched in the nucleotides cytosine and guanine) as compared with those genes repressed by MeCP2. Surprisingly, although MeCP2 binds to a similar extent to the regulatory regions of both types of targeted genes, the activated promoters are lightly methylated, whereas the promoters of genes repressed by MeCP2 are more heavily methylated. These data suggest that the function of MeCP2 as an activator or a repressor, rather than the recruitment of MeCP2, may depend on the DNA methylation status of the relevant genomic region.

Consistent with the idea that MeCP2 functions as a context-specific transcriptional modulator, both MeCP2 and CREB are phosphorylated in response to neuronal activation, suggesting that dynamic regulation of these two factors may control activity-dependent gene expression (*8*, *9*). Activity-dependent transcription coordinates the processes of postnatal synaptic development and plasticity that may be disrupted in disorders of cognitive function (*10*, *11*). Indeed, MeCP2-deficient neurons form fewer excitatory synapses than wild-type neurons, whereas MeCP2-overexpressing neurons form more excitatory synapses (*12*), suggesting that activity-dependent processes of neuronal development are regulated by MeCP2. If so, the products of the genes identified by Chahrour *et al.* may function in aspects of late neuronal development that depend on neuronal activity, such as synapse formation. Alternatively, genetic mutations that alter MeCP2 protein abundance or function might first affect synapse number through a distinct mechanism, and the broad transcriptional changes seen in the MeCP2-deficient or -overexpressing animals may instead reflect an indirect effect secondary to perturbation of neuronal function. Future work that links the binding of MeCP2 to specific sequences of DNA in regulating particular genes will be required to distinguish between direct and indirect effects on gene expression.

The finding by Chahrour *et al.* that MeCP2 may have a role beyond simple transcriptional repression comes at a time when the dynamic nature of epigenetic modifications of the genome, including DNA methylation, is being newly appreciated (*13*, *14*). Deeper understanding of the modulatory role of MeCP2 on gene expression will help elucidate the molecular nature of the neuronal dysfunction that underlies neurodevelopmental disorders such as Rett syndrome and other autism spectrum disorders.



A choice for MeCP2. Whether MeCP2 functions as a promoter-specific transcription factor or is involved in coordinating higher-order chromatin structure, both the cellular and molecular context may influence how it modulates transcriptional repression or activation. Neuronal activity may play a central role in determining the transcriptional function of MeCP2.

References and Notes

1. J. Sebat *et al.*, *Science* **316**, 445 (2007).
2. M. Chahrour, H. Y. Zoghbi, *Neuron* **56**, 422 (2007).
3. J. D. Lewis *et al.*, *Cell* **69**, 905 (1992).
4. M. Chahrour *et al.*, *Science* **320**, 1224 (2008).
5. M. Tudor, S. Akbarian, R. Z. Chen, R. Jaenisch, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15536 (2002).
6. P. T. Georgel *et al.*, *J. Biol. Chem.* **278**, 32181 (2003).
7. S. Horike, S. Cai, M. Miyano, J. F. Cheng, T. Kohwi-Shigematsu, *Nat. Genet.* **37**, 31 (2005).
8. J. M. Kornhauser *et al.*, *Neuron* **34**, 221 (2002).
9. Z. Zhou *et al.*, *Neuron* **52**, 255 (2006).
10. H. Y. Zoghbi, *Science* **302**, 826 (2003).
11. S. W. Flavell, M. E. Greenberg, *Annu. Rev. Neurosci.* **31** (2008).
12. H. T. Chao, H. Y. Zoghbi, C. Rosenmund, *Neuron* **56**, 58 (2007).
13. S. Kangaspeska *et al.*, *Nature* **452**, 112 (2008).
14. R. Metivier *et al.*, *Nature* **452**, 45 (2008).

15. This work was supported by NIH grant NS048276 (M.E.G.).

10.1126/science.1159146

SCIENCE POLICY

As Election Nears, Policy Experts Plan–Cautiously–for S&T Turnaround

To many of the experts who gathered this month for the AAAS Forum on Science & Technology Policy, the challenge is clear: With the U.S. presidential and congressional elections 6 months away, scientists must engage with the political system to galvanize action on innovation, climate change, energy, and other critical issues.

Politics is not a natural venue for most scientists and engineers, but in 2 days of Forum speeches and presentations, the rough outlines emerged for a pre-election strategy and a post-election agenda. While many expressed frustration at the inaction on some issues in recent years and at the failure of this year's candidates to address science issues, a sense of guarded optimism emerged that the months ahead could yield a fundamental change of course.


Lewis M. Branscomb

Washington's "sustained complacency" must yield to a more comprehensive global vision, veteran science and technology policy expert Lewis M. Branscomb said in AAAS's annual William D. Carey Lecture. "We must have new leadership in the executive branch which recognizes that a broad range of government policies directly affect the nation's power to innovate," Branscomb said. Beyond education and R&D funding, he urged the White House to use trade and economic policy and infrastructure development to advance innovation.

Former U.S. Rep. John E. Porter, an Illinois Republican who now serves as chairman of the Research!America health research advocacy group, was bluntly critical of the administration of President George W. Bush. Reduced real funding for key federal science agencies and political interference in research have had a "devastating effect," Porter said. He urged scientists and engineers to redouble efforts to convey the importance of their work to elected officials of both parties and to the public.

The 33rd annual AAAS Forum on Science and Technology Policy, held just a few blocks from the White House on 8 and 9 May, attracted more than 500 policy-makers from government, education, industry, and other fields,

as well as more than two dozen journalists. While topics ranged from science and the new media to the frontiers of human enhancement, discussion returned inevitably to innovation, the 2008 campaign, and scientific engagement.

"The Forum has long provided an occasion for examining the federal investment in R&D," said Al Teich, director of Science & Policy Programs at AAAS. "Increasingly, however, we have been going well beyond the numbers and taking an in-depth look at the trends and personalities that are shaping the future policy environment for science."

White House Science Adviser John H. Marburger III defended federal S&T funding, saying in a keynote address that "there cannot be any question that this country has significantly boosted spending on research during this administration." He also offered advice for the next president, saying that his successor should be appointed as early as possible as the new team begins work.

Porter and others said the S&T community should be at work now to find candidates for top S&T posts in the new administration. Some proposed that the office of the science adviser be a cabinet-level position.

The 2005 AAAS president Gilbert S. Omenn, professor of Internal Medicine, Human Genetics, and Public Health at the University of Michigan, said the next president "must re-order the priorities of our nation" and "address long-deferred needs in energy, global environment, the economy, the workforce, education, health, and infrastructure."

The nation's fiscal condition—and especially its mounting bills for Medicare and Medicaid—will make such re-ordering difficult, said Peter Orszag, director of the Congressional Budget Office. A key problem, Orszag said, is that the nation does little to assess which policies actually work and which don't. Innovative, evidence-based policy could help shape effective initiatives on health care, education, and climate change, he said.

But while much discussion centered on policy and governance, several Forum speakers explored the possibility of a tidal shift in relations between science, political leaders, and the public.

John Kao, author of *Innovation Nation: How America is Losing Its Innovation Edge, Why It Matters, and What We Can Do to Get It Back,* suggested that the next administration must take a leadership role in developing a national narrative that makes research and innovation compelling, even sexy. The United States needs to "mobilize national energy to pursue this agenda, especially among the young," he said.


John Kao

R. James Woolsey, a foreign policy expert who served as head of the CIA under President Bill Clinton, detailed how climate change and terrorism—both profound security risks—arise from dependence on oil. In his breakfast address, he imagined a conversation between conservationist John Muir and World War II General George Patton. Where those titans might be expected to clash, the ghosts channeled by Woolsey found much agreement on how a green approach to energy could improve national security.

In the end imagined by Woolsey, Patton and Muir agreed that it could be "the beginning of a beautiful friendship."

See Forum presentations at www.aaas.org/spp/rd/forum.

—Earl Lane contributed to this report

INTERNATIONAL

UK Innovation Policy Explored at AAAS

As the United States and the United Kingdom grapple with issues of innovation and ethics in the changing global economy, AAAS is using its decades-long partnership with the British S&T community to explore how research can help address the challenges of climate change, global security, and new medical technologies.

John Denham, head of the UK's new Department of Innovation, Universities and Skills, said at a 21 April AAAS lecture that his government will allocate one-fifth of its science

budget over the next 3 years to "grand challenges" such as alternative energy research and new medical training centers. But he emphasized that Britain's push to become an "innovation nation" goes beyond domestic budget maneuvers, saying that international collaborations with the United States will combine "the strength of our research base with those of our partners overseas."

Sponsored by the AAAS International Office and the British Embassy, Denham's talk was one of the most recent engagement efforts undertaken by AAAS to underscore the importance of the relationship between UK and U.S. scientists and policy-makers. On 15 April, AAAS, the British Embassy, and the Johns Hopkins Berman Institute of Bioethics sponsored a meeting by the Hinxton Group, an international ethics consortium, to discuss their recent report on eggs and sperm grown from multipurpose stem cells.

"We value AAAS as a key partner as we continue to deepen the UK-U.S. relationship on science policy issues and in promoting scientific collaboration," said Brian Ferrar, first secretary for Science and Innovation at the British Embassy, noting that the recent talks "illustrate the breadth of activities where we work together to inform the scientific community about key issues."

"AAAS values our dynamic and robust relationship with both the British government and its broader science community," said Vaughan Turekian, AAAS's chief international officer. "Our work together helps highlight areas where international cooperation between and among scientists can help produce innovative discoveries that benefit society."

The creation of Denham's department last July by UK Prime Minister Gordon Brown "signals a government that is excited by science," the innovation minister said. His visit to AAAS was part of a recent tour to learn more about American science and innovation policy. While in the United States, Denham also visited Boston and the Research Triangle in North Carolina to see firsthand how research-rich geographic regions develop and sustain themselves.

The new department will also promote public understanding of science, so that people can participate fully in the moral and ethical deliberations surrounding genetically modified organisms and similarly controversial research, Denham said. "If people have the confidence to engage with new technology, they can drive innovation from below," he suggested.

The meeting of the Hinxton Group steering committee offered a glimpse of one such looming controversy: the possibility of sperm and egg cells grown from multipurpose stem cells within 5 to 15 years, according to a consensus statement released by the group on 14 April in the United Kingdom.

At the AAAS discussion, the group's members said the research could provide new insights into the basic biology of reproductive cells and perhaps lead to new forms of assisted conception. But the work also raises the possibility that a baby might one day be conceived from two laboratory-grown stem cells, one manipulated to become a sperm and the other an egg.

Ruth Faden, executive director of the Berman Institute of Bioethics, said the stem cell research should not be prevented or unnecessarily restricted by moral disagreements about its outcome. Instead, society should do what she called "the hard work to reach sufficient consensus" about the road forward, based on "defensible and public reasons."

—*Molly McElroy, Earl Lane, and Becky Ham*

ART OF SCIENCE

Siberian Students Commit to Climate Research

In one watercolor, children pluck icicles hanging from a cabin roof, and in another a dog watches as fishermen collect their catch from a net. In a third painting, a bear eyes a plump chipmunk in a vividly green forest. The works, like more than a dozen others displayed at AAAS, share a theme: They focus on Arctic lands and culture, and their creators are Siberian schoolchildren participating in climate research.

R. Max Holmes and Anya Suslova

The exhibit grew out of a propitious 2003 meeting between Earth system scientist R. Max Holmes of Woods Hole Research Center and 13-year-old Anya Suslova during a research expedition in the remote village of Zhigansk, along Siberia's Lena River. Anya became interested in Holmes' research and he enlisted her help with water sampling. "She blew me away by how quickly she picked things up," Holmes said.

When the expedition ended, Holmes left bottles behind for Anya to continue the sampling. Soon Anya's classmates—ranging in age from 9 to 14—heard about her work and wanted to help. The endeavor became the Student Partners Project, an initiative funded by the U.S. National Science Foundation to advance scientific understanding of Arctic rivers and their role in the changing ecosystem.

Holmes is amazed by the curiosity of the Siberian students, despite their barren classrooms and tattered books. Education is prized there and teachers are valued, he said.

Other Siberian communities, as well as Arctic-dwelling Canadian and Alaskan schools and U.S. schools in Vermont, Massachusetts, and Tennessee have joined the project, which is funded through mid-2009.

Watercolor on paper by Senya Koyakin, a middle school student in Zhigansk, Siberia.

Their measurements of the volume and chemical composition of Arctic rivers provide baselines for possible changes due to global warming. Already, Holmes' research team has found a 7% increase since the 1930s in the amount of water emptying from the rivers into the Arctic Ocean. The cause of the increase has not been firmly established, but Holmes contends that it is due to more warm, moisture-filled air from the tropics going north and ending up in the Arctic rivers.

To thank Holmes, the schoolchildren gave him a collection of their artwork, which is on display through 6 June at AAAS. The exhibit is co-sponsored by Woods Hole and is part of the AAAS Art of Science and Technology Program, which showcases the art of science. During a 14 May video-teleconference, the young artists briefly described their work to a AAAS audience.

Anya told the audience that she and her classmates drew frequently while at the Zhigansk school. "We draw what we see, it's our nature," she explained, referring to the dramatic Siberian landscape. Holmes received another 102 paintings when he visited Siberia this spring, and 11 of them currently are displayed at the United Nations. When the AAAS exhibit closes, the paintings will travel to a new show in Alaska.

Anya, now an 18-year-old college student, attended the AAAS event in person. She described how her climate research with Holmes has shaped her decision to study world economics at Yakutsk State University. "I was good at math and economics, and I liked to listen to policy," Anya said. "Our nature depends on politics."

—*Molly McElroy*

PUBLIC ENGAGEMENT

Science Breakthrough on Late Night TV

With science-related issues such as climate change and rising energy costs among the top issues of the 2008 U.S. presidential campaign, former and current AAAS leaders took to the soundstages of late-night television in April to

CREDIT: TOP: COURTESY OF THE STUDENT PARTNERS PROJECT

discuss the topics before a national audience on *The Late Show with David Letterman* and *The Charlie Rose Show.*

John P. Holdren, director of the Woods Hole Research Center and the Teresa and John Heinz Professor of Environmental Policy at Harvard University, addressed the topic of climate change in between jokes about Letterman's "boiling" backyard fishpond during the 17 April airing of *The Late Show* on CBS. The 2006 AAAS president warned of the accelerating pace of global climate change and dismissed the dwindling number of climate change skeptics who say the Earth's warming is nothing to worry about.

"You wouldn't say, 'Gee, you know, my usual temperature is 98.6, I've now got 104, but that's only a few degrees, why should I worry about that?'" Holdren said. "The Earth has a fever."

On the 7 April *Charlie Rose Show* on PBS, a roundtable of prominent researchers, including Shirley Ann Jackson, president of Rensselaer Polytechnic Institute and 2004 AAAS president, and *Science* Editor-in-Chief Bruce Alberts, discussed the "imperative of science" as a way to keep the U.S. competitive in a global economy. In particular, the roundtable agreed that the United States' dominance in S&T research could disappear within a few short decades as a result of what Alberts called "the mess of science education" that emphasizes memorization over dynamic problem-solving skills.

Holdren's appearance on *The Late Show* can be viewed at www.aaas.org/go/lateshow, and the full episode of the Charlie Rose science roundtable is available at www.aaas.org/go/rose/.

—*Becky Ham*

COMMUNICATION

Nanotech: A Critical Need For Public Engagement

As the promise and potential risks of nanotechnology emerge into public awareness, science leaders should undertake a concerted communication effort to build public understanding and acceptance, says Shirley Malcom, head of Education and Human Resources at AAAS.

In remarks before the Congressional Nanotechnology Caucus, Malcom said the challenge facing nanotech is one that often accompanies scientific and technological advances: When a breakthrough promises rich potential rewards but also raises ethical and social issues, scientists have to build a connection with the public to provide trustworthy information and to address the concerns.

Malcom served on former U.S. President Bill Clinton's Committee of Advisors on Science and Technology when the National Nanotechnology Initiative was being developed. Today, she told the caucus, "the public's need to understand the broader impacts and support the research [must be brought] into the center of the discussion."

Nanotechnology is a process requiring the manipulation of atoms and the assembly of molecules—a process that allows building products or machines in the scale of nanometers, or billionths of a meter. While nanotechnology has applications from health care and energy to clothing and cosmetics, some researchers have raised questions about environmental problems, social impacts, and other potential consequences.

For members of the baby-boom generation, Malcom said, the current cultural reference point for nanotech might be "The Blob," a 1958 horror movie about a malicious ball of space protoplasm that tries to take over a small Pennsylvania town.

"Are the powerful stories about the science and engineering available to counter balance the sci-fi films and presentations related to the risk?" she asked. "More recent statements about gray goo, coupled with regulatory failures and the global nature of the science, raise legitimate concerns even among those of us who have worked to understand more about the initiative."

Malcom said that the Human Genome Project may provide a model for public engagement that could be used for nanotechnology. AAAS in the late 1990s developed a video and published *Your Genes, Your Choices,* which combined science and personal vignettes in clear, accessible language to convey the meaning and importance of the research.

"The book and video traveled far beyond our anticipated destinations—every educational level from middle school through college, informal venues, and around the world," Malcom told the Caucus on 31 March. "Requests were received to translate the materials into many other languages, including Icelandic.

"It seems that by bringing people into the story, we struck a chord. They were willing to work a little harder to understand the underlying scientific concepts and to consider that this science could affect their own lives."

The Congressional Nanotechnology Caucus seeks to promote the development of nanotechnology and to assure that the United States stays competitive in the field.

INTERNATIONAL

Chinese Academy and AAAS Release High-Impact Science Translations

The Chinese Academy of Sciences (CAS) and AAAS have released a collection of 31 major papers that were published in *Science* and then translated into Chinese, the first major project completed as part of an emerging engagement between the two organizations.

The book was launched during a ceremony and news conference on 7 May at the Academy's headquarters in Beijing. Published by China Science Press, it features high-impact *Science* papers from the past decade, selected by distinguished committees from AAAS and CAS that included *Science* editors and CAS scholars who also were critical to ensuring the quality of the translations. The articles represented fields such as cellular and molecular biology, biochemistry, biomedicine, chemistry, and materials science.

Leaders of the two organizations hailed the publication as a benefit to global science education and as an augur of future cooperative ventures between the Academy and AAAS.

"This achievement is a milestone of the cooperation between CAS and AAAS, and will definitely inspire and promote the innovative research of Chinese young scientists," Academy President Lu Yongxiang wrote in a letter to Alan I. Leshner, the chief executive officer of AAAS. "I look forward to more collaborative opportunities between our two organizations."

Leshner, who also serves as executive publisher of *Science,* wrote to Lu: "I hope this book will lead to more cutting-edge and innovative science and challenge the minds of those who read it… It is a true indication of the strong relationship that exists between our organizations, and demonstrates the enormous potential for our other collaborative efforts."

The Academy's group at the ceremony was led by Vice Secretary General Tan Tieniu. AAAS was represented by Chief International Officer Vaughan Turekian and Tom Wang, director for International Cooperation.

The publishing project emerged from extended discussions, culminating last September in a memorandum of understanding signed by Lu and Leshner at a meeting in Beijing.

With China emerging as an engine of world research and development, AAAS leaders have seen it as vital for the global scientific enterprise to establish a constructive, long-term engagement with its S&T leaders and organizations. The Chinese Academy of Sciences, founded in 1949, is a powerhouse in Chinese S&T. It oversees 108 scientific research institutes, over 200 science and technology enterprises, a university, a graduate school, and various other units. It publishes more than 200 journals.

CREDIT: COURTESY OF THE CHINESE ACADEMY OF SCIENCES (CAS)

How We See Ourselves and How We See Others

Emily Pronin

People see themselves differently from how they see others. They are immersed in their own sensations, emotions, and cognitions at the same time that their experience of others is dominated by what can be observed externally. This basic asymmetry has broad consequences. It leads people to judge themselves and their own behavior differently from how they judge others and those others' behavior. Often, those differences produce disagreement and conflict. Understanding the psychological basis of those differences may help mitigate some of their negative effects.

Have you ever attended a long and technical lecture and seen yourself as easily distracted while thinking that everyone else seemed thoroughly engaged? Or, maybe you disagreed with the message of the lecture and saw the speaker's interpretation as one-sided but your own interpretation as objective. These differences in how you perceived yourself versus others may have reflected something about the specific circumstances or people involved—such as your unique distraction during the lecture, or the particularly biased perspective of the speaker. Often, though, such differences in how people see themselves versus others are systematic and predictable, and rooted in basic processes of human perception.

In 1972, the social psychologists Edward Jones and Richard Nisbett theorized about the basic mechanics involved in perceiving oneself versus others (*1*, *2*). Their analysis sprung from the finding that people often view their own actions as caused by situational constraints, while viewing others' actions as caused by those others' internal and stable dispositions. An example would be a person arriving late for a job interview and ascribing that lateness to bad traffic while his interviewer attributed it to personal irresponsibility. Although this difference in attributions may appear self-serving, Jones and Nisbett pointed out that the difference does not always promote self-flattery and suggested that it in part reflects basic (and nonmotivational) qualities of perception.

They supported their argument with two key facts. First, people possess different information when perceiving themselves versus others. They have far more information about the feelings and intentions that precede, accompany, and follow from their own actions. As a result, they know when those actions fail to match their inner thoughts and desires due to situational constraints (as when they want the job, but miss the interview due to bad traffic). Second, people's attention focuses on different things when perceiving self versus others. Because of the structure of the human visual system, people can devote far less visual attention to themselves and their actions (which they cannot easily see without a mirror) than to others and others' actions.

Recent research has built upon Jones and Nisbett's theorizing. That research begins with the fact that we generally have access to internal inputs when perceiving ourselves and our own behavior (inputs that others lack access to), and access to external inputs from the senses, especially vision, when perceiving others and their behavior (inputs that we lack access or attention to in perceiving ourselves). As a result, we tend to perceive ourselves via "introspection" (looking inwards to thoughts, feelings, and intentions) and others via "extrospection" (looking outwards to observable behavior). In short, we judge others based on what we see, but ourselves based on what we think and feel.

My Thoughts, Your Behavior

This distinction in the information that people possess when perceiving themselves versus others affects how people evaluate their own and others' behavior. This review begins by describing some dramatic examples. After next exploring the depth and underpinnings of differences in how people see themselves versus others, the review closes with the hope that greater insight into these differences may help people to better understand themselves and each other and thereby may alleviate some aspects of social misunderstanding and conflict.

Positive illusions. People tend to have inflated views of themselves and their futures (*3*). For example, they think that they are more likely to become wealthy, and less likely to contract contagious diseases, than those around them (*4*). This unrealistic optimism partially stems from people's attentional focus on their own (but not others') internal desires and intentions (*5*, *6*). In one series of studies, people predicted how quickly they (or others) would complete various work projects and whether they would meet their deadlines. They were overoptimistic about themselves (but not others) because they focused on their industrious motives and intentions rather than their past behavior or the behavior of others in similar situations (*6*).

People's unrealistic positivity also extends to judging their traits (*3*). People's impressions of themselves typically correlate with how others perceive them (*7*, *8*), but people also tend to rate themselves more positively than do others (*3*, *7*). For example, people who see themselves as considerate are seen as more considerate than those who see themselves as selfish, but they also see themselves as more considerate than others do. People's overly positive views of their traits (such as considerateness) in part reflect their acute awareness of their internal desires and intentions when judging themselves combined with their acute awareness of outward behavior when judging others (*9*).

Interpersonal knowledge. People overestimate how much they can learn about others from brief encounters such as job interviews (*10*). At the same time, they think others can get only a glimpse of them from such encounters. As a result, people generally feel they know others better than others know them (*11*). During social interactions, people are aware most of their own internal thoughts and feelings and others' observable behavior (*12*). Consistent with this asymmetry, people infer that others' expressions and actions tell the essential story of who they are, even while they believe that in their own case it is their inner feelings, beliefs, and goals that tell that story (*11*, *13*).

Pluralistic ignorance. People often misconstrue the thoughts and motives of others. In cases of "pluralistic ignorance," those misconstruals occur even though others share one's own motives and beliefs and act in the same way as oneself (*14*). An example, suggested at the outset of this article, occurs when an audience of people all succeed in concealing their distraction and boredom during a long lecture and they then assume that they are the only ones not interested and engaged. In another example, college students often forgo trying to make friends with students of other races (even though they would like to be friends) because they interpret those others' lack of trying as indicating lack of interest (*15*). Both these examples involve people judging others based on overt behavior (e.g., failing to make social overtures) but themselves based on internal states (e.g., wanting friendship but fearing rejection) (*14*, *15*).

Miscommunications. People often fail miserably in their efforts to communicate. These communication breakdowns (whether they involve negotiating peace agreements, giving driving directions, or navigating romantic relationships) often reflect the fact that people know what they intend or mean to communicate, while others focus on what they actually say (*16*). For example, negotiators can fail to outwardly express their interest in cooperating, because their

Department of Psychology and Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08540, USA. E-mail: epronin@princeton.edu

internal awareness of that interest (gained through introspection) blinds them to the fact that the other side sees only their behavior, which often lacks clear signs of that motive (*17*, *18*).

Conformity. People are influenced by those around them (and by the mass media) in everything from fashion tastes to political views; but, they generally deny that and see themselves as alone in a crowd of sheep (*19*, *20*). In one study, Californians mimicked the positions of their political party in deciding how to vote on a series of alleged ballot initiatives. They were blind to that influence, while observers saw it (*19*). This divergence arose from the voters' focusing on their thoughts, which revealed no signs of conformity, while the observers focused on the voters' behavior. Voters' thoughts obscured their conformity since it occurred indirectly, such as by altering their interpretations of the initiatives themselves (e.g., knowing that Democrats supported a labor-related initiative made it seem worker-friendly rather than corporate-friendly). More generally, a problem with looking inwards for signs of conformity is that it often occurs nonconsciously. For example, people in conversation conform their gestures to one another without realizing it (*21*), and shoppers conform to the appeals of advertising campaigns—even when they are unaware of having seen those campaigns (*22*).

The foregoing discussion illustrates some consequences of the fact that people's impressions of themselves versus others are based on very different information. For self-assessments, that information is largely introspective (based on looking to internal thoughts and feelings). For others, it is largely extrospective (based on looking to external behavior).

Although these two sources of information differ, they share something in common for the person relying on them: Each involves seemingly immediate and direct data. Information about one's own internal states or about others' external appearances can generally be gleaned more directly than information about others' mental states or one's own external appearances. One way to gather this less direct information is by asking others to report their perceptions, but, as discussed next, we may have less faith in others' perceptions than our own.

I'm Objective, You're Biased

People tend to assume that information that comes to them through their perceptions directly reflects what is true in "reality" (*2*, *23*, *24*). Of course, this assumption is inaccurate. Perceptions only indirectly reflect reality; they are colored and shaped by influences ranging from the imperfections of vision to the distorting pressures of hopes and desires. For example, people assume that their perceptions of events ranging from military conflicts to basketball games objectively reflect who the victors and the villains are in those events. However, studies have shown that people's affiliations with particular political agendas or sports teams dramatically influence those perceptions (*25*, *26*). Because people do not consciously experience the operation of such distorting influences (*27*, *28*), they assume their absence (even while outside observers, focused on the people's actions, readily impute those influences).

People's lack of awareness of the processes that shape and distort their perceptions leads them to view those perceptions as objective. In one study, participants considered a male and a female candidate for a police-chief job and then assessed whether being "streetwise" or "formally educated" was more important for the job. The result was that participants favored whichever background they were told the male candidate possessed (e.g., if told he was "streetwise," they viewed that as more important). Participants were completely blind to this gender bias; indeed, the more objective they believed they had been, the more bias they actually showed (*29*).

Because people often do not recognize when personal biases and idiosyncratic interpretations have shaped their judgments and preferences, they often take for granted that others will share those judgments and preferences (*23*). When others do not, people's faith in their own objectivity often prompts them to view those others as biased. Indeed, people show a broad and pervasive tendency to see (and even exaggerate) the impact of bias on others' judgments while denying its influence on their own (*23*). For example, people think that others' policy opinions are biased by self-interest (*30*), that others' social judgments are biased by an inclination to rely on dispositional (rather than situational) explanations for behavior (*31*), and that others' perceptions of interpersonal conflicts are biased by their personal allegiances (*32*). At the same time, people are blind to each of these biases in their own judgments (*30*–*32*). Such divergent perceptions of bias are bolstered by the fact that people evaluate their own bias by introspecting about thoughts and motives but evaluate others' bias by considering external behavior (e.g., "My motive was to be fair; his actions only helped himself."). People place less emphasis on others' introspections even when those others proffer them (*19*, *33*)—a finding that is perhaps unsurprising in light of people's skepticism about the accuracy of others' perceptions.

In the face of disagreement, beliefs in one's own objectivity and the other side's bias can produce and exacerbate conflict (*23*). For example, American students favor bombing terrorists after being led to view them as biased and irrational, whereas they favor negotiating with terrorists after being led to view them as objective and rational (*34*). People also behave more conflictually toward those whom they suspect will be biased by self-interest. Participants in one study were instructed to consider the perspective of their adversaries in a conflict over limited resources (*35*). That instruction had the ironic effect of leading them to expect that their adversaries would be biased by self-interest, which, in turn, led the participants themselves to act more competitively and selfishly. Acts of competitiveness and aggression are likely to engender a vicious cycle, as the recipients of those acts are likely to view them as unwarranted by the objective situation and, therefore, as signaling their perpetrators' bias.

Origins of the Difference

People's attention to their conscious introspections has clear benefits. It allows people to simulate and therefore predict others' actions (*36*). It also allows people to learn from and sometimes override the action patterns that humans have evolved to display automatically (*37*). Even when introspections lead us astray, as when they promote inflated self-views, relying on them may be adaptive, as suggested by the finding that positive illusions are predictive of physical and mental health (*3*, *38*).

It also may be adaptive to place less emphasis on others' introspections than on our own. We can be certain of our own introspective contents but not of others', so it is wise to treat theirs with greater skepticism. This does not mean that we do not care about others' thoughts, feelings, motives, and intentions—indeed, we likely have evolved to wonder what others' want and intend for us (*39*, *40*). It just means that it may be safer to attend to others' actions rather than their reported mental states when the two conflict. Indeed, people seem to automatically look to others' actions as a way of inferring their "true" attitudes and dispositions (*41*, *42*).

The self-other difference in attention to introspections versus actions also has roots in human development. Young infants overvalue their own introspections in the sense that they fail to appreciate that their hopes and wishes do not directly translate into external outcomes (*43*). They show a different error when it comes to others. Because they fail to appreciate that others possess desires, beliefs, and other mental states that differ from their own, they underappreciate the role of others' real internal states in guiding those others' actions (*44*). This appreciation develops through interaction and maturation; its failure to develop is a characteristic of autism (*45*). However, even adults do not fully overcome these errors, indicating that their persistence may in part be a holdover from infancy.

Self and Other in the Brain

Experiments in neuroscience have directly examined regions of human brain activity and patterns of neuronal firing that are recruited when people (and other nonhuman primates) perceive or judge themselves versus others. The results suggest a possible neuroscientific basis for the processes described here.

Most notable are experiments that have identified neural activity specifically involved when individuals perceive both themselves and others. Areas of the medial prefrontal cortex (mPFC) have been shown to activate when

people make judgments about both their own internal states (feelings and intentions) and those of others (*46*). Monkeys have been shown to possess mirror neurons that fire both when they perform an action and when they perceive another perform the same action (*47*), and it is possible (but not definitively established) that humans also possess mirror neurons (*47*, *48*).

These data point to common brain processes uniquely involved in the perception of self and others. They suggest (albeit tentatively on the basis of current evidence) that when observing others, people automatically simulate the mental processes behind those others' actions. At a very basic level, people may satisfy their interest in knowing others' thoughts and feelings by thinking about how they themselves would think or feel were they that other person, rather than by relying on that other's introspections. This idea is consistent with behavioral experiments indicating that people infer others' mental states by first recruiting their own and then adjusting from there (*49*) and that, in the absence of contrary information, people project their own traits and attitudes onto others (*50*).

If people look to their own minds in order to simulate those of others, then why do they see themselves as objective but others as biased? One possible answer is that people show similar mPFC activation for judgments of self and others only insofar as they view those others as similar to themselves (*51*). This suggests a potential neural substrate for the observation that people are especially likely to perceive others as biased when those others disagree with them. People may be less inclined to project their own mental states (for example, their objectivity) onto those whose views are dissimilar. In such cases, people may instead judge others according to stereotypes or lay theories about humans "in general" (e.g., "Given the opportunity, people will be self-serving.") (*52*, *53*).

Fig. 1. In *Las Meninas*, the painter Velazquez sees himself in the image he is painting. When people recall past events involving themselves (or imagine future ones), they similarly often see themselves in the images they form. Such observer-like images are associated with having more difficulty recalling or accessing one's internal thoughts and feelings during the event.

CREDIT: DIEGO VELAZQUEZ/MUSEO DEL PRADO, MADRID, SPAIN/SUPERSTOCK

The Self in Time

Philosophers have long debated questions of personal identity regarding whether a given person is the same person across time. Derek Parfit famously argued that people are nothing more than successions of different overlapping selves (*54*, *55*). Regardless of whether this is the case, people often treat their past and future selves as though they are other people (*56*–*58*). This occurs, for example, when people spend money on present selves rather than save for future ones, or when people consign future selves to unpleasant experiences (e.g., painful surgeries) rather than undergo them in the present.

This article has described key differences in the underlying processes involved in people's perceptions of themselves versus others. In light of those differences, it makes sense that people sometimes treat past and future selves like other people. When it comes to temporally distant selves, whether at a week's distance or a decade's, people cannot introspectively access those selves' thoughts and feelings [a fact that in itself produces important errors in judgment (*59*)]. Indeed, studies have linked people's tendency to perceive past and future selves as others to their inattention to the internal states of those selves (*57*, *58*). This inattention also affects people's decision-making. In one study, people were faced with the prospect of drinking a murky mix of soy sauce, ketchup, and water for the benefit of science (the more they drank the better, the experimenter explained, because she was studying "disgust"). When deciding how much to drink right then, participants chose about two tablespoons; when deciding how much to drink at a future session, they chose about half a cup—the same amount they chose for a peer (*58*).

Recent neuroscientific evidence suggests that different regions of the brain are activated when people think about outcomes affecting themselves in the present versus the future (*60*), such that the limbic system is uniquely involved when

contemplating present rewards. This unique activation of a brain system strongly linked to emotional experience supports the notion that people are more attuned to internal feelings and concerns in the case of present than future selves (*60*).

People's perceptions of past and future selves resemble their perceptions of others not only in terms of their decreased attention to internal states but also in terms of their increased focus on outward behavior. For example, when people visualize their past and future experiences (as opposed to present ones), the images they form often take the perspective of an external observer (Fig. 1), such that they see their behavior in the image (*57*). Those images are especially common when people feel that they have changed since the past and used to be a "different person" (*61*). They also are especially common in the case of memories characterized by low recall of internal states such as emotions, thoughts, and physical sensations (*57*, *61*–*63*).

The presence of differences in how people perceive present versus future selves shows that people do not perceive all of their own "selves" the same. This raises a related question: Do people perceive all "others" the same—or, do they use different information when thinking about friends, relatives, strangers, and enemies? Although this question is not fully settled, studies have revealed some differences, generally suggesting that close others are perceived in more "self-like" ways than more distant others (*53*, *56*, *64*). The foregoing review suggests that this generalization is likely to be particularly true to the extent that people have greater access to and appreciation for close others' internal thoughts, feelings, and intentions, and to the extent that people who are close to each other see the world in similar ways.

Conclusions

It is almost axiomatic that as long as people are in a position to perceive themselves and to perceive others, differences in those perceptions will exist and will engender disagreement, misunderstanding, and conflict. When people judge themselves based on their good intentions but others based on their less-good behavior (or based on cynical assumptions about human nature), they are likely to feel resentful and disappointed over others' failure to meet them halfway. When people view their own perceptions and beliefs as objective reflections of reality but others' as distorted by bias, they are likely to feel frustrated and angry over others' unwillingness to be fair and reasonable. And, such feelings are likely to breed aggression and conflict.

This picture may sound dismal, but there is hope. Misunderstandings can be averted by those aware of the psychological processes involved in self and social perception. Those individuals can be mindful that it is not only their own behavior that is sensitive to the constraints of the situation, but others' as well. Perhaps this could prompt them to show more charity when others fail to meet expectations. Those individuals also can recognize that others' mistakes and errors may not be the result of conscious malice but rather of unintended influences that those others would themselves decry. And, those individuals might remind themselves that there often is a wide gulf between intention and action, but that it is only reasonable and fair to apply the same standard of judgment to others as to oneself. Following these guidelines would not just be socially charitable—it would also be scientifically informed.

References and Notes

1. The term "perceive" is used throughout to connote processing and interpreting information via the senses or mind.
2. E. E. Jones, R. E. Nisbett, in *Attribution: Perceiving the Causes of Behavior*, E. E. Jones *et al.*, Eds. (General Learning, Morristown, NJ, 1972), pp. 79–94.
3. S. E. Taylor, J. D. Brown, *Psychol. Bull.* **103**, 193 (1988).
4. N. D. Weinstein, *J. Pers. Soc. Psychol.* **39**, 806 (1980).
5. N. Epley, D. Dunning, *J. Pers. Soc. Psychol.* **79**, 861 (2000).
6. R. Buehler, D. Griffin, M. Ross, *J. Pers. Soc. Psychol.* **67**, 366 (1994).
7. D. A. Kenny, L. Albright, *Psychol. Bull.* **102**, 390 (1987).
8. T. E. Malloy, L. Albright, S. Scarpati, *Int. J. Behav. Dev.* **31**, 603 (2007).
9. J. Kruger, T. Gilovich, *Pers. Soc. Psychol. Bull.* **30**, 328 (2004).
10. L. Ross, R. E. Nisbett, *The Person and the Situation* (McGraw-Hill, New York, 1991).
11. E. Pronin, J. Kruger, K. Savitsky, L. Ross, *J. Pers. Soc. Psychol.* **81**, 639 (2001).
12. B. F. Malle, in *The New Unconscious*, R. R. Hassin, J. S. Uleman, J. A. Bargh, Eds. (Oxford Univ. Press, New York, 2005), pp. 225–255.
13. S. M. Andersen, L. Ross, *J. Pers. Soc. Psychol.* **46**, 280 (1984).
14. D. A. Prentice, D. T. Miller, *Adv. Exp. Soc. Psychol.* **28**, 161 (1996).
15. J. N. Shelton, J. A. Richeson, *J. Pers. Soc. Psychol.* **88**, 91 (2005).
16. B. Keysar, A. S. Henly, *Psychol. Sci.* **13**, 207 (2002).
17. J. D. Vorauer, S. D. Claude, *Pers. Soc. Psychol. Bull.* **24**, 371 (1998).
18. L. Van Boven, T. Gilovich, V. H. Medvec, *Negot. J.* **19**, 117 (2003).
19. E. Pronin, J. Berger, S. Molouki, *J. Pers. Soc. Psychol.* **92**, 585 (2007).
20. K. M. Douglas, R. M. Sutton, *Br. J. Soc. Psychol.* **43**, 585 (2004).
21. T. L. Chartrand, J. A. Bargh, *J. Pers. Soc. Psychol.* **76**, 893 (1999).
22. X. Fang, S. Singh, R. Ahluwalia, *J. Consum. Res.* **34**, 97 (2007).
23. E. Pronin, T. Gilovich, L. Ross, *Psychol. Rev.* **111**, 781 (2004).
24. G. Ichheiser, *Am. J. Sociol.* **55**, Suppl. (1949).
25. R. P. Vallone, L. Ross, M. R. Lepper, *J. Pers. Soc. Psychol.* **49**, 577 (1985).
26. A. H. Hastorf, H. Cantril, *J. Abnorm. Soc. Psychol.* **49**, 129 (1954).
27. T. D. Wilson, D. B. Centerbar, N. Brekke, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, T. Gilovich, D. Griffin, D. Kahneman, Eds. (Cambridge Univ. Press, New York, 2002), pp. 185–200.
28. J. A. Bargh, E. L. Williams, *Curr. Dir. Psychol. Sci.* **15**, 1 (2006).
29. E. Uhlmann, G. L. Cohen, *Psychol. Sci.* **16**, 474 (2005).
30. D. T. Miller, R. K. Ratner, *J. Pers. Soc. Psychol.* **74**, 53 (1998).
31. L. Van Boven, K. White, A. Kamada, T. Gilovich, *J. Pers. Soc. Psychol.* **85**, 249 (2003).
32. C. M. Frantz, *Basic Appl. Soc. Psych.* **28**, 157 (2006).
33. E. Pronin, M. B. Kugler, *J. Exp. Soc. Psychol.* **43**, 565 (2007).
34. E. Pronin, K. Kennedy, S. Butsch, *Basic Appl. Soc. Psych.* **28**, 385 (2006).
35. N. Epley, E. Caruso, M. H. Bazerman, *J. Pers. Soc. Psychol.* **91**, 872 (2006).
36. R. Leakey, R. Lewin, *Origins Reconsidered: In Search of What Makes Us Human* (Doubleday, New York, 1992).
37. J. A. Bargh, E. Morsella, *Perspect. Psychol. Sci.* **3**, 73 (2008).
38. S. E. Taylor *et al.*, *J. Pers. Soc. Psychol.* **85**, 605 (2003).
39. R. Dawkins, *The Selfish Gene* (Oxford Univ. Press, New York, 1976).
40. S. Pinker, *How the Mind Works* (Norton, New York, 1997).
41. D. T. Gilbert, P. S. Malone, *Psychol. Bull.* **117**, 21 (1995).
42. A. Todorov, J. S. Uleman, *J. Pers. Soc. Psychol.* **87**, 482 (2004).
43. J. P. Piaget, *The Child's Conception of the World* (Routledge and Kegan Paul, London, 1929).
44. H. Wellman, in *Blackwell Handbooks of Developmental Psychology*, U. Goswami, Ed. (Blackwell, Malden, MA, 2002), pp. 167–187.
45. S. Baron-Cohen, A. M. Leslie, U. Frith, *Cognition* **21**, 37 (1985).
46. K. N. Ochsner *et al.*, *J. Cogn. Neurosci.* **16**, 1746 (2004).
47. G. Rizzolatti, L. Craighero, *Annu. Rev. Neurosci.* **27**, 169 (2004).
48. I. Dinstein, C. Thomas, M. Behrmann, D. J. Heeger, *Curr. Biol.* **18**, R13 (2008).
49. N. Epley, B. Keysar, L. Van Boven, T. Gilovich, *J. Pers. Soc. Psychol.* **87**, 327 (2004).
50. J. M. Robbins, J. I. Krueger, *Pers. Soc. Psychol. Rev.* **9**, 32 (2005).
51. J. P. Mitchell, M. R. Banaji, C. N. Macrae, *J. Cogn. Neurosci.* **17**, 1306 (2005).
52. R. Saxe, *Trends Cogn. Sci.* **9**, 174 (2005).
53. D. R. Ames, *J. Pers. Soc. Psychol.* **87**, 340 (2004).
54. D. Parfit, *Philos. Rev.* **80**, 3 (1971).
55. S. Frederick, in *Time and Decision*, G. Loewenstein, D. Reed, R. Baumeister, Eds. (Russell Sage, New York, 2003), pp. 89–113.
56. N. Liberman, Y. Trope, E. Stephan, in *Social Psychology: Handbook of Basic Principles*, A. W. Kruglanski, E. T. Higgins, Eds. (Guilford, New York, 2007), pp. 353–384.
57. E. Pronin, L. Ross, *J. Pers. Soc. Psychol.* **90**, 197 (2006).
58. E. Pronin, C. Y. Olivola, K. A. Kennedy, *Pers. Soc. Psychol. Bull.* **34**, 224 (2008).
59. D. T. Gilbert, T. D. Wilson, *Science* **317**, 1351 (2007).
60. S. M. McClure, D. I. Laibson, G. Loewenstein, J. D. Cohen, *Science* **306**, 503 (2004).
61. L. K. Libby, R. P. Eibach, *J. Pers. Soc. Psychol.* **82**, 167 (2002).
62. G. Nigro, U. Neisser, *Cognit. Psychol.* **15**, 467 (1983).
63. H. K. McIsaac, E. Eich, *Psychol. Sci.* **15**, 248 (2004).
64. D. A. Prentice, *J. Pers. Soc. Psychol.* **59**, 369 (1990).

65. Supported by the Department of Psychology and the Woodrow Wilson School of Public and International Affairs at Princeton University and by NSF research grant BCS-0742394. I thank J. Bargh, G. Cohen, G. Love, D. Pager, L. Ross, B. Strauss, and two anonymous reviewers for comments.

10.1126/science.1154199

Ecosystem Feedbacks and Nitrogen Fixation in Boreal Forests

Thomas H. DeLuca,[1,2]* Olle Zackrisson,[2] Michael J. Gundale,[2] Marie-Charlotte Nilsson[2]

Feedback mechanisms among organisms that regulate fundamental ecosystem processes are integral to our ability to model or predict long-term outcomes in natural and disturbed environments (*1*). Biological nitrogen (N) fixation in terrestrial ecosystems is a process of great ecological importance, but one whose regulation at the ecosystem level remains poorly understood (*2*). This lack of clarity is a substantial impediment to effectively modeling ecosystem N budgets, net primary productivity, and ecosystem carbon storage (*3*).

In boreal forests, N is the primary growth-limiting nutrient; however, there is limited understanding of ecosystem factors that regulate forest N dynamics and specifically N fixation. Given that the boreal biome accounts for 17% of Earth's land surface, a more thorough understanding of feedback mechanisms in this biome has broad-scale implications.

Associative N-fixing cyanobacteria that colonize pleurocarpus feather mosses are the main source of ecosystem N in northern boreal forest ecosystems (*4*). Wildfire in boreal forest ecosystems (a primary driver of ecosystem processes in boreal forests) consumes much of the moss layer and forest floor, greatly eliminating N fixation while temporarily increasing N availability for regenerating vegetation. In late secondary succession, boreal forests retrogress as indicated by reduced productivity, reduced N availability, and increased canopy gaps (*5*). As N availability decreases, N fixation in feather moss carpets increases (*6*); however, a mechanism for this increase has not been revealed.

Herein, we report a close relationship between increasing N throughfall and decreasing N fixation, suggesting an ecosystem-level feedback in which high N bioavailability in early succession yields increased N deposition on the moss carpets via canopy throughfall and litter, which down-regulates N fixation.

In the summer of 2006, we installed replicate resin lysimeters in the moss bottom layer at 12 separate forest stands (*6*) varying in time since last fire (from 39 to 359 years since the most recent fire). The lysimeters were left in place for 1 year (June 2006–June 2007), allowing them to collect throughfall N from both the overstory and understory over the course of 1 year. During the summer of 2007, we measured N fixation in replicated plots at all 12 sites during three separate sampling periods by using a ^{15}N calibrated acetylene reduction assay (*7*).

Nitrogen fixation rates increased linearly with time since last fire (Fig. 1A). Throughfall N decreased with time since last fire (Fig. 1B), showing that early successional sites have higher N throughfall rates relative to late successional sites in this region (*6*). Given the low rates of anthropogenic N deposition recorded in this region of northern Sweden [0.1 to 0.15 g m^{-2} $year^{-1}$ (*7*)], it is certain that most of the N deposited was from overstory and understory throughfall and litter. Nitrogen fixation in the summer of 2007 was negatively related to throughfall N collected from June 2006 to June 2007 [$r^2 = 0.76$, $P < 0.01$ nonlinear regression, N fixation = 1.71 $e^{-4.4(\text{throughfall N})}$] across the 12 chronosequence sites (Fig. 1C).

These results demonstrate an internal driver for the down-regulation of N fixation in the northern boreal forest ecosystem: Throughfall N contributed by overstory and understory canopies is deposited on feather moss carpets, influencing N fixation by associative cyanobacteria. Elevated bioavailability of N in early successional forests controls the ecosystem N budget via down-regulation of cyanobacterial N fixation. We previously showed that the application of as little as 0.45 g N m^{-2} as NH_4NO_3 (in distilled water) to moss carpets in late succession (high N fixation) forests reduces N fixation rates to near detection limits, demonstrating the direct negative effect of inorganic N deposition on N fixation in moss carpets (*6*).

The relationship reported herein reveals tight internal control of N cycling in boreal forest ecosystems wherein the up- or down-regulation of N fixation by throughfall N creates a system that is somewhat buffered against natural and anthropogenic stressors. This regulated N fixation pattern generates a durable N resource throughout recurrent cycles of fire and secondary succession.

References and Notes

1. J. D. Bever, *New Phytol.* **157**, 465 (2003).
2. P. M. Vitousek *et al.*, *Biogeochemistry* **57/58**, 1 (2002).
3. P. B. Reich *et al.*, *Nature* **440**, 922 (2006).
4. T. H. DeLuca, O. Zackrisson, M.-C. Nilsson, A. Sellstedt, *Nature* **419**, 917 (2002).
5. D. A. Wardle, G. Hörnberg, O. Zackrisson, M. Kalela-Brundin, D. A. Coomes, *Science* **300**, 972 (2003).
6. O. Zackrisson, T. H. DeLuca, M.-C. Nilsson, A. Sellstedt, L. Berglund, *Ecology* **85**, 3327 (2004).
7. Materials and methods are available on *Science* Online.
8. This work was supported by a grant from Formas. Thanks to D. Wardle for suggestions on this paper.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1181/DC1
Materials and Methods
References

4 January 2008; accepted 6 March 2008
10.1126/science.1154836

[1]Ecology and Economics Research Department, The Wilderness Society, 503 West Mendenhall, Bozeman, MT 59715, USA. [2]Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, 901-83 Umeå, Sweden.

*To whom correspondence should be addressed. E-mail: tom_deluca@tws.org

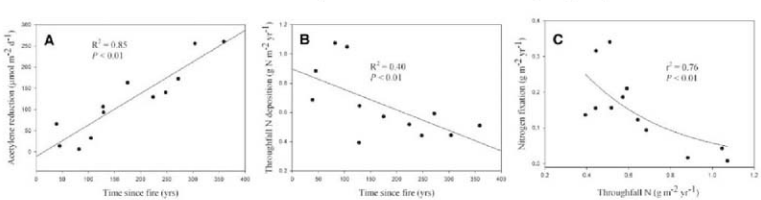


Fig. 1. (A) Relationship between N fixation (measured as acetylene reduction) and time since fire for the moss bottom layer. Acetylene reduction values represent average of 12 subplots at each of 12 sites in northern Sweden that range in age from 39 to 359 years since last fire and averaged across three sample periods in the summer of 2007. **(B)** Relationship between throughfall N deposition (resin lysimeter) and time since fire. Throughfall inorganic N (combined NO_3^--N and NH_4^+-N concentrations) averaged across 12 subplots per site and collected from June 2006–June 2007. **(C)** Relationship between N fixation (^{15}N calibrated acetylene reduction) and throughfall N deposition in northern Sweden.

Mars North Polar Deposits: Stratigraphy, Age, and Geodynamical Response

Roger J. Phillips,[1,2]* Maria T. Zuber,[3] Suzanne E. Smrekar,[4] Michael T. Mellon,[5] James W. Head,[6] Kenneth L. Tanaka,[7] Nathaniel E. Putzig,[1] Sarah M. Milkovich,[4] Bruce A. Campbell,[8] Jeffrey J. Plaut,[4] Ali Safaeinili,[4] Roberto Seu,[9] Daniela Biccari,[9] Lynn M. Carter,[8] Giovanni Picardi,[9] Roberto Orosei,[10] P. Surdas Mohit,[11,12] Essam Heggy,[13,14] Richard W. Zurek,[4] Anthony F. Egan,[1] Emanuele Giacomoni,[9] Federica Russo,[9] Marco Cutigni,[9] Elena Pettinelli,[15] John W. Holt,[16] Carl J. Leuschen,[17] Lucia Marinangeli[18]

The Shallow Radar (SHARAD) on the Mars Reconnaissance Orbiter has imaged the internal stratigraphy of the north polar layered deposits of Mars. Radar reflections within the deposits reveal a laterally continuous deposition of layers, which typically consist of four packets of finely spaced reflectors separated by homogeneous interpacket regions of nearly pure ice. The packet/interpacket structure can be explained by approximately million-year periodicities in Mars' obliquity or orbital eccentricity. The observed ~100-meter maximum deflection of the underlying substrate in response to the ice load implies that the present-day thickness of an equilibrium elastic lithosphere is greater than 300 kilometers. Alternatively, the response to the load may be in a transient state controlled by mantle viscosity. Both scenarios probably require that Mars has a subchondritic abundance of heat-producing elements.

The north polar region of Mars is part of an immense sedimentary basin that records over 3 billion years of deposition and erosion of ice, dust, lava, and other lithic material of varying composition and includes the plateau of Planum Boreum, covering an area of ~800,000 km^2. Beneath a thin residual cap of bright, nearly pure water ice, the north polar layered deposits (as a single unit: NPLD) (*1*) make up the bulk of Planum Boreum and are probably the result of cyclical variations in the orbit and rotation of Mars that affected insolation at the poles and elsewhere. The layering results largely from different fractions of dust and ice (*2*, *3*), which affect the visible-wavelength albedo. In the last decade, high-resolution images have revealed a basal unit (BU) (*4*) beneath the NPLD. The BU is considerably different than the NPLD (*5*–*7*), principally in its lower albedo, but it also has thicker layers (*8*). Images show that outcrops in the upper portions of the BU sometimes display a platy, interbedded sequence of ice-rich and sand-rich layers (*9*).

Understanding the origin and evolution of both the NPLD and the BU will provide important information on the recent climate history of Mars. These deposits also act as loads on the underlying lithosphere, and the amount of downward deflection of the substrate constrains the thermal and rheological state of the present-day martian interior. The Shallow Radar (SHARAD) and Mars Advanced Radar for Subsurface and Ionospheric Sounding (MARSIS) [see supporting online material (SOM) text] penetrate these deposits, providing a three-dimensional view of north polar stratigraphy and the deformation of its base. Here we present results from these radar sounders and discuss the genesis and evolution of these north polar units, the relation of these deposits to climate signals, and the geodynamical state of the interior.

Radar stratigraphy. Figure 1 shows a radargram of a SHARAD traverse (orbit 5192) across Planum Boreum, from the outlying plateau of Olympia Planum near 140°E to the minor lobe of Planum Boreum, Gemina Lingula, near 5°E. Here, radar reflections arise from boundaries between layers differing in their fractions of ice, dust, and sand. The radar reflections in Planum Boreum are closely spaced to a depth of ~500 m and then are more widely spaced at greater depths, though they are clustered into distinct packets of reflectors. The volume containing all of these reflectors corresponds to the NPLD, whereas the diffusely reflective zone below ~2 km of depth (mostly absent under Gemina Lingula) corresponds to the BU. Orbit 5192 crosses a number of spiral troughs that are up to several hundred meters deep. The troughs interrupt the shallow radar reflectors, though the timing of layer deposition and trough formation is ambiguous (*10*).

Mapping of BU exposures in chasmata on the periphery of Planum Boreum led to the hypothesis that the BU is absent beneath Gemina Lingula (*5*, *6*) and that the NPLD here rests directly on the few-billion-year-old Vastitas Borealis Formation (VBF) of largely sedimentary origin; our examination of many SHARAD tracks across Planum Boreum confirms this hypothesis and shows that the BU is largely confined to the major lobe of Planum Boreum. Additionally, where the radar-mapped BU unambiguously intersects the floor of Chasma Boreale, BU is observed in those locales in surface images. A radargram from SHARAD orbit 5220 (Fig. 2) shows that in this locale the BU actually extends across Chasma Boreale, where it is exposed in the walls and forms a bench on the Gemina Lingula side before terminating abruptly. The BU increases to a thickness of ~1000 m beneath the highest portions of Planum Boreum, whereas the NPLD has a roughly constant thickness, except near its boundaries. Thus, elevation variations across Planum Boreum are largely attributable to thickness variations in the BU, and it follows that depositional and erosional processes have been largely uniform across the cap since the bottommost layers of the NPLD were emplaced. This conclusion is corroborated by the observation that some of the stronger radar reflectors can be traced completely across the cap (Fig. 1), a distance of nearly 1000 km.

SHARAD signals locally penetrated the BU to reveal a substrate reflection (presumably the top of the VBF), but this reflection in many places is obscured by volume scattering of the radar waves within the BU (Fig. 1). The BU does show internal reflectors, as expected from limited surface exposures of this unit (*8*). The substrate reflector can be traced ~100 km beneath Olympia Planum in Fig. 1, which implies that the material making up the bulk of Olympia Planum was deposited on a common substrate with the BU. The lower-frequency MARSIS radar, which is able to penetrate more deeply and is apparently less sensitive to volume scattering, shows these relations more clearly (fig. S1). As projected in earlier work (*6*), the SHARAD results show that the maximum thickness of the BU beneath the NPLD is approximately the elevation of Olympia Planum above its lowland surroundings, suggesting a common link. However, the BU appears to terminate on the boundary of Planum Boreum, lacking continuity with Olympia Planum (Fig. 1). The substrate beneath the NPLD and the BU is remarkably flat, showing essentially no deflection from the mass load of the polar deposits. This lack of deflection was noted in an early MARSIS radar profile, but low range resolution provided only a crude bound on elastic lithosphere thickness (*11*), whereas SHARAD data place tighter geodynamical constraints because of the instrument's superior range resolution.

SHARAD data for Orbit 5297 (Fig. 3) show more detail in the Gemina Lingula region. The marked flatness of the basal interface is apparent, as is a repetition of the fine-scale layering observed at

[1]Southwest Research Institute, Boulder, CO 80302, USA. [2]Washington University, St. Louis, MO 63130, USA. [3]Massachusetts Institute of Technology, Cambridge, MA 02139–4307, USA. [4]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA. [5]University of Colorado, Boulder, CO 80309–0392, USA. [6]Brown University, Providence, RI 02912, USA. [7]U.S. Geological Survey, Flagstaff, AZ 86001, USA. [8]Smithsonian Institution, Washington, DC 20013–7012, USA. [9]Dipartimento InfoCom, Università di Roma "La Sapienza," I-00184 Rome, Italy. [10]Istituto Nazionale di Astrofisica, I-00133 Rome, Italy. [11]Scripps Institution of Oceanography, La Jolla, CA 92093, USA. [12]University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z4. [13]Lunar and Planetary Institute, Houston, TX 77058, USA. [14]Institute de Physique du Globe de Paris, Paris, France. [15]Università Roma Tre, 00146 Rome, Italy. [16]Jackson School of Geosciences, University of Texas, Austin, TX 78712, USA. [17]University of Kansas, Lawrence, KS 66045–7612, USA. [18]Universita' d'Annunzio 65127 Pescara, Italy.

*To whom correspondence should be addressed. E-mail: roger@boulder.swri.edu

the surface. Separated by interpacket regions of few reflections, there are four such layer packets of closely spaced reflections (also seen in Fig. 1), whose signal strength becomes weaker at depth, as expected (fig. S2). Inspection of other radargrams confirms the persistence of four packet structures across the cap, with local variation due to heterogeneous erosion and deposition leading in places to missing packets or extra packets (for instance, Fig. 3, bottom center). This packet/interpacket structure represents a second, larger scale of layering in addition to the finer-scale reflectors within the packets.

The observed dielectric transparency of the NPLD to radar waves in the frequency band for SHARAD (Figs. 1 and 3) and also for MARSIS data (*11*) implies that the amount of dust in the total NPLD column must be small: Electrical loss is high in non-ice analog materials (fig. S3) and in SHARAD non-ice targets on Mars. A low dust content is consistent with spectral mapping (*12*, *13*), which has shown that ice dominates the NPLD, and by analogy with a bulk density estimate of the south polar layered deposits (SPLD) (*14*). Following a well-established approach (*15*), we found that the weakest reflectors detected by SHARAD could be ice layers as thin as 10 cm containing ~10% dust or 1-m-thick layers containing only ~2% dust. The strongest reflectors can be explained by a 2-m layer of ice containing ~30% dust. We conclude that reflections from within the packets are caused by variations of small dust fractions contaminating an ice matrix, and that the relatively homogeneous interpacket regions are largely ice with a minimal dust component.

Climate implications. Three well-understood orbital and rotational periodicities affect insolation at Mars' poles (*16*): (i) an obliquity variation (120,000 years), (ii) a climatic precession (*17*) (51,000 years), and (iii) an eccentricity variation (95,000 to 99,000 years); the first has the largest influence on martian climate (*18*). It has long been held that fine-scale layering observed in images of the NPLD is tied somehow to these climate forcings (*2*). The fine-scale intrapacket reflections revealed by SHARAD may also be tied to one or more of these periodicities, but it is clear that the radar cannot resolve the finest scale of layering observed in images and that some of this layering is related to surficial processes alone (*9*). The actual layering, whether observed in images or radargrams, cannot yet be uniquely tied to specific climate forcings, and the finest scale of layering could conceivably reflect individual, very intense dust storms. Because physical layers mapped by SHARAD are continuous laterally up to 1000 km, whatever mechanisms that were operating to create layering must have been homogeneous over that scale, and climate forcing is the leading candidate. The packets represent widespread processes that added and removed ice, added dust, and concentrated dust during times of ice sublimation because the mere presence of radar reflections of variable strength almost certainly requires the ice/dust ratio to vary with depth. Climate modeling suggests that the NPLD postdates the last transition from high to low mean obliquity (LMO) ~5 million years ago (Ma), before which ice at the pole was unstable against complete loss by sublimation (*19*). This conclusion is consistent with several results that correlate variations in depth-dependent layer brightness with the time series of an insolation model (*16*, *20*), but it is not possible to establish this connection uniquely at present.

From the SHARAD constraints, the three regions between the packets with few radar reflections represent times of substantial ice deposited at the pole and/or little dust delivered to the pole. One simple model that yields the packet/interpacket structure is a variable dust supply over a background of overall ice growth during the last ~5 million years (My). Dust storms may be more prevalent during high obliquity at solstice (*21*), and this effect is probably enhanced at times of high eccentricity; conversely, at low obliquities and eccentricities, dust storm activity and intensity should diminish (*22*). Packets could form during times of high obliquity (and eccentricity), and sublimation events during short-period excursions in

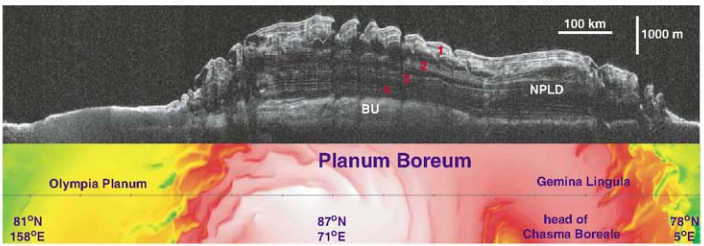


Fig. 1. (Top) Radargram from SHARAD orbit 5192. Range time delay, the usual ordinate in a radargram, has been converted to depth by assigning real permittivities of 1 and 3 above and below the detected ground surface, respectively. NPLD and the BU are labeled. Packet regions are numbered (see text). **(Bottom)** Ground track of orbit 5192 shown on a digital elevation model (DEM) derived from Mars Orbiter Laser Altimeter (MOLA) data. Elevation range is ~−4.5 (green) to −2 km (white).

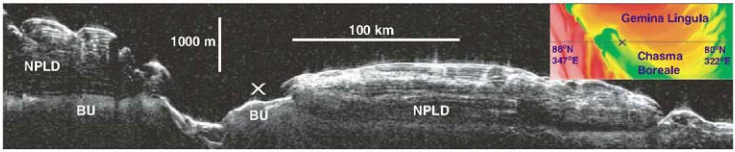


Fig. 2. Radargram from SHARAD orbit 5220 with time-to-depth algorithm applied (see Fig. 1 legend). Image contrast has been truncated to emphasize the NPLD/BU contact, so the depth extent of the BU is not apparent beneath the main portion of Planum Boreum. **(Inset)** Ground track location on MOLA DEM. The "×" marks location in radargram and on DEM of a bench of BU on the Gemina Lingula side of Chasma Boreale; the thickness of the BU rapidly terminates within Gemina Lingula.

insolation would concentrate the dust in layers (*19*). Interpacket regions could correspond to times of lows in the 2.4-My solar system dynamical resonance modulation of eccentricity that have occurred ~1.4, ~3.8, and ~6.2 Ma (*16*), with the earliest age slightly older than the calculated transition to the present LMO epoch of stable north polar ice. A second model proposes that the interpacket regions correspond to relatively low-obliquity phases at ~0.8, ~2.0, and ~3.2 Ma within the overall LMO epoch of the last 5 My. These would be times of relatively high rates of ice return to the poles within an overall net-accretion scenario (*19*). The low-obliquity scenario is perhaps more consistent with models of NPLD construction within the last 5 My (*19*), though growth rate and timing in such models depend on, among other things, assumptions regarding sublimation processes.

If the NPLD somehow survived epochs of high obliquity, each interpacket region could then represent a distinct, earlier epoch of LMO that is not predictable, as obliquity solutions become chaotic beyond 10 to 20 Ma (*23*). A problem with this hypothesis is that it does not explain the structure at the top of the NPLD column, which should be an interpacket region during the present LMO era. Instead, this section of the NPLD contains abundant reflective layers, and no interpacket region is observed with SHARAD to within its 10 to 20 m range resolution near the surface. Stratigraphic relations and geologic unit ages inferred from crater densities allow the NPLD below the uppermost several layers to range in age from a few million years to ~1 billion years (*8*), which is not a strong constraint.

Deflection of the substrate—geodynamical implications. The deflection of the substrate boundary in Gemina Lingula, where the NPLD is presumably in direct contact with the VBF, is at most 100 m (see SOM text). We used a spherical-shell loading model (*24*) to calculate deflection of the elastic lithosphere in response to the cap load (*25*), seeking the minimum value of elastic thickness, T_e, that is consistent with the 100-m value. A T_e value of 300 km satisfies this criterion (fig. S4) and can be taken as the lower bound on T_e in this region of Mars for the current era.

This exercise can be repeated for the SPLD load, where MARSIS data have been used to effectively strip off the volume of the deposits, revealing the topography of the substrate (*26*). It is difficult to estimate a flexural signal here because the substrate consists of rugged highlands topography. There is a tendency in some longitudinal sectors for a decrease in substrate elevation from the periphery toward the center of the SPLD load, although this could be a result of large impact basins near the pole. Smooth-surface fits to the data suggest that the elevation decreases as much as ~200 to 250 m, corresponding to a lower bound on T_e of 275 to 300 km. This value could reflect solely the noise of the substrate topography, which obscures any smaller flexural signal. Independent of this work, a gravity-topography spectral estimate of the SPLD load response has yielded a T_e best-fit estimated value of 140 km, though any $T_e > 102$ km satisfies the data constraints (*27*). From the substrate deflections, the lower bound for both polar loads, ~300 km, is at least a factor of 3 higher than nearly all previous estimates of T_e on Mars (*28*).

At present, T_e estimates are essentially the only way to infer heat flow and temperatures for the martian shallow interior at points during its history. The most straightforward way to estimate these thermal parameters is to equate, for a given deflection curvature, the bending moment for a given T_e to the bending moment of an elastic-plastic lithosphere (*29*), where for the latter the temperature dependence enters through a specified viscous creep relation. To calculate the elastic-plastic bending moment, we used a relatively low basal yield stress of 10 MPa, which produces considerable strain at high temperatures in the mantle (*30*) and errs on the side of producing either higher temperatures for a given lithospheric thickness or thicker lithospheres for the same temperatures. The temperature estimate was adopted from a standard chondritic thermal model for mantle heat flux into the base of the lithosphere and crustal heat-source enrichment (*31*); the model can be scaled up and down from its chondritic values. For a diabase crust (*32*) and an olivine mantle (*30*), a 300-km lithosphere corresponds to a wet or dry olivine mantle containing ~70 or ~80% of chondritic heat sources, respectively (see SOM text).

We tested the plausibility of this result using an inverse approach. We questioned what T_e value would be obtained if the polar loads were emplaced with the same chondritic fractions earlier in martian history (say at 3 Ga) and whether this number is consistent with T_e values obtained elsewhere on the planet for features of this age. The answer obtained for both wet and dry olivine at the same subchondritic fractions as above is T_e ~ 140 km. This value is on the high side relative to most other estimates made for this epoch (*28*), although such estimates carry considerable uncertainty (*33*), and most of these T_e values arise from areas that are associated with magmatism (e.g., shield volcanoes). Subsurface heat flow has certainly been spatially heterogeneous over the history of Mars, having been anomalously high at the large magmatic centers such as Tharsis and Elysium but less than the planetary average elsewhere, including the polar regions. This is an alternative explanation to the hypothesis of planetwide subchondritic heat sources, though it may be less plausible if these volcanic regions were generated by mantle plumes (*34*) driven by core energy.

It is possible that the small amount of polar substrate deflection observed results from a transient state of the planet's interior in response to the polar loads. That is, not enough time has passed since the loads were emplaced for the deflection to reach elastic equilibrium. In that case, the 300-km bound on T_e is not a valid result; instead, the SHARAD constraint (≤100-m deflection) provides information about the viscosity structure of the mantle. From the discussion above, we suggest that a lower bound on the age of the NPLD is on the order of 1 Ma, taking into account the growth time of the load. We used chondritic thermal models (*31*, *35*) to set up viscoelastic loading models (*36*) and

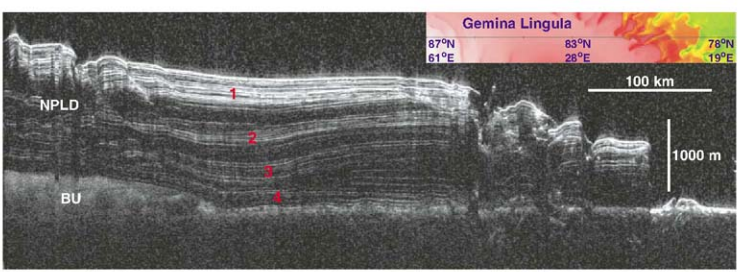


Fig. 3. Radargram from SHARAD orbit 5297 with time-to-depth algorithm applied (see Fig. 1 legend). (**Inset**) Ground track location on MOLA DEM. Packet regions are numbered (see text).

found that they generated too much deflection (see SOM text and fig. S5). This is an alternative way of arriving at the conclusion that Mars probably has subchondritic heat sources.

References and Notes

1. K. R. Blasius, J. A. Cutts, A. D. Howard, *Icarus* **50**, 140 (1982).
2. J. A. Cutts, B. H. Lewis, *Icarus* **50**, 216 (1982).
3. P. C. Thomas, S. W. Squyres, K. E. Herkenhoff, A. D. Howard, B. C. Murray, in *Mars*, H. H. Kieffer, B. M. Jakosky, C. W. Snyder, M. S. Matthews, Eds. (Univ. of Arizona Press, Tucson, AZ, 1992), pp. 767–795.
4. M. C. Malin, K. S. Edgett, *J. Geophys. Res.* **106**, 23429 (2001).
5. K. E. Fishbaugh, J. W. Head, *Icarus* **174**, 444 (2005).
6. S. Byrne, B. C. Murray, *J. Geophys. Res.* **107**, 5044 (2002).
7. K. S. Edgett, R. M. E. Williams, M. C. Malin, B. A. Cantor, P. C. Thomas, *Geomorphology* **52**, 289 (2003).
8. K. L. Tanaka *et al.*, *Icarus*, published online 29 February 2008; 10.1016/j.icarus.2008.01.021.
9. K. E. Herkenhoff, S. Byrne, P. S. Russell, K. E. Fishbaugh, A. S. McEwen, *Science* **317**, 1711 (2007).
10. Features in the radargram beneath the troughs may or may not be real structures. The discontinuities in reflections might arise, for example, from poor radar-signal penetration because of surface scattering from the trough floors.
11. G. Picardi *et al.*, *Science* **310**, 1925 (2005), published online 29 November 2005; 10.1126/science.1122165.
12. W. M. Calvin *et al.*, *Lunar Planet. Sci. XXXIX*, 1939 (abstr.) (2008).
13. Y. Langevin *et al.*, *Science* **307**, 1581 (2005), published online 17 February 2005; 10.1126/science.1109438.
14. M. T. Zuber *et al.*, *Science* **317**, 1718 (2007).
15. D. C. Nunes, R. J. Phillips, *J. Geophys. Res.* **111**, E06S21 (2006).
16. J. J. Laskar, B. Levrard, J. F. Mustard, *Nature* **419**, 375 (2002).
17. The climate precession is the insolation effect of the precession of the spin axis modulated by the orbital eccentricity. The result is that the calendar date of perihelion precesses.
18. M. T. Mellon, B. M. Jakosky, *J. Geophys. Res.* **100**, 11781 (1995).
19. B. Levrard, F. Forget, F. Montmessin, J. Laskar, *J. Geophys. Res.* **112**, E06012 (2007).
20. S. M. Milkovich, J. W. Head III, *J. Geophys. Res.* **110**, E01005 (2005).
21. R. M. Haberle, J. R. Murphy, J. Schaeffer, *Icarus* **161**, 66 (2003).
22. O. B. Toon, J. B. Pollack, W. Ward, J. A. Burns, K. Bilski, *Icarus* **44**, 552 (1980).
23. J. Laskar *et al.*, *Icarus* **170**, 343 (2004).
24. W. B. Banerdt, *J. Geophys. Res.* **91**, 403 (1986).
25. C. L. Johnson *et al.*, *Icarus* **144**, 313 (2000).
26. J. J. Plaut *et al.*, *Science* **316**, 92 (2007), published online 14 March 2007; 10.1126/science.1139672.
27. M. A. Wieczorek, *Icarus*, published online 10 January 2008; 10.1016/j.icarus.2007.10.026.
28. M. Grott, D. Breuer, *Icarus* **193**, 503 (2008).
29. M. K. McNutt, *J. Geophys. Res.* **89**, 11180 (1984).
30. S.-i. Karato, P. Wu, *Science* **260**, 771 (1993).
31. S. A. Hauck II, R. J. Phillips, *J. Geophys. Res.* **107**, 5052 (2002).
32. Y. Caristan, *J. Geophys. Res.* **87**, 6781 (1982).
33. V. Belleguic, P. Lognonné, M. Wieczorek, *J. Geophys. Res.* **110**, E11005 (2005).
34. C. L. Johnson, R. J. Phillips, *Earth Planet. Sci. Lett.* **230**, 241 (2005).
35. D. Breuer, T. Spohn, *J. Geophys. Res.* **108**, 5072 (2003).
36. P. S. Mohit, R. J. Phillips, *J. Geophys. Res.* **111**, E12001 (2006).

37. SHARAD was provided by the Italian Space Agency, and its operations are led by Dipartimento InfoCom, Università di Roma "La Sapienza." Thales Alenia Space Italia is the prime contractor for SHARAD. The Mars Reconnaissance Orbiter mission is managed by the Jet Propulsion Laboratory, California Institute of Technology, for the NASA Science Mission Directorate, Washington, DC. Lockheed Martin Space Systems, Denver, CO, is the prime contractor for the orbiter. We thank three anonymous reviewers for very constructive comments. S. Hauck provided very useful advice.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1157546/DC1
Materials and Methods
Figs. S1 to S5
References

10 March 2008; accepted 7 May 2008
Published online 15 May 2008;
10.1126/science.1157546
Include this information when citing this paper.

Widespread Translational Inhibition by Plant miRNAs and siRNAs

Peter Brodersen,[1] Lali Sakvarelidze-Achard,[1] Marianne Bruun-Rasmussen,[1] Patrice Dunoyer,[1] Yoshiharu Y. Yamamoto,[2] Leslie Sieburth,[3] Olivier Voinnet[1]*

High complementarity between plant microRNAs (miRNAs) and their messenger RNA targets is thought to cause silencing, prevalently by endonucleolytic cleavage. We have isolated *Arabidopsis* mutants defective in miRNA action. Their analysis provides evidence that plant miRNA–guided silencing has a widespread translational inhibitory component that is genetically separable from endonucleolytic cleavage. We further show that the same is true of silencing mediated by small interfering RNA (siRNA) populations. Translational repression is effected in part by the ARGONAUTE proteins AGO1 and AGO10. It also requires the activity of the microtubule-severing enzyme katanin, implicating cytoskeleton dynamics in miRNA action, as recently suggested from animal studies. Also as in animals, the decapping component VARICOSE (VCS)/Ge-1 is required for translational repression by miRNAs, which suggests that the underlying mechanisms in the two kingdoms are related.

MicroRNAs are 20- to 24-nucleotide (nt) RNAs that regulate eukaryotic gene expression posttranscriptionally. Bound to ARGONAUTE (AGO) proteins, miRNAs guide RNA-induced silencing complexes (RISCs) to partly or fully complementary mRNAs (*1*). Two modes of negative regulation by RISC exist: (i) translational repression, sometimes coupled to accelerated mRNA decay, and (ii) RISC-catalyzed endonucleolytic mRNA cleavage ("slicing"). The degree of miRNA-mRNA complementarity is a key determinant of the mechanism used, such that perfect complementarity enables cleavage, whereas central mismatches exclude slicing to promote translational repression (*2*, *3*). Unlike most animal miRNAs, most plant miRNAs show near-perfect or perfect complementarity to their targets, and slicing is believed to be their predominant, or exclusive, mode of action (*4*). Accordingly, *Arabidopsis* AGO1 binds miRNAs and displays slicer activity toward miRNA targets, and strong *ago1* loss-of-function mutants overaccumulate miRNA target transcripts (*5*–*7*). Nonetheless, many questions regarding miRNA-RISC composition, loading, and target identification persist, mostly because AGO1 remains the only factor known to be implicated in plant miRNA action. It is also unclear whether near-perfect complementarity within plant miRNA–target pairs actually excludes translational inhibition, as is commonly inferred, or whether it allows slicing to occur in addition to translational inhibition.

To address these issues, we carried out a forward genetic screen for *Arabidopsis* mutants defective in silencing of a constitutively expressed green fluorescent protein (GFP) mRNA containing a miR171 target site immediately downstream of the stop codon (Fig. 1A) (*8*). In seedlings, GFP is silenced by endogenous miR171, except in the vasculature and in the roots where miR171 expression is low (*8*). Twenty-one recessive mutants defective in miR171-guided silencing were identified by gain of GFP expression in leaves (Fig. 1B); none had mutations in the miR171 target site. Eight nonallelic mutants showing consistently higher GFP expression than the parental line GFP171.1 [wild type (WT)] after two backcrosses (fig. S1) were studied in further detail.

***mbd* and *mad* mutants.** Two mutants (class I) had strongly reduced levels of several miRNAs and were referred to as *microRNA biogenesis deficient* (*mbd1* and *mbd2*) (fig. S1 and Fig. 1C). miRNA biogenesis in *Arabidopsis* involves processing of primary miRNA transcripts by a nuclear-localized complex of DICER-LIKE1 (DCL1), the double-stranded RNA binding protein HYL1, and the zinc finger protein SE. Excised miRNA/miRNA* duplexes are then stabilized through HEN1-catalyzed 2′-*O*-methylation [reviewed in (*9*)]. A G-to-A transition in *mbd1* (renamed *dcl1-12*) disrupts a splice donor site in the *DCL1* gene, which strongly reduces the accumulation of correctly spliced DCL1 mRNA (fig. S1). *mbd2* (renamed *hen1-7*) has a missense mutation resulting in a Gly-to-Glu change in the HEN1 *S*-adenosyl methionine–binding motif, which is predicted to abolish small RNA methylation (fig. S1). The remaining six mutants exhibited normal miRNA levels and were classified as *microRNA action deficient mad1* to *mad6* (Fig. 1C). *mad1* to *mad6* showed low miR160*

[1]Institut de Biologie Moléculaire des Plantes du CNRS, Unité Propre de Recherche 2357, 12 rue du Général Zimmer, 67084 Strasbourg Cedex, France. [2]Center for Gene Research, Nagoya University, Furo-che, Chikusa-ku, Nagoya Aichi, 464-8602, Japan. [3]Department of Biology, University of Utah, Salt Lake City, UT 84112, USA.

*To whom correspondence should be addressed. E-mail: olivier.voinnet@ibmp-ulp.u-strasbg.fr

accumulation, indicating intact strand separation and miRNA* degradation (Fig. 1C). All *mad* mutants map to loci not previously implicated in RNA silencing (table S1). Thus, our screen identifies known miRNA biogenesis genes and unknown factors required for miRNA action.

Two classes of *mad* mutants. Because miR171 is perfectly complementary to its target site and guides slicing (*8*, *10*), GFP mRNA levels were expected to be elevated in *mad* mutants as compared with WT plants. Indeed, *dcl1-12*, *hen1-7*, and *mad1* to *mad4* mutants exhibited higher levels of GFP mRNA and protein than WT plants, whereas the mRNA and protein levels of the non-miRNA target Hsc70 were unchanged (Fig. 2, A and B). The elevated GFP mRNA levels resulted from defective slicing, because the ratios of GFP full-length mRNA to GFP 3′-cleavage fragments were higher than in WT plants (fig. S1). By contrast, *mad5* and *mad6* mutants had low GFP mRNA levels similar to those in WT plants, yet had much higher GFP protein levels (Fig. 2, A and B, and fig. S1). The overaccumulation of GFP protein was specifically due to defective miR171-directed repression because an mRNA lacking the miR171 target site ($GFP_{no\ miR}$) produced similar GFP levels in *mad5*, *mad6*, and WT plants (Fig.

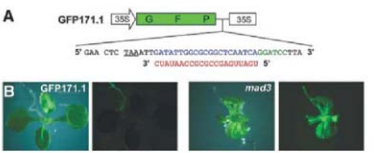
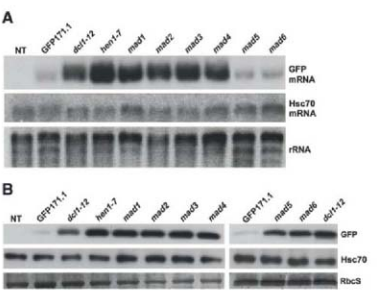
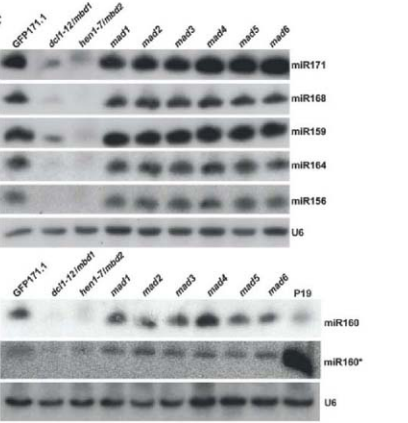


Fig. 1. Isolation of *mbd* and *mad* mutants. **(A)** Schematic representation of the GFP171.1 construct reporting miR171 activity. The miR171 target sequence was inserted downstream of the stop codon (underlined). **(B)** Example of loss of GFP silencing in a *mad* mutant (*mad3*, 15 days of growth). From left to right, the first and third images are in transmitted light, whereas the second and fourth images show GFP fluorescence upon blue light excitation. **(C)** Northern analysis of distinct miRNAs (top panels) and comparison of miR160 and miR160* accumulation (bottom panels). RNA from transgenic seedlings expressing the tombusviral P19 protein, which sequesters small RNA duplexes, provides a positive control for miRNA* overaccumulation.





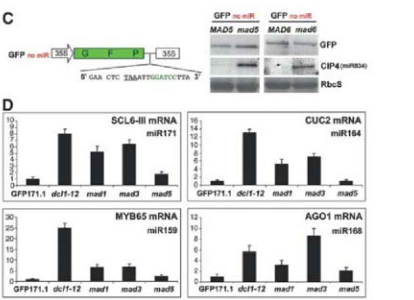


Fig. 2. Molecular analysis of mutants. **(A)** RNA blot analysis with GFP- and Hsc70-specific probes. Ribosomal RNA (rRNA) was stained with ethidium bromide. NT, nontransgenic WT plant. **(B)** Western analysis of GFP and Hsc70. Coomassie-stained RbcS provides a loading control. **(C)** Representation of the $GFP_{no\ miR}$ transgene devoid of a miR171 target site and GFP protein accumulation from $GFP_{no\ miR}$ introduced into *mad5* and *mad6* by crossing (see the SOM materials and methods). Differential accumulation of CIP4 (miR834 target) confirms the presence of the mutations (Fig. 3E). **(D)** Quantitative RT-PCR analysis of mRNA accumulation of four endogenous miRNA targets. cDNA inputs were normalized to 18S rRNA, and expression ratios for each mRNA are given relative to the level in the GFP171.1 parental line. Data are displayed as averages ± SD (n = 3 PCR replicates).

2C). Accordingly, accumulation of the non-miRNA targets RbcS, HSC70 (Fig. 2B), and CDC2A (fig. S1) was unchanged in the two mutants as compared with WT plants. Quantitative reverse transcription polymerase chain reaction (RT-PCR) analyses of representative mutants revealed that *mad1* and *mad3* overaccumulate endogenous transcripts known to be targeted by miRNAs other than miR171, whereas little or no difference was observed in *mad5* (Fig. 2D). These data indicate that *mad1→mad4* mutants carry lesions in genes required for miRNA-guided transcript degradation (class II mutants). By contrast, miRNA-guided slicing occurs normally in *mad5* and *mad6* (class III mutants), yet both fail to silence the synthetic GFP171.1 target at the protein level. We propose that the low GFP accumulation in the transgenic line GFP171.1 is due to at least two distinct mechanisms mediated by the perfectly complementary miR171: (i) slicing leading to reduced transcript levels and (ii) inhibition of protein production from the remaining unsliced mRNAs. We further propose that class III mutants are specifically defective in the second process.

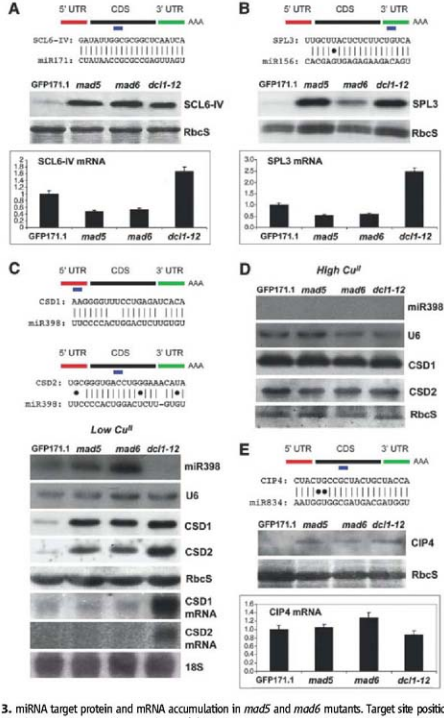


Fig. 3. miRNA target protein and mRNA accumulation in *mad5* and *mad6* mutants. Target site positions and miRNA-mRNA complementarities are shown. **(A)** Western and real-time RT-PCR analyses of SCL6-IV protein and mRNA, as in Fig. 2D. **(B)** SPL3 protein and mRNA analyzed as in (A). **(C)** mRNA and protein accumulation of the miR398 targets CSD1 and CSD2 under low Cu(II) availability. **(D)** Same as in (C), but under high Cu(II) availability. The miR398 and U6 signals shown in (C) and (D) are from the same exposure of the same membrane containing both low- and high-Cu(II) samples. **(E)** CIP4 protein and mRNA accumulation, as in (A).

Widespread translational repression by plant miRNAs. Many plant miRNAs, including miR171, target mRNAs within coding sequences, whereas most animal miRNA sites are located in 3′ untranslated regions (3′UTRs). Moreover, SPL3—one of only two examples of plant mRNAs translationally repressed by miRNAs—is targeted in the 3′UTR (*11*–*13*). Consequently, a concern was that the strong effects on protein levels might have been a result of the artificial miR171 target site position in the 3′UTR of GFP171.1, which, moreover, is a constitutively expressed transgene. Therefore, we compared mRNA and protein accumulation of several endogenous miRNA targets, representing all possible target site locations within mRNAs [5′UTR, coding sequence (CDS), and 3′UTR]: two transcription factors targeted by evolutionarily conserved miRNAs (SCL6-IV/miR171/CDS and SPL3/miR156/3′UTR), three stress-related genes also targeted by conserved miRNAs (CSD1/miR398/5′UTR, CSD2/miR398/CDS, and APS1/miR395/CDS), and one target of a nonconserved miRNA known only in *Arabidopsis* (CIP4/miR834/CDS).

mad5 and *mad6* mutants displayed increased protein levels from all these target transcripts. By contrast, both mutants accumulated WT mRNA levels of SCL6-IV, SPL3, CIP4, CSD1, and CSD2 (Fig. 3, A to C and E), although elevated CSD2 mRNA levels were detected in one of three experiments (fig. S2). *mad6*, but not *mad5*, also consistently showed about fourfold higher APS1 mRNA levels (fig. S2). Accumulation of miR171, miR156, and miR398 was unchanged or slightly increased in class III mutants as compared with WT plants (Figs. 2A and 3C), whereas miR395 and miR834 were below the detection limit of Northern analyses (*14*, *15*). Overaccumulation of CSD1 and CSD2, as compared with that in WT plants, was observed in mutant seedlings grown under low Cu(II) availability, allowing miR398 accumulation but not in the presence of high Cu(II) levels, which prevents miR398 expression [Fig. 3, C and D (*16*)]. The correlation between presence of miR398 and deregulation of its targets indicates that class III mutants are specifically affected in miRNA-guided regulation. We note that *dcl1-12* did not exhibit much stronger overaccumulation of CSD1 and SPL3 protein as compared with that observed in *mad5* and *mad6*, despite clear reduction of the corresponding miR398 and miR156 levels. We conclude that (i) translational repression is a widespread mode of plant miRNA action, irrespective of the degree of complementarity or location of target sites within mRNAs; (ii) this process can be genetically uncoupled from miRNA-directed slicing, as in *mad5* and *mad6*; and (iii) certain plant miRNAs exhibiting perfect or near-perfect complementarity to a single target site can repress mRNA expression predominantly at the translation level.

***MAD5* encodes the microtubule-severing enzyme KATANIN.** Positional cloning showed that *mad5* carries a G-to-A transition in the start codon of *KTN1* (Arabidopsis Genome Initiative number:

AT1G80350), causing strongly reduced KTN1 protein levels (Fig. 4A). *KTN1* encodes the P60 subunit of the microtubule-severing enzyme KATANIN. Transformation of *mad5* with a genomic *KTN1* fragment (*17*) restored GFP171.1 silencing (Fig. 4, A and B). Molecular analyses of three previously characterized *ktn1* mutant alleles—*fra2*, *lue1*, and *erh3-3*—showed that each overaccumulates CSD2 and SPL3 proteins without corresponding increases in mRNA levels (Fig. 4C). These results demonstrate that *MAD5* is allelic to *KTN1*. Adenosine Triphosphate (ATP)–dependent microtubule severing is the only known function of KTN1, and because the mutant ktn1 protein produced in *erh3-3* (Fig. 4C) carries a missense mutation in the ATP binding site (*18*, *19*), our data suggest that microtubule dynamics play a role in miRNA-guided translational inhibition but not in miRNA-guided cleavage.

Requirement of the mRNA decapping factor VCS. Components of the decapping complex—DCP1, DCP2, and Ge-1—are required for miRNA-guided translational repression in animals (*20*). Mutations in *Arabidopsis DCP1*, *DCP2*, and the Ge-1 homolog *VCS* are lethal in ecotype Col-0, but as-yet-unidentified modifier loci in ecotype Ler suppress seedling lethality of decapping-deficient *vcs* alleles (*21*, *22*). To test whether plant miRNA–guided translational repression involves components similar to those required in animals, we examined these mutant alleles for miRNA target accumulation. Our analysis included both the homozygous, viable *vcs-1* in Ler, and heterozygous individuals of the seedling-lethal *vcs-7* mutant in Col-0.

vcs-1 mutants exhibited elevated levels of SPL3, SCL6-IV, and CIP4 protein with little or no increase in corresponding mRNA levels (Fig. 5A). No effect was observed on CSD1 and CSD2 protein levels (fig. S3). The *vcs-7* mutation was dominant at the molecular level, despite the recessive seedling lethality phenotype. Thus, *vcs-7* heterozygotes showed overaccumulation of SPL3 and SCL6-IV proteins (fig. S3). In addition, as compared with WT plants, *vcs-7* heterozygotes showed CSD2 protein overaccumulation under low Cu(II) availability, without increases in CSD2 mRNA levels (Fig. 5B). The increased protein levels of miRNA targets in *vcs* mutants were the result of defective miRNA action, because (i) non-miRNA targets (Hsc70, CDC2A) accumulated normally in *vcs* mutants (fig. S3); (ii) CSD2 overaccumulation in *vcs-7* correlated with the presence of miR398 (Fig. 5B); and (iii) upon introgression into the GFP171.1 line, but not into the $GFP_{no\ miR}$ line (both in accession C24), *vcs-7* led to increased GFP protein accumulation in seedlings (Fig. 5, C and D). We conclude that, as in animals, the decapping component VCS is required for miRNA-guided translational repression in plants.

Involvement of AGO1 and AGO10. Next, we asked what AGO protein(s) might be responsible for miRNA-guided translational repression in plants. Four of 10 *Arabidopsis* AGO proteins play established roles in small RNA–directed functions: AGO4 and AGO6 mediate DNA methylation, and AGO7 is involved in biosynthesis of some trans-acting siRNAs, whereas AGO1 can function as a miRNA-guided slicer [reviewed in (*9*)]. AGO2, AGO5, and AGO7 do not interact with most miRNAs (*23*, *24*). Mutations in *AGO10/PNH/ZLL*—the closest paralog of *AGO1*—have been isolated based on their defective shoot apical meristem phenotype (*25*). *ago1* and *ago10* mutants show overlapping developmental defects, and strong alleles are synthetically lethal, suggesting their involvement in similar pathways (*26*). To test the possible role of AGO10 in translational repression, we used the frameshift *ago10* mutant allele, *zll-15* (ecotype Ler). CSD2 mRNA levels were slightly higher in *zll-15* than in WT plants, but CSD2 protein levels were disproportionately higher under low Cu(II) availability (Fig. 6A). CSD2 overaccumulation correlated with the presence of miR398 [as conditioned by Cu(II) availability], and the non-miRNA targets Hsc70 and CDC2A were unaffected, which suggests that this effect is miRNA-dependent (fig. S3). SCL6-IV also showed elevated protein levels in *zll-15*, with SCL6-IV mRNA and miR171 accumulation remaining unchanged (Fig. 6A). The levels of SPL3 were unaffected in *zll-15* (fig. S3), and we did not detect CIP4 in either Ler or *zll-15*. These results suggest that AGO10 is involved in translational repression of only some miRNA targets and that one or several additional AGO protein(s) might contribute to this process.

Hypomorphic *ago1-27* mutants (ecotype Col-0) exhibit near WT accumulation of many miRNA target transcripts, yet display morphological defects similar to *dcl1* mutants (*7*, *27*), suggesting that AGO1 could be involved in translational repression. CSD2 accumulation was increased in *ago1-27*, but this correlated with higher CSD2 mRNA and distinctly lower miR398 levels (fig. S3). By contrast, miR156 and miR171 levels were only moderately reduced in *ago1-27*, whereas the SPL3 and SCL6-IV mRNA levels showed corresponding moderate increases, as compared with those in WT plants (Fig. 6B). Nonetheless, *ago1-27* exhibited disproportionately higher SPL3 and SCL6-IV protein levels. *ago1-27* mutants also displayed elevated CIP4 protein levels with no appreciable difference in mRNA accumulation, but we could not verify any possible effect on miR834 accumulation due to its low abundance. Protein levels of the non-miRNA targets Hsc70 and CDC2A remained unaltered (fig. S3). These results suggest that AGO1 may indeed contribute to miRNA-directed translational repression of SPL3, SCL6-IV, and CIP4 [see also supporting online material (SOM) text].

RNA interference (RNAi) has a translational component in plants. Having established that perfect or near-perfect complementarity is generally compatible with single miRNA species guiding translational repression in plants, we asked whether silencing by populations of siRNAs could also involve translational repression in addition to slicing. To this end, *mad6* and *ago1-27* were introduced into the SUC-SUL (SS) silencing system. In this system, phloem-specific expression of an inverted-repeat (IR) construct triggers non–cell autonomous RNAi of the endogenous mRNA SULFUR (SUL), resulting in a vein-centered chlorotic phenotype. SUL silencing is strictly contingent upon loading of DCL4-dependent 21-nt siRNAs into AGO1

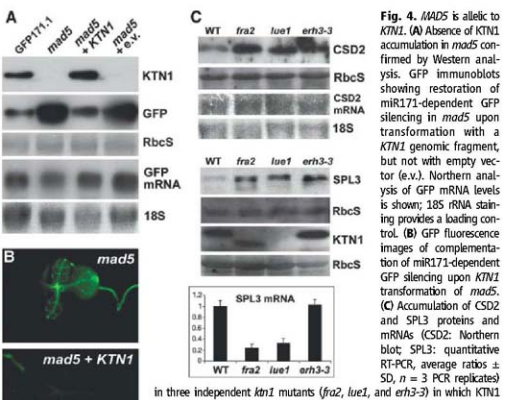


Fig. 4. *MAD5* is allelic to *KTN1*. **(A)** Absence of KTN1 accumulation in *mad5* confirmed by Western analysis. GFP immunoblots showing restoration of miR171-dependent GFP silencing in *mad5* upon transformation with a *KTN1* genomic fragment, but not with empty vector (e.v.). Northern analysis of GFP mRNA levels is shown; 18S rRNA staining provides a loading control. **(B)** GFP fluorescence images of complementation of miR171-dependent GFP silencing upon *KTN1* transformation of *mad5*. **(C)** Accumulation of CSD2 and SPL3 proteins and mRNAs (CSD2: Northern blot; SPL3: quantitative RT-PCR, average ratios ± SD, $n = 3$ PCR replicates) in three independent *ktn1* mutants (*fra2*, *lue1*, and *erh3-3*) in which KTN1 protein levels were assayed. The experiments were performed as in Fig. 3.

(*28*). Therefore, the strong *dcl4-6* mutant was used as a control in those experiments.

The *mad6* and *ago1-27* mutations suppressed SUL silencing (Fig. 6C and fig. S4). Molecular analyses of *mad6* showed a strong decrease in SUL siRNA accumulation, accompanied by a mild increase in SUL mRNA. Nonetheless, SUL protein levels were appreciably higher than those in the SS reference line (fig. S4), which suggests that SUL silencing does not rely exclusively on mRNA degradation. In *ago1-27*, the SUL siRNA levels were unchanged as compared with those in the parental SS line (Fig. 6D), and SUL mRNA levels were nearly as low in *ago1-27* as they were in the parental SS line. SUL protein levels, however, were clearly higher in *ago1-27* mutants (Fig. 6D). This represented suppression of SUL silencing because, without the SS transgene, SUL protein levels were unchanged in *ago1-27* plants as compared with WT plants (fig. S4). We conclude that the hairpin-derived SUL siRNA population mediates translational repression in addition to mRNA degradation.

Discussion. Imperfect pairing with central mismatches in small RNA–target hybrids promotes translational repression because it excludes slicing. It is a common inference that, conversely, near-perfect complementarity excludes translational repression because it enables slicing. This has contributed to the notion that plant and animal miRNAs act in fundamentally different ways. Our finding of a general translational component in plant miRNA and siRNA action demonstrates this inference to be erroneous and provides a genetic foundation to several key observations. First, translational repression by near-perfectly matched miRNAs has been reported twice in *Arabidopsis* (*11*–*13*). Although regarded as exceptions, those examples may well define a stereotype of plant small RNA action. Second, artificial miRNAs can produce phenotypes indistinguishable from genetic knockouts of their targets, despite incomplete target mRNA reduction (*29*). Third, experiments with an inducible RNAi construct identified an extended time window during which target mRNA had returned to its original levels, while protein activity remained suppressed (*30*).

We propose that translational repression is the default mechanism by which small RNAs silence messages, both in plants and animals. Near-perfectly matched small RNAs may in addition engage in slicing such that their regulatory output results from a combination of both mechanisms. The CIP4-miR834 interaction demonstrates that, while necessary, near-complete pairing is not sufficient for slicing to contribute substantially to silencing in plants. miR834 is part of a large group of recently identified nonconserved ("young") miRNAs, many of which are presumed to be nonfunctional because their putative target mRNA levels are unchanged in *dcl1* and *hen1* mutants (*31*). Rather, inspection of the CIP4 protein levels suggests that such young miRNAs might be primarily channeled to translational inhibitory pathways. Nonetheless, these examples do not undermine the importance of slicing in plant biology: The pronounced morphological defects of slicer-deficient *ago1* mutants (SOM text) and *mad1*→*mad4* mutants, as compared with the mild developmental phenotype of *mad5* and *mad6*, suggest that miRNA-guided slicing, not translational inhibition, is indispensable for plant development (fig. S5).

Differences in position, number, and pairing degree of miRNA target sites have been used to substantiate contrasted views of plant and animal miRNA action. However, the unbiased genetic analysis conducted here shows that these features have little or no influence on either the mode or efficiency of miRNA-directed repression. The premise that miRNAs in plants act mostly via slicing has prompted the use of near-perfect complementarity as the exclusive criterion for identification of plant miRNA targets (*14*, *31*, *32*). Our results suggest that the existence of extensively mismatched miRNA targets regulated mostly at the protein level now warrants serious consideration (*33*). Studies of many rice "orphan" miRNAs with no obvious complementary target transcripts could provide a means to investigate this important issue (*34*). Finally, because hairpin-derived siRNA populations in plants act partially via translational inhibition, it is conceivable that the same is true of plant viral siRNAs, which are thought to confer immunity mainly via slicing. The possibility also emerges that RNAi in animals might generally involve translational repression: In one example of RNAi in Sertoli cells, the siRNA target was indeed repressed exclusively at the protein level (*35*).

The finding that AGO1 may concurrently slice and translationally inhibit a given mRNA pool raises the question of how slicing is avoided during translational inhibition. It also remains unclear whether the two mechanisms coexist within the same cells, or whether they can be spatially and/or temporally separated. Tissue specificity of plant miRNA and target expression could also influence the prevalence of one process over the other. The identification of *MAD5* as *KTN1* suggests that dynamic reorganization of the microtubule network is important for miRNA-directed translational re-

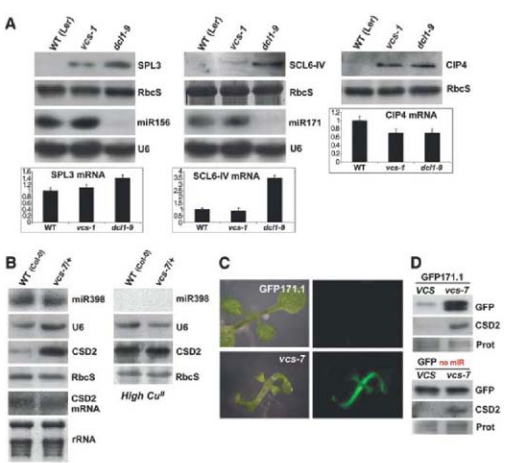


Fig. 5. miRNA target protein and mRNA accumulation in *vcs* mutants. **(A)** Accumulation of target proteins and mRNAs (SPL3, SCL6-IV, and CIP4) corresponding to miR156, miR171, and miR834 in *vcs-1* homozygotes as compared with WT and *dcl1-9* plants. Coomassie-stained RbcS provides a loading control. For real-time RT-PCR analysis of SPL3, SCL6-IV, and CIP4 mRNAs, cDNA inputs were normalized to 18S rRNA. Gene expression ratios are relative to the WT levels (Ler) and shown as averages ± SD (n = 3 PCR replicates). **(B)** Accumulation of miR398, as well as mRNA and protein of its target CSD2, as in Fig. 3, C and D. **(C)** Phenotypes (left images) and GFP fluorescence (right images) of VCS and *vcs-7* homozygotes introgressed into the GFP171.1 background at 14 days of growth. **(D)** (Top panel) GFP accumulation in seedlings depicted in (C). (Bottom panel) Same as in Fig. 2C upon introgression of $GFP_{no\ miR}$. Differential accumulation of CSD2 (miR398 target) confirms the presence of the *vcs-7* mutation [low Cu(I)]. Prot, Coomassie-stained bands that serve as a loading control.

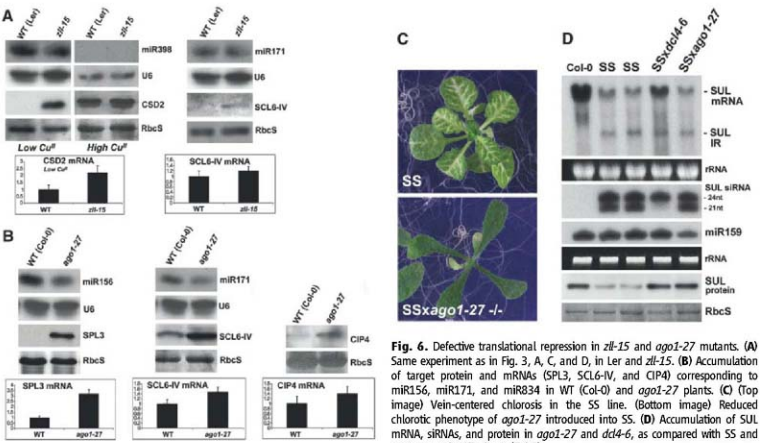


Fig. 6. Defective translational repression in *zil-15* and *ago1-27* mutants. **(A)** Same experiment as in Fig. 3, A, C, and D, in Ler and *zil-15*. **(B)** Accumulation of target protein and mRNAs (SPL3, SCL6-IV, and CIP4) corresponding to miR156, miR171, and miR834 in WT (Col-0) and *ago1-27* plants. **(C)** (Top image) Vein-centered chlorosis in the SS line. (Bottom image) Reduced chlorotic phenotype of *ago1-27* introduced into SS. **(D)** Accumulation of SUL mRNA, siRNAs, and protein in *ago1-27* and *dcl4-6*, as compared with SS and nontransgenic WT plants (Col-0).

pression and is supported by several lines of evidence: RNAi of tubulins in *Caenorhabditis elegans* compromises target regulation by distinct miRNAs (*36*), whereas *Drosophila* Armitage, which is required for RISC assembly, is a microtubule-associated protein (*37*), as are FMR (necessary for miRNA-directed translational activation) and the ribosome-interacting AGO-like protein Seawi (*38*–*40*). Finally, many mRNA decay factors that colocalize with the decapping complex in cytoplasmic processing (P) bodies interact with tubulin or with microtubule polymers in yeast (*41*, *42*). Identifying the P-body component VCS as integral to nondegradative translational repression in plants is important in two respects. First, it reveals some level of mechanistic similarity between plant and animal miRNA-mediated repression. Second, it suggests that RNA decay could be coupled to the action of at least some plant miRNAs, as established in animal cells (*20*, *43*). An outstanding question pertains to the biological importance of translational inhibition in plants. Studies on human cells suggest that one key aspect lies in the reversible nature of this type of regulation (*44*). This may be particularly adapted to coordination and resetting of stress-responsive gene expression, an emerging function of many plant miRNAs (*45*).

References and Notes

1. D. P. Bartel, *Cell* **116**, 281 (2004).
2. G. Hutvagner, P. D. Zamore, *Science* **297**, 2056 (2002).
3. J. J. Song, S. K. Smith, G. J. Hannon, L. Joshua-Tor, *Science* **305**, 1434 (2004).
4. M. W. Jones-Rhoades, D. P. Bartel, B. Bartel, *Annu. Rev. Plant Biol.* **57**, 19 (2006).
5. Y. Qi, A. M. Denli, G. J. Hannon, *Mol. Cell* **19**, 421 (2005).
6. N. Baumberger, D. C. Baulcombe, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11928 (2005).
7. H. Vaucheret, F. Vazquez, P. Crete, D. P. Bartel, *Genes Dev.* **18**, 1187 (2004).
8. E. A. Parizotto, P. Dunoyer, N. Rahm, C. Himber, O. Voinnet, *Genes Dev.* **18**, 2237 (2004).
9. E. J. Chapman, J. C. Carrington, *Nat. Rev. Genet.* **8**, 884 (2007).
10. C. Llave, Z. Xie, K. D. Kasschau, J. C. Carrington, *Science* **297**, 2053 (2002).
11. X. Chen, *Science* **303**, 2022 (2004).
12. M. J. Aukerman, H. Sakai, *Plant Cell* **15**, 2730 (2003).
13. M. Gandikota *et al.*, *Plant J.* **49**, 683 (2007).
14. M. W. Jones-Rhoades, D. P. Bartel, *Mol. Cell* **14**, 787 (2004).
15. R. Rajagopalan, H. Vaucheret, J. Trejo, D. P. Bartel, *Genes Dev.* **20**, 3407 (2006).
16. H. Yamasaki *et al.*, *J. Biol. Chem.* **282**, 16369 (2007).
17. T. Bouguin, O. Mattsson, H. Naested, R. Foster, J. Mundy, *J. Cell Sci.* **116**, 791 (2003).
18. V. Stoppin-Mellet, J. Gaillard, M. Vantard, *Plant J.* **46**, 1009 (2006).
19. M. Webb, S. Jouannic, J. Foreman, P. Linstead, L. Donan, *Development* **129**, 123 (2002).
20. A. Eulalio *et al.*, *Genes Dev.* **21**, 2558 (2007).
21. J. Xu, J. Y. Yang, Q. W. Niu, N. H. Chua, *Plant Cell* **18**, 3386 (2006).
22. D. C. Goeres *et al.*, *Plant Cell* **19**, 1549 (2007).
23. S. Ni *et al.*, *Cell* **133**, 116 (2008).
24. T. A. Montgomery *et al.*, *Cell* **133**, 128 (2008).
25. B. Moussian, H. Schoof, A. Haecker, G. Jurgens, T. Laux, *EMBO J.* **17**, 1799 (1998).
26. K. Lynn *et al.*, *Development* **126**, 469 (1999).
27. J.-B. Morel *et al.*, *Plant Cell* **14**, 629 (2002).
28. P. Dunoyer, C. Himber, V. Ruiz-Ferrer, A. Alioua, O. Voinnet, *Nat. Genet.* **39**, 848 (2007).
29. R. Schwab, S. Ossowski, M. Riester, N. Warthmann, D. Weigel, *Plant Cell* **18**, 1121 (2006).
30. C. Lu, N. Wang, E. Lam, *FEBS Lett.* **579**, 1498 (2005).
31. N. Fahlgren *et al.*, *PLoS One* **2**, e219 (2007).
32. Y. Zhang, *Nucleic Acids Res.* **33**, W701 (2005).
33. D. V. Dugas, B. Bartel, *Plant Mol. Biol.* published online 8 April 2008; 10.1007/s11103-008-9329-1.
34. J. F. Wang, H. Zhou, Y. Q. Chen, Q. J. Luo, L. H. Qu, *Nucleic Acids Res.* **32**, 1688 (2004).
35. M. K. Rao *et al.*, *Genes Dev.* **20**, 147 (2006).
36. D. H. Parry, J. Xu, G. Ruvkun, *Curr. Biol.* **17**, 2013 (2007).
37. H. A. Cook, B. S. Koppetsch, J. Wu, W. E. Theurkauf, *Cell* **116**, 817 (2004).
38. S. Vasudevan, Y. Tong, J. A. Steitz, *Science* **318**, 1931 (2007).
39. A. J. Rodriguez *et al.*, *RNA* **11**, 646 (2005).
40. H. Wang *et al.*, *Mol. Biol. Cell* **19**, 105 (2007).
41. J. Liu *et al.*, *Nat. Cell Biol.* **7**, 1261 (2005).
42. A. C. Gavin *et al.*, *Nature* **440**, 631 (2006).
43. A. J. Giraldez *et al.*, *Science* **312**, 75 (2006).
44. S. N. Bhattacharyya, R. Habermacher, U. Martine, E. I. Closs, W. Filipowicz, *Cell* **125**, 1111 (2006).
45. R. Sunkar, V. Chinnusamy, J. Zhu, J. K. Zhu, *Trends Plant Sci.* **12**, 301 (2007).

46. This work was supported by grants from the European Molecular Biology Organization (ALTF-251 2004) and the European Union (EIF 25064 2005) to P.B., as well as by grants "MicroTRAC" from l'Agence Nationale de la Recherche/Genoplante, "SIROCCO" LSHG-CT-2006-037900 from the European Union, and an Action Thématique Incitative sur Programme Plus from CNRS to O.V. We thank D. Klebenstein (CSD1, CSD2), P. Huijser (SPL3), T. Leustek (APS), P. E. Jensen (Chl/SUL), and M. Vantard (CTN1) for antibodies; Z. Ye (*fas2*), J. Mundy (*fue1*), L. Dolan (*lrh3-3*), and T. Laux (*zll-15*, *ago1-27*) for seeds; C. Himber for help with *ago1-27* × SS experiments and plant pictures; and R. Wagner's team for plant care.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1159151/DC1
Materials and Methods
SOM Text
Figs. S1 to S6
Table S1
References

15 April 2008; accepted 1 May 2008
Published online 15 May 2008;
10.1126/science.1159151
Include this information when citing this paper.

Predicting Human Brain Activity Associated with the Meanings of Nouns

Tom M. Mitchell,[1]* Svetlana V. Shinkareva,[2] Andrew Carlson,[1] Kai-Min Chang,[3,4] Vicente L. Malave,[5] Robert A. Mason,[3] Marcel Adam Just[3]

The question of how the human brain represents conceptual knowledge has been debated in many scientific fields. Brain imaging studies have shown that different spatial patterns of neural activation are associated with thinking about different semantic categories of pictures and words (for example, tools, buildings, and animals). We present a computational model that predicts the functional magnetic resonance imaging (fMRI) neural activation associated with words for which fMRI data are not yet available. This model is trained with a combination of data from a trillion-word text corpus and observed fMRI data associated with viewing several dozen concrete nouns. Once trained, the model predicts fMRI activation for thousands of other concrete nouns in the text corpus, with highly significant accuracies over the 60 nouns for which we currently have fMRI data.

The question of how the human brain represents and organizes conceptual knowledge has been studied by many scientific communities. Neuroscientists using brain imaging studies (*1*–*9*) have shown that distinct spatial patterns of fMRI activity are associated with viewing pictures of certain semantic categories, including tools, buildings, and animals. Linguists have characterized different semantic roles associated with individual verbs, as well as the types of nouns that can fill those semantic roles [e.g., VerbNet (*10*) and WordNet (*11*, *12*)]. Computational linguists have analyzed the statistics of very large text corpora and have demonstrated that a word's meaning is captured to some extent by the distribution of words and phrases with which it commonly co-occurs (*13*–*17*). Psychologists have studied word meaning through feature-norming studies (*18*) in which participants are asked to list the features they associate with various words, revealing a consistent set of core features across individuals and suggesting a possible grouping of features by sensory-motor modalities. Researchers studying semantic effects of brain damage have found deficits that are specific to given semantic categories (such as animals) (*19*–*21*).

This variety of experimental results has led to competing theories of how the brain encodes meanings of words and knowledge of objects, including theories that meanings are encoded in sensory-motor cortical areas (*22*, *23*) and theories that they are instead organized by semantic categories such as living and nonliving objects (*18*, *24*). Although these competing theories sometimes lead to different predictions (e.g., of which naming disabilities will co-occur in brain-damaged patients), they are primarily descriptive theories that make no attempt to predict the specific brain activation that will be produced when a human subject reads a particular word or views a drawing of a particular object.

We present a computational model that makes directly testable predictions of the fMRI activity associated with thinking about arbitrary concrete nouns, including many nouns for which no fMRI data are currently available. The theory underlying this computational model is that the neural basis of the semantic representation of concrete nouns is related to the distributional properties of those words in a broadly based corpus of the language. We describe experiments training competing computational models based on different assumptions regarding the underlying features that are used in the brain for encoding of meaning of concrete objects. We present experimental evidence showing that the best of these models predicts fMRI neural activity well enough that it can successfully match words it has not yet encountered to their previously unseen fMRI images, with accuracies far above those expected by chance. These results establish a direct, predictive relationship between the statistics of word co-occurrence in text and the neural activation associated with thinking about word meanings.

Approach. We use a trainable computational model that predicts the neural activation for any given stimulus word w using a two-step process, illustrated in Fig. 1. Given an arbitrary stimulus word w, the first step encodes the meaning of w as a vector of intermediate semantic features computed from the occurrences of stimulus word w within a very large text corpus (*25*) that captures the typical use of words in English text. For example, one intermediate semantic feature might be the frequency with which w co-occurs with the verb "hear." The second step predicts the neural fMRI activation at every voxel location in the brain, as a weighted sum of neural activations contributed by each of the intermediate semantic features. More precisely, the predicted activation y_v at voxel v in the brain for word w is given by

$$y_v = \sum_{i=1}^{n} c_{vi} f_i(w) \qquad (1)$$

where $f_i(w)$ is the value of the ith intermediate semantic feature for word w, n is the number of semantic features in the model, and c_{vi} is a learned scalar parameter that specifies the degree to which the ith intermediate semantic feature activates voxel v. This equation can be interpreted as predicting the full fMRI image across all voxels for stimulus word w as a weighted sum of images, one per semantic feature f_i. These semantic feature images, defined by the learned c_{vi}, constitute a basis set of component images that model the brain activation associated with different semantic components of the input stimulus words.

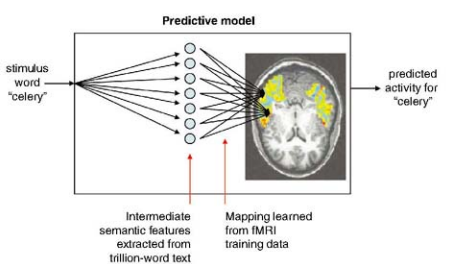


Fig. 1. Form of the model for predicting fMRI activation for arbitrary noun stimuli. fMRI activation is predicted in a two-step process. The first step encodes the meaning of the input stimulus word in terms of intermediate semantic features whose values are extracted from a large corpus of text exhibiting typical word use. The second step predicts the fMRI image as a linear combination of the fMRI signatures associated with each of these intermediate semantic features.

[1]Machine Learning Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA. [2]Department of Psychology, University of South Carolina, Columbia, SC 29208, USA. [3]Center for Cognitive Brain Imaging, Carnegie Mellon University, Pittsburgh, PA 15213, USA. [4]Language Technologies Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA. [5]Cognitive Science Department, University of California, San Diego, La Jolla, CA 92093, USA.

*To whom correspondence should be addressed. E-mail: Tom.Mitchell@cs.cmu.edu

To fully specify a model within this computational modeling framework, one must first define a set of intermediate semantic features $f_1(w)$ $f_2(w)$…$f_n(w)$ to be extracted from the text corpus. In this paper, each intermediate semantic feature is defined in terms of the co-occurrence statistics of the input stimulus word w with a particular other word (e.g., "taste") or set of words (e.g., "taste," "tastes," or "tasted") within the text corpus. The model is trained by the application of multiple regression to these features $f_i(w)$ and the observed fMRI images, so as to obtain maximum-likelihood estimates for the model parameters c_{vi} (*26*). Once trained, the computational model can be evaluated by giving it words outside the training set and comparing its predicted fMRI images for these words with observed fMRI data.

This computational modeling framework is based on two key theoretical assumptions. First, it assumes the semantic features that distinguish the meanings of arbitrary concrete nouns are reflected in the statistics of their use within a very large text corpus. This assumption is drawn from the field of computational linguistics, where statistical word distributions are frequently used to approximate the meaning of documents and words (*14*–*17*). Second, it assumes that the brain activity observed when thinking about any concrete noun can be derived as a weighted linear sum of contributions from each of its semantic features. Although the correctness of this linearity assumption is debatable, it is consistent with the widespread use of linear models in fMRI analysis (*27*) and with the assumption that fMRI activation often reflects a linear superposition of contributions from different sources. Our theoretical framework does not take a position on whether the neural activation encoding meaning is localized in particular cortical regions. Instead, it considers all cortical voxels and allows the training data to determine which locations are systematically modulated by which aspects of word meanings.

Results. We evaluated this computational model using fMRI data from nine healthy, college-age participants who viewed 60 different word-picture pairs presented six times each. Anatomically defined regions of interest were automatically labeled according to the methodology in (*28*). The 60 randomly ordered stimuli included five items from each of 12 semantic categories (animals, body parts, buildings, building parts, clothing, furniture, insects, kitchen items, tools, vegetables, vehicles, and other man-made items). A representative fMRI image for each stimulus was created by computing the mean fMRI response over its six presentations, and the mean of all 60 of these representative images was then subtracted from each [for details, see (*26*)].

To instantiate our modeling framework, we first chose a set of intermediate semantic features. To be effective, the intermediate semantic features must simultaneously encode the wide variety of semantic content of the input stimulus words and factor the observed fMRI activation into more primitive com-

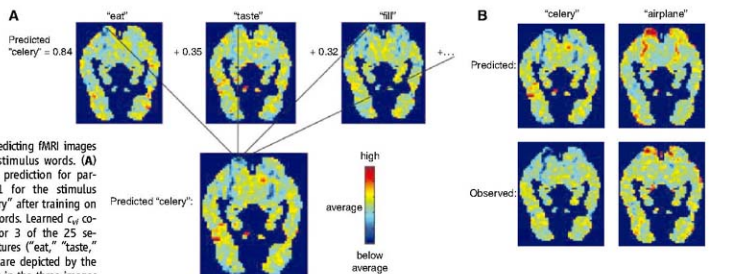


Fig. 2. Predicting fMRI images for given stimulus words. **(A)** Forming a prediction for participant P1 for the stimulus word "celery" after training on 58 other words. Learned c_{vi} coefficients for 3 of the 25 semantic features ("eat," "taste," and "fill") are depicted by the voxel colors in the three images at the top of the panel. The co-occurrence value for each of these features for the stimulus word "celery" is shown to the left of their respective images [e.g., the value for "eat (celery)" is 0.84]. The predicted activation for the stimulus word [shown at the bottom of (A)] is a linear combination of the 25 semantic fMRI signatures, weighted by their co-occurrence values. This figure shows just one horizontal slice [z = −12 mm in Montreal Neurological Institute (MNI) space] of the predicted three-dimensional image. **(B)** Predicted and observed fMRI images for "celery" and "airplane" after training that uses 58 other words. The two long red and blue vertical streaks near the top (posterior region) of the predicted and observed images are the left and right fusiform gyri.

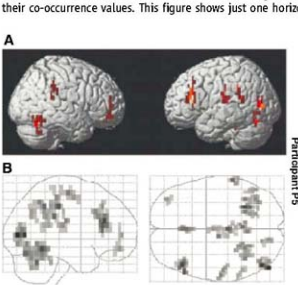


Fig. 3. Locations of most accurately predicted voxels. Surface **(A)** and glass brain **(B)** rendering of the correlation between predicted and actual voxel activations for words outside the training set for participant P5. These panels show clusters containing at least 10 contiguous voxels, each of whose predicted-actual correlation is at least 0.28. These voxel clusters are distributed throughout the cortex and located in the left and right occipital and parietal lobes; left and right fusiform, postcentral, and middle frontal gyri; left inferior frontal gyrus; medial frontal gyrus; and anterior cingulate. **(C)** Surface rendering of the predicted-actual correlation averaged over all nine participants. This panel represents clusters containing at least 10 contiguous voxels, each with average correlation of at least 0.14.

ponents that can be linearly recombined to successfully predict the fMRI activation for arbitrary new stimuli. Motivated by existing conjectures regarding the centrality of sensory-motor features in neural representations of objects (*18*, *29*), we designed a set of 25 semantic features defined by 25 verbs: "see," "hear," "listen," "taste," "smell," "eat," "touch," "rub," "lift," "manipulate," "run," "push," "fill," "move," "ride," "say," "fear," "open," "approach," "near," "enter," "drive," "wear," "break," and "clean." These verbs generally correspond to basic sensory and motor activities, actions performed on objects, and actions involving changes to spatial relationships. For each verb, the value of the corresponding intermediate semantic feature for a given input stimulus word w is the normalized co-occurrence count of w with any of three forms of the verb (e.g., "taste," "tastes," or "tasted") over the text corpus. One exception was made for the verb "see." Its past tense was omitted because "saw" is one of our 60 stimulus nouns. Normalization consists of scaling the vector of 25 feature values to unit length.

We trained a separate computational model for each of the nine participants, using this set of 25 semantic features. Each trained model was evaluated by means of a "leave-two-out" cross-validation approach, in which the model was repeatedly trained with only 58 of the 60 available word stimuli and associated fMRI images. Each trained model was tested by requiring that it first predict the fMRI images for the two "held-out" words and then match these correctly to their corresponding held-out fMRI images. The process of predicting the fMRI image for a held-out word is illustrated in Fig. 2A. The match between the two predicted and the two observed fMRI images was determined by which match had a higher cosine similarity, evaluated over the 500 image voxels with the most stable responses across training presentations (*26*). The expected accuracy in matching the left-out words to their left-out fMRI images is 0.50 if the model performs at chance levels. An accuracy of 0.62 or higher for a single model trained for a single participant was determined to be statistically significant ($P < 0.05$) relative to chance, based on the empirical distribution of accuracies for randomly generated null models (*26*). Similarly, observing an accuracy of 0.62 or higher for each of the nine independently trained participant-specific models would be statistically significant at $P < 10^{-11}$.

The cross-validated accuracies in matching two unseen word stimuli to their unseen fMRI images for models trained on participants P1 through P9 were 0.83, 0.76, 0.78, 0.72, 0.78, 0.85, 0.73, 0.68, and 0.82 (mean = 0.77). Thus, all nine participant-specific models exhibited accuracies significantly above chance levels. The models succeeded in distinguishing pairs of previously unseen words in over three-quarters of the 15,930 cross-validated test pairs across these nine participants. Accuracy across participants was strongly correlated (r = –0.66) with estimated head motion (i.e., the less the participant's head motion, the greater the prediction accuracy), suggesting that the variation in accuracies across participants is explained at least in part by noise due to head motion.

Visual inspection of the predicted fMRI images produced by the trained models shows that these predicted images frequently capture substantial aspects of brain activation associated with stimulus words outside the training set. An example is shown in Fig. 2B, where the model was trained on 58 of the 60 stimuli for participant P1, omitting "celery" and "airplane." Although the predicted fMRI images for "celery" and "airplane" are not perfect, they capture substantial components of the activation actually observed for these two stimuli. A plot of similarities between all 60 predicted and observed fMRI images is provided in fig. S3.

The model's predictions are differentially accurate in different brain locations, presumably more accurate in those locations involved in encoding the semantics of the input stimuli. Figure 3 shows the model's "accuracy map," indicating the cortical regions where the model's predicted activations for held-out words best correlate with the observed activations, both for an individual participant (P5) and averaged over all nine participants. These highest-accuracy voxels are meaningfully distributed across the cortex, with the left hemisphere more strongly represented, appearing in left inferior temporal, fusiform, motor cortex, intraparietal sulcus, inferior frontal, orbital frontal, and the occipital cortex. This left hemisphere dominance is consistent with the generally held view that the left hemisphere plays a larger role than the right hemisphere in semantic representation. High-accuracy voxels also appear in both hemispheres in the occipital cortex, intraparietal sulcus, and some of the inferior temporal regions, all of which are also likely to be involved in visual object processing.

It is interesting to consider whether these trained computational models can extrapolate to make accurate predictions for words in new semantic categories beyond those in the training set. To test this, we retrained the models but this time we excluded from the training set all examples belonging to the same semantic category as either of the two held-out test words (e.g., when testing on "celery" versus "airplane," we removed every food and vehicle stimulus from the training set, training on only 50 words). In this case, the cross-validated prediction accuracies were 0.74, 0.69, 0.67, 0.69, 0.64,

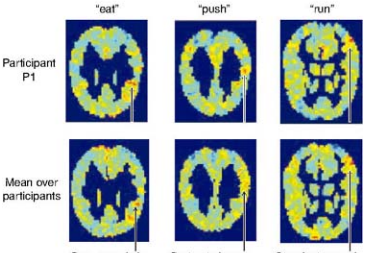


Fig. 4. Learned voxel activation signatures for 3 of the 25 semantic features, for participant P1 (top panels) and averaged over all nine participants (bottom panels). Just one horizontal z slice is shown for each. The semantic feature associated with the verb "eat" predicts substantial activity in right pars opercularis, which is believed to be part of the gustatory cortex. The semantic feature associated with "push" activates the right postcentral gyrus, which is believed to be associated with premotor planning. The semantic feature for the verb "run" activates the posterior portion of the right superior temporal sulcus, which is believed to be associated with the perception of biological motion.

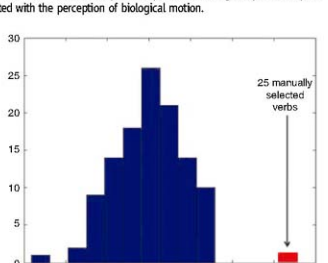


Fig. 5. Accuracies of models based on alternative intermediate semantic feature sets. The accuracy of computational models that use 115 different randomly selected sets of intermediate semantic features is shown in the blue histogram. Each feature set is based on 25 words chosen at random from the 5000 most frequent words, excluding the 500 most frequent words and the stimulus words. The accuracy of the feature set based on manually chosen sensory-motor verbs is shown in red. The accuracy of each feature set is the average accuracy obtained when it was used to train models for each of the nine participants.

0.78, 0.68, 0.64, and 0.78 (mean = 0.70). This ability of the model to extrapolate to words semantically distant from those on which it was trained suggests that the semantic features and their learned neural activation signatures of the model may span a diverse semantic space.

Given that the 60 stimuli are composed of five items in each of 12 semantic categories, it is also interesting to determine the degree to which the model can make accurate predictions even when the two held-out test words are from the same category, where the discrimination is likely to be more difficult (e.g., "celery" versus "corn"). These within-category prediction accuracies for the nine individuals were 0.61, 0.58, 0.58, 0.72, 0.58, 0.77, 0.58, 0.52, and 0.68 (mean = 0.62), indicating that although the model's accuracy is lower when it is differentiating between semantically more similar stimuli, on average its predictions nevertheless remain above chance levels.

In order to test the ability of the model to distinguish among an even more diverse range of words, we tested its ability to resolve among 1000 highly frequent words (the 1300 most frequent tokens in the text corpus, omitting the 300 most frequent). Specifically, we conducted a leave-one-out test in which the model was trained using 59 of the 60 available stimulus words. It was then given the fMRI image for the held-out word and a set of 1001 candidate words (the 1000 frequent tokens, plus the held-out word). It ranked these 1001 candidates by first predicting the fMRI image for each candidate and then sorting the 1001 candidates by the similarity between their predicted fMRI image and the fMRI image it was provided. The expected percentile rank of the correct word in this ranked list would be 0.50 if the model were operating at chance. The observed percentile ranks for the nine participants were 0.79, 0.71, 0.74, 0.67, 0.73, 0.77, 0.70, 0.63, and 0.76 (mean = 0.72), indicating that the model is to some degree applicable across a semantically diverse set of words [see (*26*) for details].

A second approach to evaluating our computation model, beyond quantitative measurements of its prediction accuracy, is to examine the learned basis set of fMRI signatures for the 25 verb-based signatures. These 25 signatures represent the model's learned decomposition of neural representations into their component semantic features and provide the basis for all of its predictions. The learned signatures for the semantic features "eat," "push," and "run" are shown in Fig. 4. Notice that each of these signatures predicts activation in multiple cortical regions.

Examining the semantic feature signatures in Fig. 4, one can see that the learned fMRI signature for the semantic feature "eat" predicts strong activation in opercular cortex (as indicated by the arrows in the left panels), which others have suggested is a component of gustatory cortex involved in the sense of taste (*30*). Also, the learned fMRI signature for "push" predicts substantial activation in the right postcentral gyrus, which is widely assumed to be involved in the planning of complex, coordinated movements (*31*). Furthermore, the learned signature for "run" predicts strong activation in the posterior portion of the right superior temporal lobe along the sulcus, which others have suggested is involved in perception of biological motion (*32*, *33*). To summarize, these learned signatures cause the model to predict that the neural activity representing a noun will exhibit activity in gustatory cortex to the degree that this noun co-occurs with the verb "eat," in motor areas to the degree that it co-occurs with "push," and in cortical regions related to body motion to the degree that it co-occurs with "run." Whereas the top row of Fig. 4 illustrates these learned signatures for participant P1, the bottom row shows the mean of the nine signatures learned independently for the nine participants. The similarity of the two rows of signatures demonstrates that these learned intermediate semantic feature signatures exhibit substantial commonalities across participants.

The learned signatures for several other verbs also exhibit interesting correspondences between the function of cortical regions in which they predict activation and that verb's meaning, though in some cases the correspondence holds for only a subset of the nine participants. For example, additional features for participant P1 include the signature for "touch," which predicts strong activation in somatosensory cortex (right postcentral gyrus), and the signature for "listen," which predicts activation in language-processing regions (left posterior superior temporal sulcus and left pars triangularis), though these trends are not common to all nine participants. The learned feature signatures for all 25 semantic features are provided at (*26*).

Given the success of this set of 25 intermediate semantic features motivated by the conjecture that the neural components corresponding to basic semantic properties are related to sensory-motor verbs, it is natural to ask how this set of intermediate semantic features compares with alternatives. To explore this, we trained and tested models based on randomly generated sets of semantic features, each defined by 25 randomly drawn words from the 5000 most frequent words in the text corpus, excluding the 60 stimulus words as well as the 500 most frequent words (which contain many function words and words without much specific semantic content, such as "the" and "have"). A total of 115 random feature sets was generated. For each feature set, models were trained for all nine participants, and the mean prediction accuracy over these nine models was measured. The distribution of resulting accuracies is shown in the blue histogram in Fig. 5. The mean accuracy over these 115 feature sets is 0.60, the SD is 0.041, and the minimum and maximum accuracies are 0.46 and 0.68, respectively. The random feature sets generating the highest and lowest accuracy are shown at (*26*). The fact that the mean accuracy is greater than 0.50 suggests that many feature sets capture some of the semantic content of the 60 stimulus words and some of the regularities in the corresponding brain activation. However, among these 115 feature sets, none came close to the 0.77 mean accuracy of our manually generated feature set (shown by the red bar in the histogram in Fig. 5). This result suggests the set of features defined by our sensory-motor verbs is somewhat distinctive in capturing regularities in the neural activation encoding the semantic content of words in the brain.

Discussion. The results reported here establish a direct, predictive relationship between the statistics of word co-occurrence in text and the neural activation associated with thinking about word meanings. Furthermore, the computational models trained to make these predictions provide insight into how the neural activity that represents objects can be decomposed into a basis set of neural activation patterns associated with different semantic components of the objects.

The success of the specific model, which uses 25 sensory-motor verbs (as compared with alternative models based on randomly sampled sets of 25 semantic features), lends credence to the conjecture that neural representations of concrete nouns are in part grounded in sensory-motor features. However, the learned signatures associated with the 25 intermediate semantic features also exhibit significant activation in brain areas not directly associated with sensory-motor function, including frontal regions. Thus, it appears that the basis set of features that underlie neural representations of concrete nouns involves much more than sensory-motor cortical regions.

Other recent work has suggested that the neural encodings that represent concrete objects are at least partly shared across individuals, based on evidence that it is possible to identify which of several items a person is viewing, through only their fMRI image and a classifier model trained from other people (*34*). The results reported here show that the learned basis set of semantic features also shares certain commonalities across individuals and may help determine more directly which factors of neural representations are similar and different across individuals.

Our approach is analogous in some ways to research that focuses on lower-level visual features of picture stimuli to analyze fMRI activation associated with viewing the picture (*9*, *35*, *36*) and to research that compares perceived similarities between object shapes to their similarities based on fMRI activation (*37*). Recent work (*36*) has shown that it is possible to predict aspects of fMRI activation in parts of visual cortex based on visual features of arbitrary scenes and to use this predicted activation to identify which of a set of candidate scenes an individual is viewing. Our work differs from these efforts, in that we focus on encodings of more abstract semantic concepts signified by words and predict brain-wide fMRI activations based on text corpus features that capture semantic aspects of the stimulus word, rather than visual features that capture perceptual aspects. Our work is also related to recent research that uses machine learning algorithms to train classifiers of mental states based on fMRI data (*38*, *39*), though it differs in that our models are capable of extrapolating to predict fMRI images for mental states not present in the training set.

This research represents a shift in the paradigm for studying neural representations in the brain,

moving from work that has cataloged the patterns of fMRI activity associated with specific categories of words and pictures to instead building computational models that predict the fMRI activity for arbitrary words (including thousands of words for which fMRI data are not yet available). This is a natural progression as the field moves from pretheoretical cataloging of data toward development of computational models and the beginnings of a theory of neural representations. Our computational models can be viewed as encoding a restricted form of predictive theory, one that answers such questions as "What is the predicted fMRI neural activity encoding word *w*?" and "What is the basis set of semantic features and corresponding components of neural activation that explain the neural activations encoding meanings of concrete nouns?" Although we remain far from a causal theory explaining how the brain synthesizes these representations from its sensory inputs, answers even to these questions promise to shed light on some of the key regularities underlying neural representations of meaning.

References and Notes

1. J. V. Haxby *et al.*, *Science* **293**, 2425 (2001).
2. A. Ishai, L. G. Ungerleider, A. Martin, J. L. Schouten, J. V. Haxby, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9379 (1999).
3. N. Kanwisher, J. McDermott, M. M. Chun, *J. Neurosci.* **17**, 4302 (1997).
4. T. A. Carlson, P. Schrater, S. He, *J. Cogn. Neurosci.* **15**, 704 (2003).
5. D. D. Cox, R. L. Savoy, *Neuroimage* **19**, 261 (2003).
6. T. Mitchell *et al.*, *Mach. Learn.* **57**, 145 (2004).
7. S. J. Hanson, T. Matsuka, J. V. Haxby, *Neuroimage* **23**, 156 (2004).
8. S. M. Polyn, V. S. Natu, J. D. Cohen, K. A. Norman, *Science* **310**, 1963 (2005).
9. A. J. O'Toole, F. Jiang, H. Abdi, J. V. Haxby, *J. Cogn. Neurosci.* **17**, 580 (2005).
10. K. Kipper, A. Korhonen, N. Ryant, M. Palmer, *Proceedings of the 5th International Conference on Language Resources and Evaluation*, Genoa, Italy, 24 to 26 May 2006.
11. G. Miller, R. Beckwith, C. Fellbaum, D. Gross, K. J. Miller, *Int. J. Lexicography* **3**, 235 (1990).
12. C. Fellbaum, Ed., *WordNet: An Electronic Lexical Database* (Massachusetts Institute of Technology Press, Cambridge, MA, 1998).
13. K. W. Church, P. Hanks, *Comput. Linguist.* **16**, 22 (1990).
14. T. K. Landauer, S. T. Dumais, *Psychol. Rev.* **104**, 211 (1997).
15. D. Lin, S. Zhao, L. Qin, M. Zhou, *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, Acapulco, Mexico, August 2003 (Morgan Kaufmann, San Francisco, 2003), pp. 1492–1493.
16. D. M. Blei, A. Y. Ng, M. I. Jordan, *J. Mach. Learn. Res.* **3**, 993 (2003).
17. R. Snow, D. Jurafsky, A. Ng, *Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics*, Sydney, Australia, 17 to 21 July 2006.
18. G. S. Cree, K. McRae, *J. Exp. Psychol. Gen.* **132**, 163 (2003).
19. A. Caramazza, J. R. Shelton, *J. Cogn. Neurosci.* **10**, 1 (1998).
20. S. J. Crutch, E. K. Warrington, *Brain* **126**, 1821 (2003).
21. D. Samson, A. Pillon, *Brain Lang.* **91**, 252 (2004).
22. A. Martin, L. L. Chao, *Curr. Opin. Neurobiol.* **11**, 194 (2001).
23. R. F. Goldberg, C. A. Perfetti, W. Schneider, *J. Neurosci.* **26**, 4917 (2006).
24. B. Z. Mahon, A. Caramazza, in *The Encyclopedia of Language and Linguistics*, K. Brown, Ed. (Elsevier Science, Amsterdam, ed. 2, 2005).
25. T. Brants, A. Franz, www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T13 (Linguistic Data Consortium, Philadelphia, PA, 2006).
26. See Supporting Online Material.
27. K. J. Friston *et al.*, *Hum. Brain Mapp.* **2**, 189 (1995).
28. N. Tzourio-Mazoyer *et al.*, *Neuroimage* **15**, 273 (2002).
29. A. Martin, L. G. Ungerleider, J. V. Haxby, in *The New Cognitive Neurosciences*, M. S. Gazzaniga, Ed. (Massachusetts Institute of Technology Press, Cambridge, MA, ed. 2, 2000), pp. 1023–1036.
30. B. Cerf, D. LeBihan, P. F. Van de Moortele, P. Mac Leod, A. Faurion, *Ann. N.Y. Acad. Sci.* **855**, 575 (1998).
31. K. A. Pelphrey, J. P. Morris, C. R. Michelich, T. Allison, G. McCarthy, *Cereb. Cortex* **15**, 1866 (2005).
32. L. M. Vaina, J. Solomon, S. Chowdhury, P. Sinha, J. Belliveau, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11656 (2001).
33. K. Sakai *et al.*, *Magn. Reson. Med.* **33**, 736 (1995).
34. S. V. Shinkareva *et al.*, *PLoS One* **3**, e1394 (2008).
35. D. R. Hardoon, J. Mouraо-Miranda, M. Brammer, J. Shawe-Taylor, *Neuroimage* **37**, 1250 (2007).
36. K. N. Kay, T. Naselaris, R. J. Prenger, J. L. Gallant, *Nature* **452**, 352 (2008).
37. S. Edelman, K. Grill-Spector, T. Kushnir, R. Malach, *Psychobiology* **26**, 309 (1998).
38. J. D. Haynes, G. Rees, *Nat. Rev. Neurosci.* **7**, 523 (2006).
39. K. A. Norman, S. M. Polyn, G. J. Detre, J. V. Haxby, *Trends Cogn. Sci.* **10**, 424 (2006).
40. This research was funded by grants from the W. M. Keck Foundation, NSF, and by a Yahoo! Fellowship to A.C. We acknowledge Google for making available its data from the trillion-token text corpus. We thank W. Cohen for helpful suggestions regarding statistical significance tests.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1191/DC1
Materials and Methods
SOM Text
Figs. S1 to S5
References

12 November 2007; accepted 3 April 2008
10.1126/science.1152876

The Cassiopeia A Supernova Was of Type IIb

Oliver Krause,[1]* Stephan M. Birkmann,[1] Tomonori Usuda,[2] Takashi Hattori,[2] Miwa Goto,[1] George H. Rieke,[3] Karl A. Misselt[3]

Cassiopeia A is the youngest supernova remnant known in the Milky Way and a unique laboratory for supernova physics. We present an optical spectrum of the Cassiopeia A supernova near maximum brightness, obtained from observations of a scattered light echo more than three centuries after the direct light of the explosion swept past Earth. The spectrum shows that Cassiopeia A was a type IIb supernova and originated from the collapse of the helium core of a red supergiant that had lost most of its hydrogen envelope before exploding. Our finding concludes a long-standing debate on the Cassiopeia A progenitor and provides new insight into supernova physics by linking the properties of the explosion to the wealth of knowledge about its remnant.

The supernova remnant Cassiopeia A is one of the most-studied objects in the sky, with observations from the longest radio waves to gamma rays. The remnant expansion rate indicates that the core of its progenitor star collapsed around the year 1681 ± 19, as viewed from Earth (*1*). Because of its youth and proximity of $3.4^{+0.3}_{-0.1}$ kpc (*2*), Cas A provides a unique opportunity to probe the death of a massive star and to test theoretical models of core-collapse supernovae. However, such tests are compromised because the Cas A supernova showed at most a faint optical display on Earth at the time of explosion. The lack of a definitive sighting means that there is almost no direct information about the type of the explosion, and the true nature of its progenitor star has been a puzzle since the discovery of the remnant (*3*).

The discovery of light echoes due both to scattering and to absorption and re-emission of the outgoing supernova flash (*4*, *5*) by the interstellar dust near the remnant raised the possibility of conducting a postmortem study of the last historic Galactic supernova by observing its scattered light. Similarly, the determination of a supernova spectral type long after its explosion using light echoes was recently demonstrated for an extragalactic supernova (*6*).

We have monitored infrared echoes around Cas A at a wavelength of 24 μm with use of the multiband imaging photometer (MIPS) instrument aboard the Spitzer Space Telescope (*4*). The results confirm that they arise from the flash emitted in the initial explosion of Cas A (*5*). An image taken on 20 August 2007 revealed a bright (flux density $F_{24\mu m} = 0.36 \pm 0.04$ Jy, 1 Jy = 10^{-26} W m^{-2} Hz^{-1}) and mainly unresolved echo feature located 80 arc min northwest of Cas A (position angle 311° east of north). It had not been detected ($F_{24\mu m} < 2$ mJy; 5-σ) on two previous images of this region obtained on 2 October 2006 and 23 January 2007 (Fig. 1).

An image obtained on 7 January 2008 shows that the peak of the echo has dropped in surface brightness by a factor of 18 and shifted toward the west. Transient optical emission associated with the infrared echo was detected in an *R*-band image obtained at a wavelength of 6500 Å at the Calar Alto 2.2-m telescope on 6 October 2007

[1]Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany. [2]National Astronomical Observatory of Japan, 650 North A'ohoku Place, Hilo, HI 96720, USA. [3]Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721, USA.
*To whom correspondence should be addressed. E-mail: krause@mpia.de

with a peak surface brightness $R = 23.4 \pm 0.2$ mag arc sec^{-2}. No optical emission feature down to surface brightness limit $R = 25.1$ mag arc sec^{-2} (3-σ) was detected toward this position in a previous R-band image obtained at the Steward Observatory 90-inch telescope on 18 September 2006.

We have acquired a deep R-band image of the echo with the faint object camera and spectrograph (FOCAS) instrument at the Subaru telescope on 9 October 2007 (Fig. 2). The morphology of the optical emission in this image closely matches the mid-infrared one observed 50 days earlier, with the echo resolved into compact emission knots of about 2 arc sec diameter in the R-band image. A long-slit spectrum covering the northern one of these compact knots (Fig. 2B) was obtained with FOCAS on the same night, covering the wavelength range from 4760 to 9890 Å with a spectral resolution of 24 Å.

This echo spectrum unambiguously shows light of a supernova origin (Fig. 3): Broad emission lines with P-Cygni absorption components from neutral and singly ionized elements are detected, all of which are commonly observed in core-collapse supernovae. A prominent feature of the Cas A supernova spectrum is an Hα emission line with a full width at half maximum (FWHM) of 17,000 km s^{-1} and a blue-shifted absorption minimum at –11,000 km s^{-1}. Other strong lines are Na I D, the Ca II infrared triplet, and permitted emission lines of neutral oxygen at 7774 and 9264 Å. In addition, lines of He I 7065 Å and likely also He I 5876 Å (blended with Na I D) and He I 7821 Å (blended with [Ca II] 7291, 7324 Å) are detected.

Whereas the presence of the hydrogen line classifies Cas A as a type II supernova (SN), the additional appearance of weak helium lines is characteristic of the rare class of type IIb supernovae (SNe). They originate from the core collapse of massive stars that have lost most of their hydrogen envelopes before exploding and consist of a nearly bare helium core at the time of collapse. SNe IIb initially show a type II spectrum dominated by their hydrogen-poor envelope and gradually transform into a SN Ib spectrum from the inner helium core (*7*–*9*). A type IIb for the Cas A supernova is supported by comparing its spectrum with that of the prototypical type SN IIb, 1993J (*10*) (Fig. 3), the collapse of a red supergiant (*11*) with a main sequence mass of 13 to 20 $M_\odot$ (*12*, *13*) in the nearby galaxy M81. The spectra are remarkably similar in terms of the presence of important spectral features and their strengths. The echo spectrum visible on Earth represents supernova light over an interval of time around maximum brightness because of the spatial extent of the interstellar cloud; therefore, the comparison spectrum was derived as the time-average of brightness-weighted spectra of SN 1993J obtained during days 1 to 83 after collapse (*10*, *14*), consistent with the geometrical constraints of the echo (fig. S1).

The identity of the Cas A progenitor has been the subject of tremendous debate (*15*). A type II or IIb has been suggested for the Cas A supernova (*16*, *17*) on the basis of detection of a few remnant ejecta knots containing some hydrogen at space velocities between 9000 and 10,300 km s^{-1} (*17*, *18*). The flat-top shape of Hα emission in our Cas A spectrum is consistent with a thin hydrogen-rich shell above the photosphere and expanding at about 10,000 km s^{-1}. In contrast, the overall lack of hydrogen emission in most knots and the nitrogen enrichment in the remnant were widely interpreted as signatures of the collapse of a Wolf-Rayet star in a type Ib supernova (*18*). The SN Ib 2005bf was likely produced by a nitrogen-rich Wolf-Rayet star and has been proposed for a possible template for the Cas A supernova (*19*). However, the spectrum of SN 2005bf does not provide a good match to our spectrum of the Cas A supernova.

Further evidence for a red supergiant progenitor comes from the comparison of SN 1993J with Cas A (*16*): Radio and x-ray observations of SN1993J imply a mass-loss rate of $4 \times 10^{-5}\ M_\odot$ year^{-1} and a wind velocity of 10 km s^{-1} (*20*), close to an estimate of $2 \times 10^{-5}\ M_\odot$ year^{-1} and wind velocity of 10 km s^{-1} consistent with the hydrodynamical state of the Cas A remnant (*16*). The presence of broad absorption components in the Cas A spectrum and the absence of prominent unresolved lines differ from typical spectra of SNe IIn whose ejecta are directly interacting with dense circumstellar gas. In the case of Cas A, the dense circumstellar wind may therefore not have extended directly to the surface, and Cas A might have exploded into a bubble created by a temporary phase of enhanced wind velocity. In addition, evidence for CNO processing in the envelope of a red supergiant has been found in the remnants of both SN 1993J (*21*) and Cas A (*22*). A helium core mass of 3 to 6 $M_\odot$ inferred for SN 1993J (*12*, *13*) matches the total mass at core collapse of Cas A based on the observational constraints for its remnant (*15*).

The optically bright emission of a SN IIb results from the production of a substantial amount of ^{56}Ni, which for the SN 1993J supernova has been determined to be between 0.07 and 0.15 $M_\odot$ (*12*, *13*, *23*). Models of nucleosynthesis predict an associated amount of 0.7×10^{-4} to $1.7 \times 10^{-4}\ M_\odot$

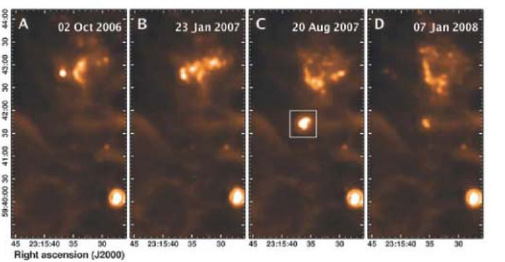


Fig. 1. Infrared images of the echo region. (**A** to **D**) MIPS 24-μm images of the same area of 2.5 arc min by 5 arc min with the corresponding observing epoch labeled on each panel. The bright infrared echo is visible at the center of (C). Other infrared echoes ~60 arc sec north of this feature are indicated by strong morphological changes in the time series; note for example the compact echo in (A), which disappeared later, in contrast to the smooth and unchanged interstellar cirrus emission. The white rectangle in (C) denotes the size of the optical images of the bright echo region shown in Fig. 2.

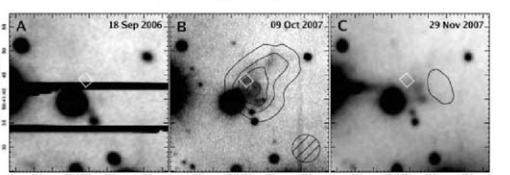


Fig. 2. Optical images of the echo region. (**A** to **C**) show R-band images of the same area of 33 arc sec by 33 arc sec with the corresponding observing epoch labeled on each panel. The white rectangle shown in all three panels denotes the 2 arc sec–by–2.1 arc sec extraction aperture of the Cas A supernova spectrum. Seeing was 1.6, 0.7, and 1.3 arc sec, FWHM, for (A), (B), and (C), respectively. The contours in (B) denote 24-μm emission from Fig. 1C in steps of 25 MJy sr^{-1} starting at 55 MJy sr^{-1}. The size of the MIPS 24-μm beam is indicated in the lower right (circle). The contours in (C) display 24-μm emission from Fig. 1D with the same levels plotted in (B).

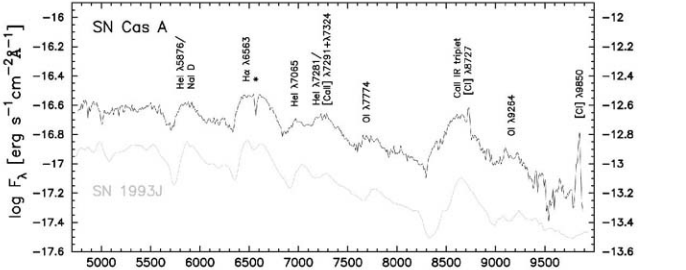


Fig. 3. Spectrum of the Cas A supernova and SN 1993J. Spectral features are labeled with their rest wavelength given in Å. The spectrum was extracted from the aperture shown in Fig. 2 and binned to 11.2 Å per pixel. Details of the data reduction are described in the SOM text. The comparison spectrum of SN1993J was dereddened by using $E(B-V) = 0.2$ mag (*28*) and shifted by $\log(F_\lambda) = 4$.

^{44}Ti produced for a stellar mass of 15 to 20 $M_\odot$ (*23*). A ^{44}Ti mass of $1.6^{+0.6}_{-0.3} \times 10^{-4}$ $M_\odot$ has been measured in the Cas A remnant (*24*). This is consistent with the Cas A supernova being an optically bright supernova such as SN 1993J.

Because the density and composition of the interstellar cloud giving rise to the echo is relatively unconstrained, it is difficult to accurately determine the peak brightness and light curve of Cas A from the scattered light. The circumstances regulating the infrared emission are simpler, and because the dust cooling time in the infrared echo is short (*5*) the rate of fading at 24 μm should be similar to the rate of fading of the heating energy pulse. The surface brightness at the peak position of the infrared (IR) echo within 140 days between 20 Aug 2007 and 7 January 2008 faded by a factor of 18 ± 3. This can be compared to a brightness decrease by a factor of 17 in the exponentially decaying bolometric light curve of SN 1993J (*25*) between day 33 and day 173.

It is a historical enigma whether Astronomer Royal Flamsteed witnessed the Cas A supernova on 16 August 1680 at sixth magnitude (*26*). For the peak visual brightness of –17.5 mag for SN 1993J (*25*) and a foreground extinction of $A_V \sim 8$ mag, a maximum visual brightness of 3.2 mag would be predicted. This value and the rapid decay (e.g., to sixth magnitude in only 2 months) are consistent both with the lack of widespread reportage, indicating the peak was fainter than third magnitude (*15*), and Flamsteed's observation. The visual extinction varies across the remnant, and we have considered here the most likely extinction value within the plausible range at the center of Cas A (*27*).

One aspect about Cas A remains puzzling: The progenitor of SN 1993J was a binary star as now confirmed by the detection of a companion (*28*). A progenitor of 15 to 25 $M_\odot$ that loses its hydrogen envelope to a binary companion and undergoes an energetic explosion is more consistent with the theoretical models and observational constraints for Cas A than an evolved single star undergoing core collapse (*15*). Although there is at present no evidence for a companion that has survived the explosion (*27*), it might be speculated that two binary companions merged during a common envelope phase before the explosion (*29*). The observed asymmetric distribution of the quasi-stationary flocculi (*18*, *30*) near Cas A might originate from the loss of such an envelope.

Lastly, we address one difference between the Cas A supernova spectrum compared with that of SN 1993J: the presence of two unresolved emission lines at 8727 and 9850 Å, which we suggest are from neutral carbon in the interstellar echo cloud. The cloud is located at a distance of 266 ± 23 light years to Cas A [supporting online material (SOM) text]. For the light curve of SN 1993J (*25*), the peak visual brightness of the Cas A supernova at the location of the cloud is –12.9 mag, slightly brighter than full moon and leading to a flux density of $F_{5600A} \sim 4.9 \times 10^{-4}$ erg cm^{-2} s^{-1} A^{-1}, which is about 1000 times stronger than the interstellar radiation field in the solar vicinity. The color temperature of SN 1993J near maximum light was close to 10,000 K (*25*). Although the relative flux of Lyman photons capable of ionizing hydrogen is low at such a color temperature, the excitation of carbon lines can still be substantial (*31*): Carbon is, by a factor of 10, the most abundant atom that can be ionized by photons less energetic than the Lyman limit. Emission lines of carbon are therefore ubiquitous in the predominantly neutral shielded environments of dense interstellar clouds irradiated with ultraviolet (UV)-optical radiation (*31*, *32*). The [C I] transitions at 8727 and 9850 Å lines are known to be the brightest carbon emission lines of photon-dominated regions in the optical wavelength range (*32*); however, these lines have so far never been observed as the most prominent lines in supernovae. Thus, it appears likely that the carbon lines are not intrinsic to the supernova spectrum but excited by the UV-optical supernova flash in the echo cloud.

References and Notes

1. R. A. Fesen *et al.*, *Astrophys. J.* **645**, 283 (2006).
2. J. E. Reed, J. J. Hester, A. C. Fabian, P. F. Winkler, *Astrophys. J.* **440**, 706 (1995).
3. S. E. Woosley, T. A. Weaver, *Annu. Rev. Astron. Astrophys.* **24**, 205 (1986).
4. O. Krause *et al.*, *Science* **308**, 1604 (2005).
5. J. Kim *et al.*, *Astrophys. J.*, in press http://arxiv.org/abs/0801.3267.
6. A. Rest *et al.*, *Astrophys. J.*, in press http://arxiv.org/abs/0801.4762.
7. D. A. Swartz, A. Clocchiatti, R. Benjamin, D. F. Lester, J. C. Wheeler, *Nature* **365**, 232 (1993).
8. A. V. Filippenko, T. Matheson, L. C. Ho, *Astrophys. J.* **415**, L103 (1993).
9. A. V. Filippenko, *Astron. J.* **96**, 1941 (1988).
10. T. Matheson *et al.*, *Astron. J.* **120**, 1487 (2000).
11. G. Aldering, R. M. Humphreys, M. Richmond, *Astron. J.* **107**, 662 (1994).
12. K. Nomoto *et al.*, *Nature* **364**, 507 (1993).
13. S. E. Woosley, R. G. Eastman, T. A. Weaver, P. A. Pinto, *Astrophys. J.* **429**, 300 (1994).
14. A. Clocchiatti *et al.*, *Astrophys. J.* **446**, 167 (1995).
15. P. Young *et al.*, *Astrophys. J.* **640**, 891 (2006).
16. R. A. Chevalier, J. Oishi, *Astrophys. J.* **593**, L23 (2003).
17. R. A. Fesen, R. H. Becker, *Astrophys. J.* **371**, 621 (1991).
18. R. A. Fesen, *Astrophys. J. Suppl. Ser.* **133**, 161 (2001).
19. N. Tominaga *et al.*, *Astrophys. J.* **633**, L97 (2005).
20. C. Fransson, P. Lundqvist, R. A. Chevalier, *Astrophys. J.* **461**, 993 (1996).
21. C. Fransson *et al.*, *Astrophys. J.* **622**, 991 (2005).
22. R. A. Chevalier, R. P. Kirshner, *Astrophys. J.* **219**, 931 (1978).
23. F.-K. Thielemann, K. Nomoto, M.-A. Hashimoto, *Astrophys. J.* **460**, 408 (1996).
24. M. Renaud *et al.*, *Astrophys. J.* **647**, L41 (2006).
25. M. W. Richmond *et al.*, *Astron. J.* **107**, 1022 (1994).
26. D. W. Hughes, *Nature* **285**, 132 (1980).
27. R. A. Fesen, G. G. Pavlov, D. Sanwal, *Astrophys. J.* **636**, 848 (2006).
28. J. R. Maund, S. J. Smartt, R. P. Kudritzki, P. Podsiadlowski, G. F. Gilmore, *Nature* **427**, 129 (2004).
29. P. Podsiadlowski, P. C. Joss, J. J. L. Hsu, *Astrophys. J.* **391**, 246 (1992).
30. S. S. Lawrence *et al.*, *Astron. J.* **109**, 2635 (1995).
31. M. Spaans, A. G. G. M. Tielens, E. F. van Dishoeck, E. L. O. Bakes, *Astrophys. J.* **437**, 270 (1994).
32. D. J. Hollenbach, A. G. G. M. Tielens, *Annu. Rev. Astron. Astrophys.* **35**, 179 (1997).

33. This report has its basis in data collected at the NAOJ Subaru Telescope, the Spitzer Space Telescope, and the German-Spanish Astronomical Center, Calar Alto. We appreciate allocations of Director's Discretionary Time at the Calar Alto observatory and thank U. Thiele and the Calar Alto observers for their support. G.H.R. acknowledges support for this work provided by NASA through Jet Propulsion Laboratory/Caltech contract 1255094.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1195/DC1
SOM Text
References

29 January 2008; accepted 8 April 2008
10.1126/science.1155788

Interfacial Polygonal Nanopatterning of Stable Microbubbles

Emilie Dressaire,[1] Rodney Bee,[2] David C. Bell,[1] Alex Lips,[2] Howard A. Stone[1]*

Micrometer-sized bubbles are unstable and therefore difficult to make and store for substantial lengths of time. Short-term stabilization is achieved by the addition of amphiphilic molecules, which reduce the driving force for dissolution. When these molecules crystallize on the air/liquid interface, the lifetime of individual bubbles may extend over a few months. We demonstrated low gas-fraction dispersions with mean bubble radii of less than 1 micrometer and stability lasting more than a year. An insoluble, self-assembled surfactant layer covers the surface of the microbubbles, which can result in nanometer-scale hexagonal patterning that we explain with thermodynamic and molecular models. The elastic response of the interface arrests the shrinkage of the bubbles. Our study identifies a route to fabricate highly stable dispersions of microbubbles.

In two-phase systems, such as foams, emulsions, pastes, and dispersions, the structures of the dispersed and continuous phases play a critical role in determining the properties of the material (*1*). Correlations have been established between the properties and the size, shape, volume fraction, spatial distribution, and connectivity of the dispersed phase (*2*–*4*). The metastability of these two-phase systems adds time dependence to this list of complex functions: The microstructure of the dispersed phase continuously evolves toward lower-energy configurations by minimizing the interfacial area, which diminishes the long-term usefulness of two-phase systems. The development of models predictive of the time evolution during coarsening (*5*, *6*) and of slowly aging systems with jammed interfaces (*7*, *8*) are among the approaches used to overcome this limitation.

The most important and rapid coarsening occurs in gas-liquid two-phase systems with fine structuring (1 to 10 μm), such as contrast agents for ultrasound imaging, aerated food and personal care products, and foamed construction materials (*9*–*11*). Because the air/liquid surface tension produces a pressure that drives bubble dissolution by gas diffusion into the liquid phase, the initial structure evolves and larger bubbles grow at the expense of smaller ones (Ostwald ripening). The time scale of this separation prevents the preparation of bubbles in the colloidal range because a micrometer-sized bubble dissolves in less than 1 s in pure water (*12*). The lifetime can be increased to hundreds of seconds by the addition of surfactant molecules that lower the interfacial tension and decrease the pressure difference (*13*). Further improvements in stabilization have been achieved by coating the air/liquid interface with ordered nanoparticles (*8*, *14*) or with gelled lipids (*15*–*17*); isolated bubbles with radii mostly greater than 10 μm were stable over a period of a few months.

We report very stable gas dispersions obtained using a standard multiphase mixing technique to trap air into surfactant shells within a viscous bulk phase (*18*). The surfactant solution was prepared by mixing at 70°C a highly viscous glucose syrup [75 weight percent (wt %)] with water (23 wt %) and sucrose stearate as the surfactant (2 wt %; principally a mixture of mono- and diesters). The formation of the dispersion was achieved by aeration, via shearing, at room temperature for 2 hours. The mechanical entrainment of air was followed by comminution; that is, the reduction in bubble size by shear-induced breakup.

The gas phase, which occupies about half the volume of the resulting foam (gas volume fraction ~ 0.45), is divided into surfactant-covered bubbles, or gas microcells, whose size ranged from hundreds of nanometers to tens of micrometers (Fig. 1A). The majority of the bubbles have radii R_b of about 1 μm and are characterized by a regular nanometer-scale structuring: Hexagonal domains that buckle outward from the bubble fully cover the air/liquid interface (Fig. 1B). These features were best observed with freeze-fracture transmission electron microscopy (TEM), as illustrated in Fig. 1C. A distribution of pentagons and heptagons disturb the lattice, which had typical dimensions such as scale ~50 nm. The pattern is not an imaging artifact or an isolated phenomenon because large bubble populations have been screened by cryogenic scanning electron microscopy (cryo-SEM), showing that the structure is systematically visible on these microbubbles (Fig. 1D) and can appear on larger ones. We characterized the inner bubble structure by cryo-TEM (Fig. 1E). Direct imaging induces a partial melting of the air/liquid interface, which reveals a spherical gas-filled cavity [supporting online material (SOM) text and fig. S1]. Taken together, the images of Fig. 1 provide evidence that, in this system, every microbubble has a nanometer-scale, nearly hexagonal surface pattern. Individual bubbles with some similar qualitative interfacial characteristics, but lacking regularity, have been observed in phospholipid-coated systems (*16*, *17*).

We describe here the dependence of the pattern size and the bubble radius on the processing conditions. Results were obtained by analyzing a few tens of TEM micrographs of freeze-fracture replicas for each dispersion (*18*). By sampling the mixture at different times during the aeration process, the evolution of the radii of the bubbles and the domains were monitored (SOM text and fig. S2). Figure 2, A and B, shows the

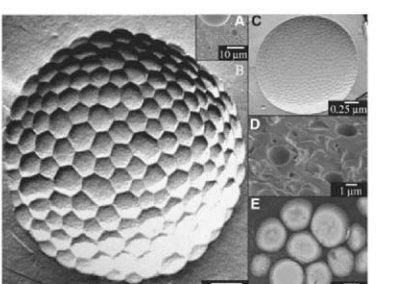


Fig. 1. Low gas-fraction dispersion of microbubbles with nanotextured surfaces. **(A)** Cryo-SEM image of the dispersion produced by shearing the surfactant mixture for 2 hours at a shear rate $\dot{\gamma} \approx 47 s^{-1}$. The highly polydisperse population of bubbles is characterized by radii ranging between hundreds of nanometers and tens of micrometers. **(B)** TEM image of a micrometer-sized bubble covered with hexagons ~50 to 100 nm in diameter. The platinum shadowing of the surface replica, under a 45° angle, reveals that the domains are buckled outward. **(C)** TEM image of a microbubble (radius $R_b \approx 1.5$ μm). The surface structure is regular both in shape (mostly hexagons) and in size (~50 nm); it includes a distribution of heptagons and pentagons. Patterns may look distorted because they sit on a curved interface. **(D)** Cryo-SEM image of the bubbles. The three-dimensional structuring of the air/glucose syrup interface appears on all of the micrometer-sized bubbles. **(E)** Cryo-TEM image of bubbles. The plunge freezing leads to nonspherical shapes. The electron beam partially melts the surfactant layer, which enables direct visualization of the hollowness of the bubbles.

[1]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. [2]Unilever Research and Development, 40 Merritt Boulevard, Trumbull, CT 06611, USA.

*To whom correspondence should be addressed. E-mail: has@seas.harvard.edu

evolution of the mean bubble and domain radii upon shearing and subsequent storage over a 45-day period. After 2 hours of shearing, the average bubble radius $\langle R_b \rangle$ is reduced from 10 μm to about 1 μm. The average scale of the surfactant patterns $\langle r \rangle$ continues to increase during the early times of storage until it reaches a value of ~50 nm (Fig. 2B). The final bubble radii vary with the shear rate, with smaller bubbles being produced at higher shear rates. Nevertheless, the final pattern sizes are ~50 nm, independent of the bubble size (Fig. 2C). Corresponding TEM images of the surface patterns when the bubble radii vary from 500 nm to 3 μm are shown in Fig. 2D. We also made dispersions using different surfactant mixtures with monoester content varying from 50 to 75 wt%. Over this composition range, the average bubble radius remains constant while the pattern size changes by a small amount (Fig. 2C).

The interface between air and liquid is covered with surfactant molecules, both mono- and diesters, which are irreversibly pinned because of their low solubility in glucose syrup. The hydrophilic headgroups of the sucrose stearate sit in the aqueous phase, whereas the hydrophobic carbonyl chains lie inside the microbubbles. The observed bulging domains (Fig. 1B) suggest that the surfactant molecules pack on the interface, with headgroups occupying substantially more surface area than the hydrophobic chains. This interpretation is consistent with the disproportionation process that results in the shrinkage of the bubble and an increase of the interfacial density of surfactant molecules. At the end of the aeration process, the mono- and diester molecules form a highly ordered phase that is naturally bent away from the air phase.

An early stage of the pattern formation is shown in Fig. 3A. We did not observe individual condensed phase domains (*17*). This indicates that the polygonal shapes do not result from the compression or close packing of originally isolated structures. Instead, the smooth surface initially buckles in a few locations with creases that initiate domain boundaries growing at 120°; the hexagonal patterning eventually covers the entire surface of the microbubbles (Fig. 3, A to C). Thus, the buckling is posterior to the interfacial condensation. As illustrated by Fig. 3B, the smooth regions (that is, the regions not yet covered with hexagons) project above the plane defined by the surrounding patterned area. The interfacial structuring is therefore associated with the reduction of the volume enclosed in a bubble of fixed surface area (Fig. 3D). This shrinkage of the microbubbles is consistent with Ostwald ripening, because these bubbles are the smallest ones of the dispersion. At room temperature, the conservation of the initial bubble surface area is due to the condensation of the interfacial layer and the irreversible pinning of the surfactant molecules.

More quantitatively, the surface state is considered to result from the thermodynamic evolution of an initially smooth, surfactant-covered interface. The domain size can be estimated by minimizing the energy of the microbubble upon shrinkage, driven by the Laplace pressure p. Each pattern is modeled as a spherical cap of radius R_c, which intersects the bubble surface with a circle of radius a (Fig. 3E). The total number of domains n, each of surface area A (a, R_c), is determined by conservation of the initial surface area of the bubble. We assume that the pattern geometry results from the energetic competition between the bending elasticity of the interface (bending rigidity κ), the formation of domain boundaries (line tension λ), and the normal pressure-volume work pV that accompanies shrinkage (*19*, *20*):

$$E(a, R_c) = n\left(A\frac{\kappa}{2}\left(\frac{2}{R_c} - \frac{2}{R_{sp}}\right)^2 + \lambda\pi a\right) - pV \quad (1)$$

The bending rigidity of condensed monolayers is typically $\kappa \approx 200\ k_B T$ (where k_B is the

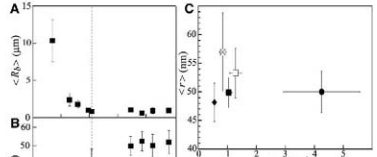


Fig. 2. TEM study of the evolution of bubble and pattern sizes. The plots show the evolution over 45 days of **(A)** the average bubble radius $\langle R_b \rangle$ (*18*) and **(B)** the average pattern size $\langle r \rangle$. The dashed line indicates the end of the 2-hour aeration process, after which the sample is left at rest at 4°C. During the shearing, $\langle R_b \rangle$ decreases and reaches its final value, whereas $\langle r \rangle$ monotonically increases. After 2 days, $\langle r \rangle$ remains ≈50 nm. t denotes time. Error bars indicate the SD of the size distributions. **(C** and **D)** Chemical and mechanical studies of the pattern size. Square symbols correspond to dispersions obtained by shearing different commercial sucrose stearate mixtures ($\dot{\gamma} \approx 47\ s^{-1}$). The monoester concentration is equal to ⊠, 50%; □, 70%; and ■, 75%. The subsequent decrease in sucrose distearate concentration does not affect $\langle R_b \rangle$, but $\langle r \rangle$ continually decreases. For the largest monoester concentration (75%), we applied different shear rates, ◆, $\dot{\gamma} \approx 62.7\ s^{-1}$; ■, $\dot{\gamma} \approx 47\ s^{-1}$; and ●, $\dot{\gamma} \approx 31.4\ s^{-1}$: $\langle r \rangle$ remains constant and $\langle R_b \rangle$ shows a significant decrease upon increase of the $\dot{\gamma}$, as illustrated by TEM micrographs in (D). Scale bars, 500 nm. Some replicas show macroscopic folds that are an artifact of the experimental protocol.

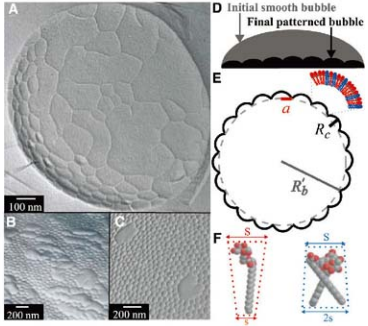


Fig. 3. Formation of the surface pattern. **(A)** TEM of a microbubble sampled after 20 min of aeration. The surface is not fully patterned but shows no isolated domains. Flat areas show short creases oriented away from existing patterns. **(B)** TEM image of a microbubble with nonpatterned areas, which, according to the shadowing, sit above the patterned areas. **(C)** TEM of a bubble prepared in a less viscous matrix (56 wt % glucose syrup). After the 2-hour aeration, flat areas remain. **(D)** Evolution from a smooth to a patterned surface. With a constant interfacial area, the buckling corresponds to a decrease in enclosed volume. **(E)** Schematic of the interfacial structure. The domains are modeled with spherical caps whose geometry (a, R_c) results from the packing of sucrose mono- and distearate (red and blue, respectively). **(F)** Structure of sucrose mono- and distearate (drawn using Chem 3D Ultra 9 software). At the interface, the molecules can be modeled by a truncated cone with a hydrophilic base or headgroup of surface area (s) and a hydrophobic base of surface area that is proportional to the number of C_{18} chains (s, monoester; $2s$, diester).

Boltzmann constant and T the absolute temperature) $\approx 8.10^{-19}$ J (*21*). The spontaneous radius of curvature of the monolayer R_{sp} depends on the composition of the layer and the geometry of the surfactant molecules (*22*), which are modeled by truncated cones (Fig. 3F). The composition of the interface can be estimated from the presence of lamellae in nonaerated solutions of the same composition (fig. S3), which yields for the condensed phase the value of the packing parameter (*23*), the composition, and $R_{sp} \approx 80$ nm (SOM text). In addition, the line-tension energy is defined as the product of the line tension ($\lambda \approx k_B T\, l^{-1} \approx 2.10^{-12}$ J m^{-1}, where l is the length of the surfactant molecule) and the domain perimeter; the reduction by a factor of 2 in Eq. 1 arises because each boundary is shared by two domains. Finally, the pV term is approximated with the Laplace pressure on a micrometer-scale bubble with a surface tension $\gamma \approx 5 \times 10^{-4}$ J m^{-2}, which is on the order of the values measured for highly compressed monolayers such as pulmonary surfactants (*24*), so $p \approx 10^3$ Pa. The volume V of the buckled bubble is estimated as that of the spherical core ($V \approx \frac{4}{3}\pi R_b^3$) (Fig. 3E).

The energy minimization procedure for a bubble of $R_b = 1$ μm leads to an estimate for the radius of the domains, $a \approx 40$ nm (SOM text), which lies within the range of the experimentally measured radii. The model also qualitatively captures the influence of the surfactant composition and bubble radius on the pattern size. For example, increasing the amount of diester in the initial mixture is expected to lead to an increased concentration of diester molecules at the gas/liquid interface, and therefore to a larger spontaneous radius of curvature. The model in turn predicts larger pattern sizes (fig. S4). Similarly, the radius of the microbubble is expected to influence the number of patterns rather than their size (fig. S5), which is also consistent with the data (solid symbols in Fig. 2C). Finally, according to the experimental observations and to predictions of the model, the presence of patterns with increasing size during the preparation of the dispersion and the early time of the storage (Fig. 2B) can be rationalized either by evolution of the interfacial composition, which leads to changes in R_{sp}, or by changes in interfacial properties.

The time evolution of the distribution of bubble sizes has been characterized over 1 year, using SEM images (Fig. 4A) to estimate bubble radii (inset in Fig. 4A). The average bubble radius, R_b, increases from about 0.7 to 3.5 μm while the distribution broadens, with the standard deviation increasing by a factor of 10. These observations are consistent with destabilization by Ostwald ripening, which becomes significant only over long periods of time.

The shrinkage of the microbubbles is driven by the Laplace pressure due to the presence of narrow regions between domains that are hypothesized to be in a liquid state. Mass transfer of the gas can be characterized by a permeability or average speed of transport that, through the bulk phase around a bubble of $R_b \approx 1$ μm, is approximately $D/R_b \approx 10^{-5}$ cm s^{-1}, where the diffusion constant of air in glucose syrup D is estimated using the Stokes-Einstein relation (SOM text). The surfactant monolayer is expected to have a higher permeability (10^{-2} to 10^{-3} cm s^{-1}) (*25*) and therefore not limit the shrinkage. Thus, the time scale for the diffusion-controlled dissolution of micrometer-size bubbles in glucose syrup is estimated to be about 5 hours [SOM text (*13*)], which is indeed the observed lifetime of SDS-coated microbubbles in glucose syrup but is orders of magnitude shorter than our observations (Fig. 4A). Therefore, reduced diffusion due to a viscous bulk phase cannot explain why the lifetime of microbubbles coated with mono- and diesters extends to more than a year.

Previous studies have shown that interfacial elasticity can stop or substantially reduce bubble dissolution (*26*, *27*). In our system, the bending elasticity of the domains resists the compression of the interface upon shrinkage. Hence, we approximate the arrest of dissolution by diffusion as being when the bubble reaches a radius R at which the increase in bending energy away from the initial state balances the work of the Laplace pressure (SOM text): $\frac{R}{R_0} \approx 0.94$. This estimate shows that the buildup of elastic stress in the membrane arrests the shrinkage of the microbubbles very early and rationalizes the observed longevity of the dispersion.

According to the molecular model that we propose, the packing of the surfactant molecules controls the interfacial pattern structure and the stability of the bubbles. In order to test this idea further, we experimentally modified the composition of the original amphiphilic mixture by adding stearic acid. This surfactant molecule has a smaller headgroup, 0.2 nm^2 in water (*28*), but the same carbonyl chain, C_{18}, as the sucrose esters. When the acid concentration increases, the domains get larger ($a \approx 80$ nm for a stearic-acid concentration equal to 0.04 wt %) then lose their hexagonal shape, and finally disappear as the surface of the bubble becomes smooth (Fig. 4, B to E). Upon addition of stearic acid, the average headgroup size of the surfactant mixture decreases, which is expected to result in an increase of R_{sp}. The thermodynamic model predicts an increase in pattern size, which is consistent with our observations (Fig. 4, B to E, and fig. S4).

Our preparation and long-term study of microbubbles in a viscous continuous phase demonstrate that it is possible to achieve stabilization of such a dispersion for more than 1 year. The observed regular surface patterning is the thermodynamic signature of the formation of an elastic, condensed surfactant phase, which also correlates with the extended stability of the system. The nanometer-scale features can be tuned by modification of the chemical composition of the interface. Beyond the experimental observations reported here, our approach will be useful for designing more complex surfaces by controlling the self-assembly of the surfactant and for studying the behavior of individual stable bubbles of colloidal dimensions.

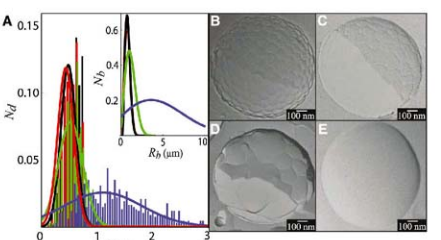


Fig. 4. Stability of microbubbles and chemical modification of surface patterns. **(A)** Evolution of distributions of microbubble radii over 1 year. For each histogram, we used SEM images to measure the radii of about 800 disks (R_d), which are the intersections of the bubbles with the plane along which the sample has been fractured. N_d is the fraction of bubbles that intersect the plane with a radius R_d which is less than or equal to the actual radius of the bubble (R_b). These distributions are fitted with Gaussian curves, which are used to extrapolate the distributions of the radii (R_b) of the bubbles (inset) (*29*). N_b is the fraction of bubbles with radius R_b. Samples of different ages were studied: 2 days (red), 1 week (black), 4 months (green), and 1 year (blue). The distributions show a modest shift and broadening over time. **(B to E)** Evolution of pattern structure upon addition of stearic acid. The TEM micrographs correspond to 1-week-old samples prepared by addition of (B) 0 wt %, (C) 0.04 wt %, (D) 0.08 wt %, and (E) 0.2 wt % of stearic acid. The pattern evolves, first showing a size increase (C), then losing the regularity (D), before disappearing (E).

References and Notes

1. D. Myers, *Surfaces, Interfaces and Colloids* (Wiley-VCH, New York, 1999).
2. R. G. Larson, *The Structure and Rheology of Complex Fluids* (Oxford Univ. Press, Oxford, 1999).
3. M. Wang, J. Wang, N. Pan, S. Chen, J. He, *J. Phys. D Appl. Phys.* **40**, 260 (2007).

4. S. Subramonian, L. Remy, D. Schroer, *Cell. Polym.* **23**, 349 (2004).
5. S. Hilgenfeldt, S. A. Koehler, H. A. Stone, *Phys. Rev. Lett.* **86**, 4704 (2001).
6. R. D. MacPherson, D. J. Srolovitz, *Nature* **446**, 1053 (2007).
7. K. Stratford, R. Adhikari, I. Pagonabarraga, J. C. Desplat, M. E. Cates, *Science* **309**, 2198 (2005).
8. M. Abkarian *et al.*, *Phys. Rev. Lett.* **99**, 188301 (2007).
9. S. H. Bloch, R. E. Short, K. W. Ferrara, E. R. Wisner, *Ultrasound Med. Biol.* **31**, 439 (2005).
10. S. Cohen-Addad, H. Hoballah, R. Hohler, *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **57**, 6897 (1998).
11. O. Faruk, A. K. Bledzki, L. M. Matuana, *Macromol. Mater. Eng.* **292**, 113 (2007).
12. P. S. Epstein, M. S. Plesset, *J. Chem. Phys.* **18**, 1505 (1950).
13. P. B. Duncan, D. Needham, *Langmuir* **20**, 2567 (2004).
14. E. Dickinson, R. Ettelaie, T. Kostakis, B. Murray, *Langmuir* **20**, 8517 (2004).
15. D. H. Kim, M. J. Costello, P. B. Duncan, D. Needham, *Langmuir* **19**, 8455 (2003).
16. E. Talu, M. M. Lozano, R. L. Powell, P. A. Dayton, M. L. Longo, *Langmuir* **22**, 9487 (2006).
17. M. A. Borden *et al.*, *Langmuir* **22**, 4291 (2006).
18. Materials and methods are available as supporting material on *Science* Online.
19. T. Baumgart, S. T. Hess, W. W. Webb, *Nature* **425**, 821 (2003).
20. H. Diamant, T. A. Witten, A. Gopal, K. Y. C. Lee, *Europhys. Lett.* **52**, 171 (2000).
21. J. Daillant, J. J. Benattar, L. Bosio, *J. Phys. Condens. Matter* **2**, 405 (1990).
22. G. Garofalakis, B. S. Murray, *Langmuir* **18**, 4765 (2002).
23. J. N. Israelachvili, *Intermolecular and Surface Forces* (Academic Press, New York, 1992).
24. L. M. Y. Yu *et al.*, *J. Appl. Physiol.* **97**, 704 (2004).
25. M. A. Borden, M. L. Longo, *J. Phys. Chem. B* **108**, 6009 (2004).
26. W. Kloek, T. Van Vliet, M. Meinders, *J. Colloid Interface Sci.* **237**, 158 (2001).
27. A. R. Cox, F. Cagnol, A. B. Russell, M. K. Izzard, *Langmuir* **23**, 7995 (2007).
28. Y. Wang, X. Du, L. Guo, H. Liu, *J. Chem. Phys.* **124**, 134706 (2006).
29. S. F. Edwards, D. R. Wilkinson, *J. Phys. D Appl. Phys.* **13**, 209 (1980).

30. The authors acknowledge the following individuals for preliminary work: Industrial training students from the University of Wageningen ran the first aeration experiments at Unilever Research (Colworth, UK); M. Fitzgerald from the University of Reading, UK, and S. Brodbank from the University of Bristol, UK, studied the surfactant mixtures during their Ph.D. studies. We thank A. Weaver, J. Brigham, M. Kirkland, and T. Belmar of Unilever Research; R. Schaleck and E. Hodges of Center for Nanoscale Systems at Harvard University; B. Graham and A. Graham of Bibs Lab for their help with the imaging; M. Brenner, L. Courbin, J. Hutchinson, J. Lucassen, and D. Weitz for helpful conversations; and Unilever Research for support of this work.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1198/DC1
Materials and Methods
SOM Text
Figs. S1 to S5
References

26 December 2007; accepted 25 April 2008
10.1126/science.1154601

The Sensitivity of Polar Ozone Depletion to Proposed Geoengineering Schemes

Simone Tilmes,[1]* Rolf Müller,[2] Ross Salawitch[3]

The large burden of sulfate aerosols injected into the stratosphere by the eruption of Mount Pinatubo in 1991 cooled Earth and enhanced the destruction of polar ozone in the subsequent few years. The continuous injection of sulfur into the stratosphere has been suggested as a "geoengineering" scheme to counteract global warming. We use an empirical relationship between ozone depletion and chlorine activation to estimate how this approach might influence polar ozone. An injection of sulfur large enough to compensate for surface warming caused by the doubling of atmospheric CO_2 would strongly increase the extent of Arctic ozone depletion during the present century for cold winters and would cause a considerable delay, between 30 and 70 years, in the expected recovery of the Antarctic ozone hole.

Geoengineering schemes have been proposed to alleviate the consequences of global warming (*1*–*3*) by continuous injection of sulfur into the stratosphere. Volcanic eruptions in the past have shown that strongly enhanced sulfate aerosols in the stratosphere result in a higher planetary albedo, leading to surface cooling (*4*). On the other hand, the potential for exceedingly high Arctic ozone depletion resulting from the simultaneous presence of high surface area density (SAD) of sulfate aerosols and cold conditions in the polar stratosphere is known (*5*, *6*) but was not quantified in the content of geoengineering (*1*–*3*).

In this report, the impact of enhanced sulfate aerosol (due to geoengineering) on future chemical polar ozone depletion is quantified. Our analysis is based on the past dependence of ozone loss on aerosol content derived via the combination of a model and measurement for both the Antarctic and Arctic and by taking into account the expected future stratospheric halogen loading. We describe this dependence by the empirical relation between observed chemical ozone loss and the potential for the activation of chlorine (PAC!). PAC! accounts for year-to-year variations in temperature, sulfur burden, and the halogen content of the stratosphere. This relation is shown to be valid based on past observations, including 4 years of volcanically enhanced aerosol loading in the stratosphere.

Severe chemical loss of ozone over the Arctic and Antarctica is caused by anthropogenic halogens. The combination of very low temperatures and increasing sunlight after the polar night results in a strong transformation of chlorine from reservoir forms to reactive radicals, leading to the rapid destruction of polar ozone (*7*). Since the 1990s, most of the available ozone has been destroyed in the Antarctic lower stratosphere (between 12 and 20 km in altitude), which corresponds to a loss in column ozone of about 120 to 150 DU (or Dobson units, one of which equals 2.687×10^{16} molecules per cm^2) (*8*). In the Arctic, the interannual variability of temperatures, and therefore of ozone depletion, is much larger. Chemical depletion of ozone in recent cold Arctic winters exceeded 100 DU [see the supporting online material (SOM)] (*8*–*11*). Over the next half-century, the stratospheric halogen loading—commonly quantified by a measure referred to as Effective Equivalent Stratospheric Chlorine (EESC)—is projected to slowly decline (*12*). Close to the year 2070, EESC is predicted to reach values last seen in 1980, a benchmark for the recovery of polar ozone (*12*, *13*). Our work is motivated by the concern that elevated SAD attributable to geoengineering will lead to additional ozone depletion that may delay the recovery of polar ozone.

We demonstrate the importance of chlorine activation on cold, liquid sulfate aerosols for polar ozone depletion by comparing two relations between chemical ozone loss and chlorine activation based on past observations (Fig. 1). The first relation is based on a previous description of polar chlorine activation: the polar stratospheric cloud (PSC) formation potential or PFP (*8*). PFP describes the fraction of the polar region that is below the threshold temperature for existence of PSCs (T_{PSC}). For cold Arctic winters, more than 10% of the vortex region is cold enough to support PSCs, whereas for warm winters, PFP is close to zero. The derived linear relation is compact, except for values derived for the four winters after the eruption of Mount Pinatubo in June 1991.

The second relation (Fig. 1B) accounts additionally for stratospheric aerosol loading. Tilmes *et al.* (*14*) defined the potential for the activation of chlorine, which is similar to PFP but uses a threshold temperature for chlorine activation (T_{ACl}) as described by Drdla (*15*). A description of the calculation of this temperature can be found in (*14*), which is a function of temperature, ambient H_2O, and SAD. For background levels of SAD (e.g., an atmosphere not perturbed by large volcanic eruptions) and present-day values of EESC, PAC! is comparable to PFP. After a strong

[1]National Center for Atmospheric Research (NCAR), Boulder, CO 80307, USA. [2]Research Center Jülich, 52425 Jülich, Germany. [3]University of Maryland, College Park, MD 20742, USA.

*To whom correspondence should be addressed. E-mail: tilmes@ucar.edu

volcanic eruption, PACl is larger than PFP because the resulting sulfate aerosols (fig. S1) provide surfaces that lead to efficient chlorine activation (*5*, *15*). The linear relation between chemical loss of Arctic ozone and PACl (Fig. 1B) is compact, including data collected during the four winters that followed the eruption of Mount Pinatubo (1992 to 1995).

For Antarctica, an empirical relation between chemical ozone loss and PACl cannot be inferred, because available observations do not span an appreciable range of ozone loss. Therefore, the Antarctic relation is defined based on results from the NCAR Whole Atmosphere Chemistry Climate Model 3 (WACCM3), for changing halogen content of the stratosphere between 1960 and the present (fig. S2). The model results are in good agreement with recent data, supporting the validity of the approach (*14*). Model simulations of Antarctic ozone loss (fig. S2) combined with recent data suggest that the relation between chemical ozone loss and PACl is linear until loss saturation occurs, which supports the validity of the approach used to estimate Arctic ozone loss for the geoengineering scenarios.

The empirical relation between chemical ozone loss and PACl (which incorporates the stratospheric sulfur burden) provides a tool to assess the risk of future ozone loss caused by geoengineering. We consider three different hypothetic future SAD scenarios:

1) The background case assumes no volcanic perturbation and is based on SAD data from the year 2000 (*16*).

2) The geoengineering scenario discussed by Crutzen (*1*) is termed "GeoEng_Large_Aerosol." Crutzen estimated that roughly 5.3 Tg (1 Tg = 10^{12} g) of stratospheric sulfur (S) would counteract surface warming due to doubled atmospheric CO_2. He considered volcanic-sized particles that require injections of 2 Tg S/year to maintain. For this case, we estimate SAD by multiplying observed SAD (*16*) in 1992 by 0.53 [5.3 Tg S (Crutzen)/ 10 Tg S (observed Pinatubo)] for all years.

3) The geoengineering scenario presented by Rasch *et al.* (*2*) is denoted "GeoEng_Small_Aerosol." They found that an injection of 1.5 Tg S/ year, using particles considerably smaller than those assumed by Crutzen (*1*), would achieve the same radiative effect. Smaller aerosols are expected to cool more efficiently than large aerosols because

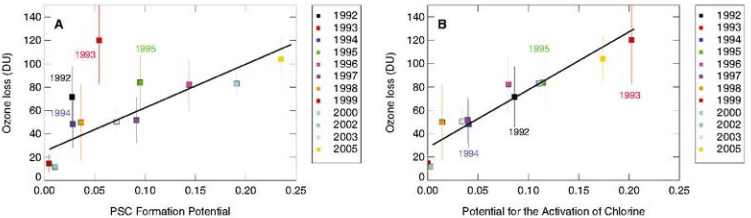


Fig. 1. (A) Relationship between chemical loss of Arctic ozone in DU and PFP, averaged between 380 to 550 K potential temperature and mid-December and March, for several winters between 1992 to 2005. A linear fit (black line) was derived, excluding the years 1992 and 1993 after the Mount Pinatubo eruption. **(A)** is adapted from (*6*), and the ozone loss value for 2005 is taken from (*9*). **(B)** As in **(A)** but with PACl (*14*) instead of PFP as the abscissa. PACl includes SAD in its formulation. Data for winters 1992 to 1995 (denoted within the panels), which had high SAD as a result of the eruption of Mount Pinatubo, now fall along the compact, near-linear relation once the effect of SAD on chlorine activation is considered.

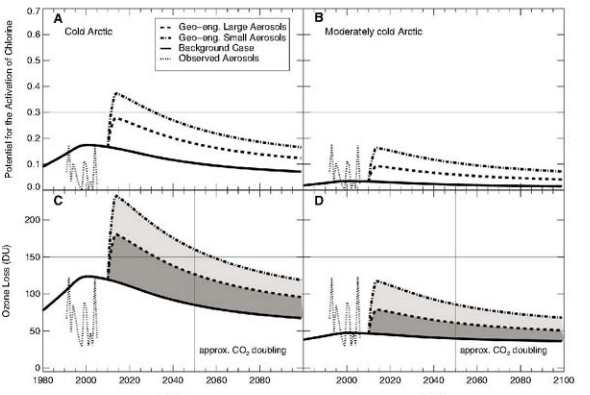


Fig. 2. (A and **B)** The temporal development of PACl between 2010 and 2050, taking into account changing EESC for two geoengineering cases, with observed temperatures for a very cold Arctic winter **(A)** and a moderately cold Arctic winter **(B)**. The temporal evolution of PACl for background SAD (solid line), taking into account changing EESC, is also shown. Finally, values of PACl based on observed SAD, temperature, and EESC, are shown (dotted lines). **(C** and **D)** Chemical ozone loss versus time, derived from PACl (top panels) for the various SAD cases (dark gray, GeoEng_Large_Aerosol case; light gray, GeoEng_Small_Aerosol case), is shown for meteorological conditions corresponding to a very cold Arctic winter (C) and a moderately cold Arctic winter (D). The ozone loss estimates are based on the linear relationship between chemical loss and PACl for the Arctic.

of the dependence of backscattering on particle size. Furthermore, smaller aerosols have a smaller effect on stratospheric temperature. The GeoEng_Small_Aerosol case has a perturbation to SAD three times as large as the GeoEng_Large_Aerosol perturbation, owing to the dependence of SAD on the particle radius (*17*). The resulting SAD used here is based on the mean of the Rasch *et al.* (*2*) perturbations resulting from injections of 1 and 2 Tg S/year, because an injection of ~1.5 Tg S/year was found to effectively counteract global warming for doubled atmospheric CO_2 (*2*).

The impact of these three sulfate loadings on PACl and ozone loss for past meteorological conditions is discussed in the SOM (see also fig. S3).

To assess the risk of future polar ozone loss caused by enhanced sulfate attributable to geoengineering, we select meteorological conditions for a recent very cold Arctic winter, a moderately cold Arctic winter, and a typical Antarctic winter, and apply these conditions to a model constrained by an estimate of future EESC (*13*). PACl and ozone depletion are quantified for the various cases (Figs. 2 and 3). Our results should be viewed as a "generic assessment" of the impact of geoengineering on future polar ozone, assuming a yearly injection of stratospheric sulfate. We assume geoengineering to begin, hypothetically in the year 2010, with a steady rise of SAD over the first 5 years until saturation is achieved. Results for other start dates can be visualized by simply connecting the background PACl case to the two geoengineering scenarios, for any specific start year or duration of SAD rise time.

The scenario that uses meteorological conditions for a recent very cold Arctic winter characterizes the maximum perturbation to PACl and Arctic ozone loss due to geoengineering (Fig. 2, A and C). PACl in the Arctic would exceed the maximum value of PACl for background conditions until about 2055 for GeoEng_Large_Aerosol and through the end of this century for GeoEng_Small_Aerosol. The estimate for a moderately cold winter (e.g., winter 2003) illustrates the response to geoengineering for meteorological conditions that are representative of about half of the past 15 Arctic winters (Fig. 2, B and D).

Injection of sulfur in the near future (during the next 20 years) can have a strong impact on Arctic ozone depletion. If small-sized aerosols are used, ozone loss between 100 DU (moderately cold winters) and 200 to 230 DU (very cold winters) can be reached. For very cold winters, which occurred 25% of the time in the past 15 years, the estimated ozone depletion is comparable to the total amount of available ozone in the Arctic lower stratosphere (fig. S2). Under these conditions, the SAD perturbation associated with the GeoEng_Small_Aerosol case could possibly result in a saturation of chemical loss of Arctic ozone, leading to a drastically thinner ozone layer than presently observed. For the SAD perturbation associated with GeoEng_Large_Aerosol, chemical loss of Arctic ozone could exceed 150 DU during very cold winters and 70 DU for moderately cold winters.

The doubling of atmospheric CO_2 with respect to preindustrial values is expected to occur between 2050 and 2100 (*18*). In this time frame, chemical ozone loss attributable to geoengineering reaches 125 to 150 DU for very cold Arctic winters, as compared with 80 DU for the background case. For either geoengineering case, ozone depletion would exceed presently observed values, because the recent spate of very cold Arctic winters occurred for low (i.e., near background) values of SAD. A moderately cold winter would reach 60 to 80 DU of ozone depletion, a value observed in the past only for very cold Arctic winters. Therefore, 75% of all winters would result in ozone loss of at least 60 to 80 DU, and possibly as high as 150 DU, if geoengineering through stratospheric sulfur injections is used to mitigate global warming.

The predicted future evolution of Antarctic ozone and PACl is depicted in Fig. 3. We show results for two different assumptions: either future chemical loss of Antarctic ozone due to geoengineering will continue to saturate at present-day values of 150 DU (*8*) (Fig. 3B), or else future chemical loss will saturate at higher levels, resulting from the likely strong increase of SAD at altitudes of 10 to 12 km in the Antarctic stratosphere, where ozone loss is presently not saturated (Fig. 3C). Indeed, enhanced Antarctic ozone loss at 10- to 12-km altitudes was observed after the eruption of Mount Pinatubo (*19*). For Fig. 3C, we assume that geoengineering will lead to a downward extension of ozone loss, adding another 15 DU (the amount between altitudes of 10 and 12 km) to the total column abundance of ozone that could be lost.

The primary effect of geoengineering on Antarctic ozone would be to delay the recovery of the Antarctic ozone hole. The time when Antarctic ozone loss drops below saturation in late winter would be delayed by 15 years for

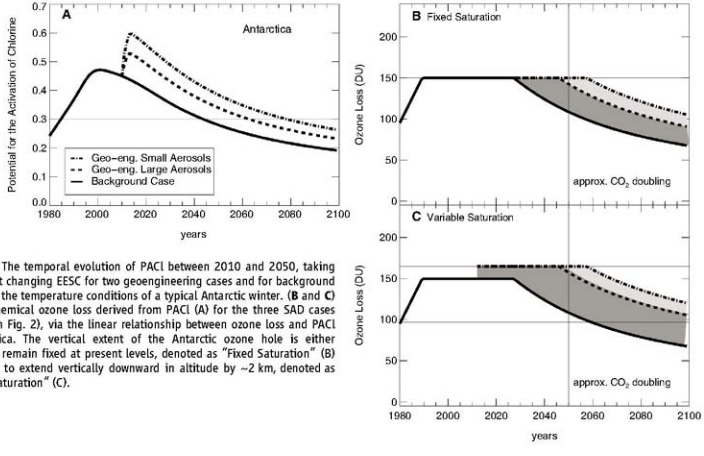


Fig. 3. (A) The temporal evolution of PACl between 2010 and 2050, taking into account changing EESC for two geoengineering cases and for background aerosol, for the temperature conditions of a typical Antarctic winter. **(B** and **C)** Antarctic chemical ozone loss derived from PACl **(A)** for the three SAD cases (colors as in Fig. 2), via the linear relationship between ozone loss and PACl for Antarctica. The vertical extent of the Antarctic ozone hole is either assumed to remain fixed at present levels, denoted as "Fixed Saturation" **(B)** or assumed to extend vertically downward in altitude by ~2 km, denoted as "Variable Saturation" **(C)**.

GeoEng_Large_Aerosol and by more than 30 years for GeoEng_Small_Aerosol (Fig. 3B). Assuming a vertical expansion of the ozone hole, the first stage of recovery would be delayed by another 30 years for both cases (Fig. 3C). The recovery of Antarctic ozone to conditions that prevailed in 1980 would be delayed until the last decade of this century by geoengineering, assuming no vertical expansion of the ozone loss region (Fig. 3B). However, if the vertical extent of the ozone hole were to increase, this stage of recovery would not occur within this century (Fig. 3C).

Our estimates of the risk of geoengineering do not consider several important effects. A possible consequence of a stratospheric sulfur injection, which was not considered here, is a strengthening of the polar vortex, because of the stronger temperature gradient between high and low latitudes caused by enhanced stratospheric aerosol (*20*). This could increase the frequency of cold polar winters, leading to even greater ozone loss than estimated. On the other hand, volcanic aerosols might result in dynamic instabilities, causing earlier major stratospheric warming at the end of Arctic winters, which results in smaller PACl and therefore less ozone loss. Also, enhanced sulfate aerosols could suppress denitrification in the polar vortices in a manner that might affect the linear relation between ozone loss and PACl. Enhanced ozone loss through geoengineering would also likely result in a thinning of the ozone layer at mid-latitudes due to the export of polar processed air, as observed during the period of rising halogen loading (1979 to 1995) (*21*). Further, the enhanced stratospheric aerosol levels would disturb ozone photochemistry in mid-latitudes, resulting from the suppression of stratospheric NO_x, leading to even further ozone depletion (*22*). Additional uncertainty results from our use of idealized aerosol size and perturbation. Finally, the impact on ozone of any future major volcanic eruption would likely exceed, by large amounts, the ozone loss that was observed after the eruption of Mount Pinatubo and El Chichón, because additional volcanic aerosols will be acting on an already perturbed stratospheric SAD layer. Comprehensive chemistry–climate model simulations, considering these effects and perhaps others, are needed to fully evaluate the impact of geoengineering on atmospheric ozone.

A substantial increase of stratospheric sulfate aerosol densities caused by various geoengineering approaches will likely result in strongly enhanced chemical loss of polar ozone during the next several decades, especially in the Arctic. The expected recovery of the Antarctic ozone hole due to a reduction in stratospheric halogen loading, brought about by the implementation of the Montreal Protocol, could be delayed by between 30 and 70 years by the aerosol perturbation associated with geoengineering.

References and Notes

1. P. J. Crutzen, *Clim. Change* **77**, 211 (2006).
2. P. J. Rasch, P. J. Crutzen, D. B. Coleman, *Geophys. Res. Lett.* **35**, L02809 (2008).
3. T. M. L. Wigley, *Science* **314**, 452 (2006).
4. D. A. Randall *et al.*, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2007), chap. 8.
5. A. Tabazadeh, K. Drdla, M. R. Schoeberl, P. Hamill, O. B. Toon, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2609 (2002).
6. S. Tilmes *et al.*, *Atmos. Chem. Phys.* **8**, 1897 (2008).
7. S. Solomon, *Rev. Geophys.* **37**, 275 (1999).
8. S. Tilmes, R. Müller, A. Engel, M. Rex, J. M. Russell III, *Geophys. Res. Lett.* **33**, L20812 (2006).
9. M. von Hobe *et al.*, *Geophys. Res. Lett.* **33**, L17815 (2006).
10. J. J. Jin *et al.*, *Geophys. Res. Lett.* **33**, L15801 (2006).
11. M. Rex *et al.*, *Geophys. Res. Lett.* **33**, L23808 (2006).
12. EESC is an index that represents the temporal evolution of the chlorine- and bromine-containing (i.e., halogen-containing) gases in the stratosphere. The index accounts for projections of future abundances, the atmospheric lifetime of organic source molecules, and the relative efficiency of ozone loss by chlorine and bromine compounds (*13*). Here, a mean age of 5.5 years, appropriate for the polar regions, has been used (*13*).
13. P. A. Newman, J. S. Daniel, D. W. Waugh, E. R. Nash, *Atmos. Chem. Phys.* **7**, 4537 (2007).
14. S. Tilmes *et al.*, *J. Geophys. Res.* **112**, D24301 (2007).
15. K. Drdla, paper presented at the American Geophysical Union Fall Meeting, San Francisco, 5 to 9 December 2005.
16. L. Thomason, T. Peter, Eds. *Assessment of Stratospheric Aerosol Properties* (report no. 4, Stratospheric Process and Their Role in Climate, World Climate Research Programme 124, World Meteorological Organization/ TD-No. 1295, Toronto, 2006).
17. SAD is proportional to sulfate mass/r_{mode}, where r_{mode} is particle mode radius. The large-aerosol case assumes a similar perturbation to sulfate mass as the small-aerosol case. Therefore, the r_{mode} of the latter case, which is one-third as small as that of the former case, will result in a three times larger perturbation to SAD (*21*). Our estimate of SAD is derived from scaling the Mount Pinatubo perturbation based on information in (*2*, *3*).
18. G. A. Meehl *et al.*, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2007), chap. 10.
19. D. J. Hofmann, S. J. Oltmans, *J. Geophys. Res.* **98**, 18,555 (1993).
20. G. Stenchikov *et al.*, *J. Geophys. Res.* **107**, 4803 (2002).
21. M. P. Chipperfield, V. E. Fioletov, in *Scientific Assessment of Ozone Depletion: 2006* (report no. 50, Global Ozone Research and Monitoring Project, World Meteorological Organization, Geneva, 2007), chap. 3.
22. S. Solomon *et al.*, *J. Geophys. Res.* **101**, 6713 (1996).

23. We acknowledge K. Drdla for helpful discussions and for sharing her formula to calculate T_{NAT} and PACl (*14*). P. Rasch for a very helpful discussion about the SAD distributions, and the UK Meteorological Office for providing meteorological analyses. The University Corporation for Atmospheric Research (UCAR) supported the portion of this work conducted at NCAR and the NASA Atmospheric Composition, Modeling, and Analysis Program supported the portion of this work conducted at the University of Maryland.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1153966/DC1
SOM Text
Figs. S1 to S3
References

10 December 2007; accepted 9 April 2008
Published online 24 April 2008;
10.1126/science.1153966
Include this information when citing this paper.

Water Activity and the Challenge for Life on Early Mars

Nicholas J. Tosca,[1]* Andrew H. Knoll,[1] Scott M. McLennan[2]

In situ and orbital exploration of the martian surface has shown that acidic, saline liquid water was intermittently available on ancient Mars. The habitability of these waters depends critically on water activity (a_{H_2O}), a thermodynamic measure of salinity, which, for terrestrial organisms, has sharply defined limits. Using constraints on fluid chemistry and saline mineralogy based on martian data, we calculated the maximum a_{H_2O} for Meridiani Planum and other environments where salts precipitated from martian brines. Our calculations indicate that the salinity of well-documented surface waters often exceeded levels tolerated by known terrestrial organisms.

Because liquid water is required by all organisms on Earth, evidence of current or past water has been viewed as a first-pass filter for habitable environments on Mars. Further evaluation of habitability, however, requires that we move beyond the mere presence of water to consider its properties. Water may be required for life on Earth, but not all terrestrial waters are habitable; the limits of terrestrial life are defined by additional parameters, including temperature, pH, and salinity (*1*).

Layered martian rocks exposed on the Meridiani plain have been interpreted as sandstones containing Mg sulfate, Ca sulfate, jarosite [(K, Na, H_3O)$Fe_3$$_{-x}$$Al_x$)$(SO_4)_2(OH)_6$, where $x < 1$], and hematite (*2*–*5*). The geochemical, textural, and sedimentological features of these rocks are consistent with deposition in an arid, predominantly eolian setting, with intermittent percolation of acidic, oxidizing, and saline groundwater (*4*, *5*). The orbiting spectrometers OMEGA and CRISM have identified other localities across the martian surface that boast thick successions of layered rock rich in Mg and Ca sulfates (*6*). Evaporite minerals in martian meteorites and Fe-sulfate–rich soils analyzed at Gusev Crater provide still further evidence for acid saline conditions (*7*, *8*). By itself, the low pH inferred for martian brines need not have imposed a barrier to life (*1*). What, however, were the consequences of salinity?

At the high solute concentrations needed to form evaporite minerals, water is, to a large degree, made chemically unavailable by ion hydration. As a result, with increasing salinity, biological activity sharply decreases. Quantitatively, the chemical availability of liquid water can be expressed in terms of its activity (a_{H_2O}) (*9*). The a_{H_2O} of pure water is 1.0, a value that decreases with increasing

salinity. The a_{H_2O} of terrestrial seawater is 0.98, and the a_{H_2O} of most waters weathering Earth's continental crust falls between 0.98 and 1.0. Most organisms cannot grow at $a_{H_2O} \sim 0.9$, and few are known to tolerate $a_{H_2O} < 0.85$ (*10*). A few extreme eukaryotes fungi and archaea can live in NaCl-saturated solutions with an $a_{H_2O} = 0.75$ (*10*), and a single fungus, *Xeromyces bisporus*, has been shown to grow at $a_{H_2O} = 0.61$ in a high-sugar food (*10*, *11*). This represents the effective habitability limit for terrestrial organisms, but the biological stresses of organic solutes differ from those imposed by inorganic salts, and so the limiting a_{H_2O} could be higher for organisms in inorganic saline systems (*10*). Recently improved thermodynamic models allow us to calculate a_{H_2O} for martian surface environments and so constrain the probability that life arose and persisted there (*12*).

As evaporation proceeds, precipitating saline minerals act as robust markers for ambient a_{H_2O}. The sequence of precipitating minerals is itself controlled by the chemistry of initially dilute waters at the outset of evaporation (*13*). Because fluids on Mars chemically weathered a basaltic crust, the chemistry of dilute martian waters would have been distinct from that of most waters on Earth (*14*). Constraints on the cation chemistry associated with this process come largely from experimental and theoretical studies on the chemical weathering of martian basalt (*15*) [supporting online material (SOM) text].

Our results show that the evaporation of dilute martian waters results in distinct saline mineral assemblages, causing some minerals that are common in terrestrial settings to precipitate at lower a_{H_2O} values than they do on Earth (*16*). For example, halite does not precipitate from martian waters until $a_{H_2O} \leq 0.50$ because of the low abundance of Na released from basaltic weathering. On this basis, the identification of halite at the martian surface (*17*) indicates extreme salinity and brines that would not be inhabitable by known terrestrial organisms.

Temperature and anion content also control the a_{H_2O} of martian brines. The effect of temperature is minor near a temperature (T) > 0°C; a_{H_2O} might increase by ~0.03 to 0.05 over the range of 25° to 0°C. When freezing occurs and equilibrium between ice and surrounding brine is maintained, a_{H_2O} will be controlled by temperature and the amount of ice formed. Thus, in water ice–bearing areas on Mars, the a_{H_2O} of brine phases can be predicted by temperature alone (*18*) (fig. S4).

The proportional anion abundances of dilute evaporating waters exert one of the largest controls on a_{H_2O}. Unlike the principal cations in martian brines, major anions originate ultimately from volcanic degassing. At Meridiani, in situ chemical analyses suggest that the initial S/Cl molar ratio of groundwater brines ranged from ~6.0 to 30.0 (*19*). Here we focus on three important anion species and how changing their relative proportion before evaporation affects a_{H_2O}. Our calculations show that the effect of the SO_4/Cl ratio is most important, so that a decreasing ratio leads to progressively lower a_{H_2O} upon evaporative concentration (*20*) (fig. S1). We calculate that for martian waters with an initial molar SO_4/Cl = 6.0, halite precipitates when $a_{H_2O} = 0.48$. At a SO_4/Cl ratio of 1.0, halite precipitates when the $a_{H_2O} = 0.39$.

For Meridiani Planum, a_{H_2O} can be constrained separately for evaporation that occurred before sediment transport and deposition (Fig. 1) and for brine evolution during diagenesis. Evaporating fluids were initially generated by basaltic weathering (*14*); all major products initially formed by evaporation are probably represented in the Meridiani outcrop because its bulk composition is largely basaltic (*4*, *14*).

The initially precipitated Mg sulfate phase is not precisely identified in Meridiani rocks. Orbital spectroscopy of the surrounding region has identified kieserite ($MgSO_4 \cdot H_2O$) and other polyhydrated sulfates, for which a number of Mg sulfate hydrates are candidates (*6*). Because of this uncertainty, we considered the a_{H_2O} associated with the precipitation of various Mg sulfate hydrates (*21*). Epsomite precipitation places a maximum on a_{H_2O} at 0.78 (Fig. 1), with hexahydrite and kieserite precipitating at 0.62 and 0.51, respectively (SOM text). Because it represents the most conservative case with respect to habitability, we focus on epsomite here.

Diagenetic events at Meridiani are recorded by detailed textures viewed with the Microscopic Imager. Secondary crystal moldic porosity (or "vugs") in some beds records the syndepositional growth and subsequent dissolution of a soluble mineral. This mineral was preferentially removed by infiltrating groundwater, but much of the primary sedimentary fabric was preserved (Fig. 2). From these textures and the observation of abundant Mg sulfate in the outcrop, we infer that groundwater was at or near saturation with respect to a soluble Mg sulfate mineral (*4*, *14*). The infiltration of dilute water would largely have removed this soluble component, obliterating sedimentary features, which did not occur (Fig. 2). Instead, a variety of recrystallization textures and multiple generations of cement were left behind (*4*).

It has been suggested that the moldic porosity in Meridiani sediments reflects oxidation of a syndepositional Fe^{2+}-sulfate mineral such as melanterite ($Fe^{2+}SO_4 \cdot 7H_2O$) (*4*, *14*). Our calculations show that at 25°C, saturation with respect to epsomite and melanterite in this system occurs at an $a_{H_2O} = 0.78$, similar to the evaporation sequence described above. Oxidation of the soluble Fe^{2+} component to Fe^{3+}-bearing oxides and sulfates (i.e., jarosite) has the overall effect of increasing the a_{H_2O} to 0.86 while still maintaining epsomite saturation. Alternatively, the incongruent melting of meridianite ($MgSO_4 \cdot 11H_2O$) to epsomite and water has been proposed as a mechanism for creating moldic porosity (*22*). The presence of this mineral before deposition and diagenesis

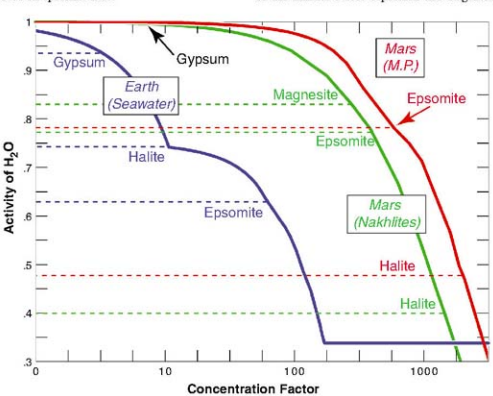


Fig. 1. Calculated mineral precipitation and a_{H_2O} values as a function of evaporative concentration. The blue line represents modern terrestrial seawater evaporation. The red line represents the evaporation of a basaltic-weathering derived fluid most representative of inferred evaporation processes at Meridiani Planum. The green line represents the evaporation of a similar fluid but with an anion concentration that gives rise to the saline mineral assemblage observed in the Nakhla meteorite (and other members of the nakhlite meteorite class).

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [2]Department of Geosciences, State University of New York, Stony Brook, NY 11794, USA.

*To whom correspondence should be addressed. E-mail: ntosca@fas.harvard.edu

requires the freezing of aqueous fluids. Water ice is likely to have formed in association with this process, appearing at temperatures near 0°C (*18*). If water ice melts, the brine composition approaches a chemical state essentially identical to the epsomite calculation described above, with an $a_{H_2O} = 0.78$ at epsomite and melanterite saturation.

Late-stage chloride precipitates have also been explored as vug precursor minerals on the basis of their relatively high solubility (*19*). Regardless of its identity, the dissolution of a chloride phase would require that infiltrating pore water became less saline, but still attained saturation with respect to Mg sulfate, with epsomite again fixing a_{H_2O} at 0.78.

Taken together, our results indicate that the evaporation of Meridiani groundwater led to a sustained $a_{H_2O} \leq 0.78$ to 0.86, values too low for all but a handful of terrestrial microorganisms to survive. Moreover, additional evidence suggests that chlorides are present in the sediments, pointing to still lower a_{H_2O} with continuing brine evolution. A range of Br/Cl ratios in different rocks suggests that highly soluble chlorides formed from evaporation and were incorporated into the wind-blown sediments (*4*, *19*). Progressive increases in Cl in rocks deeper in Endurance Crater suggest a similar process or one that differentially affected Cl during diagenesis (*19*). In addition, there is chemical and textural evidence for halite on rinds and rock coatings analyzed at Meridiani Planum, most likely reflecting late-stage fluid movement or remnants left when altering fluids finally evaporated to dryness (*23*). Accordingly, the maximum a_{H_2O} at chloride formation (0.48, Fig. 1) may reflect initial evaporite formation, syndepositional fluid chemistry, and/or very late-stage diagenesis.

On a planetary scale, Mars Odyssey Gamma Ray Spectrometer results show Cl-H_2O correlation across much of the martian surface (*24*), consistent with recent spectral evidence for widespread mid- to late-Noachian chloride-bearing lithologies in the southern highlands (*25*). Because late-stage martian brines would have evolved toward Cl-rich compositions as sulfates were precipitated (*20*), these observations probably represent the signature of low-a_{H_2O} brines. "Paso Robles"–class Fe-sulfate–rich soils at Gusev Crater represent another investigated locality characterized by extreme acidity and high salinity (*8*). Mössbauer and reflectance spectroscopic measurements, supported by α-proton x-ray spectrometer analyses, show that candidate Fe-sulfate minerals include ferricopiapite and rhomboclase (*8*, *26*, *27*). Mg sulfate is also indicated in some analyses, occurring as several percent $MgSO_4$ by weight. The a_{H_2O} for a Fe^{3+}-rich brine in equilibrium with ferricopiapite, rhomboclase, and epsomite is 0.61, with a pH (SOM text) of ~0.7: a remarkably harsh environment uninhabitable by known microorganisms (*10*, *18*).

Some martian meteorites also contain evaporite assemblages of martian origin, most likely reflecting subsurface geochemical processes. Members of the nakhlite class have carbonate minerals, including siderite and magnesite, as well as gypsum and epsomite (*7*). The Nakhla meteorite also contains late-stage halite; its saline assemblage represents the highest recorded extent of evaporative concentration on Mars. This evaporite mineral assemblage can be generated from an acidic solution with the same cation proportions as at Meridiani Planum (setting an initial pH of 4.8 and a SO_4/Cl ratio of 2.05) (*20*). After siderite and gypsum precipitate, magnesite appears at an a_{H_2O} of 0.83, then epsomite at $a_{H_2O} = 0.78$. Halite precipitates at an a_{H_2O} of 0.40 (Fig. 1). These calculations imply that the a_{H_2O} of at least some fluids percolating through the martian subsurface dipped well below known limits of terrestrial habitability.

Our calculations suggest that a number of martian localities widely distributed in time and space hosted fluids that were 10 to 100 times more saline than halite-saturated terrestrial seawater. A number of terrestrial organisms have evolved mechanisms to accommodate the osmotic stresses imposed by high salinities (*28*). However, all known salt-tolerant organisms are descended from ancestors that could not survive low a_{H_2O}, indicating that although life has evolved mechanisms to stretch the envelope of habitability, such brines were not a cradle for life on Earth (*1*, *10*, *28*). Accordingly, life that could originate and persist in the presence of extremely low a_{H_2O} on Mars would require biochemistry distinct from any known in even the most robust halophiles on Earth.

Precipitation is the sedimentary signature of brines with low a_{H_2O}; dissolution records more dilute waters. Although more dilute solutions must have carried ions to the sites where Meridiani salts precipitated, there is little evidence locally that subsequent water table regeneration introduced groundwaters dilute enough to dissolve Mg-sulfate minerals. Globally, we cannot rule out the possibility that more habitable waters existed elsewhere on the planet or in the earliest epoch of martian history. However, empirical evidence for dilute water at the martian surface remains meager. Geochemical constraints on Meridiani diagenesis suggest that the process was limited by the availability of water (*29*), which is consistent with a global surface geochemistry controlled for more than 3.5 billion years by water-limited processes (*30*). At the same time, hydrological models suggest that Meridiani Planum was one of the few regions on Mars likely to be characterized by groundwater upwelling about four billion years ago (*31*). This slow, global-scale deep hydrology would have focused water/rock interactions on the subsurface, leading to chemically distinct, concentrated groundwaters (*31*). When migrating groundwater breached the surface, continued upwelling and evaporation would have allowed standing water bodies to evolve toward greater ionic concentrations while keeping total fluid volume more or less constant. Concentrated standing water left its mark at Meridiani Planum in the form of diagnostic subaqueous cross-beds (*2*, *4*, *5*), and these same rocks display some of the highest MgO and SO_3 abundances measured, implying Mg-sulfate saturation and low a_{H_2O}. Moreover, the search for larger, more dilute bodies of water earlier in the Noachian has proven difficult (*25*, *32*). New observations have abbreviated the era when dilute surface waters could have been more common, pushing it further into Mars' planetary infancy.

Observations at Meridiani and other salt-bearing localities, then, may not reflect snapshots taken as local Mars waters evaporated to dryness, but rather a long surface history dominated by rock instead of water. On a planet where saline minerals were continuously recycled at the surface instead of removed by subduction and plate tectonics, dilute fluids may have been transient exceptions. The sobering conclusion is that Meridiani Planum, an extreme environment in terms of terrestrial habitability, may have been among the last, best places for life on the early martian surface.

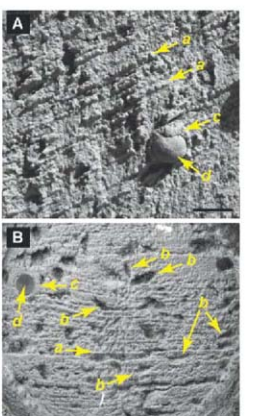


Fig. 2. Microscopic images taken by the Mars Exploration Rover Opportunity in Endurance Crater. **(A)** Portion of image taken of the undisturbed surface of Cobble Hill on martian day (sol) 144. **(B)** Portion of a mosaic taken on sol 149 of the feature London after being ground with the rock abrasion tool (RAT). Distinct grain boundaries are observable throughout both images. Primary laminations (a) and bedding structures are clearly visible. Secondary moldic porosity (b) can be seen throughout the RAT-abraded target, providing evidence for the selective removal of a syndepositionally formed soluble mineral while preserving primary sedimentary fabric. These features, observable in most places throughout Meridiani outcrop rocks, suggest groundwater saturation with respect to a Mg-sulfate phase during diagenesis. Secondary recrystallization is evident in some places, as shown by a generation of cement (c) surrounding hematitic concretions (d). Scale bars, 5 mm.

References and Notes

1. A. H. Knoll *et al.*, *Earth Planet. Sci. Lett.* **240**, 179 (2005).
2. S. W. Squyres *et al.*, *Science* **306**, 1709 (2004).
3. G. Klingelhöfer *et al.*, *Science* **306**, 1740 (2004).
4. S. M. McLennan *et al.*, *Earth Planet. Sci. Lett.* **240**, 95 (2005).
5. J. P. Grotzinger *et al.*, *Earth Planet. Sci. Lett.* **240**, 11 (2005).
6. A. Gendrin *et al.*, *Science* **307**, 1587 (2005).
7. J. C. Bridges, M. M. Grady, *Earth Planet. Sci. Lett.* **176**, 267 (2000).
8. D. W. Ming *et al.*, *J. Geophys. Res. Planets* **111**, E02S12 (2006).
9. At equilibrium, the vapor pressure of water above an aqueous solution is equal to p_{gw}. For pure water at a standard state, the vapor pressure is equal to p_w^0. The activity of the solvent water, a_{H_2O}, is given by $a_{H_2O} = \frac{p_{gw}}{p_w^0}$.
10. W. D. Grant, *Philos. Trans. R. Soc. London Ser. B* **359**, 1249 (2004).
11. The food industry has expended considerable effort to characterize biological limits to a_{H_2O} in order to retard spoilage by microbial growth. The paucity of organisms found at low a_{H_2O} represents biological limitation rather than sampling or research bias.
12. Methods are described in the SOM text.
13. L. A. Hardie, H. P. Eugster, *Mineral. Soc. Am. Spec. Pub.* **3**, 273 (1970).
14. N. J. Tosca *et al.*, *Earth Planet. Sci. Lett.* **240**, 122 (2005).
15. N. J. Tosca *et al.*, *J. Geophys. Res. Planets* **109**, E05003 (2004).
16. Our calculations simulate the evaporation of a dilute basaltic-weathering derived fluid. Equilibrium with a 1-bar CO_2 atmosphere is maintained; the SO_4/Cl is specified; and pH, initially 3.9 unless otherwise stated, is calculated by charge balance.
17. Sedimentary redistribution may complicate estimations of a_{H_2O} for specific localities.
18. G. M. Marion, J. S. Kargel, *Cold Aqueous Geochemistry with FREZCHEM* (Springer-Verlag, Berlin, 2008).
19. B. C. Clark *et al.*, *Earth Planet. Sci. Lett.* **240**, 73 (2005).
20. N. J. Tosca, S. M. McLennan, *Earth Planet. Sci. Lett.* **241**, 21 (2006).
21. Relative humidity (RH) or brine chemistry may affect the hydration state of some evaporitic minerals. Here we exclude postdepositional RH effects and consider only fluid evaporation.
22. R. C. Peterson, R. Y. Wang, *Geology* **34**, 957 (2006).
23. A. H. Knoll *et al.*, *J. Geophys. Res. Planets* **113**, E06S16 (2008).
24. J. M. Keller *et al.*, *J. Geophys. Res. Planets* **111**, E03S08 (2006).
25. M. M. Osterloo *et al.*, *Science* **319**, 1651 (2008).
26. J. R. Johnson *et al.*, *Geophys. Res. Lett.* **34**, L13202 (2007).
27. Mössbauer spectra show that candidate minerals must be Fe-sulfates distinct from jarosite. At such highly concentrated and acidic conditions, the a_{H_2O} is relatively insensitive to the exact Fe-sulfate phase (see SOM).
28. A. Oren, *J. Ind. Microbiol. Biotechnol.* **28**, 56 (2002).
29. M. E. Elwood-Madden, R. J. Bodnar, J. D. Rimstidt, *Nature* **431**, 821 (2004).
30. J. A. Hurowitz, S. M. McLennan, *Earth Planet. Sci. Lett.* **260**, 432 (2007).
31. J. C. Andrews-Hanna, R. J. Phillips, M. T. Zuber, *Nature* **446**, 163 (2007).
32. M. H. Carr, *The Surface of Mars* (Cambridge Univ. Press, New York, 2007).
33. We thank the MER science and engineering teams for measurements of martian materials that made this paper possible. Research supported in part by NASA and a Harvard Origins postdoctoral fellowship to NJT.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1204/DC1
SOM Text
Figs. S1 to S4
Tables S1 to S6
References

18 January 2008; accepted 18 April 2008
10.1126/science.1155432

A Cytosolic Iron Chaperone That Delivers Iron to Ferritin

Haifeng Shi,[1] Krisztina Z. Bencze,[2] Timothy L. Stemmler,[2] Caroline C. Philpott[1]*

Ferritins are the main iron storage proteins found in animals, plants, and bacteria. The capacity to store iron in ferritin is essential for life in mammals, but the mechanism by which cytosolic iron is delivered to ferritin is unknown. Human ferritins expressed in yeast contain little iron. Human poly (rC)–binding protein 1 (PCBP1) increased the amount of iron loaded into ferritin when expressed in yeast. PCBP1 bound to ferritin in vivo and bound iron and facilitated iron loading into ferritin in vitro. Depletion of PCBP1 in human cells inhibited ferritin iron loading and increased cytosolic iron pools. Thus, PCBP1 can function as a cytosolic iron chaperone in the delivery of iron to ferritin.

Ferritins are iron storage proteins that are ubiquitously expressed in animals, plants, and bacteria. They serve both to sequester excess iron taken up by the cell and to release stored iron to meet the cell's metabolic needs during iron scarcity (*1*). In animals, ferritin is a cytosolic heteropolymer consisting of 24 subunits of heavy (H) and light (L) isoforms that assemble into a hollow sphere into which iron is deposited. Ferritin H chains contain the iron-binding and ferroxidase activities that are required for mineralization of the ferritin core. Deletion of the H-ferritin gene is lethal in mice (*2*) and in flies (*3*).

In cells, metallochaperones deliver metals to their cognate enzymes and transporters. Although cytosolic copper and nickel chaperones have been described (*4*–*7*), no cytosolic iron chaperones have been identified, despite the presence of numerous iron-dependent enzymes in the cytosol. Frataxin—the protein lacking in the neurological disease Friedreich's ataxia—functions as a mitochondrial iron chaperone for iron-sulfur cluster and heme biosynthesis (*8*, *9*).

Fungi are anomalous among eukaryotes in that they do not express ferritins. We expressed human H and L ferritins in the yeast *Saccharomyces cerevisiae*. The peptides assembled into multimeric complexes, with properties similar to those of native human ferritins, but contained only small amounts of iron (fig. S1, A and B). We hypothesized that yeast might also lack the requisite iron chaperones needed for delivery of iron to ferritin and designed a genetic screen to identify human genes that, when expressed in yeast, could increase the amount of iron loaded into ferritin. We introduced an iron-regulated *FeRE/HIS3* reporter construct (*10*) into a yeast strain expressing H and L ferritin (Fig. 1A). This construct confers histidine prototrophy to cells when the reporter is bound and transcriptionally activated by Aft1p, the major iron-dependent transcription factor in yeast. Aft1p is activated during periods of cytosolic iron depletion (*11*), which could occur if substantial amounts of cytosolic iron were diverted into ferritin.

Yeast cells containing ferritin and the iron-responsive reporter were transformed with a library synthesized from human liver cDNA engineered into a yeast expression vector. Transformants that exhibited growth on plates lacking histidine were selected for further analysis. We isolated multiple copies of PCBP1, as well as proteins encoded by other unrelated genes, including H ferritin. Plasmids containing PCBP1 or the empty vector were retransformed into reporter yeast strains lacking or expressing H and L ferritins (Fig. 1B). Expression of PCBP1 did not activate the *FeRE/HIS3* reporter in cells lacking ferritin, as indicated by a lack of growth on media without histidine. But expression of PCBP1 did activate the *FeRE/HIS3* reporter in the yeast strain expressing ferritins, resulting in growth on media lacking histidine. Thus, expression of human PCBP1 activated the iron-responsive reporter only in the presence of ferritin. To confirm that reporter activation was due to delivery of cytosolic iron into ferritin, we directly measured the incorporation of iron into ferritin by growing yeast in the presence of $[^{55}Fe]Cl_3$, isolating ferritin on nondenaturing gels, and measuring the amount of ^{55}Fe in the ferritin heteropolymer (Fig. 1, C and D). Substantial amounts of iron-containing protein were detected only in cells expressing ferritin, and iron was detected in a single species that comigrated with the ferritin heteropolymer. Coexpression of PCBP1 in these cells resulted in a 2.3-fold increase in the amount of iron in ferritin. This increase was not due to changes in the overall amount of ferritin (Fig. 1E) or in the relative ratio of H and L chains (fig. S1A). Similarly, the total amount of ^{55}Fe taken up by the cells expressing ferritin alone was not different from the amount taken up by cells expressing both PCBP1 and ferritin (fig. S1C).

The delivery of cytosolic iron to ferritin in the presence of PCBP1 activated the *FeRE/HIS3* reporter. We asked whether other proteins ex-

[1]Liver Diseases Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA. [2]Department of Biochemistry and Molecular Biology, Wayne State University School of Medicine, Detroit, MI 48201, USA.

*To whom correspondence should be addressed. E-mail: carolinep@intra.niddk.nih.gov

pressed during yeast iron deficiency, such as the plasma membrane ferric reductases, were also activated by expression of PCBP1 (Fig. 1F). Ferric reductase activity was low in cells that did not express ferritins, regardless of whether the cells contained pPCBP1 or the empty vector. The ferritin-expressing strain exhibited slightly greater reductase activity than the nonferritin strain when transformed with vector alone and four times the reductase activity when the ferritin strain also expressed PCBP1. Thus, when human PCBP1 was expressed in yeast cells containing human ferritins, iron was diverted into ferritin, and the cellular iron deficiency response was activated.

PCBP1 is an RNA binding protein that is ubiquitously expressed in mammalian cells and is located in both the cytosol and the nucleus (*12*). We tested whether PCBP1 was involved in ferritin iron loading in human cells by depleting cellular PCBP1, loading cells with ^{55}Fe, and measuring the amount of ^{55}Fe loaded into endogenous cytosolic ferritin. Huh7 cells were transfected with PCBP1 or control small interfering RNAs (siRNAs), and partial depletion of PCBP1 mRNA and protein was confirmed (fig. S2). Transfected cells were loaded with [^{55}Fe]Cl_3, and ferritin was examined (Fig. 2, A to C). PCBP1-depleted cells exhibited a 63% reduction in the amount of ^{55}Fe incorporated into ferritin when compared with the amount in control cells at 6, 12, and 24 hours. This reduction in ferritin iron loading was not due to lowered levels of ferritin protein, because these levels did not change significantly when PCBP1 was depleted (Fig. 2B). The reduction in ferritin iron loading was also not due to loss of ^{55}Fe uptake in the PCBP1-depleted cells, because uptake of both [^{55}Fe]Cl_3 and [$^{55}Fe_2$]transferrin was equivalent in cells transfected with control or PCBP1 siRNAs (fig. S3). The loss of ferritin iron loading in cells treated with PCBP1 siRNA was not due to off-target effects of the siRNA. Ferritin iron loading was restored in cells cotransfected with a plasmid expressing human PCBP1 containing silent mutations (fig. S4). The ferritin mineralization that remained after PCBP1 depletion might have been due to the activity of residual PCBP1. Alternatively, paralogs of PCBP1—such as PCBP2, which also activated the *FeRE/HIS3* reporter in yeast (fig. S5)—may contribute to ferritin iron loading.

To determine whether the loss of ferritin iron loading during PCBP1 depletion also resulted in an increase in cytosolic iron, we measured the levels of iron regulatory protein 2 (IRP2). The half-life of IRP2 is inversely related to cytosolic iron levels (*13*), with IRP2 levels increasing when iron is scarce and decreasing when iron is abundant (fig. S6). We transfected human embryonic kidney–293 (HEK293) cells stably overexpressing IRP2 (*14*) with control and PCBP1 siRNAs and measured the levels of IRP2 (Fig. 2D). Loss of PCBP1 was associated with a decrease in IRP2, which is consistent with the loss of PCBP1 leading to an increase in the cytosolic iron pool. The relative levels of the chelatable cytosolic iron pool can be measured with the use of fluorescent iron chelators (*15*), and depletion of PCBP1 in Huh7 cells led to a 67% increase in the chelatable iron pool (Fig. 2E).

PCBP1 may facilitate ferritin iron loading by directly interacting with ferritin or by an indirect mechanism that requires other cellular factors. We tested for a direct, in vivo, interaction between ferritin and PCBP1 by coimmunoprecipitation in yeast cells (Fig. 3). PCBP1 coimmunoprecipitated with ferritin in cells expressing PCBP1 and H and L ferritin (Fig. 3, A and B). No PCBP1 was detected in immune complexes from cells lacking either PCBP1 or ferritin. When cells expressed both ferritins and PCBP1, PCBP1 was not detected in immunoprecipitates collected in buffer without iron (Fig. 3B), but when ferrous iron was added to the buffer, PCBP1 was detected in antiferritin immunoprecipitates (Fig. 3, A and B). The addition of bathophenanthroline disulfonate

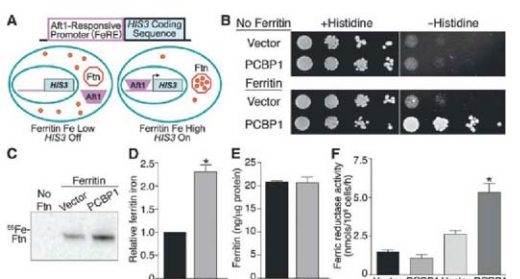


Fig. 1. PCBP1 delivers iron to ferritin in *S. cerevisiae*. **(A)** Activation of an iron-regulated, Aft1-responsive promoter (*FeRE*) fused to the *HIS3* coding sequence when cytosolic iron is transferred into ferritin (Ftn). **(B)** PCBP1-dependent activation of the *FeRE/HIS3* reporter in yeast expressing ferritin, but not in yeast without ferritin. Transformed strains were plated in serial dilutions on media with and without histidine. **(C** and **D)** PCBP1 increases ferritin mineralization. Strains transformed as in (B) were grown with [^{55}Fe]Cl_3, ferritin was isolated by native gel electrophoresis, and iron within ferritin was detected by autoradiography. In (D), replicates were normalized to the vector-transformed strain (n = 5 independent replicates). *P < 0.002. **(E)** PCBP1 did not affect ferritin protein levels (n = 10 samples). **(F)** PCBP1-dependent increase in surface ferric reductase activity in yeast expressing ferritin, but not in yeast without ferritin (n = 5 samples). *P < 0.002. This and subsequent P values were determined via a two-tailed t test. Error bars in (D) to (F) indicate SEM.

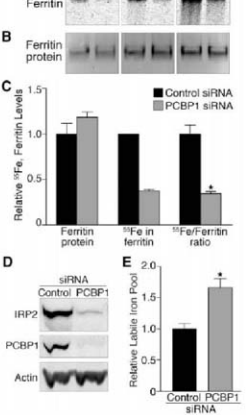


Fig. 2. Depletion of PCBP1 impairs ferritin iron loading in Huh7 cells. Huh7 cells transfected with PCBP1 or control siRNAs were labeled with [^{55}Fe]Cl_3 for the indicated times. Ferritin was detected by autoradiography **(A)** or immunoblotting **(B)**. **(C)** Relative quantitation of ferritin protein, iron, and iron/protein ratio (n = 5 independent replicates). *P < 0.0002. **(D)** Degradation of IRP2 in cells lacking PCBP1. HEK293 cells expressing IRP2 were transfected with PCBP1 and control siRNAs. IRP2, PCBP1, and actin were detected by immunoblotting. **(E)** Increased labile iron pool in cells lacking PCBP1. PCBP1 was depleted in Huh7 cells, and the relative amounts of the chelatable intracellular iron pool were measured (n = 6 samples). *P = 0.003. Error bars in (C) and (E) indicate SEM.

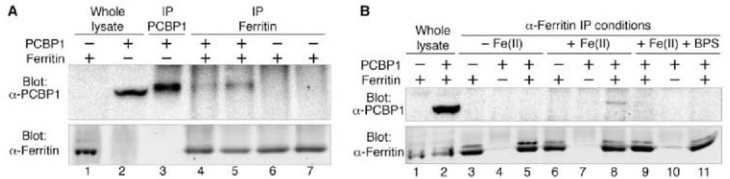


Fig. 3. PCBP1 binds to ferritin in vivo. (**A** and **B**) Yeast cells expressing ferritin H and L chains, PCBP1, or both were lysed and subjected to immunoprecipitation (IP) with antibodies to PCBP1 or ferritin in buffers containing ferrous iron [both (A) and (B)], no iron (B), or ferrous iron and a ferrous iron chelator [BPS, (B)]. Immune complexes were detected with antibodies against PCBP1 and ferritin. Note that PCBP1 migrates more rapidly in whole lysates than after IP (A).

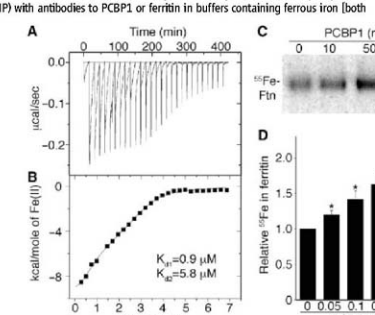


Fig. 4. PCBP1 binds ferrous iron and increases ferritin iron loading in vitro. (**A** and **B**) PCBP1 binds ferrous iron with micromolar affinity by ITC. (A), raw ITC spectrum; (B), heat evolved per addition of titrant. Metal binding was enthalpically favorable (enthalpy $\Delta H \cong -9.9$ kcal/mol) with an overall favorable free energy ΔG. Binding was entropically unfavorable [entropy $\Delta S \cong -10.7$ cal/(K·mol)]. (**C** and **D**) Increased ferritin mineralization in the presence of PCBP1. Equine apoferritin was incubated with ^{55}Fe(II) and PCBP1 or bovine serum albumin (BSA). Ferritin iron was detected by autoradiography. (D) Relative ^{55}Fe incorporated into ferritin. Samples were normalized to the sample without added PCBP1 ($n = 6$ samples). *$P < 0.007$. Error bars indicate SEM.

(BPS), a ferrous iron chelator, to buffer containing iron blocked the coimmunoprecipitation of PCBP1 with ferritin (Fig. 3B). Thus, PCBP1 physically interacted with ferritin in the presence of iron and might directly bind iron.

We used isothermal titration calorimetry (ITC) to directly measure interactions between PCBP1 and iron. PCBP1 overexpressed and purified from *Escherichia coli* was folded with high helical content (fig. S7). Titration of ferrous iron into solutions of PCBP1 under anaerobic conditions produced negative peaks in the raw thermogram, which indicated ferrous iron bound to PCBP1 in an exothermic process (Fig. 4A). An integration of each individual titration peak gave rise to the processed spectrum (Fig. 4B). PCBP1 bound a total of three iron atoms with a dissociation constant (K_d) of 0.9 ± 0.1 μM for the first and an average K_d of 5.8 ± 0.3 μM for the remaining two metal ions.

Although ferritin mineralization occurs in vitro in the presence of only ferrous iron and oxygen, we tested whether PCBP1 could enhance mineralization of ferritin at low iron concentrations (Fig. 4, C and D). Addition of purified PCBP1 to solutions of apoferritin and ^{55}Fe(II) increased the amount of ^{55}Fe detected in ferritin in a dose-dependent manner, with mineralization increasing twofold at the higher concentrations of PCBP1. Albumin, which binds iron with low affinity, did not significantly alter the amount of iron incorporated into ferritin, indicating that PCBP1 specifically and directly delivered iron for ferritin mineralization in vitro.

Human H-ferritin homopolymers bind ferrous iron with an affinity of 15 μM (*16*), a concentration far above the levels of "free" ferrous iron thought to be present in cytosol, which raises the likelihood that a specific iron carrier is required for delivery of iron to ferritin. PCBP1 bound three atoms of ferrous iron with dissociation constants of 0.9 to 5.8 μM, which is similar to the binding affinity of yeast frataxin: a mitochondrial iron chaperone that binds two ferrous iron atoms with dissociation constants of 2 to 3 μM (*17*). Similarly, the cytosolic copper chaperone Atox1 binds copper with a K_d of ~10 μM (*18*). PCBP1 bound to ferritin only in the presence of Fe(II). Similarly, the interaction between yeast Atx1p and the copper transporter Ccc2p only occurs in the presence of copper (*19*).

PCBP1 is a member of a family of four homologous RNA binding proteins belonging to the heterogeneous nuclear ribonucleoprotein K-homology domain superfamily and is widely expressed and highly conserved among mammals. PCBP1 and PCBP2 bind specifically to sequences within multiple cellular or viral mRNA species with a consequent increase in the stability of the messages or alteration of their translation efficiency (*12*). The bifunctional nature of PCBP1 as both an iron-binding and an RNA binding protein is reminiscent of the mammalian IRP1. IRP1 functions as cytosolic aconitase when it contains an iron-sulfur cluster and binds to mRNA transcripts involved in iron homeostasis (such as ferritin) when it does not (*13*). We propose that PCBP1 acts as a cytosolic iron chaperone, directly binding iron and loading ferritin.

References and Notes

1. P. M. Harrison, P. Arosio, *Biochim. Biophys. Acta* **1275**, 161 (1996).
2. E. C. Theil, M. Matzapetakis, X. Liu, *J. Biol. Inorg. Chem.* **11**, 803 (2006).
3. F. Missirlis *et al.*, *Genetics* **177**, 89 (2007).
4. J. Kuchar, R. P. Hausinger, *Chem. Rev.* **104**, 509 (2004).
5. T. V. O'Halloran, V. C. Culotta, *J. Biol. Chem.* **275**, 25057 (2000).
6. R. A. Pufahl *et al.*, *Science* **278**, 853 (1997).
7. T. D. Rae, P. J. Schmidt, R. A. Pufahl, V. C. Culotta, T. V. O'Halloran, *Science* **284**, 805 (1999).
8. A. L. Bulteau *et al.*, *Science* **305**, 242 (2004).
9. M. Pandolfo, *J. Neurol Transm. Suppl.* **70**, 143 (2006).
10. A. Dancis *et al.*, *Cell* **76**, 393 (1994).
11. C. C. Philpott, O. Protchenko, *Eukaryot. Cell* **7**, 20 (2008).
12. A. V. Makeyev, S. A. Liebhaber, *RNA* **8**, 265 (2002).
13. T. A. Rouault, *Nat. Chem. Biol.* **2**, 406 (2006).
14. F. Bourdon *et al.*, *Blood Cells Mol. Dis.* **31**, 247 (2003).
15. S. Epstejn, O. Kakhlon, H. Glickstein, W. Breuer, I. Cabantchik, *Anal. Biochem.* **248**, 31 (1997).
16. F. Bou-Abdallah *et al.*, *Biochemistry* **41**, 11124 (2002).
17. J. D. Cook *et al.*, *Biochemistry* **45**, 7767 (2006).
18. A. K. Wernimont, L. A. Yatsunyk, A. C. Rosenzweig, *J. Biol. Chem.* **279**, 12269 (2004).
19. L. Banci *et al.*, *Nat. Chem. Biol.* **2**, 367 (2006).

20. The authors thank S. Liebhaber for the PCBP1 plasmid, protein samples, and helpful discussions; H. Levin for helpful discussions and ferritin plasmids; P. Ponka for salicylaldehyde isonicotinoyl hydrazone; and T. Rouault for helpful discussions, cell lines, and the IRP2 antibody. These studies were supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases (H.S. and C.C.P.) and by NIH R01 DK068139 (K.Z.B. and T.L.S.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/320/5880/1207/DC1
Materials and Methods
Figs. S1 to S7
References

11 March 2008; accepted 18 April 2008
10.1126/science.1157643

Massive Horizontal Gene Transfer in Bdelloid Rotifers

Eugene A. Gladyshev,[1] Matthew Meselson,[1,2]* Irina R. Arkhipova[1,2]*

Horizontal gene transfer in metazoans has been documented in only a few species and is usually associated with endosymbiosis or parasitism. By contrast, in bdelloid rotifers we found many genes that appear to have originated in bacteria, fungi, and plants, concentrated in telomeric regions along with diverse mobile genetic elements. Bdelloid proximal gene-rich regions, however, appeared to lack foreign genes, thereby resembling those of model metazoan organisms. Some of the foreign genes were defective, whereas others were intact and transcribed; some of the latter contained functional spliceosomal introns. One such gene, apparently of bacterial origin, was overexpressed in *Escherichia coli* and yielded an active enzyme. The capture and functional assimilation of exogenous genes may represent an important force in bdelloid evolution.

Horizontal gene transfer (HGT), the movement of genes from one organism to another by means other than direct descent (vertical inheritance), has been documented in prokaryotes (*1*) and in phagocytic and parasitic unicellular eukaryotes (*2*–*4*). Despite the large number of sequenced genomes, documented HGT is rare in metazoans, at least in part because of the sequestration of the germ line (*5*, *6*). HGT may be facilitated by long-term association with organelles or with intracellular endosymbionts and parasites (*7*, *8*), or it may involve transposable elements (TEs) (*9*, *10*).

Bdelloid rotifers are small freshwater invertebrates that apparently lack sexual reproduction and can withstand desiccation at any life stage (*11*, *12*). Their genomes contain diverse TEs, including DNA transposons and retrovirus-like *env*-containing retrotransposons, such as *Juno* and *Vesta*, possibly acquired from exogenous sources and concentrated near telomeres (*13*, *14*). We investigated TE distribution in bdelloids by sequencing clones from an *Adineta vaga* fosmid library hybridizing to *Juno1* probes. Unexpectedly, in two *Juno1* long terminal repeat (LTR)–containing clones (contigs Av240A and Av212A), we found 10 protein-coding sequences (CDS) yielding strong database hits (BLAST E-values of $8E^{-102}$ to 0.0) to bacterial and fungal genes (Fig. 1A, Table 1, fig. S1A, and table S1). Half of these CDS have no metazoan orthologs, and three apparently bacterial CDS are interrupted by canonical spliceosomal introns, which are nonexistent in bacteria.

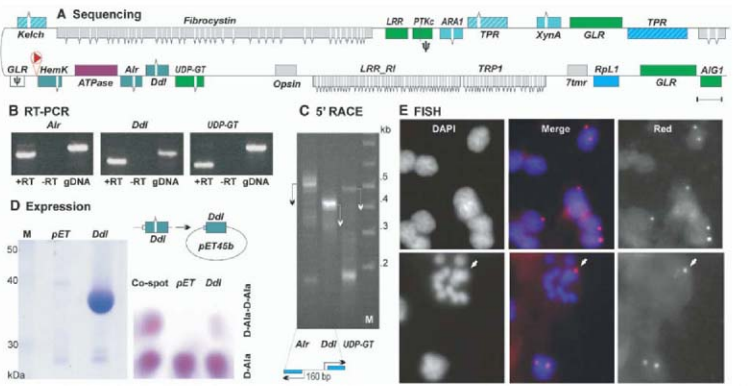


Fig. 1. Structural and functional analysis of the Av240 genomic region. (**A**) The 85-kb Av240B contig. The colinear 50-kb Av240A contig contains the *Juno1* LTR (red triangle) and extends from TPR to LRR_RI (not shown). CDS (boxes) are colored according to their putative origin: metazoan, gray; bacterial, blue; fungal, purple; plant, green; indeterminate, striped; hypothetical, white. Intron positions are indicated. Pseudogenes are denoted by ψ, and defects in their reading frames appear as vertical lines. Scale bar, 1 kb. (**B**) RT-PCR analysis of *Alr*, *Ddl*, and UDP-glycosyltransferase genes. Unspliced and spliced RNA are visible in +RT lanes; –RT, no reverse transcriptase; gDNA, genomic DNA control. (**C**) 5′ RACE analysis of transcription start sites (arrows) for genes in Fig. 1B. (**D**) Expression in *E. coli* of the *A. vaga Ddl* cloned in pET45b to yield a protein fused to the 6xHis tag. Left, SDS–polyacrylamide gel electrophoresis (PAGE) analysis of the Ddl protein after purification by metal-affinity chromatography; right, thin-layer chromatography after incubation of purified AvDdl with D-Ala; co-spot, control for D-Ala and D-Ala-D-Ala migration; pET, vector with no insert. (**E**) FISH of the Av240A probe, labeled with the red fluorophore, to *A. vaga* embryo nuclei. The arrow points to a signal in a nucleus with condensed chromosomes.

Fluorescent in situ hybridization (FISH) with a probe from Av240A confirmed that these CDS reside in the *A. vaga* genome (Fig. 1E). Their hybridization pattern resembles that of known telomeric fosmids (*15*), suggesting proximity to telomeres. The appearance of two hybridizing sites in some nuclei is consistent with the genome structure of bdelloids in which chromosomes occur as colinear pairs (*16*, *17*). Indeed, we found colinear partners of both Av240A and Av212A (Av240B and Av212B) with overall pairwise divergence (4%) similar to that in other regions of bdelloid genomes.

The cluster of foreign genes near the *Juno1* LTR in Av240 includes two divergently oriented genes for enzymes involved in bacterial cell wall peptidoglycan biosynthesis—*Alr*, encoding alanine racemase, and *Ddl*, encoding D-Ala-D-Ala ligase—adjacent to a uridine diphosphate (UDP)–glycosyltransferase apparently of plant origin (Fig. 1A). Reverse transcription polymerase chain reaction (RT-PCR) shows that all three genes are transcribed and that their introns are properly spliced (Fig. 1B). 5′ RACE (rapid amplification of cDNA ends) demonstrates that the UDP-glycosyltransferase mRNA is trans-spliced, as are numerous bdelloid genes (*18*), and that *Alr* and *Ddl* transcription initiates at oppositely oriented promoters located between them (Fig. 1C). Furthermore, the purified *A. vaga* AvDdl protein overexpressed in *E. coli* catalyzes the synthesis of the D-Ala-D-Ala dipeptide from D-Ala in vitro (Fig. 1D).

In addition to ubiquitous bacterial genes, such as *Alr* and *Ddl*, we identified genes occurring in only a few species of bacteria and fungi. The *XynB*-like gene (figs. S1A and S2B) apparently represents a fusion between two different conserved domains and is found in only 10 bacterial species. Next to the *Alr-Ddl* pseudo-operon, there is a *HemK*-like methyltransferase and a putative adenosine triphosphatase. These two genes are also rare: They are found in only four genera of proteobacteria and three genera of filamentous fungi. In the bacterium *Methylobacillus flagellatus* and in the fungus *Phaeosphaeria nodorum* they are adjacent and in the same relative orientation as in *A. vaga*, indicating that they might have been acquired from a single source.

We also found genes with similarity to those present in both metazoans and nonmetazoans. We characterized the foreignness of such genes with an alien index (AI), which measures by how many orders of magnitude the BLAST E-value for the best metazoan hit differs from that for the best nonmetazoan hit (Table 1 and table S1) (*19*). We tested the AI validity by phylogenetic analyses of all CDS with AI > 0 yielding metazoan hits, excluding those with repetitive sequences. We found that AI ≥ 45, corresponding to a difference of ≥20 orders of magnitude between the best nonmetazoan and metazoan hits, was a good indicator of foreign origin, as judged by phylogenetic assignment to bacterial, plant, or fungal clades (Fig. 2 and fig. S2A). Genes with 0 < AI < 45 were designated indeterminate, because their phylogenetic analysis may or may not confidently support a foreign origin. Four *FabG*-like genes for short-chain dehydrogenases/reductases (SDH), from two different SDH subfamilies, are most likely of bacterial origin (AI = 98/92/88/45; Fig. 2A). The *A. vaga* galacturonidase (AI = 212) appears to be of fungal origin (Fig. 2B), and the UDP-glycosyltransferase in Av240 (AI = 28, indeterminate) belongs to a plant clade (Fig. 2C). Two genes, *XynB* and *MviM*, had sufficient nucleotide sequence similarity (~70%) to bacterial homologs for phylogenetic reconstruction and determination of nonsynonymous and synonymous divergence (fig. S2B).

We extended our search for foreign genes to two pairs of contigs ending with telomeric repeats (telomeres K and L) (*15*) (fig. S1, B and C). In addition to various TEs, telomeric repeats, and *Athena* retroelements characteristic of bdelloid telomeric regions, we found additional examples of foreign genes, including a pseudogene of fungal origin (putative urea transporter; Table 1) with three frameshifts and two in-frame stop codons in one of the two colinear homologs of telomere L. Additionally, we identified foreign genes sandwiched between short stretches of telomeric repeats, suggesting their addition to deprotected chromosome ends (fig. S1C).

We also observed examples of possible loss of genes and TEs from telomeres (fig. S3A), such

[1]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA. [2]Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA 02543, USA.

*To whom correspondence should be addressed. E-mail: msm@wjh.harvard.edu (M.M.); larkhipova@mbl.edu (I.R.A.)

Table 1. Representative bdelloid CDS of foreign origin homologous to genes with known function. For a complete list and additional information on each CDS, see table S1. Data are from BLASTP similarity searches, as described in (*19*). Asterisks indicate putative pseudogenes.

Gene ID, name	Contig ID	Introns	AI	% Identity to best hit	Best hit, E-value	Best hit, metazoan	Best hit, taxonomy	Definition
AV10027_*XynB*	Av212A	0	460	63	0.00E+00	No hits	Bacteria; Bacteroidetes	Xylosidase/arabinosidase
AV10001_*NRPS*	Av110A	10	460	32	0.00E+00	No hits	Bacteria; (Proteobacteria/ Cyanobacteria)	Nonribosomal peptide synthetase
AV10134_*PheA*	161F07	0	400	61	1.00E-174	No hits	(Fungi; Bacteria)	Monooxygenase, FAD dependent
AV10002_*TrkA*	Av110A	0	379	54	1.00E-175	4.00E-11	Bacteria; Proteobacteria	Monooxygenase, NAD dependent
PR10002_*MviM*	182F10	0	327	67	1.00E-149	2.00E-07	Bacteria; (Acidobacteria/ Chloroflexi)	Oxidoreductase
PR10010_*DAP2*	182F10	0	310	27	1.00E-140	1.00E-05	Bacteria; (Acidobacteria/ Proteobacteria)	Prolyl oligopeptidase*
AV10104_*Dur3*	AvTelL.B	1	243	44	1.00E-132	4.00E-27	Eukaryota; Fungi	Urea active transporter*
PR10012_*RamA*	182J17	0	246	31	1.00E-107	No hits	(Bacteria; Fungi)	α-L-Rhamnosidase
AV10121_*NRPS*	99O7	4	237	30	1.00E-103	No hits	Bacteria; Cyanobacteria	Nonribosomal peptide synthetase
AV10153_*XghA*	210B3	0	212	50	1.00E-108	2.00E-16	Eukaryota; Fungi	Endo-xylogalacturonan hydrolase
AV10042_*HemK*	Av240B	1	199	56	2.00E-91	1.00E-04	Bacteria; Proteobacteria	HemK-like methyltransferase
AV10092_*β-Gal*	AvTelL.A	0	153	33	1.00E-105	4.00E-39	Eukaryota; Viridiplantae	β-D-Galactosidase
AV10044_*Alr*	Av240B	1	152	38	1.00E-67	No hits	Bacteria; Bacteroidetes	Alanine racemase
AV10025_*AMH*	Av212A	1	150	52	8.00E-77	2.00E-11	Eukaryota; Fungi	Amidohydrolase
AV10045_*Ddl*	Av240B	1	138	40	1.00E-60	No hits	Bacteria; Bacteroidetes	D-Alanine-D-alanine ligase
AV10140_*PLDc*	193E18	2	126	31	1.00E-55	No hits	Eukaryota; Fungi	Phospholipase-D active site motif protein*
AV10016_*FabG*	Av212A	0	98	58	1.00E-74	8.00E-32	Bacteria	Short-chain dehydrogenase/reductase
AV10109_*FabG*	AvTelL.B	0	92	57	4.00E-73	5.00E-33	Bacteria	Short-chain dehydrogenase/reductase*
AV10011_*FabG*	Av212A	0	88	54	6.00E-67	2.00E-28	Bacteria	Short-chain dehydrogenase/reductase
AV10071_*HAL*	AvTelK.A	0	77	48	2.00E-61	1.00E-27	Bacteria	Histidine ammonia-lyase
AV10095_*GCN5*	AvTelL.A	0	59	35	2.00E-27	No hits	Bacteria; Proteobacteria	GCN5-related *N*-acetyltransferase*/*
AV10158_*FabG*	210B3	2	46	41	2.00E-39	2.00E-19	Bacteria	Short-chain dehydrogenase/reductase

as the metazoan long-chain acyl–coenzyme A synthetase (ACS) gene fragment on telomere K, which has an apparently intact 5′ sequence but is 3′-truncated by telomeric repeat addition to exon 2 (fig. S3A). Single-telomere length analysis PCR (STELA-PCR) (*20*) verified its telomeric localization (fig. S3B). No colinear partner of ACS was found, and the lack of RT-PCR product suggests that it is transcriptionally inactive or that its transcript is unstable (fig. S3C).

Two other contigs containing telomeric repeats (Table 1 and fig. S1D) had apparently intact genes for two nonribosomal peptide synthetases (NRPSs), large enzymes responsible for synthesis in bacteria and fungi of biologically active peptides including antibiotics and toxins (*21*). This finding suggests that bdelloid biosynthetic activity includes the production of secondary metabolites.

We also examined sequences in the vicinity of retrovirus-like elements in *Philodina roseola*, a bdelloid that separated from *A. vaga* tens of millions of years ago (*22*). The *P. roseola* telomere P (*15*) contains a *gag* gene from a retrovirus-like element similar to *A. vaga Juno1*, named *Juno2*, which is 3′-truncated by *P. roseola* telomeric repeats (fig. S3A). We probed a *P. roseola* genomic library with this *gag* fragment and found that two of five *Juno2* copies are surrounded by foreign genes (Table 1 and fig. S1E). Thus, extensive HGT probably represents a general feature of bdelloid rotifers.

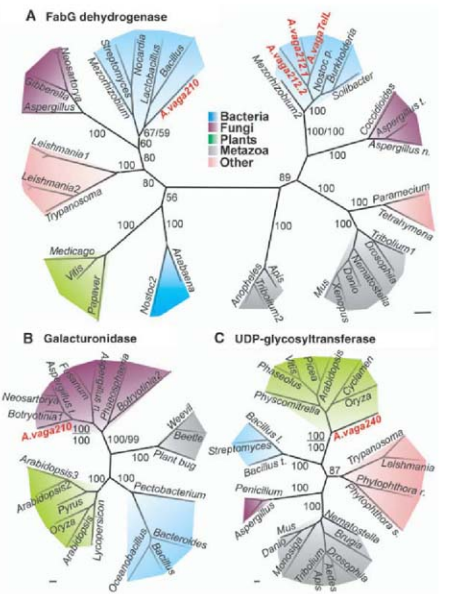


Fig. 2. Bayesian analysis of *A. vaga* CDS from three different kingdoms. Clades with different taxonomic affiliation are color-coded as in Fig. 1; CDS from *A. vaga* are in red. **(A)** Two subfamilies of *FabG* short-chain dehydrogenases/reductases, with four *A. vaga FabG*-like genes of bacterial origin (one contains an intron; three are intronless and could have entered independently or undergone duplication). **(B)** Galacturonidases and the corresponding *A. vaga* gene of fungal origin. **(C)** UDP-glycosyltransferases and the corresponding *A. vaga* gene of plant origin. Bayesian posterior probabilities are shown; for branches leading to *A. vaga*, neighbor-joining bootstrap support values are also indicated. For alignments, see table S11. Scale bars, 0.1 amino acid substitution per site.

We have not yet found a case of HGT in which the donor species could be identified by near-identity of a bdelloid gene to that of a putative donor, as in some other taxa (*7*, *8*, *23*), or in which the time of transfer could be estimated from the degree of synonymous difference, presumably because the transferred genes and their relatives are not yet represented in databases or because they have resided in bdelloids long enough to have substantially diverged. The similarity of the length distribution and sequence of introns in the putatively foreign genes, including genes of apparently bacterial origin, to that typical of other bdelloid genes (fig. S4) and the similarity of nucleotide composition and codon usage (table S1) suggest that some foreign genes were acquired long enough ago to conform to their host genomes, whereas other intronless genes such as *MvtM* and *XynB* (fig. S2B) may have arrived more recently.

Our examination of ~1 Mb of bdelloid telomeric/TE-rich sequence (about 1% of the genome) revealed dozens of genes that are of foreign (bacterial, plant, or fungal) origin, and about twice as many indeterminate genes, some of which may be foreign (Figs. 2 and 3, Table 1, and table S1). The AI values for all the CDS in bdelloid telomeric/TE-rich regions indicate that a substantial fraction resulted from HGT, with about one-third having AI > 45 and the majority having AI > 0. In contrast, inspection of comparable DNA amounts from *P. roseola* and *A. vaga* gene-rich regions surrounding *hsp82*, histone, and Hox genes (*16*, *17*, *19*) shows that these regions contain mostly genes with known metazoan relatives and are virtually devoid of TEs (Fig. 3, A and B). Indeed, the AI distribution of CDS in these regions is indistinguishable from that for translated expressed sequence tags (ESTs) from the monogonont rotifer *Brachionus plicatilis* (*24*) and for CDS from *Caenorhabditis elegans* (including those near telomeres) and *Drosophila melanogaster* (Fig. 3C).

The concentration of foreign genes near telomeres may result from preferential incorporation at or near chromosome ends, perhaps at deprotected telomeres, or from selection against more proximal insertions that disrupt essential sequences or are associated with deleterious rearrangements caused by ectopic repair of double-strand breaks via capture of foreign DNA (*15*, *25*, *26*). The lack of orientation bias with respect to telomeres indicates that their acquisition does not involve an RNA intermediate copied directly into cDNA, as occurs during terminal retrotransposition. Among ORFs with AI > 45 (Table 1), we found 5 in various degrees of decay and 17 that appear to be intact. Most of the intact foreign genes code for simple enzymatic functions such as carbohydrate decomposition (Table 1) and are not parts of multicomponent pathways that might not function properly after transfer to a distant host (*27*).

The apparent HGT cannot reasonably be explained by vertical inheritance from the com-

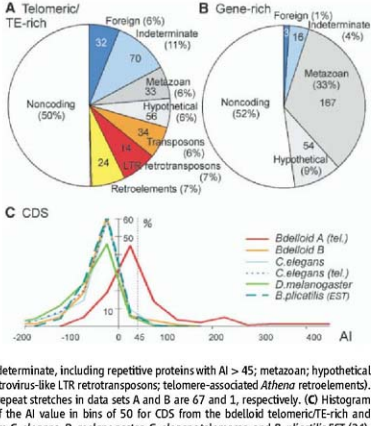


Fig. 3. Comparison of bdelloid telomeric/TE-rich regions (**A**) with bdelloid gene-rich regions (**B**) and with other model invertebrates (**C**). Pie charts were built with 921,903 base pairs (bp) (**A**) and 661,316 bp (**B**) of genomic DNA, respectively (excluding one of the two colinear partners). For TE count, ~1.3 Mb of genomic DNA in each data set was analyzed (including both colinear partners). The size of each sector corresponds to the percentage of the total length (bp) occupied by each category in a data set. Numbers correspond to the total count of entries from each category in a data set: genes (foreign; indeterminate, including repetitive proteins with AI > 45; metazoan; hypothetical ORFs) and TEs (DNA TEs; retrovirus-like LTR retrotransposons; telomere-associated *Athena* retroelements). The numbers of telomeric repeat stretches in data sets A and B are 67 and 1, respectively. (**C**) Histogram showing the distribution of the AI value in bins of 50 for CDS from the bdelloid telomeric/TE-rich and gene-rich data sets and from *C. elegans*, *D. melanogaster*, *C. elegans* telomeres, and *B. plicatilis* EST (*24*).

mon universal ancestor and subsequent loss in other kingdoms and in at least four major metazoan branches preceding Rotifera on the evolutionary tree (*28*), and is inconsistent with our finding of genes representing fusions between domains of prokaryotic and eukaryotic origin. It may be that HGT is facilitated by membrane disruption and DNA fragmentation and repair associated with the repeated desiccation and recovery experienced in typical bdelloid habitats, allowing DNA in ingested or other environmental material to enter bdelloid genomes (*12*, *29*). Whether there may also be homologous replacement by DNA segments released from related individuals remains to be seen. If there is, bdelloid rotifers may experience genetic exchange resembling that in sexual populations (*30*). Although the adaptive importance of such massive HGT remains to be elucidated, it is evident that such events have frequently occurred in the genomes of bdelloid rotifers, probably mediated by their unusual lifestyle.

References and Notes

1. Y. Boucher *et al.*, *Annu. Rev. Genet.* **37**, 283 (2003).
2. W. F. Doolittle, *Trends Genet.* **14**, 307 (1998).
3. B. Loftus *et al.*, *Nature* **433**, 865 (2005).
4. H. G. Morrison *et al.*, *Science* **317**, 1921 (2007).
5. J. O. Andersson, *Cell. Mol. Life Sci.* **62**, 1182 (2005).
6. L. A. Katz, *Int. J. Syst. Evol. Microbiol.* **52**, 1893 (2002).
7. U. Bergthorsson, K. L. Adams, B. Thomason, J. D. Palmer, *Nature* **424**, 197 (2003).
8. J. C. Dunning Hotopp *et al.*, *Science* **317**, 1753 (2007); published online 30 August 2007 (10.1126/science.1142490).
9. M. G. Kidwell, *Annu. Rev. Genet.* **27**, 235 (1993).
10. M. Syvanen, C. I. Kado, Eds., *Horizontal Gene Transfer* (Academic Press, London, 2002).
11. B. B. Normark, O. Judson, N. Moran, *Biol. J. Linn. Soc* **79**, 69 (2003).
12. J. Lapinski, A. Tunnacliffe, *FEBS Lett.* **553**, 387 (2003).
13. I. R. Arkhipova, M. Meselson, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11781 (2005).
14. E. G. Gladyshev, M. Meselson, I. R. Arkhipova, *Gene* **390**, 136 (2007).
15. E. A. Gladyshev, I. R. Arkhipova, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9352 (2007).
16. D. B. Mark Welch, J. L. Mark Welch, M. Meselson, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5145 (2008).
17. J. H. Hur, thesis, Harvard University (2006).
18. N. N. Pouchkina-Stantcheva, A. Tunnacliffe, *Mol. Biol. Evol.* **22**, 1482 (2005).
19. See supporting material on *Science* Online.
20. D. M. Baird, J. Rowson, D. Wynford-Thomas, D. Kipling, *Nat. Genet.* **33**, 203 (2003).
21. R. Finking, M. A. Marahiel, *Annu. Rev. Microbiol.* **58**, 453 (2004).
22. D. Mark Welch, M. Meselson, *Science* **288**, 1211 (2000).
23. N. Kondo, N. Nikoh, N. Ijichi, M. Shimada, T. Fukatsu, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14280 (2002).
24. K. Suga, D. Mark Welch, Y. Tanaka, Y. Sakakura, A. Hagiwara, *PLoS ONE* **2**, e671 (2007).
25. A. Haviv-Chesner, Y. Kobayashi, A. Gabriel, M. Kupiec, *Nucleic Acids Res.* **35**, 5192 (2007).
26. T. de Lange, *Genes Dev.* **19**, 2100 (2005).
27. R. Jain, M. C. Rivera, J. A. Lake, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3801 (1999).
28. C. W. Dunn *et al.*, *Nature* **452**, 745 (2008).
29. E. Gladyshev, M. Meselson, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5139 (2008).
30. S. P. Otto, T. Lenormand, *Nat. Rev. Genet.* **3**, 252 (2002).

31. We thank W. Reznikoff, M. Belfort, and D. Mark Welch for comments and J. Mark Welch, K. van Dorninck, and J. Hur for communicating results before publication. Supported by NSF grant MCB-0614142 (M.M. and I.R.A.) and NIH grant GM072708 (M.M.). Sequences obtained in this study were deposited in GenBank (accession numbers EU643473 to EU643504).

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1210/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S12
References

12 February 2008; accepted 22 April 2008
10.1126/science.1156407

Ancestral Monogamy Shows Kin Selection Is Key to the Evolution of Eusociality

William O. H. Hughes,[1]* Benjamin P. Oldroyd,[2] Madeleine Beekman,[2] Francis L. W. Ratnieks[3]

Close relatedness has long been considered crucial to the evolution of eusociality. However, it has recently been suggested that close relatedness may be a consequence, rather than a cause, of eusociality. We tested this idea with a comparative analysis of female mating frequencies in 267 species of eusocial bees, wasps, and ants. We found that mating with a single male, which maximizes relatedness, is ancestral for all eight independent eusocial lineages that we investigated. Mating with multiple males is always derived. Furthermore, we found that high polyandry (>2 effective mates) occurs only in lineages whose workers have lost reproductive totipotency. These results provide the first evidence that monogamy was critical in the evolution of eusociality, strongly supporting the prediction of inclusive fitness theory.

Eusocial behavior, exemplified by social insects, represents one of the pinnacles of sociality and is characterized by individuals altruistically helping to rear siblings rather than their own offspring (*1*). The established paradigm, based on inclusive fitness (kin selection) theory, is that eusociality evolves because of a combination of the direct benefits of altruism (i.e., helpers increase the number of individuals reared) and close relatedness between group members, such that the inclusive fitness of helpers exceeds that achievable through a solitary lifestyle (*2*–*5*). High relatedness, arising from the delayed dispersal of offspring [possibly

but not necessarily strengthened by genetic factors such as inbreeding or haplodiploidy (*5*)], has therefore long been thought to be critical to the evolution of eusociality. However, relatedness in colonies of some extant species is relatively low, and this has led E. O. Wilson, the founder of sociobiology (*1*), to propose recently that eusociality instead evolves because of direct benefits alone, with relatedness being unimportant (*6–8*). Under this alternative hypothesis, individuals that share a "eusociality allele," and which may be related or unrelated, cooperate because their direct fitness is greater when in a social group than when alone (for example, if independent nest founding is extremely risky). Limited dispersal may then tend to lead to groups being made up of kin, but, in contrast to kin selection theory, this increased relatedness is hypothesized to be a consequence, rather than a cause, of the evolution of eusociality (*6–8*).

Comparative patterns of female mating frequencies may help to resolve this controversy, because mating with multiple males (polyandry) reduces relatedness among siblings (*2*). If kin selection is important for the evolution of eusociality, then monandry, which maximizes relatedness, should be ancestral (*2*, *4*, *5*). Alternatively, if high relatedness is unnecessary for the evolution of eusociality (*6–8*), then there should be no such relationship. The mating systems present when eusociality originated can be inferred by ancestral state reconstruction on the basis of the phylogeny and mating systems of extant species. The eusocial Hymenoptera provide the ideal group for such an analysis, as they are well studied and comprise nine lineages in which eusociality has independently evolved: once in sphecid wasps (*9*, *10*), three times in halictid bees (*11*), once in allodapine bees (*12*), once in corbiculate bees (*13*), twice in vespid wasps (in the Stenogastrinae and the Polistinae+Vespinae) (*14*), and once in ants. The eusocial Hymenoptera also exhibit a wide range of mating systems, from obligate monandry to extreme polyandry (*15–17*).

We compiled a data set of female mating frequencies for 267 species of eusocial bees, wasps, and ants (table S1), covering all of the eusocial lineages except allodapine bees. Most species are monandrous, about a quarter show facultative low polyandry (<2 effective mates), and nine clades within the eusocial lineages have high levels of polyandry (>2 effective mates) (Fig. 1). We mapped these data onto a phylogeny (Fig. 1 and fig. S1) and carried out ancestral state reconstruction (*18*).

Our analysis shows that monandry was the ancestral state for all eight of the independent origins of eusociality (Fig. 1 and Table 1). All females are monandrous in the only eusocial sphecid wasp (maximum likelihood model $P < 0.0001$ of monandry not being the ancestral state). Females in the three eusocial lineages of halictid bees mate either singly or very rarely doubly ($P = 0.001$, $P = 0.001$, and $P = 0.003$ for the three lineages). Among the corbiculate bees, all stingless bees are monandrous, all honeybees are polyandrous, and bumblebees are ancestrally monandrous with facultative low polyandry being derived in, principally, the *Pyrobombus* clade ($P = 0.003$) (fig. S2). Overall, monandry is the ancestral state for the eusocial corbiculate bees ($P = 0.015$). Monandry is also the most probable ancestral state for both the stenogastrine ($P = 0.003$) and Polistinae+Vespinae ($P = 0.043$) lineages of wasps. The ants show the most evolutionary transitions in mating system, but again monandry was the ancestral state ($P = 0.034$). Thus, in all the eusocial lineages of Hymenoptera where polyandry occurs (Fig. 1), it is derived. Furthermore, most polyandrous species have very low effective mating frequencies. The phylogenetic analysis shows that high levels of polyandry (>2 effective mates) are even more clearly derived ($P < 0.001$ in all cases) (Fig. 1 and Table 1).

The data do not allow us to determine whether monandry was already present in the solitary ancestors or whether monandry and eusociality evolved concurrently, but they are clearly linked. These findings of ancestral monandry and derived polyandry strongly support the prediction of kin selection theory that high relatedness was important in the evolution of eusociality. The results are incompatible with the idea that high relatedness is derived (*6–8*). Data on mating systems in non-hymenopteran eusocial lineages are far less extensive than for the Hymenoptera, but these data indicate that monandry and high relatedness are also the norm in these groups (Table 1) (*4*, *19–23*). Furthermore, in the one group in which the comparative relationship has been examined (the thrips), high relatedness resulting from inbreeding was inferred at the origins of eusociality (*24*).

We also used our data for a second, independent test of the role of kin selection in the evolution of eusociality. In some eusocial taxa (*25*, *26*), social evolution has reached the point where workers have lost the ability to mate and reproduce sexually (that is, they are no longer reproductively totipotent). As a result, workers in these taxa cannot found colonies independently or replace the breeder queen and so are irreversibly eusocial. Inclusive fitness theory predicts that polyandry in eusocial taxa will normally evolve only after workers have lost reproductive totipotency (*2*, *4*), whereas there should be no relationship if Wilson's hypothesis is correct (*6–8*). For each of the species in our data set, we scored worker reproductive totipotency as 1 or 0 from the literature (*18*). The loss of the physiological and behavioral ability to mate almost always results from the evolution of morphologically distinct workers, so where explicit data were unavailable, we assumed that species with queen/worker dimorphism had non-totipotent workers. We analyzed the relationship between totipotency and polyandry across the eusocial Hymenoptera using the method of phylogenetically independent contrasts (*18*). Both the proportion of females mating multiply (least-squares regression $F_{1,184} = 7.53$, $P = 0.007$) and their effective mating frequency ($F_{1,174} = 6.42$, $P = 0.012$) were significantly lower in species with totipotent workers than in species where workers had lost reproductive totipotency (Fig. 2A and table S1). Worker totipotency was significantly rarer in polyandrous species than in those with monandry, and none of the species with high levels of polyandry had totipotent workers ($F_{1,217} = 5.91$, $P = 0.016$) (Fig. 2B).

In five of the eusocial lineages (sphecid wasps, the three halictid lineages, and stenogastrine wasps), all females are totipotent. However, in the three eusocial lineages in which some taxa have non-totipotent workers (corbiculate bees, polistine+vespine wasps, and ants), ancestral state reconstruction indicated that totipotency was lost prior to or concurrently with the evolution of polyandry. Non-totipotency is ancestral in the eusocial corbiculate bees (maximum likelihood model $P = 0.015$ of totipotency being the ancestral state), whereas polyandry is derived. The clade of polistine+vespine wasps is ancestrally totipotent ($P = 0.969$). However, non-totipotency is ancestral in the vespines ($P = 0.015$), so totipotency was lost before the subsequent evolution of polyandry in this clade. Finally, in the ants, totipotency is ancestral overall ($P = 0.983$) but was lost early in the evolution of polyandrous *Pachycondyla* ($P = 0.015$). Non-totipotency is the ancestral state for nonponerine ants ($P = 0.03$), indicating that it was lost before the multiple evolutions of polyandry in these ants.

In addition to polyandry, relatedness in eusocial colonies is also decreased if colonies have multiple, simultaneously breeding females (functional polygyny), as occurs in some extant species of eusocial Hymenoptera (*27*). If functional polygyny had been the ancestral state in eusocial lineages, then this would weaken the support for high relatedness being important in the evolution of eusociality. Although our data set was collected primarily to examine polyandry, we also performed an ancestral state reconstruction for functional polygyny (*18*, *27*). Colonies of eusocial sphecids normally have only a single reproducing female at any one time, as do colonies of the only *Augochlora* and *Halictus* species included in our analyses. In the third eusocial lineage of halictid bees, *Lasioglossum*, the ancestral state was uncertain ($P = 0.069$). Func-

[1]Institute of Integrative and Comparative Biology, University of Leeds, Leeds, LS2 9JT, UK. [2]School of Biological Sciences, University of Sydney, Sydney, New South Wales 2006, Australia. [3]Department of Biological and Environmental Science, University of Sussex, Falmer, Brighton, BN1 9QG, UK.

*To whom correspondence should be addressed. E-mail: w.o.h.hughes@leeds.ac.uk

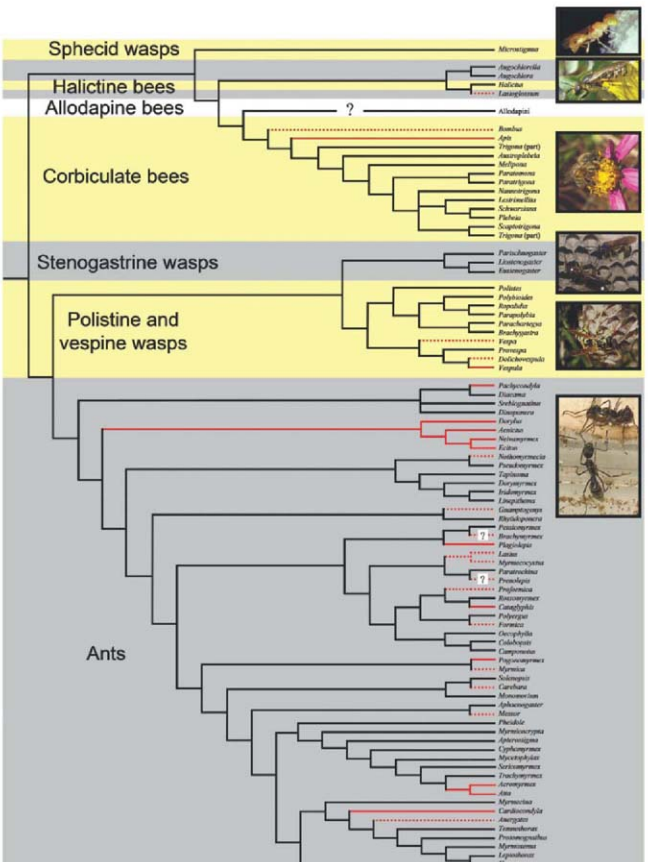


Fig. 1. Phylogeny of genera of eusocial Hymenoptera (ants, bees, and wasps) for which female mating-frequency data are available. See fig. S1 for the species-level phylogeny used in the analyses. Each independent origin of eusociality is indicated by alternately colored clades. Clades exhibiting high polyandry (>2 effective mates) have solid red branches, those exhibiting facultative low polyandry (>1 but <2 effective mates) have dotted red branches, and entirely monandrous genera have solid black branches. Mating-frequency data are not available for the allodapine bees. [Photographs from top to bottom are of (and reproduced with permission from): *Microstigmus comes* (R. Matthews), *Lasioglossum malachurum* (C. Polidori), *Apis mellifera* (F.L.W. Ratnieks), *Liostenogaster flavolineata* ( ). Fields), *Polistes dominulus* (W.O.H. Hughes), and *Diacamma* sp. (W.O.H. Hughes)]

tional monogyny was found to be ancestral in the eusocial corbiculate bees ($P = 0.001$), stenogastrine wasps ($P = 0.0005$), polistine+vespine wasps ($P = 0.006$), and ants ($P = 0.0018$). Colonies of some primitively eusocial species, such as *Polistes* wasps (*28*) and the stenogastrine wasp *Liostenogaster* (*29*), may have several mated females. However, these females reproduce consecutively, reducing the impact on relatedness relative to concurrent reproduction. Therefore, the evidence indicates that functional monogyny is ancestral in the eusocial Hymenoptera, again in keeping with inclusive fitness theory.

Our study comes to the simple conclusion that monandry was the ancestral state when eusociality arose in ants, bees, and wasps, something that also appears generally true in the non-hymenopteran eusocial taxa. Additionally, we show that worker reproductive totipotency is associated with monandry. Obligate or high levels of polyandry are invariably derived and occur only in species whose workers lack the ability to become the primary breeders. Thus, polyandry seems to have evolved after social evolution had reached a point at which reversion from eusociality was impossible. Similarly, functional polygyny is also derived. Our conclusions strongly indicate that kin selection and high relatedness have played a decisive role in the evolution of eusociality.

Table 1. Summary of evidence for monandry in the independent origins of eusociality in the Hymenoptera (ants, bees, and wasps; this study) and other eusocial lineages. *P* values relate to ancestral state reconstructions by maximum likelihood. The first value indicates the probability of monandry not being the ancestral state, whereas the second value indicates the probability of high polyandry (>2 effective mates) being the ancestral state.

Taxa	Eusocial origins	Eusocial species	Evidence
Sphecid wasps	1	1	This study; $P < 0.0001$; $P < 0.0001$
Halictid bees (*Augochlorella/Augochlora*)	1	Many of 140	This study; $P = 0.0014$; $P < 0.0001$
Halictid bees (*Halictus*)	1	Most of 217	This study; $P = 0.0014$; $P < 0.0001$
Halictid bees (*Lasioglossum*)	1	Most of 544	This study; $P = 0.003$; $P < 0.0001$
Allodapine bees	1		No data
Corbiculate bees	1	≈1000	This study; $P = 0.015$; $P = 0.0007$
Stenogastrine wasps	1	≈50	This study; $P = 0.0026$; $P < 0.0001$
Polistine and vespid wasps	1	≈860	This study; $P = 0.043$; $P < 0.0001$
Ants	1	≈12,000	This study; $P = 0.034$; $P = 0.0007$
Ambrosia beetle	1	1	Monoandry thought probable but no data (*19*)
Aphids	≈17	≈50	Eusocial colonies produced parthenogenetically by single female (*4*, *23*)
Termites	1	≈2800	Generally monoandrous, with only a few species exhibiting low polyandry (*4*)
Thrips	1	7	Normally monoandrous (*4*, *23*, *24*)
Snapping shrimps	3	6	Monoandrous (*20*)
Mole rats	2	2	Facultative low polyandry (*21*, *22*)

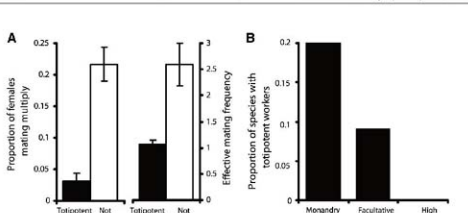


Fig. 2. (**A**) Mean (±SE, indicated by error bars) proportion of females mating multiply and effective mating frequency for species with and without totipotent workers. (**B**) Occurrence of totipotent workers in species with monandry, facultative low polyandry (>1 but <2 effective mates), or high polyandry (>2 effective mates).

References and Notes

1. E. O. Wilson, *Sociobiology: The New Synthesis* (Belknap Press of Harvard Univ. Press, Cambridge, MA, 1975).
2. W. D. Hamilton, *J. Theor. Biol.* **7**, 1 (1964).
3. D. C. Queller, J. E. Strassmann, *Bioscience* **48**, 165 (1998).
4. J. J. Boomsma, *Curr. Biol.* **17**, R673 (2007).
5. R. H. Crozier, *Aust. J. Entomol.* **47**, 2 (2008).
6. E. O. Wilson, *Bioscience* **58**, 17 (2008).
7. E. O. Wilson, B. Hölldobler, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13367 (2005).
8. E. O. Wilson, *Soc. Res. (New York)* **72**, 159 (2005).
9. R. W. Matthews, *Science* **160**, 787 (1968).
10. K. G. Ross, R. W. Matthews, *Anim. Behav.* **38**, 613 (1989).
11. S. G. Brady, S. Sipes, A. Pearson, B. N. Danforth, *Proc. R. Soc. London Ser. B Biol. Sci.* **273**, 1643 (2006).
12. M. P. Schwarz, N. J. Bull, S. J. B. Cooper, *Syst. Biol.* **52**, 1 (2003).
13. G. J. Thompson, B. P. Oldroyd, *Mol. Phylogenet. Evol.* **33**, 452 (2004).
14. H. M. Hines, J. H. Hunt, T. K. O'Connor, J. J. Gillespie, S. A. Cameron, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3295 (2007).
15. J. Strassmann, *Insectes Soc.* **48**, 1 (2001).
16. S. Sumner, W. O. H. Hughes, J. S. Pedersen, J. J. Boomsma, *Nature* **428**, 35 (2004).
17. D. R. Tarpy, R. Nielsen, D. I. Nielsen, *Insectes Soc.* **51**, 203 (2004).
18. Materials and methods are available as supporting material on *Science* Online.
19. L. R. Kirkendall, D. S. Kent, K. A. Raffa, in *The Evolution of Social Behavior in Insects and Arachnids*, J. C. Choe, B. J. Crespi, Eds. (Cambridge Univ. Press, Cambridge, 1997).
20. J. E. Duffy, in *Genes, Behavior, and Evolution in Social Insects*, T. Kikuchi, N. Azuma, S. Higashi, Eds. (Univ. of Hokkaido Press, Sapporo, Japan, 2002), pp. 217–252.
21. H. K. Reeve, D. F. Westneat, W. A. Noon, P. W. Sherman, C. F. Aquadro, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2496 (1990).
22. T. M. Burland, N. C. Bennett, J. U. M. Jarvis, C. G. Faulkes, *Proc. R. Soc. London Ser. B Biol. Sci.* **269**, 1025 (2002).
23. H. Helantera, K. Bargum, *Oikos* **116**, 217 (2007).
24. T. W. Chapman, B. J. Crespi, B. D. Kranz, M. P. Schwarz, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1648 (2000).
25. B. J. Crespi, D. Yanega, *Behav. Ecol.* **6**, 109 (1995).
26. M. Beekman, C. Peeters, M. J. O'Riain, *Behav. Ecol.* **17**, 622 (2006).
27. W. O. H. Hughes, F. L. W. Ratnieks, B. P. Oldroyd, *J. Evol. Biol.*, published online 14 April 2008; 10.1111/j.1420208.01532.x.
28. S. Sumner, E. Lucas, J. Barker, N. Isaac, *Curr. Biol.* **17**, 140 (2007).
29. J. Field, G. Shreeves, S. Sumner, M. Casiraghi, *Nature* **404**, 869 (2000).

30. We thank H. Helantera, T. Wenseleers, A. Bourke, R. Hammond, J. Boomsma, and the anonymous reviewers for comments on the manuscript. This work was supported by the European Commission via a Marie Curie Outgoing International Fellowship to W.O.H.H. B.P.O. and M.B. were supported by the Australian Research Council.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1213/DC1
Materials and Methods
SOM Text
Figs. S1 and S2
Table S1
References

5 February 2008; accepted 22 April 2008
10.1126/science.1156108

Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures

Stanislas Dehaene,[1,2,3,4]* Véronique Izard,[1,2,4,5] Elizabeth Spelke,[5] Pierre Pica[6]

The mapping of numbers onto space is fundamental to measurement and to mathematics. Is this mapping a cultural invention or a universal intuition shared by all humans regardless of culture and education? We probed number-space mappings in the Mundurucu, an Amazonian indigene group with a reduced numerical lexicon and little or no formal education. At all ages, the Mundurucu mapped symbolic and nonsymbolic numbers onto a logarithmic scale, whereas Western adults used linear mapping with small or symbolic numbers and logarithmic mapping when numbers were presented nonsymbolically under conditions that discouraged counting. This indicates that the mapping of numbers onto space is a universal intuition and that this initial intuition of number is logarithmic. The concept of a linear number line appears to be a cultural invention that fails to develop in the absence of formal education.

What then is mathematics if it is not a unique, rigorous, logical structure? It is a series of great intuitions carefully sifted, and organized by the logic men are willing and able to apply at any time.

– Morris Kline, *Mathematics: The Loss of Certainty* [(*1*), p. 312]

The mapping of numbers onto space plays an essential role in mathematics, from measurement and geometry to the study of irrational numbers, Cartesian coordinates, the real number line, and the complex plane (*1*, *2*). How does the human mind gain access to such abstract mathematical concepts? Constructivist theories view mathematics as a set of cultural inventions that are progressively refined in the history of mathematics and are slowly acquired during childhood and adolescence (*3*). However, the mental construction of mathematics may have deeper foundations. Mathematical objects may find their ultimate origin in basic intuitions of space, time, and number that have been internalized through millions of years of evolution in a structured environment and that emerge early in ontogeny, independently of education (*2*, *4*). Here we present evidence that reconciles these two points of view: Our results suggest that all humans share the intuition that numbers map onto space, but that culture-specific experiences alter the form of this mapping.

Previous psychological and neuroimaging research supports the view that a sense of number is present in humans and many other species at an early age and has a reproducible substrate in the bilateral intraparietal sulcus (*5*–*8*). This region is remarkably close to or even overlapping with areas engaged in the coding of spatial dimensions such as size, location, and gaze direction (*9*–*11*). Interactions between numerical and spatial codes in the parietal cortex may therefore occur at this level. Indeed, in human adults, the mere presentation of an Arabic numeral automatically elicits a spatial bias in both motor responding and attention orienting (*11*–*13*). Brain-lesioned patients show corresponding impairments in comparing and bisecting line segments and numbers (*14*), and some people even report a vivid experience of seeing numbers at fixed locations on an idiosyncratic spatially contiguous "number form" (*15*, *16*).

Recent experiments document a remarkable shift in the child's conception of how numbers map onto space (*17*–*19*). When asked to point toward the correct location for a spoken number word onto a line segment labeled with 0 at left and 100 at right, even kindergarteners understand the task and behave nonrandomly, systematically placing smaller numbers at left and larger numbers at right. They do not distribute the numbers evenly, however, and instead devote more space to small numbers, imposing a compressed logarithmic mapping. For instance, they might place number 10 near the middle of the 0-to-100 segment. This compressive response fits nicely with animal and infant studies that demonstrate that numerical perception obeys Weber's law, a ubiquitous psychophysical law whereby increasingly larger quantities are represented with proportionally greater imprecision, compatible with a logarithmic internal representation with fixed noise (*7*, *20*, *21*). A shift from logarithmic to linear mapping occurs later in development, between first and fourth grade, depending on experience and the range of numbers tested (*17*–*19*).

All of these observations, however, were made in Western children who all had access to mathematical education and culture at an early age. Before formal schooling, Western children may acquire the number-line concept from Arabic numerals seen on elevators, rulers, books, etc. Thus, existing studies do not reveal which aspects of the number-space mapping constitute a basic intuition that would continue to exist in the absence of a structured mathematical language and education. In particular, we do not know whether the log-to-linear shift would occur spontaneously in the course of brain maturation or whether it requires exposure to critical educational material or culture-specific devices such as rulers or graphs.

To address these issues, we gathered evidence from psychological experimentation in the Mundurucu, an Amazonian indigene culture with little access to education (*22*, *23*). Previous research has established that, although their lexicon of number words is reduced and they have little or no access to rulers, measurement devices, graphs, or maps, the Mundurucu entertain sophisticated concepts of both number and space, albeit in an approximate and nonverbal manner (*22*, *23*). We therefore asked whether they conceive of these two domains as being related by a systematic mapping and, if so, what form this number-space mapping takes.

A total of 33 Mundurucu adults and children were tested individually in a number-space task (Fig. 1) (*24*). In each trial, a line segment was displayed on a computer screen, with 1 dot at left and 10 dots at right (or, in a separate block, 10 and 100 dots, respectively). Then other numbers were presented in random order, in various forms (sets of dots, sequences of tones, spoken Mundurucu words, or spoken Portuguese words). For each number, the participant pointed to a screen location and this response was recorded by a mouse click, without feedback. Only two training trials were presented, with sets of dots whose numerosity corresponded to the ends of the scale (1 and 10). The participants were told that these two stimuli belonged to their respective ends, but that other stimuli could be placed at any location. Because training did not involve intermediate numbers, performance on all subsequent trials served to reveal whether the participants would spontaneously use systematic mapping, and if so, whether their mapping would be compressive or linear.

The Mundurucu's mean responses revealed that they understood the task. Although some participants tended to use only the end points of the scale (*24*), most used the full response continuum and adopted a consistent strategy of mapping consecutive numbers onto consecutive locations (Fig. 2). There was a significant positive correlation between stimulus number and mean response location, regardless of the modality in which the numbers were presented. The task was easy when the stimuli were sets of dots similar to the reference labels placed at the end points [numbers 1 to 10, correlation coefficient (r^2) = 92.6%, 8 df; numbers 10 to 100, r^2 = 91.9%, 8 df]. However, the Mundurucu continued to use systematic number-space mapping with stimuli they

[1]INSERM, Cognitive Neuro-Imaging Unit, Institut Fédératif de Recherche 0FR 49, Gif sur Yvette, France. [2]Commissariat à l'Energie Atomique, NeuroSpin center, IFR 49, Gif sur Yvette, France. [3]Collège de France, Paris, France. [4]Université Paris-Sud, IFR 49, F-91191 Gif sur Yvette, France. [5]Department of Psychology, Harvard University, Cambridge, MA 02139, USA. [6]Unité Mixte de Recherche 7023 "Formal Structure of Language," CNRS and Paris VIII University, Paris, France.

*To whom correspondence should be addressed. E-mail: stanislas.dehaene@cea.fr

had not been trained with, which shared with the reference labels only an abstract concept of number: sequences of tones 1 to 10 (r^2 = 92.5%, 8 df), spoken Mundurucu number words (r^2 = 91.8, 6 df), and Portuguese number words (r^2 = 91.1%, 8 df), although a small proportion of random responses tended to slightly flatten the curves. The Mundurucu stimuli included complex expressions that are very rarely uttered, such as "pũg põgbi ebadipdip bodi" [approximate translation: "one handful (and) four on the side"]. The results suggest that the Mundurucu partially understand the quantity to which these expressions refer.

Crucially, however, linear regression did not provide the best model of the Mundurucus' responses. Rather, for all modalities of presentation, the curves were negatively accelerated. A multiple regression procedure evaluated the contribution of a logarithmic regressor, over and above the linear regressor. The logarithmic compression effect was significant for all stimulus modalities, although it was only marginal with Portuguese words (one-tailed test, P = 0.04; see significance levels and regression weights in Fig. 2). Additional analyses allowed us to exclude interpretations in terms of linear responding with different slopes for small and large numbers, parallax error, experimenter bias, or bimodal responding (*24*). The Mundurucu seem to hold intuitions of numbers as a log scale where the middle of the interval 1 through 10 is 3 or 4, not 5 or 6.

Previous number-space mapping experiments with Western participants included only symbolic numerals, whereas the present experiment included nonsymbolic visual and auditory numerosities. Thus, it was important to verify whether these novel stimuli were rated linearly or logarithmically in educated Western participants. As shown in Fig. 2, American adults rated linearly all numerals presented in English and in Spanish, as well as the sets of 1 to 10 dots, which could easily be counted. However, they exhibited a significant logarithmic component with sets of 10 to 100 dots and with sequences of tones. When the two groups of participants were compared directly, the Mundurucu showed a greater compressive nonlinearity than the American participants only with sets of 1 to 10 dots (P = 0.003) and with numerals in the first language (Mundurucu versus English numerals, P = 0.033). This finding concurs with previous data suggesting that Western people estimate large numerosities in an approximate and compressive manner (*25*, *26*). Their judgments are linear only when the numbers are presented in a symbolic manner or as small sets whose numerosity can be precisely assessed.

The Mundurucu population is heterogeneous, and some of our participants, particularly the children, had received a little education. To examine the impact of this variable, we calculated, for each participant, an index of nonlinearity in the number-space mapping: the weight of the log regressor in a multiple regression of the data on linear and log regressors. For this analysis, we pooled the trials with 1 to 10 dots and number words, but excluded those with 10 to 100 dots and with tones, for which Western participants showed some nonlinearity. The index confirmed a highly significant nonlinearity in Mundurucu participants (Student's t test = 6.20, 34 df, $P < 10^{-6}$). In American participants, performance did not deviate from linearity (P = 0.08) and differed markedly from that of the Mundurucu (Welch's t test = 4.37, 48.6 df, $P < 0.0001$). Crucially, the Mundurucu's nonlinearity remained significant even when the analysis was restricted to adults (t = 4.34, 23 df, P = 0.0002), to monolingual speakers (t = 5.36, 29 df, $P < 10^{-5}$), or to uneducated participants (t = 2.60, 7 df, P = 0.035; see figs. S7 to S10 for a graphic depiction of subgroup performance) (*24*). t tests and linear and rank-order regression analyses showed no effect of gender, age, education, or bilingualism. There was only a trend toward reduced nonlinearity as a function of age (Kendall tau = –0.23, P = 0.055). Although this observation suggests that older Mundurucu may evolve a greater understanding of the linear number line, it should be noted that in Western children, the mapping becomes linear over the range from 10 to 100 by the first or second grade (*17*–*19*), whereas in our data, even the oldest Mundurucu adults (those over 40) continued to show highly significant nonlinearity over the range from 1 to 10 (t = 3.36, 11 df, P = 0.006).

Finally, we analyzed the special case of Portuguese numerals. Although overall performance was logarithmic, subdivision by education level indicated that logarithmic responding held for participants with 1 to 2 years of education (t = 3.15, 16 df, P = 0.006; fig. S9) but not for those with no education at all or with more education. In uneducated participants, performance with Portuguese numerals was highly variable and weakly correlated with number (r^2 = 39.0, P = 0.053; fig. S8), suggesting that many of these participants simply did not know the meaning of Portuguese numerals. For the most educated group, on the other hand, performance was strictly linear (r^2 = 94.5%, $P < 10^{-5}$; fig. S10). Excluding participants with no education, we found that greater education significantly changed the responses to Portuguese from logarithmic to linear (t = 2.48, 16.6 df, P = 0.024) but left responses to Mundurucu numerals and dot patterns unchanged ($P > 0.5$), thus yielding a significant interaction (P = 0.008). Strikingly, within the more educated group, performance varied significantly with number notation (t = 3.12, 9 df, P = 0.012), because it was linear for Portuguese numerals but logarithmic for Mundurucu numerals and dot patterns from 1 to 10.

Overall, these results reveal both universal and culture-dependent facets of the sense of number. After a minimal instruction period, even members of a remote culture with reduced vocabulary and education readily understand that number can be mapped onto a spatial scale. The exact form of this mapping switches dramatically from logarithmic to linear, however, depending on the ages

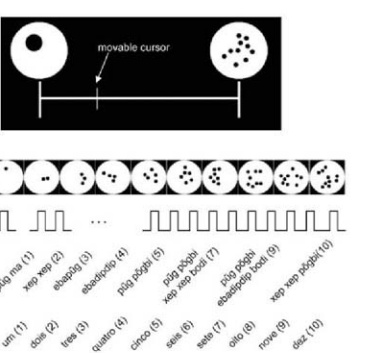


Fig. 1. Number mapping task with numbers from 1 to 10. A horizontal segment, labeled with a set of 1 dot on the left and a set of 10 dots on the right, was constantly present on screen. Numbers were presented visually as sets of dots or auditorily as sequences of tones (*24*), Mundurucu numerals, or Portuguese numerals. For Mundurucu numerals, a rough translation into Arabic numerals is provided (for example, "pũg põgbi xex xep bodi" ≈ "one handful (and) two on the side" ≈ 7; "xex xep põgbi" ≈ "two handfuls" ≈ 10). For each stimulus, participants pointed to a place on the line, and the experimenter clicked it with the computer mouse, which made a small bar appear.

at which people are tested, the education they have received, and the format in which numbers are presented.

In light of the performance of Amazonian adults, it is clear that the mental revolution in Western children's number line does not result from a simple maturation process: Logarithmic thinking persists into adulthood for the Mundurucu, even for very small numbers in the range from 1 to 10, whether presented as dots, tones, or spoken Mundurucu words. What are the sources of this universal logarithmic mapping? Research on the brain mechanisms of numerosity perception have revealed a compressed numerosity code, whereby individual neurons in the parietal and prefrontal cortex exhibit a Gaussian tuning curve on a logarithmic axis of number (*27*). As first noted by Gustav Fechner, such a constant imprecision on a logarithmic scale can explain Weber's law—the fact that larger numbers require a proportional larger difference in order to remain equally discriminable. Indeed, a recent model suggests that the tuning properties of number neurons can account for many details of elementary mental arithmetic in humans and animals (*21*). In the final analysis, the logarithmic code may have been selected during evolution for its compactness: Like an engineer's slide rule, a log scale provides a compact neural representation of several orders of magnitude with fixed relative precision.

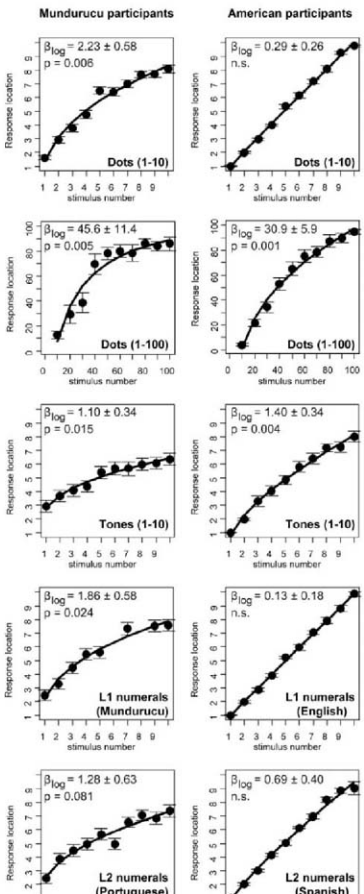


Fig. 2. Average location of numbers on the horizontal segment, separately for Mundurucu participants (left column) and for American participants (right column). β_{log} weight of the logarithmic regressor in a multiple regression analysis with linear and logarithmic predictors. Data are mean ± SE of the mean. Graphs of performance broken down by age group and education are available as supporting online material (*24*). L1, first language; L2, second language.

The simultaneous presence of linear and compressed mental representations of numbers is probably not unique to the Mundurucu. In American children, logarithmic mapping does not disappear all at once, but vanishes first for small numbers and much later for larger numbers from 1 to 1000 (up to fourth or sixth grade in some children) (*17*–*19*). In fact, a logarithmic representation may

It is not yet known which critical educational or cultural experience turns this initial representation into a linear scale. When a cultural difference in conceptual representation is observed in a remote population, Whorf's hypothesis is often invoked (*28*), according to which language determines the organization of thought. In the present case, however, the Whorfian explanation fails, because neither linguistic competence per se (present in all Mundurucu), nor numerical vocabulary and verbal counting [present in bilingual Mundurucu and in young children (*24*)] suffice to induce the log-to-linear shift (*17*–*19*). Speculatively, two factors underlying the shift may be experience with measurement, whereby a fixed numerical unit is applied to different spatial locations; and experience with addition and subtraction, ultimately yielding the intuition that all consecutive numbers are separated by the same interval + 1. The most educated Mundurucu eventually understand that linear scaling, which allows measurement and invariance over addition and subtraction, is central to the Portuguese number word system. At the same time, they still do not extend this principle to the Mundurucu number words, where perceptual similarity between quantities is still seen as the most relevant property of numbers. The system of Mundurucu number words may be a cultural device that does not emphasize measurement or invariance by addition and subtraction as defining features of number, contrary to Western numeral systems.

remain dormant in all of us for very large numbers or whenever we approximate numbers (*29*), including prices (*30*). Thus, log and linear scales may be deeply embedded in all of our mental activities.

References and Notes

1. M. Kline, *Mathematics: The Loss of Certainty* (Oxford Univ. Press, New York, 1980).
2. S. Dehaene, *The Number Sense* (Oxford Univ. Press, New York, 1997).
3. J. Piaget, *The Child's Conception of Number* (Norton, New York, 1952).
4. R. N. Shepard, *Behav. Brain Sci.* **24**, 581 (2001).
5. S. Dehaene, N. Molko, L. Cohen, A. J. Wilson, *Curr. Opin. Neurobiol.* **14**, 218 (2004).
6. B. Butterworth, *The Mathematical Brain* (Macmillan, London, 1999).
7. L. Feigenson, S. Dehaene, E. Spelke, *Trends Cogn. Sci.* **8**, 307 (2004).
8. A. Nieder, *Nat. Rev. Neurosci.* **6**, 177 (2005).
9. O. Simon, J. F. Mangin, L. Cohen, D. Le Bihan, S. Dehaene, *Neuron* **33**, 475 (2002).
10. P. Pinel, M. Piazza, D. Le Bihan, S. Dehaene, *Neuron* **41**, 983 (2004).
11. E. M. Hubbard, M. Piazza, P. Pinel, S. Dehaene, *Nat. Rev. Neurosci.* **6**, 435 (2005).
12. S. Dehaene, S. Bossini, P. Giraux, *J. Exp. Psychol. Gen.* **122**, 371 (1993).
13. M. H. Fischer, A. D. Castel, M. D. Dodd, J. Pratt, *Nat. Neurosci.* **6**, 555 (2003).
14. M. Zorzi, K. Priftis, C. Umilta, *Nature* **417**, 138 (2002).
15. F. Galton, *Nature* **21**, 252 (1880).
16. M. Piazza, P. Pinel, S. Dehaene, *Cogn. Neuropsychol.* **23**, 1162 (2006).
17. J. L. Booth, R. S. Siegler, *Dev. Psychol.* **42**, 189 (2006).
18. R. S. Siegler, J. L. Booth, *Child Dev.* **75**, 428 (2004).
19. R. S. Siegler, J. E. Opfer, *Psychol. Sci.* **14**, 237 (2003).
20. R. N. Shepard, D. W. Kilpatrick, J. P. Cunningham, *Cognit. Psychol.* **7**, 82 (1975).
21. S. Dehaene, in *Attention & Performance XXII. Sensori-Motor Foundations of Higher Cognition*, P. Haggard, Y. Rossetti, Eds. (Harvard Univ. Press, Cambridge, MA, 2007), pp. 527–574.
22. P. Pica, C. Lemer, V. Izard, S. Dehaene, *Science* **306**, 499 (2004).
23. S. Dehaene, V. Izard, P. Pica, E. Spelke, *Science* **311**, 381 (2006).
24. Materials and methods are available as supporting material on *Science* Online.
25. V. Izard, S. Dehaene, *Cognition* **106**, 1221 (2008).
26. M. P. van Oeffelen, P. G. Vos, *Percept. Psychophys.* **32**, 163 (1982).
27. A. Nieder, E. K. Miller, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7457 (2004).
28. P. Gordon, *Science* **306**, 496 (2004).
29. W. P. Banks, M. J. Coleman, *Percept. Psychophys.* **29**, 95 (1981).
30. S. Dehaene, J. F. Marques, *Q. J. Exp. Psychol.* **55**, 705 (2002).
31. This work is part of a larger project on the nature of quantification. It is based on psychological experiments and linguistics studies in the Mundurucu territory (Pará, Brazil) under the supervision of P.P., in accordance with the Conselho de Desenvolvimento Científico et Tecnologicico and the Fundacão do Indio (Funai; Processo 2857/04). We thank the Nucleo de Documentação e Pesquisa (Funai), L. Braga, A. Ramos, and C. Romeiro for useful discussion and advice and A. Amor, M. Karu, Y.-h. Liu, R. M. Sullivan, and C. Tawe for help with data collection. Supported by INSERM, the Département des Sciences Humaines et Sociales of CNRS (P.P.), NIH (E.S.), and a McDonnell Foundation centennial fellowship (S.D.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1217/DC1
Methods
SOM Text
Figs. S1 to S10
References

15 February 2008; accepted 29 April 2008
10.1126/science.1156540

Coordination of Early Protective Immunity to Viral Infection by Regulatory T Cells

Jennifer M. Lund, Lianne Hsing, Thuy T. Pham, Alexander Y. Rudensky*

Suppression of immune responses by regulatory T cells (Tregs) is thought to limit late stages of pathogen-specific immunity as a means of minimizing associated tissue damage. We examined a role for Tregs during mucosal herpes simplex virus infection in mice, and observed an accelerated fatal infection with increased viral loads in the mucosa and central nervous system after ablation of Tregs. Although augmented interferon production was detected in the draining lymph nodes (dLNs) in Treg-deprived mice, it was profoundly reduced at the infection site. This was associated with a delay in the arrival of natural killer cells, dendritic cells, and T cells to the site of infection and a sharp increase in proinflammatory chemokine levels in the dLNs. Our results suggest that Tregs facilitate early protective responses to local viral infection by allowing a timely entry of immune cells into infected tissue.

Regulatory T cells (Tregs) expressing the transcription factor Foxp3 play an essential role in controlling immune response–mediated inflammation. Their importance is emphasized by the fact that deficiency in Tregs results in a fatal autoimmune syndrome affecting multiple organs (*1*, *2*). Theoretically, the potent suppressor function of Tregs might present a serious obstacle to establishing robust protective immunity toward pathogens, and recent studies addressing a role for Tregs during infection have suggested several mutually exclusive scenarios. Some studies have suggested that by limiting late immune responses to an infectious agent, Tregs minimize associated tissue damage while at the same time preventing or diminishing pathogen clearance (*3*–*5*). Alternatively, it has been proposed that during viral infection, Tregs lose their suppressor capacity in response to engagement of virus-sensing mechanisms such as Toll-like receptor (TLR) signaling (*6*). Another study suggests that effector T cells responding to infection might become resistant to Treg-mediated suppression as a result of exposure to proinflammatory cytokines and increased costimulatory signals (*7*). Thus, with several scenarios proposed, the role for Tregs during infection remains unclear.

We examined a role for Tregs in mucosal herpes simplex virus (HSV) infection by taking advantage of *Foxp3*^{gfp} and *Foxp3*^{DTR} knock-in mice harboring Treg subsets tagged with green fluorescent protein (GFP) or a human diphtheria toxin receptor (DTR), respectively. This allowed us to track and isolate Tregs and to efficiently ablate these cells upon in vivo DT treatment (*8*, *9*). In these studies we used a well-established model of genital HSV-2 infection via a natural route (*10*). In HSV-2–infected mice, initial viral replication is limited to the vaginal mucosa (*11*), followed by spread into the central nervous system (CNS) upon retrograde transport of virions into the sacral ganglia, resulting in a fatal paralysis. The adaptive immune response to HSV-2 is dominated by virus-specific T helper 1 (T_H1) cells essential for limiting HSV-2 replication (*12*, *13*).

We first examined whether Tregs respond to viral infection by monitoring the dynamics of the GFP-tagged Treg subset relative to "effector" T cells, as defined by the absence of Foxp3, in infected *Foxp3*^{gfp} mice. After genital infection with HSV, the total numbers of both non-Tregs and Tregs drastically increased in the draining lymph nodes (dLNs) and at the site of infection with essentially identical kinetics, and both subsets displayed an increased proportion of cells expressing activation markers (Fig. 1, A to D). Furthermore, using continuous in vivo 5-bromo-2′-deoxyuridine (BrdU) labeling, we observed that CD4$^+$Foxp3$^-$ and Foxp3$^+$ Tregs that had undergone cell division could be found in both the dLNs and at the infection site within 4 days of viral challenge (Fig. 1, E and F). Finally, Tregs isolated from the dLNs of infected mice had a measurably greater potency in suppressing the virus-specific proliferative response of CD4 T cells relative to Tregs isolated from uninfected mice (Fig. 1G). Taken together, these

Department of Immunology and Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA.
*To whom correspondence should be addressed. E-mail: aruden@u.washington.edu

results show that Tregs, rather than remaining refractory, can become activated upon local viral infection.

In light of this activation, we next examined whether Tregs might fulfill a specific role during the immune response to HSV-2. Previous studies have shown enhanced $CD8^+$ T cell responses in mice treated with antibody to CD25 that were infected with HSV-1 via footpad injection, whereas depletion of $CD25^+$ Tregs in a corneal HSV-1 infection model has been reported to increase severity of T cell–mediated tissue lesions (*14*, *15*). On the basis of these studies, we anticipated that Treg depletion during genital HSV infection might enhance the immune response to the virus. However, $Foxp3^{DTR}$ mice succumbed more rapidly to genital HSV infection after DT-induced depletion of $Foxp3^+$ Tregs (Fig. 2A), developing severe lesions and hindlimb paralysis 4 to 5 days earlier than Treg-sufficient mice (Fig. 2B).

The uninfected $Foxp3^{DTR}$ mice did not develop any signs of CNS and genital tract pathology resulting from Treg ablation. Unexpectedly, infected, DT-treated $Foxp3^{DTR}$ mice carried significantly elevated viral loads at the site of infection, relative to $Foxp3^{WT}$ mice, by day 3 after infection (Fig. 2C). High titers of virus were also present in the spinal cord of $Foxp3^{DTR}$ mice 4 days after infection, a time at which $Foxp3^{WT}$ spinal cords remained virus-free (Fig. 2D). Note that viral titers from infected $Foxp3^{DTR}$ mice reconstituted with DTR-negative $Foxp3^+$ Tregs were similar to those detected in $Foxp3^{WT}$ mice (fig. S1); this result indicates that ablation of Tregs was solely responsible for the increased viral titers and that Tregs facilitate protective immunity to HSV infection.

Type I interferons, produced primarily by plasmacytoid dendritic cells (pDCs) in response to TLR9 stimulation, and interferon-γ (IFN-γ), produced by $CD4^+$ T cells and natural killer (NK) cells, are indispensable for the protective immune response to HSV (*12*, *16*). This led us to hypothesize that ablation of Tregs might in some way impair the production of these essential antiviral cytokines. However, we found sharply augmented IFN-γ production by virus-specific CD4 T cells (Fig. 3A) as well as enhanced levels of IFN-α in the dLNs of infected Treg-ablated mice (Fig. 3B). Furthermore, a comparable increase in the number of activated T cells and in the production of IFN-γ, tumor necrosis factor–α, and interleukin-2 (IL-2), potentially produced by "self"-reactive $CD4^+$ T cells, was observed both in uninfected and infected $Foxp3^{DTR}$ mice after Treg ablation (figs. S2 and S3). These results suggest that Tregs limit proinflammatory responses regardless of infection status.

Although interferon levels in the dLNs were increased, we reasoned that local rather than systemic interferon production might restrain the infection. To test this idea more directly, we examined type I and II interferon levels at the site of viral replication, the vaginal mucosa. Indeed, in contrast to augmented IFN-γ production in the dLNs, vaginal IFN-γ as well as IFN-α levels remained low in the absence of Tregs (Fig. 3, C and D). This discrepancy between interferon production in the dLNs and in infected tissue suggested that ablation of Tregs might impair immune cell trafficking to the infected tissue.

To test this, we examined the proportion of effector immune cell subsets including NK cells, $CD4^+$ T cells, pDCs, and the dominant antigen-presenting cells (APCs), DCs, in the dLNs and vaginal mucosa early during infection. On day 2 after infection, we found an increased presence of these cells in the dLNs of infected, Treg-ablated mice. Comparable increases in these cell subsets were found in the LNs of DT-treated uninfected $Foxp3^{DTR}$ mice (Fig. 3E). In contrast, the site of infection showed a marked decrease in the number of these immune cells at 2 days after infection when Tregs were absent (Fig. 3E). Because the peak of IFN-α produced by pDCs and IFNγ produced by NK cells occurs at this early time point, this result was in full agreement with the negligible local interferon production in Treg-depleted mice.

At day 4 after infection, the numbers of NK cells and pDCs were low irrespective of the presence of Tregs. Thus, it appeared that ablation of Tregs impaired the migration of these critical

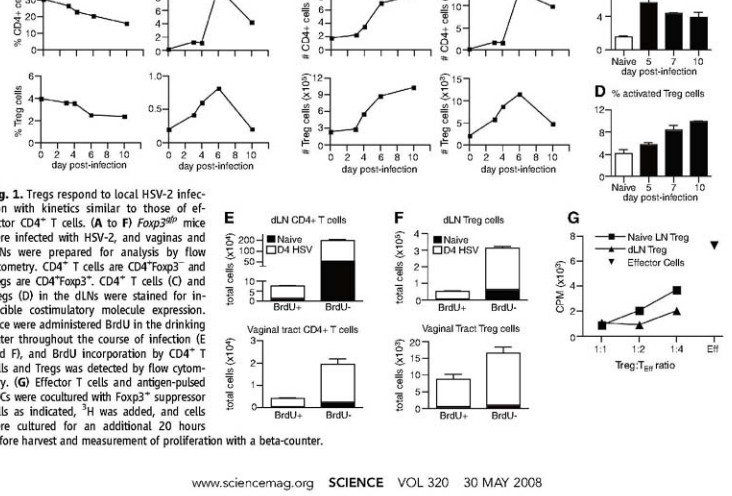


Fig. 1. Tregs respond to local HSV-2 infection with kinetics similar to those of effector $CD4^+$ T cells. (**A** to **F**) $Foxp3^{dtr}$ mice were infected with HSV-2, and vaginas and dLNs were prepared for analysis by flow cytometry. $CD4^+$ T cells are $CD4^+Foxp3^-$ and Tregs are $CD4^+Foxp3^+$. $CD4^+$ T cells (**C**) and Tregs (**D**) in the dLNs were stained for inducible costimulatory molecule expression. Mice were administered BrdU in the drinking water throughout the course of infection (**E** and **F**), and BrdU incorporation by $CD4^+$ T cells and Tregs was detected by flow cytometry. (**G**) Effector T cells and antigen-pulsed APCs were cocultured with $Foxp3^+$ suppressor cells as indicated, ^3H was added, and cells were cultured for an additional 20 hours before harvest and measurement of proliferation with a beta-counter.

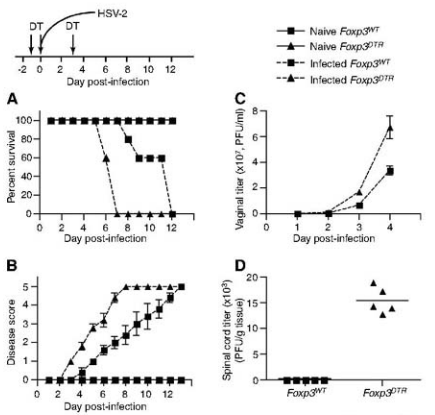


Fig. 2. Tregs are required to prevent early death from HSV-2 infection. *Foxp3*^{WT} or *Foxp3*^{DTR} mice were infected with HSV-2 and treated with DT. Survival (**A**) and disease score (**B**) were monitored daily and vaginal washes were collected daily to assess vaginal viral titer (**C**) by plaque assay on Vero cells. (**D**) Spinal cords were collected 4 days after infection and plaque assays were performed on homogenates.

antiviral effector cells to the site of infection. At this time, however, CD4$^+$ T cell numbers at the site of infection were greater in the absence of Tregs than in their presence, whereas CD11b$^+$ DC numbers were similar in both groups, indicating that arrival of these two cell subsets had been delayed by about 2 days. Together, these results suggest that early during infection, Tregs provide an important means of balancing the trafficking of effector cells between the sites of immune induction and sites of infection.

To directly address this question, we adoptively transferred carboxyfluorescein diacetate succinimidyl ester (CFSE)–labeled cell populations into HSV-infected Treg-sufficient or Treg-deficient hosts. In the absence of Tregs, the proportion of CFSE$^+$ DCs, pDCs, and CD4$^+$ T cells increased in the LNs 24 hours after transfer relative to Treg-intact mice (Fig. 4A). In contrast, there was a striking absence of the CFSE$^+$ NK, DC, and CD4 T cells and a major reduction in pDCs in the vaginal tract of the infected hosts (Fig. 4A). Thus, the absence of Tregs resulted in enhanced entry and retention of effector cells in the dLNs, coupled with a lack of migration to the site of infection. One likely scenario for this would be an alteration in essential chemotactic cues for migrating effector cells. Consistent with this idea, global disruption of cellular trafficking in HSV-infected mice treated with pertussis toxin resulted in an increase in viral titers similar to that

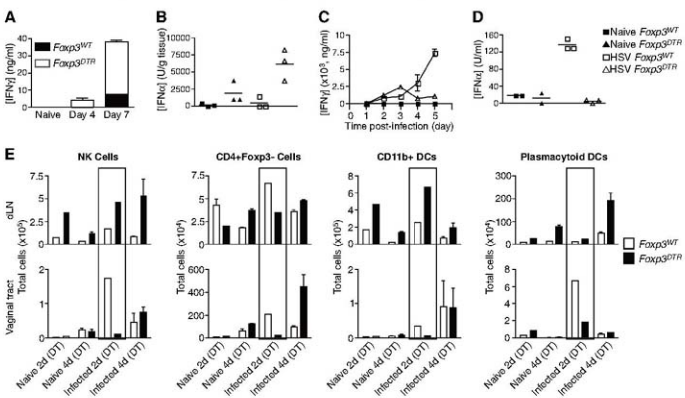


Fig. 3. Tregs are required to mount a protective immune response at the site of infection subsequent to genital HSV-2 infection. *Foxp3*^{WT} or *Foxp3*^{DTR} mice were infected with HSV-2 and treated with DT. (**A**) For ex vivo cultures, CD4$^+$ T cells were isolated from the dLNs of naïve or infected *Foxp3*^{WT} or *Foxp3*^{DTR} mice and cultured for 3 days with irradiated, heat-inactivated HSV-2–pulsed APCs. IFN-γ was detected by enzyme-linked immunosorbent assay (ELISA). (**B**) Mice were infected and treated with DT as in Fig. 2, dLNs were collected 2 days after infection, and extracts were prepared for IFN-α detection by ELISA. (**C** and **D**) IFN-γ present in vaginal washes collected at various times after infection (C) and IFN-α from vaginal washes collected 2 days after infection (D) were measured by ELISA. (**E**) At the indicated times after infection, dLNs and vaginal tracts were subjected to flow cytometric analysis. Naïve mice received DT according to the same schedule as infected mice.

seen upon ablation of Tregs (fig. S4). In parallel, we also observed that levels of chemokines CXCL10, CCL2, CXCL9, and lymphoid chemokine CXCL13 were greatly increased in the dLNs upon Treg ablation (Fig. 4B and fig. S5). Upon infection of Treg-intact mice, CCL21 and CXCL13 were decreased, in agreement with a recent report (fig. S5) (*17*).

The data thus far suggest that by exerting control over inflammatory chemokine cues, Tregs might achieve immune control indirectly by limiting CD4 T cell activation or by directly affecting chemokine-producing cell types. To distinguish between these alternatives, we measured chemokine levels in the LNs of HSV-infected *Foxp3*^{DTR} mice subjected to DT-mediated ablation of Foxp3$^+$ Tregs alone or in combination with antibody-mediated CD4$^+$ T cell depletion. Simultaneous depletion of CD4$^+$ T cells and Tregs prevented the lymphadenopathy characteristic of Treg depletion alone (fig. S6A), yet resulted in a similar increase in inflammatory chemokines in the LNs (fig. S6B). Thus, it is likely that Tregs suppress inflammatory chemokine production in the LNs directly. To identify cellular sources of LN chemokines, we measured chemokine mRNA levels in fluorescence-activated cell sorter–purified cell subsets from *Foxp3*^{WT} or *Foxp3*^{DTR} mice. We found that CXCL9 and CCL2 mRNA levels strongly increased in DCs and that CXCL10 increased in both DCs and CD45$^+$ stromal cells in the absence of Tregs (Fig. 4C). Additionally, CXCL10, CCL2, and CCL5 were elevated in LN NK cells, whereas pDCs did not show differences in the absence of Tregs (fig. S7). In contrast to increased chemokine production, CD4$^+$ T cell levels of sphingosine 1-phosphate receptor–1 ($S1P_1$) were unchanged in the absence of Tregs (fig. S8), excluding potential modulation of $S1P_1$ as a cause of delayed egress of CD4$^+$ T cells from the dLNs. Examination of chemokines in the vaginal tract of infected Treg-deficient mice revealed increased expression of CXCL10 and CCL2 relative to Treg-sufficient mice, albeit at levels considerably lower than those found in the dLNs (Fig. 4B). Unlike CXCL10 and CCL2, CCL5 was largely lacking in the vaginal tissue of Treg-depleted mice, in contrast to a markedly augmented CCL5 expression found in Treg-replete HSV infected mice (Fig. 4B). This could account for a failure of recruitment to the infected tissue of CCR5-expressing cells, including pDCs. In agreement with this notion was a previous finding that protection against lethal HSV-2 challenge afforded by administration of CpG was coincident with a wave of CCL5 as well as IFN-γ, IL-12, and IL-18 production in the genital mucosa (*18*).

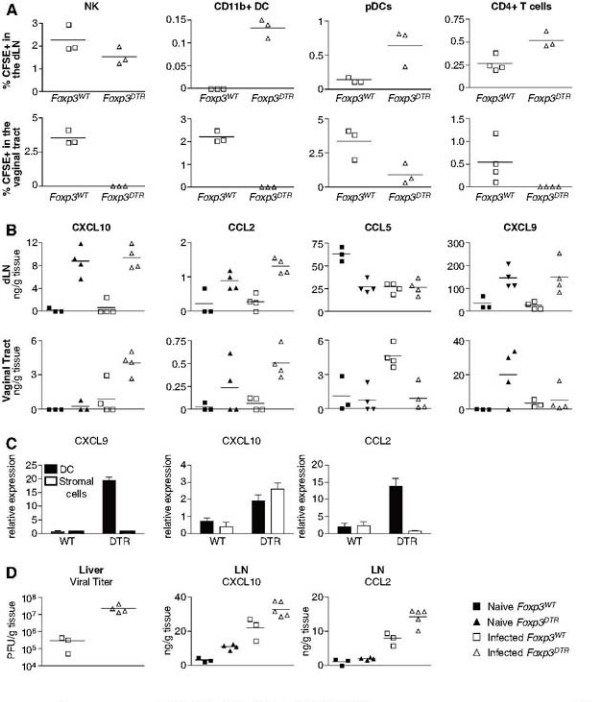


Fig. 4. Tregs modulate the chemokine gradient that controls proper effector cell homing to the dLNs and site of infection. **(A)** NK cells, CD11b$^+$ DCs, and pDCs from naïve donors or CD4$^+$ T cells isolated from the dLNs of *Foxp3*^{gfp} donor mice infected with HSV-2 for 6 days were labeled with CFSE and injected into recipient mice. Recipients were *Foxp3*^{WT} or *Foxp3*^{DTR} mice treated with DT and infected with HSV-2, and mice received CFSE-labeled cells 2 days after infection. Twenty-four hours after transfer, dLNs and vaginal tracts from recipient mice were examined for percentage of the indicated population that was CFSE$^+$ by flow cytometry. **(B)** *Foxp3*^{WT} or *Foxp3*^{DTR} mice were infected with HSV-2 and treated with DT. Two days after infection, dLNs and vaginal tract extracts were prepared for chemokine detection by Luminex bead assay. **(C)** Chemokine mRNA amounts in DCs and CD45$^-$ stromal cells were measured by real-time PCR and are shown as the mean of three independent experiments relative to hypoxanthine-guanine phosphoribosyl transferase expression. **(D)** Foxp3^{DTR} mice were treated with DT and infected with LCMV. Five days after infection, livers were collected and assayed for viral titer, and LNs were analyzed for chemokines as in (B).

It is likely that Treg-dependent suppression of proinflammatory chemokines in the lymph nodes plays an important role in restraining viral replication in nonlymphoid tissues other than mucosal epithelium. To test this possibility, we examined liver viral titers and LN chemokine levels during lymphocytic choriomeningitis virus (LCMV) Armstrong infection of Treg-replete or Treg-depleted mice. In agreement with the HSV data, in the absence of Tregs we observed an increase in LCMV liver titers on day 5 after infection, corresponding to the peak viral load (*19*), as well as increases in CXCL10 and CCL2 in the LNs (Fig. 4D). Thus, the observation of increased pathogen replication and altered chemokine milieu in secondary lymphoid organs in the absence of Tregs in mucosal HSV infection can be extended to a systemic LCMV infection.

Our results show that Tregs facilitate early immune responses to local viral infection at least in part by orchestrating a timely homing of immune effector cells to the site of infection. This unexpected finding of an immune response–promoting role for Tregs seems to be applicable to other infections resulting in pathogen replication in nonlymphoid tissues.

References and Notes

1. J. D. Fontenot, A. Y. Rudensky, *Nat. Immunol.* **6**, 331 (2005).
2. S. Sakaguchi *et al.*, *Immunol. Rev.* **212**, 8 (2006).
3. Y. Belkaid, C. A. Piccirillo, S. Mendez, E. M. Shevach, D. L. Sacks, *Nature* **420**, 502 (2002).
4. Y. Belkaid, B. T. Rouse, *Nat. Immunol.* **6**, 353 (2005).
5. B. T. Rouse, P. P. Sarangi, S. Suvas, *Immunol. Rev.* **212**, 272 (2006).
6. G. Peng *et al.*, *Science* **309**, 1380 (2005).
7. C. Pasare, R. Medzhitov, *Science* **299**, 1033 (2003).
8. J. M. Kim, J. P. Rasmussen, A. Y. Rudensky, *Nat. Immunol.* **8**, 191 (2007).
9. J. D. Fontenot *et al.*, *Immunity* **22**, 329 (2005).
10. See supporting material on *Science* Online.
11. X. Zhao *et al.*, *J. Exp. Med.* **197**, 153 (2003).
12. G. N. Milligan, D. I. Bernstein, *Virology* **212**, 481 (1995).
13. M. B. Parr, E. L. Parr, *Virology* **258**, 282 (1999).
14. S. Suvas, A. K. Azkur, B. S. Kim, U. Kumaraguru, B. T. Rouse, *J. Immunol.* **172**, 4123 (2004).
15. S. Suvas, U. Kumaraguru, C. D. Pack, S. Lee, B. T. Rouse, *J. Exp. Med.* **198**, 889 (2003).
16. J. M. Lund, M. M. Linehan, N. Iijima, A. Iwasaki, *J. Immunol.* **177**, 7510 (2006).
17. S. N. Mueller *et al.*, *Science* **317**, 670 (2007).
18. A. M. Harandi, K. Eriksson, J. Holmgren, *J. Virol.* **77**, 953 (2003).
19. E. J. Wherry, J. N. Blattman, K. Murali-Krishna, R. van der Most, R. Ahmed, *J. Virol.* **77**, 4911 (2003).
20. We thank K. Forbush, L. Karpik, and T. Chu for assistance with animal care and breeding; A. Iwasaki for viral stocks; and J. Cyster for $5I^{+}_{2}$ antibody and discussions. Supported by the Howard Hughes Medical Institute (A.Y.R.) and the Irvington Institute Fellowship Program of the Cancer Research Institute (J.M.L.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1155209/DC1
Materials and Methods
Figs. S1 to S8
References

14 January 2008; accepted 11 April 2008
Published online 24 April 2008;
10.1126/science.1155209
Include this information when citing this paper.

MeCP2, a Key Contributor to Neurological Disease, Activates and Represses Transcription

Maria Chahrour,[1] Sung Yun Jung,[2] Chad Shaw,[1] Xiaobo Zhou,[3] Stephen T. C. Wong,[3] Jun Qin,[2,4] Huda Y. Zoghbi[1,5,6,7,8]*

Mutations in the gene encoding the transcriptional repressor methyl-CpG binding protein 2 (MeCP2) cause the neurodevelopmental disorder Rett syndrome. Loss of function as well as increased dosage of the *MECP2* gene cause a host of neuropsychiatric disorders. To explore the molecular mechanism(s) underlying these disorders, we examined gene expression patterns in the hypothalamus of mice that either lack or overexpress MeCP2. In both models, MeCP2 dysfunction induced changes in the expression levels of thousands of genes, but unexpectedly the majority of genes (~85%) appeared to be activated by MeCP2. We selected six genes and confirmed that MeCP2 binds to their promoters. Furthermore, we showed that MeCP2 associates with the transcriptional activator CREB1 at the promoter of an activated target but not a repressed target. These studies suggest that MeCP2 regulates the expression of a wide range of genes in the hypothalamus and that it can function as both an activator and a repressor of transcription.

Rett syndrome (RTT, MIM 312750) is a progressive neurodevelopmental disorder caused by mutations in the X-linked gene encoding methyl-CpG binding protein 2 (MeCP2) (*1*). RTT patients appear to develop normally up to 1 year of age, after which they lose any acquired speech and replace purposeful hand use with stereotypies. *MECP2* mutations also result in a host of neurobehavioral abnormalities, ranging from mild learning disabilities to autism, X-linked mental retardation, and infantile encephalopathy. Interestingly, an increase in *MECP2* dosage is equally detrimental to the nervous system: Duplications spanning the *MECP2* locus cause features overlapping with those of RTT and related neurological disorders (*2*).

MeCP2 aberrations in mice also result in a neurological phenotype. *Mecp2*-null mice appear normal until 6 weeks of age when they develop severe neurological dysfunction resulting in premature death by 12 weeks of age (*3*). Mice that overexpress *MECP2* under the control of its endogenous promoter (*MECP2*-Tg) display neurological features similar to the human *MECP2* duplication syndrome (*4*). MeCP2 is believed to function as a transcriptional repressor by binding to methylated CpG dinucleotides and recruiting corepressors and chromatin remodeling proteins (*5*, *6*). Loss of MeCP2 and doubling of *MECP2* dosage in mice have opposing effects on excitatory synapse numbers (*7*), which suggests that similar target genes may be affected in the two mouse models, but in opposite directions. We reasoned that analysis of gene expression profiles in *MECP2*-Tg and *Mecp2*-null mice might shed light on the molecular mechanism underlying the *MECP2* duplication syndrome, reveal whether this disorder shares downstream molecular pathways with RTT, and possibly distinguish primary versus secondary gene targets of MeCP2. If the *MECP2* duplication syndrome is caused by loss or partial loss of normal MeCP2 function (due to sequestration of MeCP2 binding partners because of excess MeCP2), then the gene expression profiles are expected to go in the same direction. However, if the mechanism is due to gain of function, owing to increased MeCP2 activity, then the same primary target genes would be misregulated in opposite directions in *MECP2*-Tg versus *Mecp2*-null mice, and the secondary target genes would be specific to each mouse model.

Previous transcriptional profiling studies comparing brain tissue from *Mecp2*-null and wild-type (WT) mice revealed only subtle differences in gene expression (*8*). We hypothesized that critical disease-related alterations in gene

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. [2]Center for Molecular Discovery, Departments of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA. [3]Center for Bioinformatics, The Methodist Hospital Research Institute and Weill Cornell College of Medicine, Houston, TX 77030, USA. [4]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA. [5]Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030, USA. [6]Departments of Pediatrics and Neurology, Baylor College of Medicine, Houston, TX 77030, USA. [7]Programs in Cell and Molecular Biology and Developmental Biology, Baylor College of Medicine, Houston, TX 77030, USA. [8]Howard Hughes Medical Institute, Baylor College of Medicine, Texas Children's Hospital, Houston, TX 77030, USA.

*To whom correspondence should be addressed. E-mail: hzoghbi@bcm.edu

expression might occur only in specific brain regions or neurons and, therefore, would be masked in analyses of whole brain tissue. To circumvent this problem, we investigated the hypothalamus because a number of RTT phenotypes (anxiety, growth deceleration, sleep-wake rhythm disturbance, and autonomic abnormalities) could be attributed to hypothalamic dysfunction (*9*). We performed microarray analysis using hypothalamic RNA from four *Mecp2*-null males, four *MECP2*-Tg males, and their respective WT littermates at 6 weeks of age using the Affymetrix Mouse Exon 1.0 ST microarray (*10*, *11*). Given that MeCP2 functions as a transcriptional repressor in vitro, we were surprised to find that 2184 out of the 2582 genes misregulated in both mouse models (~85%) were up-regulated in the transgenic hypothalami and down-regulated in the *Mecp2*-null hypothalami, suggesting that many of these genes are likely activated by MeCP2. In contrast, only 377 genes were down-regulated in the transgenic hypothalami and up-regulated in the *Mecp2*-null hypothalami, suggesting that these are normally repressed by MeCP2 [false discovery rate (FDR)–adjusted $P < 0.05$] (Fig. 1A) (tables S1 and S2). The magnitude of the changes in expression levels for both activated and repressed genes was moderate (less than a factor of 5). Moreover, only 9 genes were up-regulated in both mouse models (table S3), and 12 genes were down-regulated in both models (table S4) (FDR-adjusted $P < 0.05$). We identified 1187 genes whose expression was altered only in the transgenic hypothalami and 369 genes whose expression was altered only in the null hypothalami (tables S5 and S6) (FDR-adjusted $P < 0.05$). Because these changes are specific to each mouse model, they likely represent secondary changes due either to altered expression of primary gene targets of MeCP2 or to disease processes. Interestingly, gene ontology analysis (fig. S1) revealed that the genes activated and repressed by MeCP2 fall into distinct categories in terms of biological function. For example, neuropeptides were enriched in the list of activated MeCP2 targets, whereas olfactory receptors were enriched in the list of repressed targets. Overall, the genes encoding G protein–coupled receptors (GPCRs) were the most significantly affected by MeCP2 levels; this is expected because most hypothalamic neuropeptide receptors are GPCRs. In addition, we found that the 5-kb upstream promoter regions of genes that are up-regulated by MeCP2 were significantly enriched in CpG islands, whereas the down-regulated genes were not (Fig. 1B).

To validate the gene expression changes, we performed quantitative real-time reverse transcription polymerase chain reaction (QPCR) using hypothalamic RNA samples from an independent set of *Mecp2*-null males, *MECP2*-Tg males, and their respective WT littermates. We confirmed alterations in the expression of 66 genes

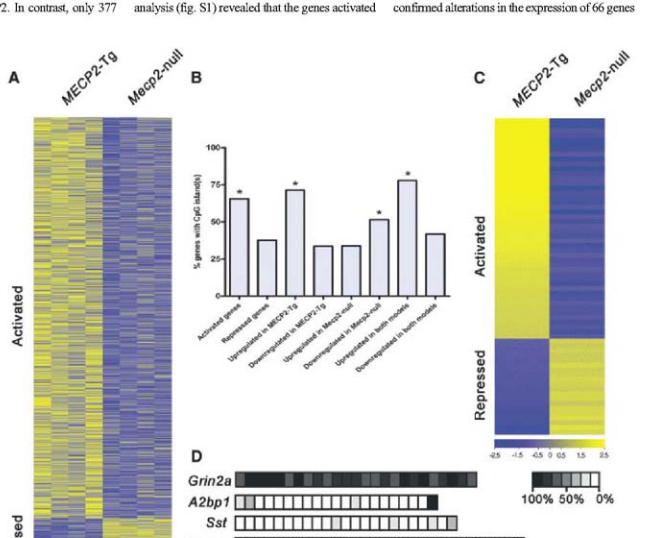


Fig. 1. Significant gene expression changes in hypothalami of MeCP2 mouse models. **(A)** Heatmap showing hypothalamic gene expression profiles in *MECP2*-Tg and *Mecp2*-null mice. Yellow and blue colors indicate increased and decreased expression, respectively, relative to WT. Each column represents one RNA sample from each genotype, and each row represents one gene. Expression levels are depicted according to the color scale at the bottom. A total of 2184 genes are activated in the presence of MeCP2, and 377 genes are repressed (FDR-adjusted $P < 0.05$). **(B)** Promoter regions of genes up-regulated by MeCP2 contain significantly more CpG islands compared with those of genes down-regulated by MeCP2 (*$P \leq 0.006$). **(C)** Validation of expression changes for 66 genes by QPCR analysis. Gene expression levels from the microarray were validated in four *MECP2*-Tg males and four *Mecp2*-null males. Data are plotted as relative up-regulation (yellow) or down-regulation (blue) over WT ($P < 0.05$, *t* test). Each row represents a single gene, and each column represents data for four samples from each genotype. Levels are depicted according to the color scale at the bottom. The complete list of validated genes is in table S7. **(D)** Bisulfite sequencing revealed that CpG sites in the promoters of the repressed target genes *Grin2a* and *A2bp1* are heavily methylated, whereas those of activated targets *Sst*, *Gprin1*, *Gamt*, and *E2F1* are not. CpG sites are depicted as squares, with the percent methylation presented according to the scale. The data were generated from three independent animals, and for each promoter ~10 clones were sequenced.

representative of the most significantly altered gene families; of these, 46 were up-regulated by MeCP2 and 20 were down-regulated ($P < 0.05$) (Fig. 1C). The complete list of validated genes is available in supporting online material (*12*) (table S7). To check the CpG island methylation status of some of the validated genes, we performed bisulfite sequencing of hypothalamic DNA from WT males. We found that CpG sites in promoters of activated MeCP2 gene targets were not heavily methylated compared with those of repressed targets (Fig. 1D), consistent with the fact that in general CpG islands are unmethylated. To determine whether both activated and repressed genes are altered because they are direct MeCP2 targets, we performed chromatin immunoprecipitation (ChIP) with antibody to MeCP2, using hypothalamic tissue from *Mecp2*-null males, *MECP2*-Tg males, and their respective WT littermates. We used QPCR to assess MeCP2 binding to the promoter regions of six candidate target genes that were either activated or repressed by MeCP2. This in vivo analysis revealed that MeCP2 bound to the promoter regions of genes encoding somatostatin (*Sst*), opioid receptor kappa 1 (*Oprk1*), guanidinoacetate methyltransferase (*Gamt*), G protein–regulated inducer of neurite outgrowth 1 (*Gprin1*) (all activated by MeCP2), myocyte enhancer factor 2C (*Mef2c*), and ataxin 2 binding protein 1 (*A2bp1*) (both repressed by MeCP2).Consistent with the microarray data, MeCP2 binding was significantly enhanced in the *MECP2*-Tg samples compared with WT ($P \leq 0.03$) (Fig. 2). To confirm that the binding of MeCP2 to its candidate primary targets (genes inversely regulated in *MECP2*-Tg and *Mecp2*-null mice) is specific, we tested MeCP2 binding to the promoter region of the gene encoding cyclin-dependent kinase-like 4 (*Cdkl4*), one of the genes down-regulated in both mouse models (table S4). We chose this gene because the GC content of its promoter region is similar to that of the confirmed target genes. We did not detect MeCP2 at the promoter of *Cdkl4*, as would be predicted because of its similar regulation in both mouse models (fig. S2).

Given that gain of MeCP2 causes significantly more activation rather than repression, whereas the loss of MeCP2 causes the reverse, we explored whether MeCP2 associates with coactivators in vivo. We immunopurified MeCP2 using an antibody to its C terminus (*13*) from brain extracts of WT mice and, as a negative control, from *Mecp2*-null mice. Mass spectrometry analysis identified CREB1, a major transcriptional activator, as one of nine proteins that copurified with MeCP2 (Fig. 3A and table S8). We confirmed the association of MeCP2 and CREB1 by performing the reciprocal immunoprecipitation with anti-CREB1 from Neuro2a cells overexpressing MeCP2 (Fig. 3B). The data suggest that the relationship between MeCP2 and CREB1 is non-stoichiometric and that not all of CREB1 is associated with MeCP2 in complexes from brain. To determine whether MeCP2 and CREB1 co-occupy any of the promoters of activated genes in vivo, we performed sequential ChIP (seqChIP) analysis using antibody to CREB1 on hypothalamic chromatin that was initially immunoprecipitated with antibody to MeCP2. QPCR analysis showed that the promoter of an activated target gene (*Sst*) was enriched for CREB1 binding, consistent with MeCP2 and CREB1 being simultaneously associated with the promoter. In contrast, the promoter of a gene repressed by MeCP2 (*Mef2c*) was not enriched for CREB1 binding (Fig. 3C). To determine whether there is functional synergy between MeCP2 and CREB1, we transfected Neuro2a cells with luciferase reporter constructs driven by the promoter of *Sst* (activated MeCP2 target) or the promoter of *A2bp1* (repressed MeCP2 target). MeCP2 and CREB1 cotransfection resulted in significant enhancement of the *Sst* reporter activity but not the *A2bp1* reporter (Fig. 3D). Interestingly, *Creb1* was one of the activated MeCP2 targets that we validated by QPCR (Fig. 1C), and using ChIP analysis we found that in vivo MeCP2 binds to the promoter region of *Creb1*, with significantly enhanced binding in *MECP2*-Tg samples compared with WT ($P < 0.05$) (Fig. 4A). We confirmed this result and assessed MeCP2 binding to neighboring genomic regions using a custom array with probes to the *Creb1* promoter (Fig. 4B). In addition, Sst and CREB1 protein levels were increased in *MECP2*-Tg hypothalami compared with WT, indicating that MeCP2 indeed enhances expression of *Sst* and *Creb1* (Fig. 4C).

In contrast to previous studies that failed to detect many transcriptional changes due to MeCP2 dysfunction, we found that interrogating a discrete brain region, the hypothalamus, is very effective for uncovering MeCP2 target genes. Consistent with the role of MeCP2 as an activity-dependent regulator of gene expression, it would be predicted that MeCP2 plays a critical role in neurons that must constantly respond to new physiologic states. It is of interest that the list of putative targets includes genes that cause one or more RTT phenotypes when disrupted. One MeCP2 repression target is the gene encoding ataxin 2 binding protein 1 (*A2bp1*), which regulates splicing of neuronal genes. Disruption of *A2BP1* has been identified in patients with mental retardation and

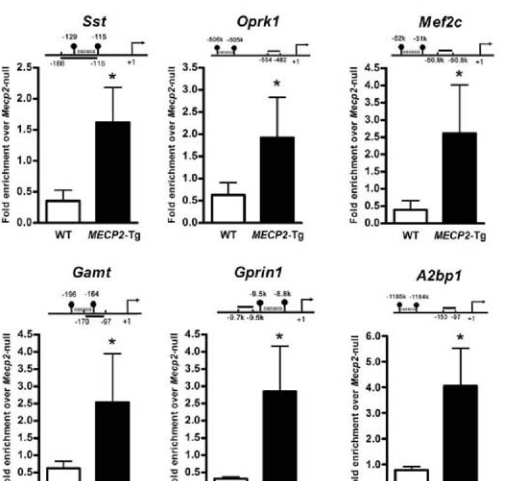


Fig. 2. MeCP2 binds to the promoter region of six target genes. ChIP with antibody to MeCP2 shows that MeCP2 binds to the promoter regions of activated targets *Sst*, *Oprk1*, *Gamt*, and *Gprin1*, and repressed targets *Mef2c* and *A2bp1*. The line indicates the location of the probe and primers [relative to the translation start site (+1), in base pairs or in kb (k)] and the closest CpG islands in the promoter region are depicted. QPCR values were normalized to the input and plotted as relative enrichment over *Mecp2*-null ($N \geq 3$, $^*P \leq 0.03$, *t* test).

epilepsy (*14*), and a recent study associated *A2BP1* disruption with autism susceptibility (*15*). *Gamt*, a target activated by MeCP2, encodes the enzyme guanidinoacetate methyltransferase, involved in creatine biosynthesis. Interestingly, patients with GAMT deficiency suffer severe mental retardation, absent or limited speech development, seizures, and hypotonia (*16*–*18*).

Beyond demonstrating that altered transcriptional regulation occurs in mouse models of *MECP2* disorders, a key finding from our study is that MeCP2 acts as both an activator and a repressor (Fig. 4D) and that, at least in the hypothalamus, it is more of an activator. Several lines of evidence support our conclusions. First, the strongest evidence comes from the genetic data, whereby all the common transcriptional changes occur in opposite directions in *Mecp2*-null versus *MECP2*-Tg mice (except for 21 genes) (tables S3 and S4). Second, we detect MeCP2 binding at the promoters of genes that are altered in opposite directions in the two models (*Sst*, *Oprk1*, *Gamt*, *Gprin1*, *Mef2c*, and *A2bp1*) but not at the promoter of a gene down-regulated in both mouse models, leading us to propose that these are direct MeCP2 targets. Third, the fact that MeCP2 binding to these targets is increased in the *MECP2*-Tg mice compared with WT, argues against the possibility that excess MeCP2 is titrating co-repressors away from promoters and supports a model in which MeCP2 directly recruits activating factors to promoters. Finally, the binding of MeCP2 to promoters of activated genes and the association of MeCP2 with the transcription factor CREB1 at an activated promoter (but not at a repressed promoter) supports a role for

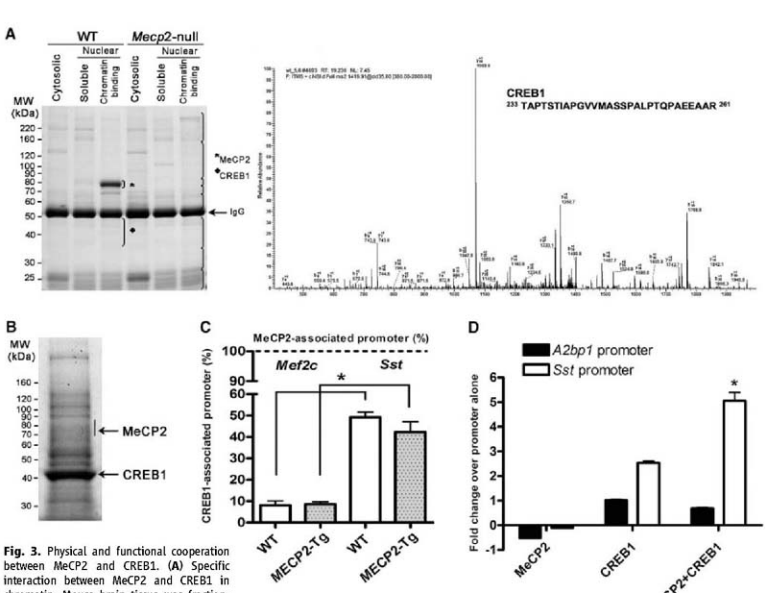


Fig. 3. Physical and functional cooperation between MeCP2 and CREB1. **(A)** Specific interaction between MeCP2 and CREB1 in chromatin. Mouse brain tissue was fractionated into cytosolic, nuclear, and chromatin fractions for protein extraction from WT and *Mecp2*-null mice. Protein extracts were immunoprecipitated with an antibody to MeCP2 and resolved by SDS–polyacrylamide gel electrophoresis (SDS-PAGE). Each lane was divided into eight sections (brackets) and analyzed by mass spectrometry. Proteins identified in the WT but not in the knockout sample are considered specific MeCP2 interacting proteins. MeCP2 was identified only in the chromatin fraction from WT mice (star). A representative nano-high-performance liquid chromatography tandem mass spectrometry spectrum is shown that identifies CREB1 as one of the specific MeCP2 binding proteins in the 35 to 50 kD region in SDS-PAGE, indicated by diamond. The complete list of interacting proteins is available in table S8. **(B)** MeCP2 was detected after immunoprecipitation with antibody to CREB1 from Neuro2a cells, using mass spectrometry analysis of the indicated band. **(C)** MeCP2 associates with CREB1 at the promoter of an activated but not a repressed target. SeqChIP analysis detects co-occupancy of MeCP2 and CREB1 at the promoter of the activated MeCP2 target *Sst* but not at the promoter of the repressed target *Mef2c*. The primary ChIP was performed with an antibody to MeCP2, and the secondary ChIP was done with an antibody to CREB1. QPCR values were normalized to the input and plotted as percent of secondary over primary ChIP [$N = 3$, $^*P < 0.001$, two-way analysis of variance (ANOVA)]. **(D)** Functional synergy between MeCP2 and CREB1 at the promoter of an activated target. Luciferase assay in Neuro2a cells reveals synergistic activation at the promoter of the activated MeCP2 target *Sst* but not the repressed target *A2bp1* ($N = 3$, $^*P < 0.002$, two-way ANOVA).

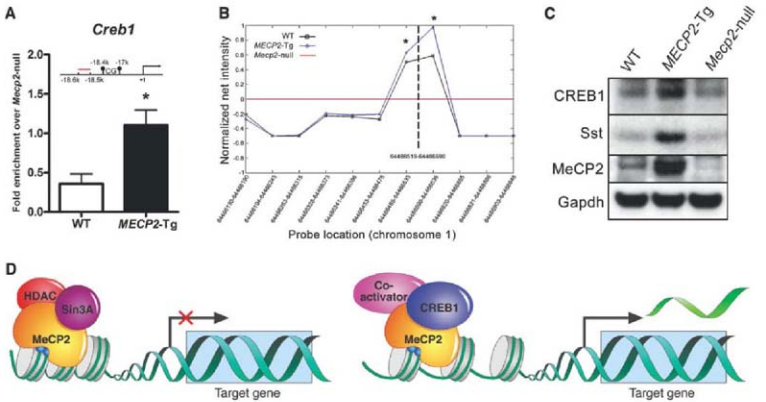


Fig. 4. *Creb1* is a direct MeCP2 target. **(A)** ChIP analysis revealed binding of MeCP2 to the promoter region of *Creb1*. QPCR values were normalized to the input and plotted as relative enrichment over *Mecp2*-null ($N = 4$, $^*P < 0.05$, *t* test). **(B)** ChIP analysis using a custom array with probes to the *Creb1* promoter. Genomic locations of the array probes are indicated on the *x* axis, and the dashed vertical line indicates location of the QPCR probe. The red line represents the normalized net signal from the *Mecp2*-null samples. For the WT and *MECP2*-Tg samples, normalized net intensities from each probe are plotted and are significantly different from the *Mecp2*-null near the predicted MeCP2 binding site ($^*P < 0.02$). **(C)** CREB1 and Sst protein levels are increased in *MECP2*-Tg hypothalami compared with WT. Western blot is representative of three animals from each genotype. **(D)** MeCP2 can function as a transcriptional activator and repressor in the hypothalamus (CpG sites are depicted as blue circles irrespective of their methylation status).

MeCP2 as an activator in certain contexts. Although it is possible that the gene expression changes shared by both mouse models are secondary to the physiological properties of hypothalamic neurons, we believe this is unlikely to explain expression changes in all 2582 genes, especially because many are not activity-regulated. Although we cannot exclude the possibility that some of the inversely altered genes we found in the two mouse models are due to MeCP2 repressing a transcriptional repressor (resulting in activation of the targets of such a repressor), this also seems unlikely, because we did not find a repressor in the list of repressed MeCP2 targets. On the other hand, we found that MeCP2 directly activates the major transcriptional activator CREB1 (Fig. 1C, Fig. 4, A to C, and table S7). A recent study showed that the CREB-induced microRNA, miR132, represses MeCP2 translation (*19*). This finding, along with our data that MeCP2 activates *Creb1*, proposes a negative regulatory loop: MeCP2 activating *Creb1* results in an increase in CREB1 levels, which in turn induces miR132, leading to decreased MeCP2 levels. In addition, ChIP-chip analysis using SH-SY5Y cells showed that MeCP2 associates more frequently with promoters that are also associated with RNA polymerase II (*20*), further supporting that the function of MeCP2 extends beyond transcriptional repression.

Our findings resolve inconsistencies in the literature regarding MeCP2 and *Bdnf*. Chen *et al.* and Martinowich *et al.* demonstrated that MeCP2 binds the *Bdnf* promoter in neuronal cultures (*21*, *22*). However, the in vivo data showing that Bdnf is down-regulated both at the RNA and protein levels in *Mecp2*-null mice (*23*) were confusing in light of the proposed role of MeCP2 as a repressor. Our finding that *Bdnf* is up-regulated in *MECP2*-Tg and down-regulated in *Mecp2*-null animals (Fig. 1C and table S7), and the ample evidence that *Bdnf* is a direct target of MeCP2 (*21*, *22*), is consistent with the role of MeCP2 as an activator on the *Bdnf* promoter and reconciles the existing data on MeCP2 and *Bdnf*.

Finally, our data provide insight into the molecular mechanism underlying *MECP2* disorders. The transcriptional changes suggest that the duplication phenotype is due to MeCP2 gain of function (hypermorph), rather than loss of function, and that at the molecular level, RTT is due mostly to loss of transcriptional activation rather than derepression. In terms of clinical relevance, our results suggest that patients with RTT will have to be treated differently than patients with *MECP2* duplications. The finding that MeCP2 regulates a large number of genes, at least in the hypothalamus, suggests a need for therapeutic strategies that focus on restoring neuronal function rather than restoring the activity of individual gene products affected by MeCP2 dysfunction. It might prove challenging to restore the level of each of these genes at the same time; thus, an alternative approach will be to identify proteins or pathways that suppress MeCP2 dysfunction phenotypes, or bypass MeCP2, and restore neuronal homeostasis.

References and Notes

1. R. E. Amir *et al.*, *Nat. Genet.* **23**, 185 (1999).
2. P. Moretti, H. Y. Zoghbi, *Curr. Opin. Genet. Dev.* **16**, 276 (2006).
3. J. Guy, B. Hendrich, M. Holmes, J. E. Martin, A. Bird, *Nat. Genet.* **27**, 322 (2001).
4. A. L. Collins *et al.*, *Hum. Mol. Genet.* **13**, 2679 (2004).
5. X. Nan *et al.*, *Nature* **393**, 386 (1998).
6. P. L. Jones *et al.*, *Nat. Genet.* **19**, 187 (1998).
7. H. Chao, H. Y. Zoghbi, C. Rosenmund, *Neuron* **56**, 58 (2007).
8. M. Tudor, S. Akbarian, R. Z. Chen, R. Jaenisch, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15536 (2002).
9. F. B. Axelrod, G. G. Chelimsky, D. E. Weese-Mayer, *Pediatrics* **118**, 309 (2006).
10. P. J. Gardina *et al.*, *BMC Genomics* **7**, 325 (2006).
11. K. Srinivasan *et al.*, *Methods* **37**, 345 (2005).
12. Materials and methods are available as supporting material on *Science* Online.
13. Z. Zhou *et al.*, *Neuron* **52**, 255 (2006).
14. K. Bhalla *et al.*, *J. Hum. Genet.* **49**, 308 (2004).
15. C. L. Martin *et al.*, *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **144**, 869 (2007).
16. H. Caldeira Araújo *et al.*, *Am. J. Med. Genet. A* **133**, 122 (2005).

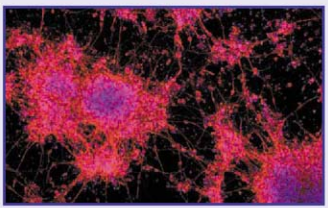
17. K. T. Verbruggen *et al.*, *Eur. J. Pediatr.* **166**, 921 (2007).
18. J. Sykut-Cegielska, W. Gradowska, S. Mercimek-Mahmutoglu, S. Stockler-Ipsiroglu, *Acta Biochim. Pol.* **51**, 875 (2004).
19. M. E. Klein *et al.*, *Nat. Neurosci.* **10**, 1513 (2007).
20. D. H. Yasui *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19416 (2007).
21. W. G. Chen *et al.*, *Science* **302**, 885 (2003).
22. K. Martinowich *et al.*, *Science* **302**, 890 (2003).
23. Q. Chang, G. Khare, V. Dani, S. Nelson, R. Jaenisch, *Neuron* **49**, 341 (2006).
24. We thank members of the Zoghbi laboratory for their thoughtful comments, Z. Yin for technical assistance, M. Shinawi and A. Tawil for technical advice, Y. Klisch for the art work, and J. Lim, J. Neul, and M. Greenberg for reagents. The Baylor College of Medicine Microarray Core Facility performed the microarray experiments. This work was funded by NIH/National Institute of Neurological Disorders and Stroke grant NS057819 (H.Z.), National Institute of Child Health and Human Development Mental Retardation and Developmental Disabilities Research Center HD024064 (H.Z.), the International Rett Syndrome Foundation, and the Simons Foundation. H.Z. is a Howard Hughes Medical Institute Investigator. The microarray data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) (www.ncbi.nlm.nih.gov/geo) and are accessible through GEO Series accession number GSE11150.

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5880/1224/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S8
References

21 November 2007; accepted 3 April 2008
10.1126/science.1153252

New Products



Stem Cell Culture

HyStem hydrogel is a synthetic matrix that can be used in a xeno-free cultivation system for the expansion of human embryonic stem cells, $CD34^+$ stem cells, and hepatic stem/progenitor cells. HyStem is suitable for in vitro and in vivo research and is transparent, enabling live cell imaging at any point during culture. Once mixed, the hydrogel forms in less than 20 minutes at physiological pH and temperature, with rigidity dependent on cross-linker quantity.

Glycosan BioSystems
For information 801-583-8212
www.glycosan.com

Terahertz Spectrometer

The Spectra 3000 is a terahertz spectrometer capable of performing both transmission and attenuated total reflection (ATR) measurements. The proprietary platform makes use of the terahertz region of the spectrum between light and radio waves and offers the advantage of being noninvasive and nondestructive. It provides rapid and accurate chemical and structural identification of materials normally invisible via conventional techniques. Terahertz radiation interrogates the fundamental vibrations between molecules, providing intermolecular information previously difficult to obtain. Applications of the new spectrometer include providing information on crystalline structure, investigating and identifying polymorphs, explosives fingerprinting, airborne agent and gas-phase detection, and investigation of pharmaceutical tablet constituents.

TeraView
For information +44 (0) 1223 435380
www.teraview.com

Cytotoxicity Assay

The CytoTox-Glo Cytotoxicity Assay provides a sensitive, simple, luminescent method for detecting small changes in cell viability. The assay measures a distinct "dead-cell" protease biomarker released from cells that have lost membrane activity, a key indicator of cytotoxicity. The assay makes use of a peptide substrate that releases luciferin after cleavage by the "dead-cell" protease. The luciferase/luciferin reaction produces light of intensity that correlates with dead-cell number.

Promega
For information 608-274-4330
www.promega.com

Fast Spectrometer

The USB2000+ Miniature Fiber Optics Spectrometer offers the ability to monitor high-speed chemical and biochemical reactions. It features a charge-coupled device (CCD) array detector with a high-speed USB 2.0 port interface that can capture and store a full spectrum into memory at 1,000 scans per second. The spectrometer measures optical properties within a wavelength band of 200–1,100 nm. It can be equipped with a fixed grating that disperses the light to the 2,048-element CCD array detector, producing results at a resolution of 0.35-nm with full-width-at-half-maximum sharpness. The USB2000+ can be custom configured with a choice of 14 gratings, six slits, and hundreds of fiber optic accessories for specific applications.

Ocean Optics
For information 727-733-2447
www.OceanOptics.com

Unstable DNA Cloning

Copy Cutter EPI400 *E. coli* cells can significantly improve cloning results for unstable or toxic DNA sequences in many commonly used vectors. The CopyCutter EPI400 cell line reduces the copy number of plasmids with difficult-to-clone inserts to less than five per cell, thus improving the likelihood of successful cloning and stable clone maintenance. The cells are available in electrocompetent or chemically competent forms.

Epicentre Biotechnologies
For information 800-284-8474
www.EpiBio.com

DNA and RNA from Plant Material

The Agencourt Chloropure System is an extraction kit for DNA and RNA isolation and purification from plant material that delivers high recovery. The Agencourt Chloropure chemistry is automated on the Beckman Coulter Biomek NXP and FXP Laboratory Automation Workstations to process three 96-well plates in just 80 minutes. This patented, paramagnetic bead-based nucleic acid purification technology does not require organic solvents, vacuum filtration, or centrifugation steps.

Agencourt Bioscience
For information 978-867-2962
www.agencourt.com

Bunsen Burner Alternative

The Fireboy is a safe and mobile alternative to the traditional Bunsen burner. Operating from a gas cartridge and powered from a rechargeable battery, the Fireboy can be used wherever needed. Advanced temperature protection, flame monitor, and alarm display features protect the operator and the operating environment. The gas ignition is rapid and safe, requiring no outside flame.

Integra Biosciences
For information +41 81 286 9530
www.integra-biosciences.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 availabilty 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.