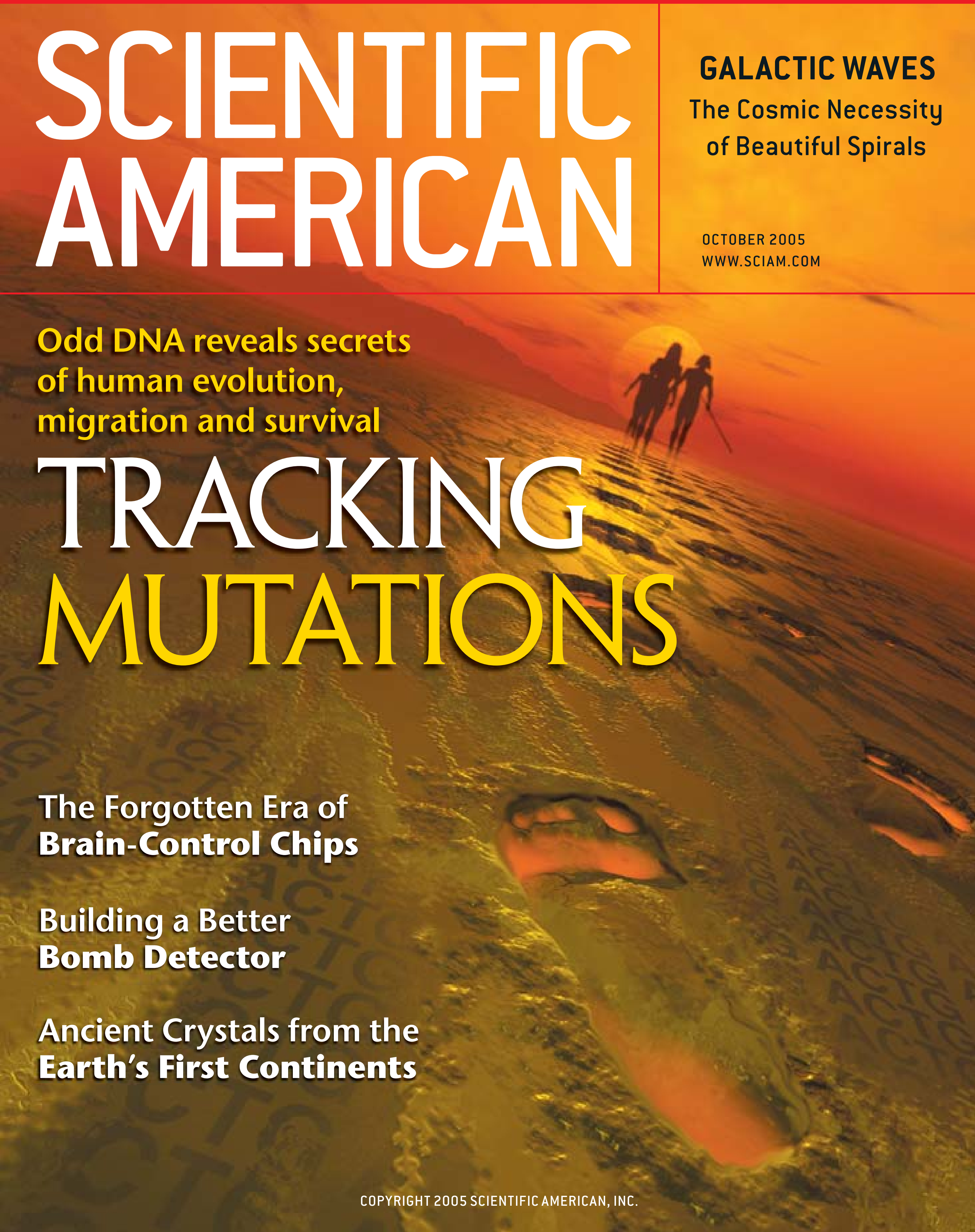
SMARTER WIRELESS FOR THE INTERNET • NEW TARGETS FOR NOVEL DRUGS

SCIENTIFIC AMERICAN

GALACTIC WAVES
The Cosmic Necessity of Beautiful Spirals

OCTOBER 2005
WWW.SCIAM.COM



Odd DNA reveals secrets of human evolution, migration and survival

TRACKING MUTATIONS

The Forgotten Era of
Brain-Control Chips

Building a Better
Bomb Detector

Ancient Crystals from the
Earth's First Continents

october 2005

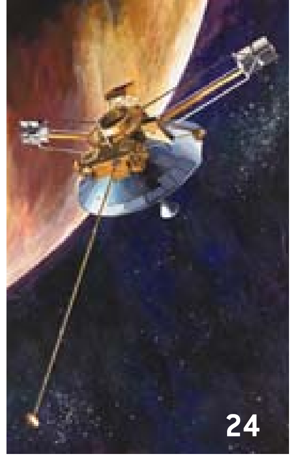
contents

SCIENTIFIC AMERICAN Volume 293 Number 4

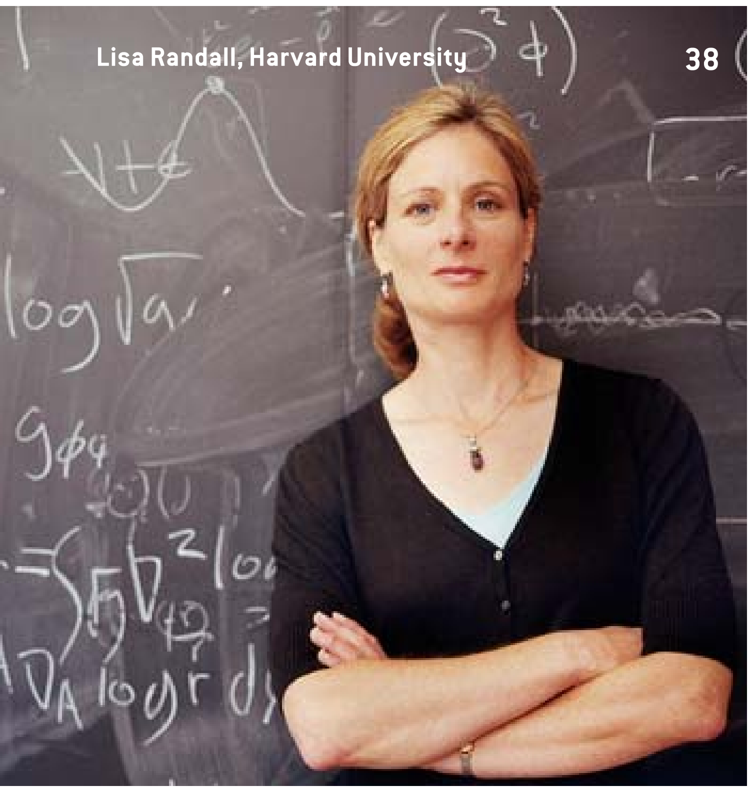
features

78 A genetic trail

departments


24


112


Lisa Randall, Harvard University 38

columns

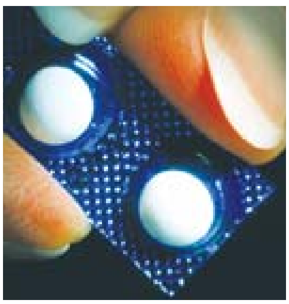
Cover image by Phil Saunders, Space Channel Ltd.; photograph at left by Kathleen Dooher; illustration on preceding page by Slim Films.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 2005 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40012504. Canadian BN No. 127387652RT; QST No. Q1015332537. Publication Mail Agreement #40012504. Return undeliverable mail to Scientific American, P.O. Box 819, Stn Main, Markham, ON L3P 8A2. Subscription rates: one year $34.97, Canada $49 USD, International $55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 248-7684. Send e-mail to sacust@sciam.com** Printed in U.S.A.

Fill This Prescription

No one's health should be hostage to a caregiver's opinion of his or her morality. In prisons, even hardened murderers are entitled to decent, prompt medical attention. A woman walking into a pharmacy with a prescription deserves no less.

Yet in at least a dozen states, pharmacists have refused to fill clients' orders for birth control on personal moral grounds. Often the prescriptions have been for emergency contraception—the "morning-after" pills marketed as Plan B and Preven that prevent a fertilized egg from implanting in the uterus if taken within 72 hours of unprotected sex. By medical definition, the pills block rather than terminate pregnancy. To the objecting pharmacists, however, it is abortion, and they want no part of it.



MORNING-AFTER pills cannot wait.

The responsibilities of pharmacists in this regard are murky. As the only licensed gatekeepers of public access to legal medications, pharmacists have an implicit duty to dispense them as needed. The American Pharmacists Association upholds the prerogative of its workers to protect their conscience if they can efficiently refer the prescription elsewhere. At least three states have laws that bar pharmacists from refusing drugs for personal reasons, but at least four others specifically permit a conscientious exception for birth control. The rest are in a gray zone.

It is tempting to wonder how far the principle of denying medicines for ethical reasons could stretch. Could one who disapproves of homosexuality refuse antiretrovirals to an HIV-positive gay man? If suffering is good for the soul, can one refuse to give out pain medication? But the pharmacists are not really fighting for a broad entitlement to morally judge which prescriptions to fill. And it is unnecessary to play "Where will this stop?" on an issue that already threatens women's vital reproductive rights.

Since 1973 physicians and nurses in the U.S. have conditionally had legal leave to abstain from abortions for personal reasons; the objecting pharmacists want that same option. But physicians are directly involved in the abortion procedure, and ethics aside, compelling them to perform it against their will is impractical and potentially unsafe. Druggists cannot make the same claim.

This fall Congress will consider a bill that would put the burden on pharmacies to fill birth-control prescriptions: an individual druggist could refuse but only if another can fill the order. That solution may be unworkable for small pharmacies in remote settings where standby pharmacists are uncommon.

The best answer is probably to make the morning-after pill available without prescription. Two Food and Drug Administration panels have recommended as much, but the agency continues to drag its feet. Over-the-counter versions of the product have nonetheless shown themselves to be safe in the U.K. and other countries.

If pharmacists can legitimately refuse to dispense the morning-after pill, then it is appropriate to consider seriously the American Medical Association's proposal that physicians be allowed to do so instead. Patients should always have confidence in their access to the drugs they need. We also suggest that if pharmacists are going to sift clients' prescriptions through the sieve of their own morals, they should prominently post signs to that effect. Let female customers know what to expect well before their health is at the mercy of their pharmacist's conscience.

THE EDITORS editors@sciam.com

DAMIEN LOVEGROVE *Photo Researchers, Inc.*

How to Contact Us

EDITORIAL

For Letters to the Editors:
Letters to the Editors
Scientific American
415 Madison Ave.
New York, NY 10017-1111
or
editors@sciam.com
Please include your name and mailing address, and cite the article and the issue in which it appeared. Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

For general inquiries:
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8200
fax: 212-755-1976
or
editors@sciam.com

SUBSCRIPTIONS

For new subscriptions, renewals, gifts, payments, and changes of address:
U.S. and Canada
800-333-1199
Outside North America
515-248-7684
or
www.sciam.com
or
Scientific American
Box 3187
Harlan, IA 51537

REPRINTS

To order reprints of articles:
Reprint Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8877
fax: 212-355-0408
reprints@sciam.com

PERMISSIONS

For permission to copy or reuse material from *SA*:
www.sciam.com/permissions
or
212-451-8546 for procedures
or
Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
Please allow three to six weeks for processing.

ADVERTISING

www.sciam.com has electronic contact information for sales representatives of Scientific American in all regions of the U.S. and in other countries.

New York
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8893
fax: 212-754-1138

Los Angeles
310-234-2699
fax: 310-234-2670

San Francisco
415-403-9030
fax: 415-403-9033

Detroit
Karen Teegarden & Associates
248-642-1773
fax: 248-642-6138

Midwest
Derr Media Group
847-615-1921
fax: 847-735-1457

Southeast
Publicitas North America, Inc.
404-262-9218
fax: 404-262-3746

Southwest
Publicitas North America, Inc.
972-386-6186
fax: 972-233-9819

Direct Response
Special Aditions Advertising, LLC
914-461-3269
fax: 914-461-3433

Australia
IMR Pty Ltd.
Tel: +612-8850-2220
Fax: +612-8850-0454

Belgium
Publicitas Media S.A.
+32-(0)2-639-8420
fax: +32-(0)2-639-8430

Canada
Derr Media Group
847-615-1921
fax: 847-735-1457

France and Switzerland
PEM-PEMA
+33-1-46-37-2117
fax: +33-1-47-38-6329

Germany
Publicitas Germany GmbH
+49-211-862-092-0
fax: +49-211-862-092-21

Hong Kong
Hutton Media Limited
+852-2528-9135
fax: +852-2528-9281

India
Convergence Media
+91-22-2414-4808
fax: +91-22-2414-5594

Japan
Pacific Business, Inc.
+813-3661-6138
fax: +813-3661-6139

Korea
Biscom, Inc.
+822-739-7840
fax: +822-732-3662

Middle East
Peter Smith Media & Marketing
+44-140-484-1321
fax: +44-140-484-1320

Scandinavia and Finland
Publicitas Nordic AB
+46-8-442-7050
fax: +46-8-442-7059

U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-9922

On the Web

WWW.SCIAM.COM

UPDATED EVERY WEEKDAY

Visit **www.sciam.com/ontheweb**
to find these recent additions to the site:

Earliest Dinosaur Embryos Recovered

The oldest fossilized dinosaur embryos yet discovered contain tantalizing clues about dinosaur development and evolution, paleontologists report. The remains indicate that some of the prehistoric creatures started out on four legs before growing into bipedal behemoths. They also support the notion that hatchlings did not fend for themselves but instead relied on their parents for nourishment.

Recipe for D-I-Y DNA Decoding Revealed

A thousand dollars can buy a lot of things. Researchers hope to soon add an individual's genetic sequence to that list. Full-genome DNA decoding, estimated to now cost $20 million, could soon be done for about $2.2 million, experts predict, and will continue to drop in price as new ways to conquer the task emerge. A recent study suggests one such method: a technique that used off-the-shelf instruments and reagents to successfully sequence the *E. coli* genome.

Why Cats Don't Cotton to Sweets Explained

The lure of sweets is the downfall of many a dieter. Cats, however, are indifferent to sugar—a trait that is rare in the mammal kingdom. Now scientists have figured out why. Felines apparently carry a defect in a gene that encodes part of the mammalian sweet taste receptor.

Ask the Experts

How do scientists turn genes on and off in living animals?
Miriam Meisler, professor of human genetics at the University of Michigan at Ann Arbor, explains.

Exclusive Online Issue: The Human Odyssey

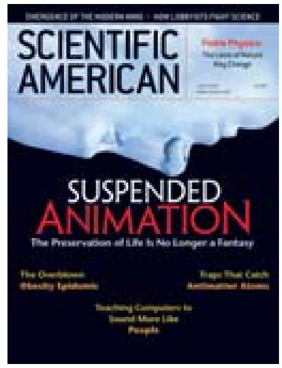
Tracing the origins of humankind.

Available today for $5 at www.sciamdigital.com

ARTWORK BY KEVIN DUPUIS/COURTESY OF UNIVERSITY OF TORONTO AT MISSISSAUGA

Letters

EDITORS@SCIAM.COM



TWO ARTICLES in June proved to be not only provocative but complementary. "Doubt Is Their Product," by David Michaels, discussed efforts by corporations and interest groups to employ their own research to undermine the science supporting product regulation. Many readers suspected they had found an example of such intentionally generated uncertainty in "Obesity: An Overblown Epidemic," W. Wayt Gibbs's report on assertions by researchers that the government, medical establishment and media are misleading the public on the health consequences of rising body weight.

For all the heat, there was a really cool article, "Buying Time in Suspended Animation," by Mark B. Roth and Todd Nystul, that prompted Elbert Pence of Kenmore, Wash., to wonder about the chilling implications of a hiatus from an active lifestyle: "Will the government cut off my Social Security payments if I decide to go into suspended animation? If so, will they be restored when I am warmed back up to 98.6?"

UNPRINCIPLED UNCERTAINTY

In "Doubt Is Their Product," David Michaels states, "Uncertainty is an inherent problem of science, but manufactured uncertainty is another matter entirely." He then draws a distinction between the efforts of government-funded scientists who generate causal inferences, presumably for the protection of the public's health, and those of industry-supported scientists who "manufacture" uncertainty by challenging those inferences, presumably for protection of the products made by industry.

In contrast, the same issue carries "Obesity: An Overblown Epidemic?" by W. Wayt Gibbs, which celebrates the uncertainty raised by academicians over the public health importance of obesity and suggests that unfounded concerns over the "epidemic of obesity" are promoted by a self-serving diet industry.

Who can judge whether scientific uncertainty raised by one group serves a greater good than that created by another? All challenges are most likely part of the crucible required to forge the knowledge needed to protect public health.

J. Michael Muhm
Woodinville, Wash.

PLUMPING THE OBESITY EPIDEMIC

In his article on obesity, Gibbs made a valiant effort to clear the air about the health consequences associated with obesity. But at least with reference to our report in the *New England Journal of Medicine,* he made a number of errors of fact and interpretation.

First, contrary to Gibbs's claim, we did not project death rates resulting from obesity: we used the observed relation between body mass index (BMI) and the conditional probability of death. Second, our assumption that everyone who is obese has a BMI of 30 or 35 was one of our ultraconservative assumptions, in that the fastest increases in obesity in the U.S. are occurring at a BMI of 35 or higher. It is unclear why Gibbs would call this a "false" presupposition.

Third, the *JAMA* article by Katherine M. Flegal et al. was unrelated to our manuscript: our focus was on childhood obesity and its effect on death rates over the next 50 years; theirs was devoted to estimating the current mortality burden of obesity. We had no need to estimate deaths attributable to obesity because we had observed data on BMI and age-sex-race-specific death rates. Fourth, scientists do not provide a "statistical analysis" of forecasts of life expectancy as Gibbs seems to think is necessary. As such, David B. Allison's statement that "we never meant for [our forecasts] to be portrayed as precise" is not a hedge; it is an accurate portrayal of our appropriate avoidance of point estimates in favor of a range of estimates.

Finally, Gibbs states that we never attempted to determine whether the number of months lost because of obesity was reliably different from zero. But that is exactly what we did, because we knew what the actual death rate was for those with BMIs ranging from 17 to 45 for the population aged 20 to 85.

It is critical to remember that never before in U.S. history have so many children been so heavy at such a young age. It is naive to believe that new drug therapies will enable an obese child to live a healthy life. If medical advances have reduced the death rate from cardiovascular diseases attributable to overweight/obesity in adults, that is wonderful news. We fear, however, that Gibbs's rosy outlook will prove misguided at best.

S. Jay Olshansky
University of Illinois at Chicago
David S. Ludwig
Children's Hospital, Boston

GIBBS REPLIES: Olshansky and Ludwig say they used the observed relation between BMI and the risk of death, not a projection. As my article stated, their paper used risk data that are more than a decade old, and its conclusions depend on those figures' validity now and for decades to come. Yet newer surveys (analyzed recently by Flegal and others) indicate that obesity's effect on mortality has been declining in the U.S. in recent decades. The relation their studies used is obsolete.

Olshansky et al. could have applied the actual distribution of BMI among American adults, as measured in recent government surveys, to perform their analysis. Instead they made estimates for an imagined U.S. where all obese adults have the same BMI. That assumption was false in the sense that it is not even approximately true. Last year Ali Mokdad and others at the CDC reported that 17 percent of the adult population had a BMI between 30 and 35 in 1999–2000, whereas only 12.6 percent had a BMI greater than 35. Assigning the entire obese population a BMI of 35 thus probably overestimates the negative effect of obesity on life expectancy. The researchers acknowledged this problem in their paper.

"Our focus was on childhood obesity," Olshansky and Ludwig say. But no nationally representative data yet exist that could test the connection between childhood obesity and adult mortality that they hypothesize, and the formal analysis in the paper was limited to adults aged 20 to 85. They claim that no statistical analysis is required to back up their speculation. Others may disagree.

They quarrel with my observation that "the study did not attempt to determine whether, given its many uncertainties, the number of months lost was reliably different from zero." Scientists perform a sensitivity analysis that checks how the output of a model changes as its inputs are tweaked together and individually. In lieu of a full analysis, Olshansky and Ludwig changed one parameter—the BMI rounding point for obese adults—by an arbitrary amount and were satisfied. Again, others may disagree.

THE OBESITY EPIDEMIC: **A gut reaction?**

PONDERING THE MODERN MIND

Kate Wong's article "The Morning of the Modern Mind" raises questions about symbolic behavior—not only whether it can be discerned in primitive hominids but what it is and why it should be considered a distinguishing characteristic of modernity. A stick used by a chimpanzee to extract termites is adventitious, and though not made by another tool, it contains information about its future utility. Through its use, knowledge about termite extraction can be passed on to future generations. Thus, the practice of hands working with the brain to create something useful for the future has been around for a very long time.

Just as human beings express a range of thinking within a population, we will find that modern thinking falls along a temporal continuum as well—and perhaps not only in higher primates. A better line of research might focus on the degree of forethought involved in creating a tool and how far into the future it implies the maker was looking. Agreed, this is a more difficult archaeological and, indeed, epistemological problem.

David A. Burack
South Burlington, Vt.

MAKING IT ANTIMATTER

In "Making Cold Antimatter," Graham P. Collins describes an exotic helium atom "in which one electron is replaced by an antiproton." Is this composition correct, as this resulting atom would not appear to be helium anymore—even if the antiproton is not in the nucleus—but a sort of hybrid He-H molecule?

Mark F. Wilks
Shrewsbury, U.K.

COLLINS REPLIES: A characteristic of a molecule is a chemical bond, produced by the sharing of electrons between the two atoms involved. In the exotic helium atoms studied by ASACUSA, the antiproton, like a very heavy electron, is in an orbital around the helium nucleus. Even though it is in a highly excited state, the antiproton is at an average distance of less than 30 picometers.

CLARIFICATIONS On pages 90 and 91 in "The Morning of the Modern Mind," by Kate Wong, the archaeological site of Quneitra is located in the Demilitarized Zone in the Golan Heights, currently occupied by Israel according to the cease-fire agreements of 1974. The city of Quneitra was returned to Syria the same year.

"Making Cold Antimatter," by Graham P. Collins, should have noted that Jeffrey S. Hangst of the University of Aarhus was the physics coordinator for ATHENA.

HUGH KRETSCHMER

50, 100 & 150 Years Ago

FROM *SCIENTIFIC AMERICAN*

Nuclear Babel ▪ Exobiology Babble ▪ Cholera Fable

OCTOBER 1955

NUCLEAR CONFERENCE—"The International Conference on the Peaceful Uses of Atomic Energy, in Geneva, was not only the largest meeting to date on nuclear energy but probably the most exciting international gathering of scientists ever held. One of the major surprises of the Conference was provided by President Dr. Homi J. Bhabha, Indian Cabinet member, in his opening address. Dr. Bhabha predicted that within 20 years it will be possible to derive energy from the controlled fusion of heavy hydrogen nuclei. Such a technological break-through would guarantee mankind a source of abundant energy essentially forever."

HEREDITY—"What right have we to assume randomness in the hereditary material, the product of a natural selection process which has operated for millions of years? In answer we can cite the fact that the sequence of digits in the number π (3.14159265...) is also random: there is no discernible system or pattern in the sequence. We may imagine a mad mathematician who, searching for 'useful numbers,' writes down one random sequence after another until, after rejecting millions of numbers as useless, he finally stumbles on the random number π and finds by test that it is very helpful indeed. Similarly in a living organism over eons of time the random mutations may once in a great while produce a sequence of nucleotides which blueprints a new and helpful enzyme.—George Gamow"

OCTOBER 1905

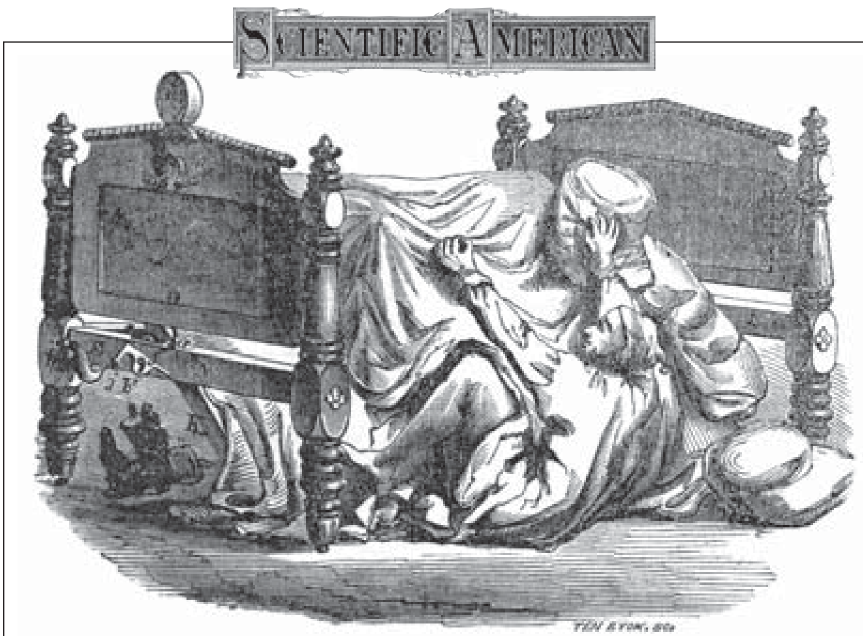
LIFE BEYOND EARTH—"The recent utterances of the venerable Dr. A. R. Wallace, fellow-discoverer with Darwin of the origin of species, tending to show that our earth is the only body in the known creation suited for life such as we find it here upon the globe, has awakened a wide interest among progressive scientists. Dr. Wallace is now a very old man, and like Lord Kelvin, he finds a Providential design in the arrangement of the material universe. One school claims that he is old and in his dotage; the other, that he has become wise in his old age."

OCTOBER 1855

CHOLERA—"J. F. Reigert, Esq. of Lancaster, Pa., attributes cholera to a small poisonous fly, and considers lime water a certain antidote. Lancaster is a limestone region, and has never had over ten cases of cholera in it, although it had daily communications with Philadelphia and Baltimore when this terrible disease prevailed in those cities in 1832 and 1852. In cholera seasons he believes they fill the atmosphere in great swarms, but are not readily observed, owing to their dust color, and that they carry disease into every place they visit by being inhaled by persons while breathing, and also by being taken unnoticed in food and drink."

MILITARY PHOTOGRAPHY—"The Glasgow (Scotland) *Herald* states that the commander of the British militia troops in Lanarkshire having lost a considerable number of men from desertion—after they have received part of their bounty—has hit upon a capital auxiliary to identification. The likenesses of the men are taken by daguerreotype; the picture is fidelity itself. When a man disappears from Lanark, the plate upon which his physog is imprinted is sent to the recruiting serjeants for the regiment, who can look after the man as if he had been an old acquaintance."

AN ALARMING BED—"The annexed engravings represent, in a forcible manner, the alarm bed for which a patent has been granted. A perspective view shows how the bed has operated upon its occupant, who recklessly dared to sleep beyond his allotted period of rest. The alarm having sounded, if the person snores on, the side rail is released, tilting the bed. Any sinner sleeping beyond a certain hour deserves to be tumbled out of the blankets in such a manner."



EJECTOR BED for curing the sin of sloth, 1855 news SCAN

GEOSCIENCE

The Biggest Dig

JAPAN BUILDS A SHIP TO DRILL TO THE EARTH'S MANTLE **BY TIM HORNYAK**

If you've ever thought about digging a hole to China, a new Japanese ship might be your best bet. Workers have just put the finishing touches on an ocean drilling vessel that is designed to bore to unprecedented depths and attain a long-held goal: penetrating the earth's rocky crust to the mantle.

The poorly understood mantle accounts for about two thirds of the planet's mass and is key in the unseen convection processes linked with tectonic plate motion. For Japan, an archipelago straddling the fractious intersection of at least three crustal plates, the issue is also earthquakes. "Japan is situated on these active planetary processes, and 30 million people actually live on one of the most dangerous or active places on the earth," says Asahiko Taira, director general of the Center for Deep Earth Exploration (CDEX) of the Japan Agency for Marine-

HEAVY DUTY: Deep-sea drilling vessel *Chikyu,* designed to penetrate through kilometers of rock to the earth's mantle, undergoes final outfitting at Mitsubishi Heavy Industries's shipyard in Nagasaki.

TIM HORNYAK

GOING DOWN FOR LIFE

The deep earth may not be sterile. For instance, peridotite, a mantle igneous rock, can produce hydrogen, methane and other compounds essential for life when it reacts with seawater during a metamorphic process called serpentinization. So as the *Chikyu* drills into the earth, researchers will keep a lookout for extremophiles, deep-sea bacteria in the crust and upper mantle that can survive intense heat and pressure. "If you drill the deep, there is virtually no influence of surface organisms," says Asahiko Taira, director general of Japan's Center for Deep Earth Exploration. "You can find the place where life originated and very early life is still evolving."

DRILLING DERRICK of the *Chikyu* rises 112 meters above the waterline.

TIM HORNYAK

Earth Science and Technology, which operates the ship.

To study the mantle, geophysicists have had to rely on indirect methods, such as looking at seismic signals and measuring gravitational field variations. They can examine mantle rocks that have been brought to the surface via volcanism or faulting, but because this material has undergone massive amounts of heating, cooling and other processes, many argue it is not truly representative of the mantle. Breaking through the border between the crust and hotter mantle—known as the Mohorovicic discontinuity, or Moho—would give scientists a direct, fresh sample of mantle as well as the fluid, gas, temperature and pressure conditions of its environment (including possible microorganisms) that are lost by the time the rock arrives at the surface naturally. Researchers from 18 countries working on the U.S. drill ship *JOIDES Resolution* recently tried to reach the mantle at the Mid-Atlantic Ridge, but they missed by less than an estimated 300 meters.

In July technicians in a Nagasaki port completed the final outfitting of the *Chikyu* (Japanese for planet "Earth") and handed over the colossal 57,500-ton, 210-meter-long white ship to CDEX. The *Chikyu,* to start crew training around Hokkaido this fall, is being deployed as part of the Integrated Ocean Drilling Program, a long-term effort begun in 2003 and whose main participants are Japan, the U.S. and the European Union.

Besides being the most sophisticated

laboratory on the seas, the science vessel boasts the tallest drilling derrick at 112 meters above the waterline and a drill pipe that is 9.5 kilometers long—22 times the height of the Empire State Building. This borer is expected to cut through some 7,000 meters of crust when the *Chikyu,* which cost about $540 million, is floating in seas up to 2,500 meters deep. Target drilling sites include areas where the mantle has been brought closest to the surface by tectonic action or where the crust is relatively thin, such as the Nankai Trough off Japan's Pacific coast.

To beat the current record drill depth of 2,111 meters, the *Chikyu* brings technology proven in the oil industry to bear. Its drilling system uses a 380-ton protective casing over the wellhead that is about the size of a six-story office building. It shields the vessel against eruptions of methane gas and pressurized fluids and allows for the secure retrieval of nine-meter-long core samples.

Another vital technology is the *Chikyu*'s dynamic positioning system, an automatic, satellite-guided location fixer that corrects against wind, wave and current forces with six 360-degree thrusters under the hull, keeping the ship over the borehole. The *Chikyu* will be involved in multiple drilling and coring projects, and some of the holes it creates could be used in the future to house on-site crust monitors that would improve quake warning systems, according to CDEX scientist Shinichi Kuramoto.

One risk in drilling to the mantle—or any seafloor drilling—is tapping into a pocket of gas hydrates. If a plume blows out and rises, it can sink the ship. Escaping gas can also spark catastrophic explosions and fires. But geophysicists believe that the risk is worth taking. Directly sampling and monitoring the mantle, asserts CDEX scientist Daniel Curewitz, "will greatly expand our understanding and could open new avenues of inquiry into the nature, history and future of the planet we call home."

Tim Hornyak is based in Tokyo.

BRAIN

One Face, One Neuron

STORING HALLE BERRY IN A SINGLE BRAIN CELL BY DIANE MARTINDALE

When you spot a celebrity on a magazine cover, your brain recognizes the image in an instant—an effect that seems to occur because of a single neuron. A recent study indicates that our brains employ far fewer cells to interpret a given image than previously believed, and the findings could help neuroscientists determine how memories are formed and stored.

BIG MEOW: The concept of Halle Berry (here disguised as Catwoman), not her visage per se, sets off a neuron that enables recognition.

Exactly how the human brain works to record and remember an image is the subject of much debate and speculation. In previous decades, two extreme views have emerged. One says that millions of neurons work in concert, piecing together various bits of information into one coherent picture, whereas the other states that the brain contains a separate neuron to recognize each individual object and person. In the 1960s neurobiologist Jerome Lettvin named the latter idea the "grandmother cell" theory, meaning that the brain has a neuron devoted just for recognizing each family member. Lose that neuron, and you no longer recognize grandma.

Experts long ago dismissed this latter view as overly simplistic. But Rodrigo Quian Quiroga of the University of Leicester in England and his colleagues decided to investigate just how selective single neurons might be. The team looked at eight patients who each had 64 tiny electrodes implanted in their brains before epilepsy surgery (a procedure to pinpoint the source of their seizures). Many of the electrodes were placed in the hippocampus, an area critical for the storage of long-term memories.

While each participant was shown a large number of images of celebrities, animals, objects and landmark buildings, electrodes recorded the brain cells' firings. This screening stage determined which images elicited a strong response in at least one neuron. The team then tested the responses to three to eight variations of those images from the narrowed list.

In one patient, a single neuron responded to seven different photographs of actor Jennifer Aniston, while it practically ignored the 80 other images of animals, buildings, famous or nonfamous people that were also presented. "The first time we saw a neuron firing to seven different pictures of Jennifer Aniston—and nothing else—we literally jumped out of our chairs," Quian Quiroga recalls.

Similar results occurred in another patient with a neuron specific for actor Halle Berry; the neuron responded not only to photographs but also to a drawing and an image of her name. What is more, even when Berry was costumed as the masked Catwoman, if the patient knew it was Berry, the neuron still fired. "This neuron is responding to the abstract concept of Halle Berry rather than to any particular visual feature. It's like, 'I won't recall every detail of a conversation, but I'll remember what it was about.' This suggests we store memories as abstract concepts," Quian Quiroga adds. Besides celebrities, famous buildings, such as the Sydney Opera House and the Tower of Pisa, elicited single-neuron firing.

"Not many scientists would have predicted such explicit single-neuron signals associated with individual people," says Charles Connor, a neuroscientist at Johns Hopkins University. "It should now be possible to look at precisely what information is represented by those cells—a clear starting point for studying how memories are encoded."

Although the "Jen" and "Halle" neurons behave much like a grandmother cell, the findings do not mean that a given brain cell will react to only one person or object, notes Christof Koch, one of the study's researchers at the California Institute of Technology. These cells probably respond to a wide range of items (some neurons responded to more than one person or object). "We are not saying that these are grandmother cells, but for familiar things, like your family or celebrities, things you see frequently, the neurons are wired up and fire in a very specific way—much more so

MIND OVER OBJECT

Think of a place, person or thing; then watch the word instantly appear on a computer screen. Such "brain reading" is far into the future. But Cyberkinetics in Foxborough, Mass., has developed an implant, called the BrainGate, that detects neural firing, permitting the control of objects with thought. In June 2004 surgeons implanted the firm's tiny chip containing 100 electrodes into the motor cortex of a 24-year-old quadriplegic. Each electrode connected directly into a neuron and allowed the patient to play computer games and check e-mail with his thoughts.

than previously thought," Koch explains.

The findings, in the June 23 *Nature*, could influence research into illnesses such as dementia, but Quian Quiroga sees a more practical application: implantable prosthetic communication devices, so-called brain readers. "We may be able to help patients communicate with the outside world, where their thoughts are interpreted by a computer," he predicts.

Diane Martindale is based in Toronto.

ASTROPHYSICS

A Force to Reckon With

WHAT APPLIED THE BRAKES ON PIONEER 10 AND 11? BY ALEXANDER HELLEMANS

One of the most intriguing mysteries in physics is the "Pioneer anomaly," the slowing down of two spacecraft by an unknown force. NASA launched Pioneer 10 and 11 in 1972 and 1973, respectively, and the craft returned stunning images of Jupiter and Saturn. But as both spacecraft continued their voyages at speeds of roughly 27,000 miles per hour, astronomer John Anderson of the Jet Propulsion Laboratory in Pasadena, Calif., noticed anomalies in telemetry data dating from as far back as 1980. With continued analysis, researchers determined that the spacecraft had been slowing down at a constant rate: each year they fell 8,000 miles short of their calculated positions. The strange behavior sparked several theories, but the lack of data made culling the ideas difficult. Now a proposal to analyze telemetry from the early years could literally point toward the correct explanation.

PIONEER PROBES zipped past Jupiter and are now well beyond Pluto's orbit. They have slowed down for some reason.

NASA AMES RESEARCH CENTER

The most obvious theory was that something on the spacecraft themselves created a braking force—leaking gas or heat radiation, perhaps. Over the years, however, researchers increasingly viewed this hypothesis as less likely, and some physicists began to explore possible flaws in Newton's laws and relativity. Others posited that dark matter was the culprit: it might exert a gravitational or drag force. A third theory embraces the idea that a minute acceleration exists in the velocity of light, which might result in the appearance that the probes are slowing down: if light travels faster, telemetry signals arrive faster, and the craft seem to be closer.

Anderson and theorist Michael M. Nieto of Los Alamos National Laboratory have proposed a way to filter the ideas, noting the interesting fact that the direction of the anomalous force would be different for each theory. If the force points toward the sun, then it should be a gravitational effect. If it points toward Earth, it should be an anomaly relating to the velocity of light. If it points in the direction of motion, it should be a drag force or a modification of inertia. And finally, if it points along the spin axis of the probes, it should indicate a force generated by the craft.

But determining the force's direction means studying telemetry when the Pioneer craft were closer than 20 astronomical units (1 AU equals the distance between Earth and

FOLLOWING THE PIONEER TRAIL

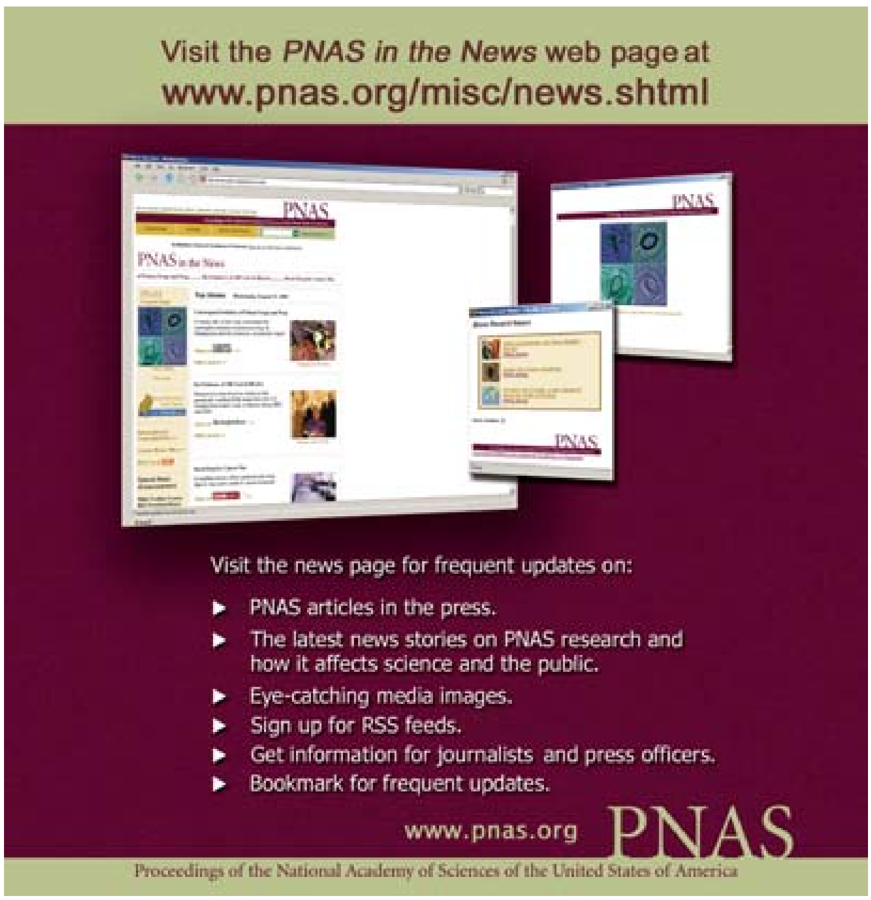
To solve the Pioneer anomaly, many scientists have been calling for a dedicated mission. (Other deep-space probes, such as Voyager 1 and 2, conducted too many thruster maneuvers to provide clear data about the anomaly.) Hopes are high that the European Space Agency (ESA) will include such a mission in its Cosmic Vision program for 2015–2025. NASA's New Horizons Pluto–Kuiper Belt mission, to be launched next year, may also furnish rough data. In fact, any future mission to Pluto or the Kuiper belt would be suitable to test the anomaly if the craft can be allowed to ride without corrections to its trajectory for longer times, says Dario Izzo of ESA's Advanced Concepts Team at the European Space Research and Technology Center in Noordwijk, the Netherlands.

the sun). Within this distance, to about Uranus's orbit, the angles between the sun, Earth and the craft's motions are sufficiently large. Until now, though, investigators have mostly analyzed telemetry covering the distance beyond 20 AU (to 70 AU so far—Pioneer 10's last useful transmission occurred at 80 AU in 2002). Experts had not bothered to study closer-in data in detail because they believed that radiation pressure of the sun and the many flight maneuvers would make it difficult to measure the anomaly.

Nieto and Anderson insist that it should be possible to correct for these factors and determine the direction of the anomalous force, especially in Pioneer 11's trajectory between Jupiter and Saturn. At that point, it traveled at practically a right angle to the direction of the sun and Earth, so any force toward the sun or Earth will be noticed by a sideways displacement of the probe. And rough measurements seem to show that the anomaly existed back to 10 AU, Nieto says. The telemetry for the early part of the Pioneer missions is available, so the analysis "is a relatively cheap thing to do, and at the very least it will give us more information and perhaps an indication," he remarks.

Especially if the mysterious force points toward the sun, then the explanation might be a deviation from Newtonian dynamics—termed modified Newtonian dynamics, or MOND—an idea originally proposed to explain why rotating galaxies do not fly apart. Dark matter may modify gravity, though as an alternative, Mordehai Milgrom of the Weizmann Institute of Science in Rehovot, Israel, proposes an additional component of gravity that should appear over large distances. In any case, MOND has become one of the more popular approaches to solving the anomaly problem. Jacob D. Bekenstein of the Hebrew University in Jerusalem applied a relativistic theory of MOND to the solar system and found that the Pioneer anomaly "is of crudely the right magnitude" to fall within MOND.

Alexander Hellemans writes about physics from Naples, Italy.



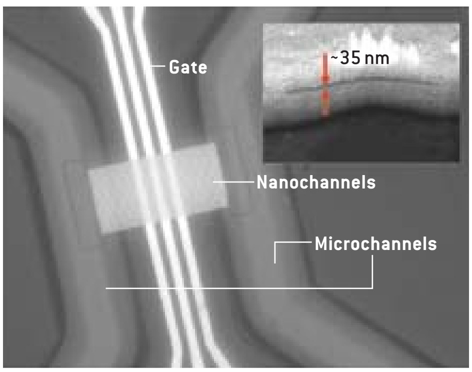
NANOTECH

Transistor Flow Control

FORGET VALVES—CONTROLLING FLUIDS WITH ELECTRIC FIELDS **BY CHARLES Q. CHOI**

At the heart of modern electronics are transistors, which act like valves to direct the flow of electrons. Now researchers at the University of California at Berkeley have created the first transistors that electrically control molecules instead. By connecting them to microscopic test tubes and petri dishes, these nanoscale transistors could lead to labs-on-a-chip that work without moving parts.

Much as 30-ton computers shrank over decades to microchip size, investigators are now miniaturizing labs to run millions of experiments simultaneously and dramatically speed analysis of DNA, proteins and other molecules. Although valves and pumps exist to control flow in microfluidic channels, they are not easy to miniaturize further for use on nanometer levels, says Berkeley mechanical engineer Arun Majumdar.



CHANNELING: Controlled by three gate electrodes, liquid can move between the microchannels via a block of 30 nanochannels 120 microns long. One nanochannel is about 35 nanometers wide (*inset*).

Instead of relying on mechanical manipulations, Majumdar and his colleagues speculated that silicon transistors might electrically control ions dissolved in fluids as well as they could electrons. Prior attempts to control ions by charging the surfaces of microfluidic channels, however, showed that ions quickly migrated to channel walls and canceled out the voltage, shielding the rest of the liquid from further electric manipulation.

With Peidong Yang, Rohit Karnik, Rong Fan and their co-workers, Majumdar found that channels less than 100 nanometers across are tiny enough to enable electric fields to breach this shielding. After constructing a 35-nanometer-high channel between two silica plates and filling it with potassium chloride saltwater, they demonstrated that voltage applied across this nanofluidic transistor could switch potassium ion flow on and off. Matching results were seen with transistors made from silica tubes 10 to 100 nanometers wide. "It's very good basic science and a clever idea. They elegantly take advantage of a physical effect that only dominates in very small channels," remarks Stanford University biophysicist Stephen Quake.

Most biomolecules are electrically charged, and the transistor could manipulate DNA fragments effectively. Majumdar envisions nanofluidic transistors that could rapidly sort the slew of molecules in cells by their mass and charge, thereby helping to purify DNA for sequencing or look for markers of disease.

Currently the transistors work on femtoliter (10^{-15} liter) amounts of fluid or less, roughly one one-hundredth the volume of a red blood cell. They could in theory prove sensitive enough to detect and manipulate single biomolecules, for exquisitely sensitive bioweapons detectors or "a lab for a single molecule, where you trap them and then study their behavior with light, force or any stimulus you want," Majumdar says.

Silicon offers the opportunity to build conventional and nanofluidic transistors onto the same chip for computerized control of chemical and biological processing. In early prototypes the voltage needed to switch ion flow on and off was 75 volts, far too high to incorporate into modern integrated circuits. But Majumdar explains they dropped the switching voltage to a sufficiently low one volt by thinning the channel walls. The team hopes to link nanofluidic transistors together into an integrated circuit within the year as the next step to harnessing massive numbers of transistors in parallel.

Charles Q. Choi is a frequent contributor.

FLUIDIC DIAGNOSTICS

Nanofluidic transistors can help analyze a few cells' worth of molecules. That could lead to diagnostics that spot a disease in its earliest stages, when only a few afflicted cells exist, and identify the exact form of the ailment, allowing for more precise therapies. For instance, with the early stages of cancer, "there are probably in a test tube of blood only two or three cancerous cells, and there is no technology at present that can tell what kind of cancer a person has from only a few cells. [The nanofluidics transistor] can change that," states Arun Majumdar, a mechanical engineer at the University of California at Berkeley who helped to develop the device.

COURTESY OF ARUN MAJUMDAR AND PEIDONG YANG *University of California, Berkeley*

PHYSICS

Quantum Bug

QUBITS MIGHT SPONTANEOUSLY DECAY IN SECONDS **BY GRAHAM P. COLLINS**

Computers that exploit the weird properties of quantum mechanics could have capabilities far exceeding those of conventional computers for certain problems, such as breaking a widely used type of encryption. Yet physicists must overcome a fundamental obstacle before quantum computers can become a practical reality: decoherence, which is the loss of the very quantum properties that such computers would rely on. Decoherence stems from the tiniest stray interactions with the ambient environment, and thus most quantum computer designs seek to isolate the sensitive working elements from their surroundings.

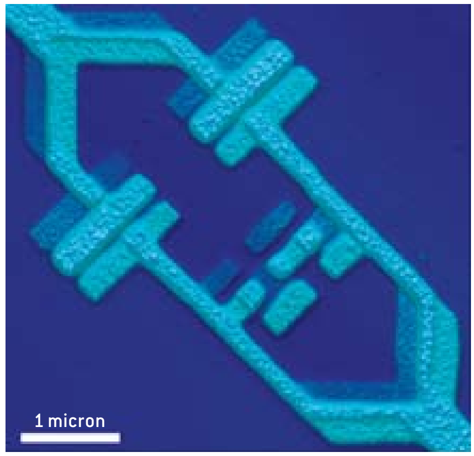
Jeroen van den Brink and his colleagues at Leiden University in the Netherlands, however, suggest that even perfect isolation would not keep decoherence at bay. A process called spontaneous symmetry breaking will ruin the delicate state required for quantum computing. In the case of one proposed device based on superconducting quantum bits (qubits), they predict that this new source of decoherence would degrade the qubits after just a few seconds.

A key feature of qubits is their ability to be in a so-called superposition; in essence, they can be 0 and 1 simultaneously, unlike bits in a standard computer, which must have a definite value. A qubit in a superposition is typically in a highly symmetrical state. For example, in a superconducting qubit a small electric current circulates in a loop both clockwise and counterclockwise at the same time. Spontaneous symmetry breaking disturbs that equanimity. The process occurs throughout physics—a ball perched on the top of a hill, for instance, tends to roll down one side or the other, ruining the symmetrical (if unstable) state of the ball balanced at the top. In the case of the superconducting loop, spontaneous symmetry breaking tends to cause the qubit to choose a definite state, ruining the superposition.

The Leiden researchers' result applies only to qubits that are composed of a large number of particles. Superconducting qubits fit that bill, because the electric current consists of many billions of electrons. The result does not apply to qubits based on single particles, such as an ion suspended in a magnetic trap or a single electron in a quantum dot on a chip. Indeed, in August physicists at the National Institute of Standards and Technology demonstrated single-ion qubits with a coherence time of more than 10 seconds.

Not everyone agrees that the constraint of a few seconds is a serious obstacle for superconducting qubits. John Martinis of the University of California at Santa Barbara says that one second "is fine for us experimentalists, since I think other physics will limit us well before this timescale." According to theorist Steven M. Girvin of Yale University, "if we could get a coherence time of one second for a superconducting qubit, that would mean that decoherence would probably not be a limitation at all." That is because quantum error correction can overcome decoherence once the coherence time is long enough, Girvin argues. By running on batches of qubits that each last for only a second, a quantum computer as a whole could continue working indefinitely.

So far superconducting qubits in the laboratory last about 500 nanoseconds before decoherence takes its toll. Girvin points out that decoherence times were just nanoseconds a few years ago, so that 500 nanoseconds "represents tremendous progress."



SUPERCONDUCTING QUBIT (*pentagonal loop at lower right*) would lose its data in seconds even if perfectly isolated. The upper left loop (referred to as a SQUID) reads out the state of the qubit.

IRINEL CHIORESCU *Delft University of Technology*

CORRECTING BIT BY BIT

Quantum error correction could solve the problem of decoherence, or the destruction of the delicate state required in quantum computation. It works by encoding each qubit of a computation in several physical qubits, so bad qubits can be detected and corrected. This process is only beneficial, however, if at least about 10,000 operations can be carried out per bad qubit; if qubits go bad more frequently, the increased computational overhead of quantum error correction introduces more errors than the error correction eliminates.

Yale University physicist Steven M. Girvin says that present-day superconducting qubits could perform an operation in about 10 nanoseconds, so about 100 million operations could occur in a second—well above the threshold for successful error correction.

HEALTH

Parsing Parasites

GENOMES OF THREE TROPICAL PARASITES ARE SEQUENCED **BY KASPAR MOSSMAN**

This past summer three international groups published the DNA sequences of the parasites that cause Chagas' disease, African sleeping sickness and leishmaniasis. These deadly ailments, transmitted by bloodsucking insects such as the "kissing bug" and the tsetse fly, take an enormous toll on tropical and subtropical populations—not only do they kill 125,000 people every year, but they can also compromise blood banks and disfigure patients so badly that they are ostracized from society. Although treatments derived from the work are still several years away, the sequencing of the genomes of the three parasites—*Trypanosoma cruzi, T. brucei* and *Leishmania major*—represents new hope in the battle against the diseases.

No vaccine exists against these parasites, collectively called the Tritryps, and current medicines are toxic, difficult to administer and not always successful, according to the World Health Organization. Based on arsenic and antimony, those therapies in widest use date back to the 1940s. Some newer treatments exist but have limited efficacy. "They are neglected parasites that afflict deeply impoverished people," says Najib El-Sayed, a parasitologist at the Institute for Genomic Research in Rockville, Md., who is the lead author of the paper on *T. cruzi* and the principal investigator of the *T. brucei* project.

To sequence *L. major,* the scientists used the "shotgun" method—chopping up the DNA from each chromosome into short bits, sequencing the fragments separately, and then assembling them into the entire chromosome like a giant, one-dimensional jigsaw puzzle. They solved *T. brucei* by a combination of the shotgun method and "chromosome walking," in which longer lengths of DNA are stitched together from the middle of the chromosome outward.

But the DNA of *T. cruzi,* the agent of Chagas' disease, proved far harder to decode. For starters, scientists did not know how many chromosomes *T. cruzi* had; moreover, at least 50 percent of the genome consists of repetitive material. The huge number of gene repeats made the chromosome-walking approach impossible. So El-Sayed and his colleagues sequenced *T. cruzi* by the more complex whole-genome shotgun strategy.

The parasites' genomes suggest good news about the prospects for new drugs. "They are so different from humans that there would be no problem identifying drug targets," says George Cross, head of the molecular parasitology department at the Rockefeller University. As parasites, the Tritryps require nutrients from their hosts—for example, they cannot manufacture sialic acid, a sugar they require to construct the complex surface molecules they use to evade the human immune response. It should be possible to devise custom drugs to block pathways such as sialic acid uptake and kill the parasites.

SEQUENCED: The parasite *Trypanosoma brucei*, shown among red blood cells, causes African sleeping sickness.

Much work remains, however, before novel therapies are invented. To know exactly where to stick a molecular wrench in the Tritryp infection machinery, scientists will need to deduce the metabolic pathways from the genomes—learning what proteins are made, what they do and how they interact. The job is akin to figuring out how a car's fuel-injection system works from a complete list of the vehicle's components broken down to the last screw, spring and bearing.

Considering the challenge of deducing the metabolic pathways—and the need for a wide spectrum of drugs to ensure that resistant strains do not emerge—big pharmaceutical firms are unlikely to pursue Tritryps therapies. "Nobody is going to make money out of these drugs," says Cross, because the victims of trypanosome diseases are so poor. For that reason, he proposes specialized research institutes for drug development, where research would be driven not by profit or by the publication of academic papers but by humanitarian motives and long-term strategic planning. The prospect of cures for these terrible diseases that afflict so many millions demands no less.

DOUBLED GENOME

As researchers sequenced the DNA of *Trypanosoma cruzi*, the protozoan that causes Chagas' disease, they realized that they were looking at not one genome but two—even though they had stringently purified the protozoan strain. Evidently, two strains of *T. cruzi* fused together at some point in recent evolutionary history (a few million years ago), so that today the species has two nearly identical halves of the genome. It is as if for a human, "instead of the mother's and father's DNA recombining, the child had both the mother's and father's entire genomes at the same time," explains Najib El-Sayed, a parasitologist at the Institute for Genomic Research in Rockville, Md.

DINOSAURS

Crawling Sensation

FROM ALL FOURS TO BIPEDAL GIANTS—AND NEEDING PARENTS BY JR MINKEL

Some dinosaurs, just like people, had to crawl before they could walk. The conclusion comes from a rare clutch of remarkably pristine dinosaur embryo fossils dating back to 190 million years ago. Although the embryos—the oldest yet discovered—were found in South Africa in 1978, analysis had to wait until the fossils could be sent to a paleontology laboratory with the right tools, including a vibrationless table and special dissecting devices.

After a year of scraping away rock and eggshell, investigators assigned the bones to the Prosauropod *Massospondylus*, a five-meter-long, plant-eating biped with a long neck and short head. The 15-centimeter-long embryos had large forelimbs and heads, horizontal necks and short tails compared with the adults—a clunky form that suggests they hatched as crawlers.

FOSSIL EMBRYO indicates that this plant eater crawled as a juvenile. Lack of teeth suggests that parents fed their young.

By comparing these two life stages with juvenile intermediates, the group concludes that this species' neck grew like a beanstalk relative to its head and forelimbs, resulting in a more bottom-heavy frame. "We were able to show the animal hatched as a quadruped and became a biped, which is very unusual," says co-author Robert Reisz of the University of Toronto at Mississauga.

Later sauropods were closely related giants that moved on all fours as adults, and their evolution may have occurred by preserving a similar juvenile state, Reisz and his colleagues speculate in the July 29 *Science*. The fossils also bolster the idea that dinosaurs watched over their young, they note: lacking teeth, the embryos may have needed their parents to provide them with food.

JR Minkel is a frequent contributor.

EYE OF SCIENCE/PHOTO RESEARCHERS, INC. (*top*); ROBERT REISZ *University of Toronto at Mississauga* (*bottom*)

PLANETS

Worlds without End

News of another "planet" beyond Pluto may become common. First came Quaoar in 2002 and then Sedna in 2003. Unlike these two worlds, the latest candidate, announced in July, is actually bigger than Pluto, by 50 percent. Designated 2003UB313 and unofficially nicknamed Xena, the mass of ice and rock currently lies three times farther out than Pluto. Investigators originally photographed it in 2003 at Palomar Observatory near Los Angeles, but its strange orbit, tilted nearly 45 degrees off that of nearly all other planets, delayed its discovery until this past January. Near-infrared images reveal a surface of mostly methane ice, remarkably similar to Pluto's. One or two more planets of like size might dwell within the same distance, says planetary scientist Michael Brown of the California Institute of Technology. An untold number of worlds might lurk beyond that, perhaps deriving from the Kuiper belt or the hypothesized Oort cloud. "No one has really probed out to that distance," he remarks. *—Charles Q. Choi*

GENETIC ENGINEERING

Missing in Maize

In 2001 a controversial *Nature* paper reported that genetically modified corn ended up where none should be in the Mexican state of Oaxaca, raising concern that the transgenic varieties might invade the natural population. The journal disavowed that paper the next year on insufficient evidence, but subsequent Mexican government studies backed it. New research supports *Nature*'s disavowal. Mexican and U.S. researchers analyzed nearly 154,000 seeds from 870 maize plants in 125 fields over two years in Oaxaca, looking for traces of engineered genes. After expecting a transgenic presence as high as 5 or 10 percent, they were surprised to find no traces. The scientists suggest transgenic maize may not have survived in the harsh mountain climate and soil in which indigenous corn mainly grows and that local farmers may have taken extra precautions with their seed stocks after they became aware of potentially uninvited genes. The *Proceedings of the National Academy of Sciences USA* published the findings online August 10. *—Charles Q. Choi*

TRANSGENES apparently did not escape in Mexico. Here maize is sold in Merida, Yucatán.

PHYSICS

Clues to the Earth's Heat

Neutrinos and antineutrinos detected in laboratories are normally generated by stars and nuclear reactors. An international team of physicists now has picked up neutrinos emanating from inside the earth. The antineutrino detector KamLAND, made up of nearly 2,000 light sensors placed around 1,000 tons of oil and fluorescent dye buried in a mine cavern in Japan, recorded roughly one antineutrino a month generated by the decay of uranium 238 and thorium 232 deep within the planet. The findings, in the July 28 *Nature*, suggest that the origin of the 30 trillion to 45 trillion watts of heat produced by the earth's interior is about evenly divided between radioactive decay and leftover heat from the planet's molten formation. Because the sound waves used to study the inner earth yield little about its chemistry, geoneutrinos should help shed light on the location and concentration of unstable elements. *—Charles Q. Choi*

DATA POINTS: SOAPING UP

Hand washing with soap significantly reduces the spread of pneumonia and diarrhea, the two leading causes of death worldwide among children younger than age five. In a study funded primarily by P&G Beauty, a division of Procter & Gamble, researchers examined 900 households in squatter settlements in Karachi, Pakistan. Although the public health benefits of clean hands are clear, whether poor communities by themselves can afford to purchase soap regularly remains uncertain.

Number of pneumonia cases (per 100 person-weeks) among households that washed:

With antibacterial soap: 2.42

With plain soap: 2.20

Without soap: 4.40

Number of diarrhea cases among households that washed:

With antibacterial soap: 2.02

With plain soap: 1.91

Without soap: 4.06

Grams of soap used each day by each person in the study: 4.4

Weekly cost of soap use: about $1

Weekly income of nearly half the households: less than $15

SOURCE: Lancet, *July 16, 2005*

ROBERT HOLMES *Corbis*; ILLUSTRATION BY MATT COLLINS

ASTROPHYSICS

Extreme Star Formation

Stars can apparently form in the extreme environment near a supermassive black hole, astrophysicists report. Near-infrared studies had revealed about 100 high-mass stars within 0.3 light-year of the Milky Way's central supermassive black hole. If those stars had formed elsewhere and spiraled in, they should be accompanied by tens of thousands of low-mass stars. Small stars around black holes should also be intensely bright x-ray sources, but the x-rays emanating from around the galactic core could represent no more than 1,000 such stars, the researchers note. Hence, they conclude, the massive stars must have formed near the black hole, presumably by fragmentation of the surrounding accretion disk. "Theorists long believed it possible," says Sergei Nayakshin of the University of Leicester in England, co-author of the paper to appear in the *Monthly Notices of the Royal Astronomical Society.* —*JR Minkel*

PSYCHOLOGY

Remember Wrong to Eat Right

Implanted memories of ice-cream-induced illness seem to create an aversion to the fattening treat. Psychologists at the University of California at Irvine had volunteers answer a questionnaire about their eating history. All subjects received the same false but plausible sounding summary of their eating patterns, which claimed that strawberry ice cream had once made them sick. After being encouraged to reflect on the alleged sickness, up to 40 percent of study participants reported they were less willing than before to eat strawberry ice cream. The deceit, also previously demonstrated to be effective with low-fat foods, did not work on more commonly eaten munchies such as potato chips, however, and still has to be tested for its long-term effects and sway during actual noshing, says Elizabeth Loftus, co-author of the August 3 *Proceedings of the National Academy of Sciences USA* online report. —*JR Minkel*

BY THE NUMBERS

More Boys Than Girls

SKEWED RATIO MAY LEAD TO THE NEXT SEXUAL REVOLUTION BY RODGER DOYLE

One of the most underappreciated forces affecting life in the U.S. is the sex ratio. Generally defined as the number of males per 100 females, it has a profound bearing on several issues, not least of which is the status of women.

Before World War I, immigrants, who tended to be predominantly male, kept the ratio high. Restrictive legislation in the 1920s and the Great Depression of the 1930s reduced the influx to a trickle. Beginning in the 1940s, lung cancer and cardiovascular disease, both of which affected men far more than women, resulted in an increasing proportion of females. The rise in the ratio since 1970 has resulted from a greater reduction in mortality among males than females. Increasing illegal immigration, which brings in more males than females, has apparently offset the effect of legal immigration, which now tends to involve more females than males.

The falling ratios between 1920 and 1970 led to delayed marriage [see "A Surplus of Women," By the Numbers, April 2004]. More important has been the effect on gender relations, according to two social psychologists, the late Marcia Guttentag of Harvard University and Paul F. Secord of the University of Houston. They see the decline in the sex ratio during this period as one of the key contributors to modern feminism.

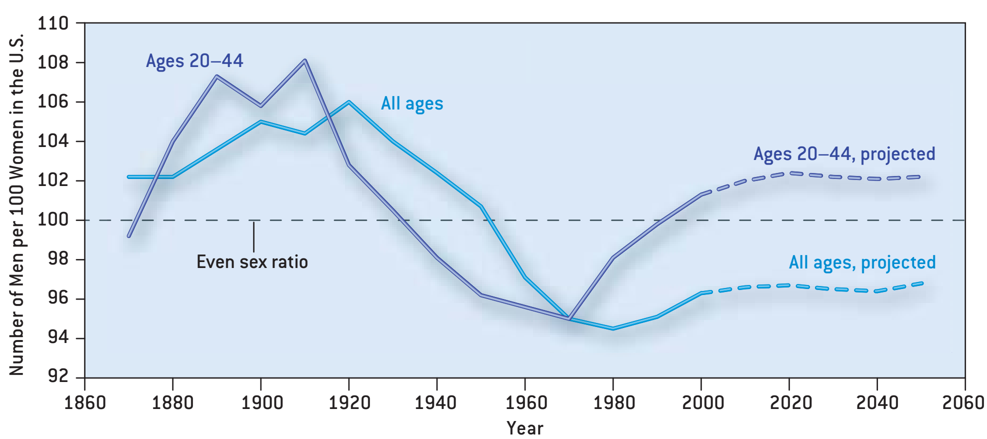
Until well past the middle of the 20th century, men dominated women because of their control of government, religion and business. They maintained a superior legal position and imposed traditional notions of patrimony. Women generally held sway only in the brief time before marriage, when they could choose among suitors.

With the decline in the number of marriageable men beginning about 1920, women were at a double disadvantage, because men could be more choosy while retaining many of the advantages of a more patriarchal era. The reaction of many women to their growing disadvantage was, according to Guttentag and Secord, to redefine their relationship to men and to strive for independence. These developments came to a head in the feminist movement of the 1960s.

The rising sex ratios of the past few decades means that American women of mating age are becoming, for the second time in U.S. history, a scarce commodity in the marriage market, although over the foreseeable future they will not be as rare as in earlier times. Women's advantageous sex ratio, together with greater legal protections than ever before and their growing superiority in educational achievement, could lead, arguably, to an unprecedented situation in which they become the dominant sex.

Next month: Sex ratio and crime.

Rodger Doyle can be reached at rodgerpdoyle@verizon.net



SOURCE: U.S. Census Bureau

RODGER DOYLE

FURTHER READING

Too Many Women? The Sex Ratio Question. Marcia Guttentag and Paul F. Secord. Sage Publications, 1983.

Consequences of Imbalanced Sex Ratios: Evidence from America's Second Generation. Josh Angrist. National Bureau of Economic Research, Working Paper No. 8042, December 2000. Available at www.nber.org/papers/W8042

The Human Dichotomy: The Changing Numbers of Males and Females. John I. Clark. Oxford University Press, 2000.

These good acts give us pleasure, but how happens it that they give us pleasure? Because nature hath implanted in our breasts a love of others, a sense of duty to them, a moral instinct, in short, which prompts us irresistibly to feel and to succor their distresses. —Thomas Jefferson, 1814

Unweaving the Heart

Science only adds to our appreciation for poetic beauty and experiences of emotional depth By MICHAEL SHERMER

Nineteenth-century English poet John Keats once bemoaned that Isaac Newton had "Destroyed the poetry of the rainbow by reducing it to a prism." Natural philosophy, he lamented, "Will clip an Angel's wings/Conquer all mysteries by rule and line/Empty the haunted air, and gnomed mine/Unweave a rainbow."

Does a scientific explanation for any given phenomenon diminish its beauty or its ability to inspire poetry and emotional experiences? I think not. Science and aesthetics are complementary, not conflicting; additive, not detractive. I am nearly moved to tears, for example, when I observe through my small telescope the fuzzy little patch of light that is the Andromeda galaxy. It is not just because it is lovely, but because I also understand that the photons of light landing on my retina left Andromeda 2.9 million years ago, when our ancestors were tiny-brained hominids. I am doubly stirred because it was not until 1923 that astronomer Edwin Hubble, using the 100-inch telescope on Mount Wilson in the hills just above my home in Los Angeles, deduced that this "nebula" was actually a distant extragalactic stellar system of immense size. He subsequently discovered that the light from most galaxies is shifted toward the red end of the electromagnetic spectrum (literally unweaving a rainbow of colors), meaning that the universe is expanding away from its explosive beginning. That is some aesthetic science.

Science and aesthetics are complementary, not conflicting.

No less awe-inspiring are recent attempts to unweave the emotions, described by anthropologist Helen Fisher of Rutgers University in her book *Why We Love* (Henry Holt, 2004). Lust is enhanced by dopamine, a neurohormone produced by the hypothalamus that in turn triggers the release of testosterone, the hormone that drives sexual desire. But love is the emotion of attachment reinforced by oxytocin, a hormone synthesized in the hypothalamus and secreted into the blood by the pituitary. In women, oxytocin stimulates birth contractions, lactation and maternal bonding with a nursing infant. In both women and men it increases during sex and surges at orgasm, playing a role in pair bonding, an evolutionary adaptation for long-term care of helpless infants.

At the Center for Neuroeconomics Studies at Claremont Graduate University, Paul J. Zak posits a relation between oxytocin, trust and economic well-being. "Oxytocin is a feel-good hormone, and we find that it guides subjects' decisions even when they are unable to articulate why they are acting in a trusting or trustworthy matter," Zak explained to me. He argues that trust is among the most powerful factors affecting economic growth and that it is vital for national prosperity for a country to maximize positive social interactions among its members by ensuring a reliable infrastructure, a stable economy, and the freedom to speak, associate and trade.

We establish trust among strangers through verification in social interactions. James K. Rilling and his colleagues at Emory University, for instance, employed a functional magnetic resonance imaging (fMRI) brain scan on 36 subjects while they played Prisoner's Dilemma. In the game, cooperation and defection result in differing payoffs depending on what the other participants do. The researchers found that in cooperators the brain areas that lit up were the same regions activated in response to such stimuli as desserts, money, cocaine and beautiful faces. Specifically, the neurons most responsive were those rich in dopamine (the lust liquor that is also related to addictive behaviors), located in the anteroventral striatum in the middle of the brain—the so-called pleasure center. Tellingly, cooperative subjects reported increased feelings of trust toward, and camaraderie with, like-minded partners.

In Charles Darwin's "M Notebook," in which he began outlining his theory of evolution, he penned this musing: "He who understands baboon would do more towards metaphysics than Locke." Science now reveals that love is addictive, trust is gratifying and cooperation feels good. Evolution produced this reward system because it increased the survival of members of our social primate species. He who understands Darwin would do more toward political philosophy than Jefferson. SA

Michael Shermer is the publisher of Skeptic *(www.skeptic.com). His latest book is* Science Friction.

BRAD HINES

The Beauty of Branes

Lisa Randall's thinking on higher dimensions, warped space and membranes catalyzed ideas in cosmology and physics. It might even unify all four forces of nature By MARGUERITE HOLLOWAY

It was the summer of 1998, recalls Harvard University physicist Lisa Randall, when extra dimensions finally pulled her in. Extra dimensions—beyond the four we encounter every day (three of space plus one of time)—have been an ingredient of theoretical physics for decades: mathematician Theodor Kaluza proposed a fifth in 1919, string theory requires 10 of them, M-theory needs 11. But Randall hadn't much use for them, she says, until that summer when she decided they might be helpful to supersymmetry, one of the conundrums she was pondering.

Randall contacted Raman Sundrum, a Boston University postdoctoral student with whom she had previously collaborated, and asked him if he would like to brainstorm about extra dimensions and membranes—"branes," as they are called for short. Branes are domains or swaths of several spatial dimensions within a higher-dimensional space. The everyday world we live in could be a three-brane, for example, and it is anyone's guess as to what dimension brane it might be embedded in. "Raman had already thought about branes and extra dimensions, and he was an obvious person to join forces with," Randall explains.

But Sundrum was a little worried. He was on his third postdoc, didn't have a job lined up and was considering leaving physics for finance. But he liked the way Randall thought and decided to set off on what might be his final physics adventure. The fruits of that collaboration, as fueled by caffeine and ice cream as by heady equations, were papers known as RS-1 and RS-2, two of the most cited in physics for the past five years.

The papers, which appeared in 1999, offered novel ways to think about gravity, branes and extra dimensions, and they suggested that the universe might have evolved differently in the beginning than it did later. "For me and a lot of people interested in cosmology and particle physics, it meant that there was this whole new set of possibilities of what could be going on in the early universe," says James Cline of McGill University. For Sundrum, now a professor at Johns Hopkins University, it meant seven job offers. "She is somebody with a marvelous instinct," he laughs.

This instinct often draws Randall to problems she knows little about. While at the renowned Stuyvesant High School in New York City, Randall decided to work on perfect versions of complex numbers called Gaussian integers for the then Westinghouse science talent search. (In perfect numbers such as 6, the fac-

KATHLEEN DOOHER

LISA RANDALL: WARPED THOUGHTS

- **With Raman Sundrum, offered a way to unify gravity with other forces. Their papers RS-1 and RS-2 are among the most cited in recent physics.**
- **Published *Warped Passages: Unraveling the Mysteries of the Universe's Hidden Dimensions* to describe the excitement of physics to lay readers. It focuses on ideas, not personalities, in part to avoid scaring off women: "If you read a book about all guys, it can do that."**

tors—in this case, 1, 2 and 3—add up to the number itself.) "The project was looking for and seeing if there were any patterns. And there aren't very many. Basically, I always do this. I don't know anything and take on a big project," she says. Nevertheless, Randall's musings on these numbers tied for first place—a fitting precedent for her subsequent mathematical forays into a host of arcane physics fields: technicolor, charged parity symmetry violation, flavor structure and baryogenesis, to mention a few.

Although they did not intend to, Randall and Sundrum ended up using extra dimensions to offer a solution to what is called the hierarchy problem. It can be framed in several ways, but the problem is essentially this: Why is gravity so puny, so many billion on billions of times weaker compared with the other forces—electromagnetism and the weak and strong nuclear forces? Discrepancy in strength makes it impossible to combine gravity with the other three forces, a unification thought to have existed during the early phase of the big bang.

Randall and Sundrum offered a new set of options of what went on in the early universe.

But rather than invoking supersymmetry—a popular solution that argues for the existence of as yet undetected partners to all the known particles—Randall and Sundrum posited that gravity could reside on a different brane than ours, one separated from us by a five-dimensional spacetime in which the extra dimension is 10^{-31} centimeter wide. In this RS-1 model, all forces and particles stick to our three-brane except gravity, which is concentrated on the other brane and is free to travel between them across spacetime, which is warped in a negative fashion called anti–De Sitter space. By the time it gets to us, gravity is weak; in the other brane it is strong, on a par with the three other forces.

String theorists had looked at the idea of confining all forces to a brane and having gravity leak, but they had not worked out the mechanism, says physicist Joseph Lykken of Fermilab in Batavia, Ill. Randall and Sundrum, he remarks, "changed people's thinking about this stuff entirely."

As Randall and Sundrum refined their idea, they realized that if the extra dimension of spacetime were warped in anti–De Sitter fashion, it could be infinitely large and what we observe about gravity could still be true. This model came to be known as RS-2. "Working that out was mind-blowing," Sundrum recalls. "We had reason to be dead scared. In each of these cases, there was a distinct fear of making complete fools of ourselves."

"It was counterintuitive," notes theorist Michael J. Duff of Imperial College London. "It came as a surprise even to those working in extra dimensions that even though the extra dimension is very large, we wouldn't be aware of it. Newton's law would still be an inverse square law, not an inverse cube law, which is what you might naively expect."

It took a while for many physicists to realize what Randall and Sundrum were suggesting, but the time was right for such thinking. Anti–De Sitter space was popping up in some models, branes were thriving, and in 1998 Nima Arkani-Hamed of Harvard, Georgi Dvali of New York University and Savas Dimopoulos of Stanford University (or ADD, for short) had postulated a three-brane within two large extra dimensions.

Some of the recent models, be they RS, elaborations of ADD or others, will be put to the test when the Large Hadron Collider (LHC) at CERN near Geneva fires up in 2007. "If there is any solution to the hierarchy problem, it should be revealed at the energies the LHC will explore," Randall enthuses. Evidence could include gravitons, supersymmetric partners or evanescent, tiny black holes. "Even if we don't know the answer, it should tell us what the answer is," she adds.

In typical fashion, Randall recently took on two things new to her. The first was writing a book about physics, released last month. The second was participating on a task force formed by Harvard president Lawrence Summers after his comments about women in science. She says she is nervous about the reception of the first project and dislikes talking about the second one. "I like to solve simple problems like extra dimensions in space," Randall declares. "Everyone thinks [women in science] is a simpler issue, but it is so much more complicated."

She should know: she was the first female captain of her high school math team, and even though Stuyvesant is famous for cultivating science and math whizzes, she did not find it supportive of girls. "There was one teacher who kept saying that Stuyvesant was much better when it was all boys, even though the two best students in his class were girls, and he liked us both. It was this weird cognitive mismatch," she says. Regarding Harvard and the task force, Randall is reticent: "I just want to see a whole bunch more women enter the field so these issues don't have to come up anymore."

The 43-year-old Randall is now collaborating with Andreas Karch of the University of Washington, investigating some of the cosmological implications of branes and extra dimensions. According to Randall, we may live in a three-brane, but "there are regions beyond the horizon that look really entirely different. And we haven't fully explored them yet."

If her ideas don't feel obvious to you, don't fret. You are in good company. "I often don't understand her," Karch confesses. "When she says things, they don't make sense and I first think 'she is crazy.' But I don't say anything, because she is usually right. Lisa just knows the answer." SA

Ripples *in a* Galactic Pond

Astronomers are coming to realize that the beautiful shapes of galaxies are not merely incidental. They are essential to the galaxies' growth and development

By Françoise Combes

SPIRAL GALAXIES are some of the most beautiful sights in the night sky. Most, such as galaxy NGC 1097 (*above*), have a central rectangle, or "bar," of stars. Others, such as Messier 51 (*opposite page*), do not. Both types of spiral galaxies consist of a flattened, rotating disk of stars, gas and dust. Bars and arms mark comparatively dense regions. Despite recent advances, the nature and origin of these shapes remain mysterious in many ways.

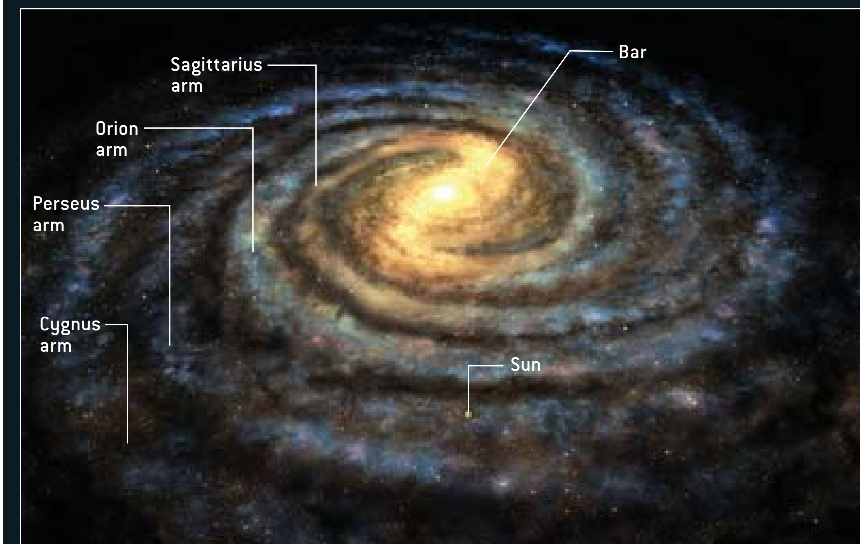
The elegant spiral shape of galaxies is one of the quintessential sights of astronomy. A classic example is the galaxy Messier 51: it resembles a giant cyclone, and one of the first names given to it was the "whirlpool." The brightest stars in this galaxy are confined like pearls on a coiled necklace. Running alongside the strands of stars are dark swaths of dust, which betray the presence of interstellar gas, from which stars are born. In Messier 51 and many other galaxies, the spiral pattern is anchored by an inner ring of stars, but in most the spiral begins in a bar—a long, luminous rectangle of stars. A barred galaxy looks like a spinning lawn sprinkler, where the water flows through a straight tube, emerges at right angles and then curves around.

Most people think of our own galaxy, the Milky Way, as a pure spiral, but astronomers now know it is actually a barred spiral. The evidence, at first indirect, began to accumulate in 1975: stars and gas tracked in the middle of the galaxy did not follow the orbits they would if the spiral pattern reached all the way in. Recent surveys of the sky in near-infrared light, which penetrates the dust clouds that block our view of the galactic core, have revealed the bar directly and dispelled the remaining doubts.

Bars and spirals may appear to be persistent bodies, like pinwheels, but in fact they are dynamic patterns—waves that sweep through the disk of stars, dust and gas, often violently redistributing the material. Observers catch them at one moment in time, as with a stroboscopic light. These waves are one aspect of a general theme of astronomy during the past decade, the realization that seemingly immutable properties of galaxies, such as their shape, change dramatically over time. The best-known shape-shifting process is galactic cannibalism: merging with a neighbor can turn an orderly spiral into a messy, beehivelike elliptical galaxy. But astronomers are coming to appreciate that internal wave processes may be even more important.

MILKY WAY BAR

Astronomers used to think that our own galaxy fell into the pure-spiral category, but today they recognize that it is actually a barred spiral. In the top image from the 2MASS survey, the Milky Way is seen edge-on, the core of the galaxy is at the center, and the colors represent the intensity of light at three near-infrared wavelengths. The galaxy is thicker than expected, and the core is wider and broader on the left than on the right. Both features are signs of the bar, as an artist's conception shows (*bottom*).



Overview/*Galactic Waves*

- Since the 1960s astronomers have realized that the iconic spiral shapes of galaxies are not fixed structures but transient oscillations in the density of material. Stars and gas clouds crowd together and then separate in a self-reinforcing orbital choreography. It is almost as though someone dropped a stone into the galaxy, sending out ripples in slow motion.
- Until fairly recently, however, theorists could not explain crucial aspects of these waves, such as the relative numbers of galaxies having different shapes. It turns out that the missing ingredient in the models was interstellar gas, which exerts an effect out of all proportion to its mass.
- Waves can transport angular momentum, giving matter a chance to clump in the middle of a galaxy; the black hole at the hub may be one beneficiary. Moreover, the waves come and go in a cycle, causing the galaxy to metamorphose from one shape to another. These theoretical findings have been getting support from observations of galaxies both near and far.

PAGE 42: NASA, EUROPEAN SPACE AGENCY, STEVEN BECKWITH *Space Telescope Science Institute* AND HUBBLE HERITAGE TEAM; PAGE 43: STEVE LEE AND DAVID MALIN *Anglo-Australian Observatory*;
THIS PAGE: 2MASS/J. CARPENTER, T. H. JARRETT AND R. HURT (*top*); DON DIXON (*bottom*)

Floppy Disks

ALL SPIRAL GALAXIES, barred or not, rotate; their stars orbit the center in a regular pattern. But they do not rotate rigidly. Their stars do not all go around in unison. In the Milky Way, the innermost stars have whirled around the center thousands of times over the past five billion years, whereas the sun, located about halfway out the disk, has made the trip only about 20 times. The variation in orbital rates is the reason that bars and spirals cannot be fixed structures. If they were, they would quickly coil up in the way a rope wraps around a winch.

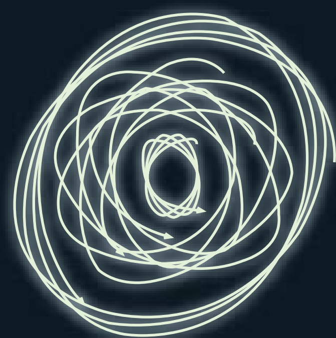
For decades, astronomers wondered what maintained these shapes. In the 1960s the puzzle was partially solved by the theory of density waves, developed by Chia-Chiao Lin and Frank Shu, both then at the Massachusetts Institute of Technology. In this theory, bars and spiral arms are waves of excess density, where stars are crowded together temporarily in a cosmic traffic jam. Stars enter and exit the wave just as cars enter and exit a jam.

The wave involves the synchronization of stellar orbits. The orbit of a star in a galaxy does not look like the orbit of a planet around the sun or a satellite around Earth, because a galaxy is not dominated by a single central body. Although most galaxies have a central black hole, it is a small fraction of the galaxy's mass. The bulk of the mass is spread out, forcing stars to trace a Spirograph-like rosette: an ellipse that does not close on itself but gets offset each time the star completes one revolution [*see box at right*]. The sun, for example, takes about 230 million years to go around its elliptical orbit. In that time, the orbit gets offset by 105 degrees; thus, the sun's ellipse does a full rotation once every 790 million years.

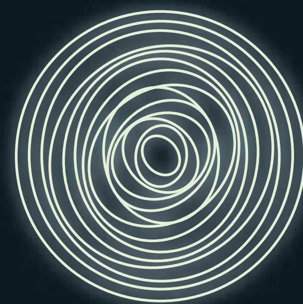
When stars' ellipses rotate at widely different speeds, the galaxy has no wave structure. Stars randomly pass close to one another, but they soon separate again, much as cars briefly bunch together even on a freely flowing highway. A wave occurs when many or all the ellipses rotate at the same rate. In a bar

BRINGING ORDER TO ORBITS

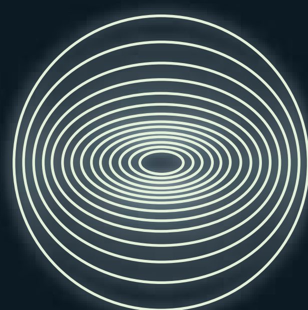
Bar and spiral patterns are thought to represent waves passing through the galaxy. As the wave front enters a region, stars move closer to one another; as it exits, they spread out again. The wave action works not by literally pushing the stars together but by subtly choreographing their orbits.



LOOP-D-LOOP ORBITS are stars' paths through the complex gravitational field of a galaxy. A star trundles around the galaxy on an ellipse, but the ellipse itself is moving. This secondary motion is the raw material for galactic waves.



NO WAVE MOTION occurs when orbits are oriented randomly. (For simplicity, this figure shows only a sample of orbits—those that appear to be closed loops when viewed from a rotating frame of reference.)



BAR WAVE arises when orbits align. Gravity causes the ellipses to move in unison, maintaining their alignment. The wave front is the area of greatest star density, which occurs along the major axis of the ellipses.

SPIRAL WAVE arises when the ellipses move in unison but are not perfectly aligned; each ellipse is slightly skewed compared with its neighbors. The density of stars is highest where the ellipses crowd together.

BARRED SPIRAL pattern arises when orbits near the center of the galaxy are aligned but those farther out are skewed.

ALISON KENDALL

wave, these ellipses are aligned, producing a region of enhanced density along their major axis. In a spiral wave, the orbits are progressively misaligned, so that the region of enhanced density is a curved line.

In short, stellar orbits can move in unison even if the stars themselves do not. What causes the ellipses to move in lockstep? It is a spontaneous gravitational instability. Because gravity in these systems is not a fixed external force but a product of the stars themselves, waves can be self-reinforcing. The process starts when stellar orbits become aligned by chance. Amplified by proximity, the gravity of the stars modifies the rotation speed of the ellipses. Faster ones slow down and slower ones speed up, so they bring themselves into sync. When a star enters the wave, gravity locks it in, but only temporarily; after a while, it becomes unlocked and exits. Stars coming in the other side of the wave ensure that the structure persists.

Breaking the Waves

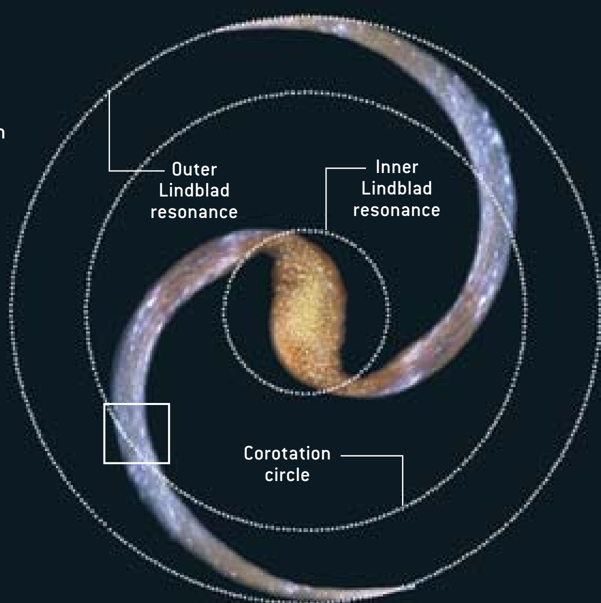
IN THE INNER PARTS of a galaxy, stars move faster than the wave, so they catch up with it; in the outer parts, they move slower, so the wave hits them from behind. In between is a corotation circle where the wave accompanies the stars at the same velocity.

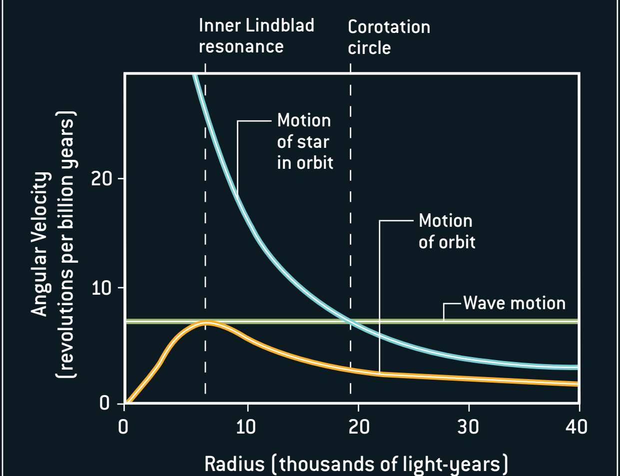
Besides the corotation circle, there are two other special locations in the disk—one inside the circle, the other outside—known as Lindblad resonances, after their discoverer, Swedish astronomer Bertil Lindblad. Stars orbiting at these positions have a certain synchrony with the wave: each time they hit (or are hit by) it, they have reached a particular point in their orbits. This consistency al-

ANATOMY OF A WAVE

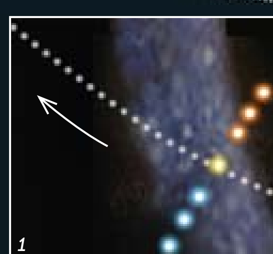
In some ways, bars and spiral arms are like the wave sweeping around a sports stadium. As the wave front passes by, individuals act in a coordinated way: spectators in a stadium stand up and down together; orbits in a galaxy arrange themselves to bring stars closer together. Galactic waves have the added complication that the stars themselves are moving.

SPIRAL ARMS seen by astronomers are a wave caught at one moment in time. In this example, the wave is propagating clockwise. Individual stars also move clockwise, but at a different speed (*insets*). Those in the inner region move faster than the wave. They catch up with it, join the wave for a certain period and then zoom out of it. Stars in the outer region of the galaxy get hit from behind, enter the wave and are left in its wake. The corotation circle defines the boundary between the two regions. The length of the spiral arms is defined by two other circles, marking the positions of so-called Lindblad resonances, where wave and stars move in phase.

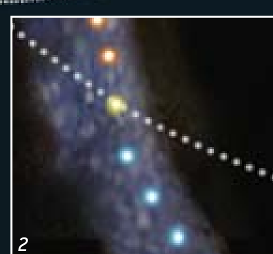




ANGULAR VELOCITY of the wave is constant throughout the galaxy, whereas stars' orbital velocity drops off with distance from the center. The two velocities are equal at the corotation circle. The stars' orbits are not fixed in space; they, too, rotate. Where their rotation rate equals the wave velocity, a Lindblad resonance occurs. (For clarity, the graph shows only one resonance.)



The wave moves clockwise (*bottom right to top left across this box*). So do the stars.



Interior stars (*orange*) move faster than the wave; exterior ones (*blue*), slower.

DON DIXON (*illustrations*); ALISON KENDALL (*graph*) lows the gentle tug exerted by the wave on the star to build up. The Lindblad and other resonances play a distinct role in shaping the orbits and delimiting the density waves. Similar processes define planetary rings [see "Bejeweled Worlds," by Joseph A. Burns, Douglas P. Hamilton and Mark R. Showalter; SCIENTIFIC AMERICAN, February 2002].

The theory of density waves explains how the bar and spiral structure can persist, but not long after Lin and Shu proposed it, it ran into problems of its own. Alar Toomre of M.I.T. observed that waves lose their energy by setting up shock waves in the interstellar gas. Something has to reinvigorate them. One idea was more complex wave motions. Instead of a single wave swooping around the galaxy, multiple waves can travel inward and outward. The corotation circle acts as an interface that can either reflect or transmit these waves, allowing them to gain energy at the expense of the global rotational kinetic energy of the galaxy. Waves can bounce back and forth across the central region, becoming successively amplified—a cosmic echo chamber.

This complex mechanism of wave amplification and reflection was a plausible hypothesis, but the equations were too complex to solve exactly except with severe approximations. So astronomers had to undertake numerical computer simulations—not an easy task in the days of punch-card computing. Early efforts indicated that the extra wave complexity, far from saving the spiral structure, actually accelerated its destruction. A spiral developed at first but rapidly faded away, leaving a bar. Theorists could not find a way to avoid bar formation—and thereby explain the galaxies that lack such a structure—without contradicting other observations.

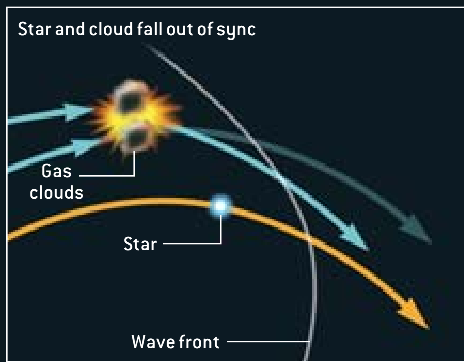
This unsatisfying state of affairs changed completely in the 1980s and 1990s when my colleagues and I included another ingredient in the simulations: gas. Because gas constitutes only a few percent of the mass of spiral galaxies, modelers had neglected it to make the simulations more tractable. But gas has a disproportionate dynamical role. Interstellar gas clouds frequently collide and convert their kinetic energy into shock waves and radiation. Thus becalmed, the clouds are more susceptible to wave instabilities. Stars, in contrast, seldom collide, so they maintain a wider range of relative velocities and put up more resistance to passing waves.

As soon as we included gas, the simulations produced a rich variety of galactic morphologies. The torque exerted by the stellar bar acted as a giant stirrer, continuously driving a spiral structure in the gas. The spiral did not fade away as it had in earlier simulations. Moreover, the gas-rich waves solved a number of other long-standing problems in galactic astronomy. To begin with, they explain the presence of dust lanes on the leading edge of spiral arms. Because of collisions, the gas (mixed with dust) is not in phase with the stars: it loses orbital energy, falls toward the center, and runs ahead of the stars in the spiral arms [*see box above*].

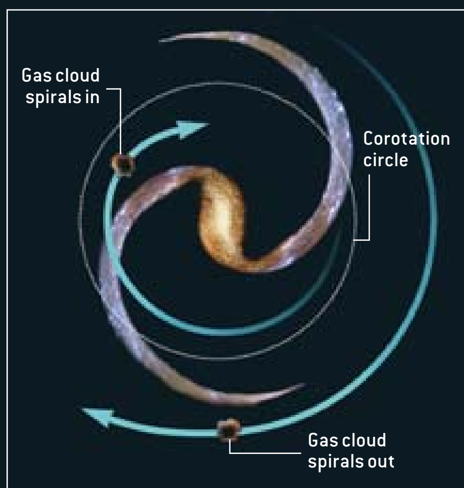
Energy dissipation pulls gas all the

GAS VS. STARS

Galactic waves turn out to be an essential process for redistributing interstellar gas within a galaxy and thereby triggering star formation. The key to this process is that although stars and gas clouds participate in the wave, they respond differently.



IF A STAR AND CLOUD begin on similar orbits, the star holds to its path, but the cloud—by virtue of its enormous size—cannot avoid hitting others of its kind, losing energy and dropping into a lower orbit. Consequently, a galactic wave cannot bring stars and gas clouds into perfect synchrony.



IMBALANCE OF FORCES, arising because the clouds and stars are out of sync, produces a torque on the clouds. A cloud inside the corotation circle outpaces the stars in the bar, so their gravity holds it back—taking away its orbital energy and angular momentum, dragging it down. A cloud outside the circle is pulled forward in its orbit—adding to its orbital energy and angular momentum, hauling it to a higher orbit.

DON DIXON

THE AUTHOR

FRANÇOISE COMBES, an astronomer at the Paris Observatory, is a leading expert on galaxy dynamics. In an age of increasing specialization, she maintains a broad range of interests, from interstellar gas to quasars to dark matter, spanning both observation and theory. Last year Combes was elected to the French Academy of Sciences—the first female astronomer so honored—and currently serves as president of the French National Committee on Astronomy. In her spare time, she does Impressionist-style oil painting.

way to the center within a few rotations of the bar, typically about one billion years. There the gas forms new stars. Thus, the waves explain the persistently high rate of star formation in galactic centers. They could also solve the mystery of what refuels central black holes. Pouring matter into a black hole is not as easy as one might think. Although the tendency of a galaxy is to minimize its gravitational potential energy and therefore to concentrate mass at its center, rotation and the corresponding centrifugal forces counterbalance gravity [see "A Universe of Disks," by Omer Blaes; SCIENTIFIC AMERICAN, October 2004]. For material to fall inward, the angular momentum must be transferred away by galactic-scale torques. Bars and spirals can do just that.

Astronomers have observed that gas falls into black holes in large puffs. Waves could perform the fueling in two steps. First, as the gas descends it reaches a resonance where it is in phase with the bar and hence immune to its torque. The gas piles up in a ring and gives birth to stars. Second, gas and stars inside the ring form their own small bar wave. This mini bar dumps gas into the black hole. In our galaxy, such a bar has been suspected from near-infrared surveys. These waves are therefore more than a beautiful decoration. They allow our galaxy to grow.

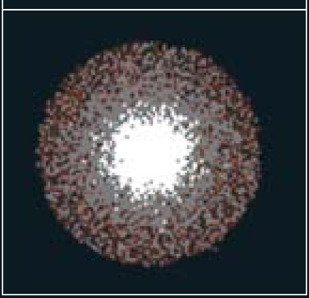
Disbarred

NOT ONLY DOES the bar pull material laterally across the disk, it can elevate it vertically out of the disk. A star can come into a resonance in which its vertical oscillations and its encounters with the bar occur at the same rate (or a multiple thereof). Then the bar can pump up the oscillations like a parent pushing a child on a swing. We found this phenomenon almost serendipitously during three-dimensional computer simula-

A GALAXY REINVENTS ITSELF

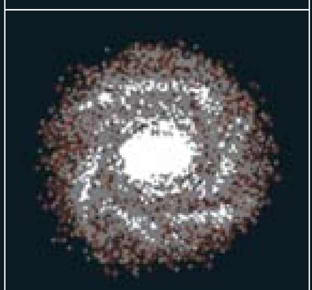
Astronomers used to regard bars and spirals as permanent features of a galaxy but now think that they come and go. The gravitational processes that make a bar ultimately destroy it and then create it anew, as this simulation shows.

START



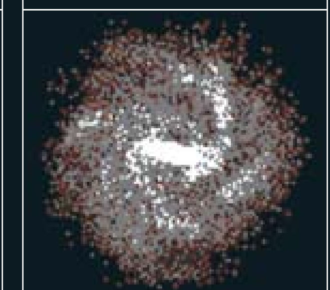
The galaxy is born as an amorphous disk of stars, gas and dust.

2 BILLION YEARS



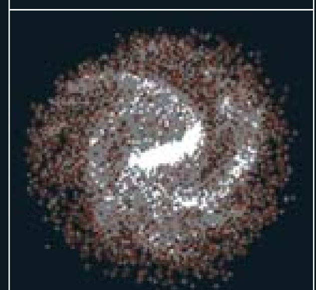
Bar and spiral waves develop. Gas trickles in from intergalactic space and will double the disk mass in 6.5 billion years.

5 BILLION YEARS



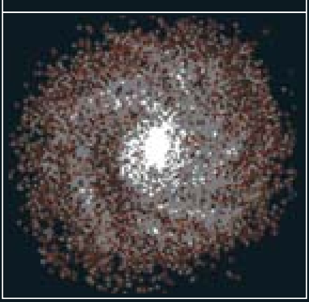
The waves strengthen. The bar sweeps up gas near the core but holds intergalactic gas at bay.

8 BILLION YEARS



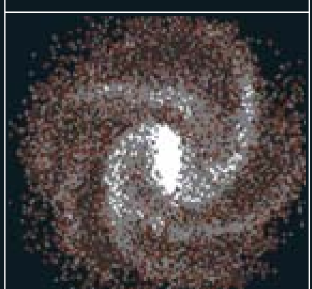
Gas accumulating in the core begins to tear apart the bar.

11 BILLION YEARS



The bar is history. No longer held back by its torques, the intergalactic gas that had been lingering in the galactic outskirts pours in.

14 BILLION YEARS

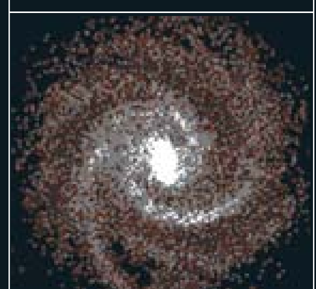


The bar reemerges. Gas infall is crucial: in simulations without it, the bar, once gone, never comes back.

17 BILLION YEARS

As before, the bar starts to peter out.

20 BILLION YEARS



The bar is all but gone. In simulations under different conditions, bars form and vanish more quickly.

ALISON KENDALL

DISTANT GALAXIES, representing an earlier epoch of cosmic history, appear to have the same basic bar and spiral shapes as closer ones—supporting the idea that these shapes have been generated by internal wave processes. These galaxies are so far away that their detailed structure is hard to see, so researchers have debated the relative abundance of different shapes.

tions. The disk of stars thickened considerably in the region of resonance, giving the inner regions of the galaxy a boxy or a peanut shape. These findings explained the odd shapes that telescope observers had been seeing for a decade or longer.

Ironically, by pulling matter into the core of a galaxy, a bar can destroy itself. The accumulated mass scatters stars and prevents them from orbiting regularly, as a bar requires. The bars that seemed too robust when only the stellar component was taken into account are quite fragile when the gas is included. But if so, what accounts for the large number of barred galaxies that astronomers see? In visible images, two thirds of galaxies have bars, and near-infrared surveys in 2002 raised the estimate to three quarters. The logical conclusion is that bars form, fall apart, and form once again in a continuous cycle.

Explaining the reformation of bars is a challenge. The galaxy has to evolve considerably from the conditions that initially destroyed the bar. In particular, the orbits have to settle back down again into a regular pattern with low relative velocities. One way to do that is for the galaxy to accrete a large amount of intergalactic gas. As the gas clouds fall in, they collide, lose energy and regularize their orbits. Their high initial angular momentum slows their descent and gives the bar a chance to regroup. The amount of gas required is truly huge: to reform bars at the right frequency, a typical galaxy has to double its mass over 10 billion years. Astronomers now know that intergalactic space contains reservoirs equal to the task [see "Our Growing, Breathing Galaxy," by Bart P. Wakker and Philipp Richter; SCIENTIFIC AMERICAN, January 2004].

One way to check this model is to look back in time. The Hubble Space Telescope is powerful enough to discern the shape of earlier generations of galaxies. The first attempt to do so, conducted by Sidney van den Bergh of Herzberg Institute of Astrophysics in Victoria, British Columbia, and his collaborators between 1998 and 2002, concluded that bars were once much less frequent than they are today. This result was very surprising; it cast doubt not just on the gas-infall model but on the entire theory of bar waves. Early galaxies were more gaseous and less centrally concentrated, so if anything bars should have been more frequent. But Shardha Jogee of the University of Texas at Austin and her colleagues recently explained van den Bergh's preliminary results as observational bias. Bars are hard to recognize in distant galaxies. Correcting for this effect reveals that they were as prevalent in the past as they are now, suggesting that they are destroyed and rebuilt in a steady state.

In short, it appears that galaxies are not born with a given shape, barred or unbarred. They metamorphose. If three quarters of galaxies are barred, a typical galaxy must spend three quarters of its life barred. During this time, the bar prevents fresh gas from entering the central part of the galaxy. The gas accumulates in the outer region and, after the bar self-destructs, pours in and rejuvenates the galaxy.

The other, better-known way to assemble mass in galaxies is cannibalism, or successive galactic mergers. This process, though also very important, is destructive. A major merger wipes out the disk and leaves an elliptical galaxy in its place. Only a minority of galaxies have been transformed so thoroughly. Gentle gas accretion from intergalactic space allows galaxies to grow while maintaining their shape. Waves distribute the fresh material and save the galaxies from slowly winding down. They keep the universe a vibrant place. SA

MORE TO EXPLORE

Galactic Rings. Ron Buta and Françoise Combes in *Fundamentals of Cosmic Physics,* Vol. 17, pages 95–281; 1996. Available at **http://nedwww.ipac.caltech.edu/level5/Rings/frames.html**

Gas Accretion on Spiral Galaxies: Bar Formation and Renewal. Frédéric Bournaud and Françoise Combes in *Astronomy & Astrophysics,* Vol. 392, No. 1, pages 83–102; September 2002. Preprint available at **arxiv.org/abs/astro-ph/0206273**

Redesigning the Milky Way. William H. Waller in *Sky and Telescope,* Vol. 108, No. 3, pages 50–56; September 2004.

Secular Evolution versus Hierarchical Merging. Françoise Combes in *Penetrating Bars through Masks of Cosmic Dust.* Edited by D. L. Block, I. Puerari, K. C. Freeman, R. Groess, E. K. Block. Springer, 2004. **astro-ph/0406306**

NEW BULL'S-EYES

A familiar class of cell-surface receptors turns out to offer an array of fresh targets that could yield new treatments for disorders ranging from HIV infection to obesity

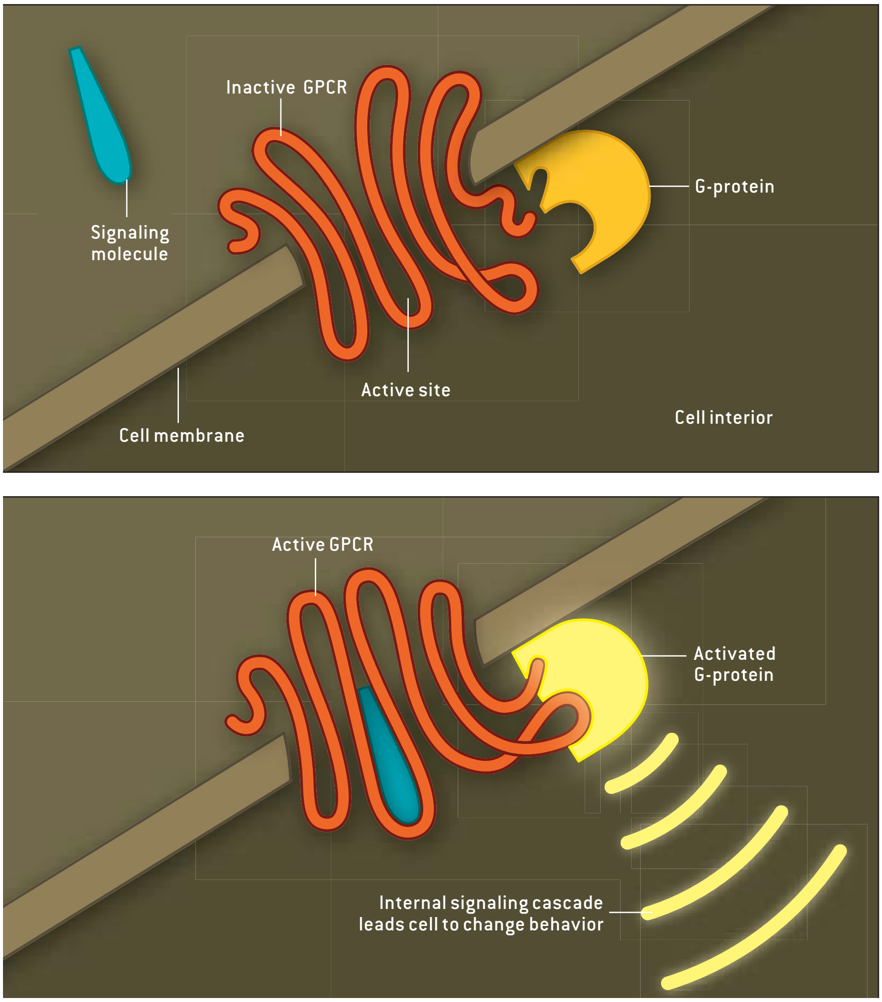
FOR DRUGS

GIACOMO MARCHESI

By Terry Kenakin

An amazing fraction—roughly half—of all the medicines prescribed today have a striking commonality. At the molecular level, they act on the same type of target: a serpentine protein that weaves seven times through the membrane that envelops the cell. External parts of each serpent serve as an antenna for molecular signals approaching the cell, and internal parts trigger the cell's responses to such cues, beginning with the activation of a signal processor called a G-protein. The serpents themselves are thus known as G-protein coupled receptors, or GPCRs.

As a group, GPCRs show far more versatility than any other class of cell-surface receptor. For instance, the natural molecules to which GPCRs respond range in size from neurotransmitters that are only a few times as massive as a single carbon atom all the way up to proteins 75 times larger than that. Moreover, GPCRs participate in just about every bodily function that sustains life, from heartbeat and digestion to breathing and brain activity. The drugs that target these receptors are equally diverse. The list includes blood pressure reducers (such as propranolol), stomach acid suppressors (such as ranitidine), bronchodi-



G-PROTEIN COUPLED RECEPTOR (GPCR), which snakes through the cell membrane seven times, typically issues no messages into a cell (*top*) until a signaling molecule, such as a hormone or a neurotransmitter, binds to a region called the active site. This binding (*bottom*) causes the receptor to activate a molecule called a G-protein, which triggers a series of intracellular interactions culminating in a change in the cell's behavior. New insights into the functioning of GPCRs suggest novel avenues for treating disease.

JEN CHRISTIANSEN

lators (such as albuterol) and antidepressants (such as paroxetine). The disorders these medicines treat include hypertension, congestive heart failure, ulcer, asthma, anxiety, allergy, cancer, migraine and Parkinson's disease.

Incredibly, today's GPCR-targeting drugs all work in one of two ways—they either attach to the "antenna" region of the receptor (also known as the active site) and mimic the effect of the natural neurotransmitter, hormone or other molecule that normally signals through the GPCR, or they interfere with a natural signaler's ability to act on the antenna. Over the past 15 years, a technological revolution has furnished investigators with new eyes with which to see GPCRs at work. Consequently, other ways of manipulating GPCR activity have emerged and are beginning to be mined for drug discovery. In other words, despite the wealth of medicines already known to act on these fascinating receptors, many more may lie ahead. The search for such pharmaceuticals is still in early stages, but a few agents, including some for HIV infection (the cause of AIDS), are now advancing through human trials.

Shape Matters

UNTIL ABOUT 10 YEARS AGO, pharmaceutical researchers thought that to influence the activity of GPCRs they would have to aim drugs at a receptor's active site. During the body's normal operation, a neurotransmitter or other information-bearing molecule (or "ligand") at the cell's outer surface essentially plays the "key" to the active site's "lock." So a substance that plugged the lock could prevent unwanted signaling through the receptor by any key and serve as an inhibitor. Conversely, something that mimicked the natural ligand could essentially open the lock and therefore take the place of the natural key if it were missing.

Scientists thought as well that the best way to evoke a selected physiological response was to choose a compound that interacted with a specific form of a receptor but ignored other variants. The neurotransmitter norepinephrine, for instance, activates two types of GPCR, called alpha and beta adrenoceptors, of which the first has four subtypes and the second has three. These various receptors, in turn, govern different life-sustaining processes. In the heart, beta$_1$ adrenoceptors quicken the heart rate and increase the force of each beat; in the lungs, beta$_2$ adrenoceptors widen the air passageways. Hence, to open constricted airways without unwanted

Overview/*New Drug Targets*

- Proteins called G-protein coupled receptors (GPCRs), which sit on the cell surface, convey signals from hormones and the like into cells by activating G-proteins—signal processors residing just under the cell membrane.
- About half of all pharmaceuticals on the market act on GPCRs, binding to the sites normally targeted by the body's own extracellular signalers.
- In the past 10 years, researchers have learned that GPCR activity can also be modulated by compounds that bind to other sites on GPCRs. This discovery opens new possibilities for treating cancer and other major disorders.

effects on the heart, pharmaceutical makers might seek an agent that mimicked norephinephrine's ability to stimulate $beta_2$ adrenoceptors but without binding to $beta_1$ adrenoceptors.

Many medicines do, in fact, function as inhibitors or agonists (mimics) by interacting with the active site of a specific GPCR. But an emerging drug development strategy has to do with the "allosteric" nature of GPCRs: the shape of one part of the receptor can affect the conformation, and thus the activity, of a distant part.

GPCRs constantly adopt somewhat different shapes, essentially sampling a library of conformations. When a natural signaling molecule binds to the active site, it stabilizes the arrangement that activates G-proteins. But it turns out that certain molecules, known as allosteric modulators, can bind elsewhere to influence form and activity. Some stabilize GPCR conformations that promote signaling, whereas others maintain shapes that impede it (say, by burying the active site so that it becomes inaccessible to its natural ligand).

The implications are profound. The entire receptor can theoretically offer binding sites, at any one of which a diminutive molecule might stabilize a shape that yields some biological effect. This property greatly enlarges the vista for therapeutic modification of GPCR function.

AIDS researchers are among those actively pursuing the potential of allosteric modulators, trying to find ones able to block HIV from infecting cells. Biologists have long known that the virus attacks cells called helper T lymphocytes by adhering to a cell-surface protein named CD4. But in the mid-1990s they learned that this protein does not act alone.

To enter cells, the virus also has to bind to an additional anchor: a GPCR known as CCR5 (or, in late-stage infection, a GPCR called CXCR4). Normally CCR5 responds to any of three chemokines, natural signals that can attract immune system cells to a site of infection. Unfortunately, it also offers a hook for the virus's coat protein (gp120). Indeed, CCR5 now appears to be a central player in HIV infection; people whose genetic makeup causes them to lack a functional form tend to be extraordinarily resistant to HIV.

Several allosteric modulators that hold CCR5 in a shape inimical to binding by HIV's gp120 have already reached human trials. Blocking the gp120-CCR5 interaction by delivering these tiny drugs is an achievement comparable to, in a geophysical analogy, an island the size of Fiji preventing two Australias from coming together. In more allegorical terms, if such drugs work they will be the David that smites Goliath.

Beyond Volume Control

THE EFFECTS produced by GPCRs depend not only on the extracellular molecules that bind to them but also on how many copies of the receptors are accessible on the cell surface. As might be expected, when extracellular signalers bind to many copies of a receptor, the cell receives a "louder" message and undergoes a more pronounced behavioral change than when few copies of the receptor are bound. But the number of receptors can do more than control "volume." It can actually influence which of several G-protein species become stimulated and can thereby lead to activation of distinct pathways (cascades of molecular interactions) inside a cell.

G-proteins come in four major forms, with subtypes in each class. Each has a different proclivity for working with any given GPCR, and for its part a GPCR may not be equally active toward all G-proteins. A scant supply of a given receptor might therefore result in activation of only the most sensitive G-protein, whereas a greater abundance might lead to responses by multiple G-proteins, eliciting a different cellular behavior.

Accordingly, a GPCR can no longer be seen as simply a toggle switch turned on by a hormone or neurotransmitter and turned off when the natural signal diffuses away from its binding site. It is a much more sophisticated information-processing unit.

Theoretically, the variety of response patterns a given GPCR can generate will

MARKETED DRUGS ACTING ON GPCRs

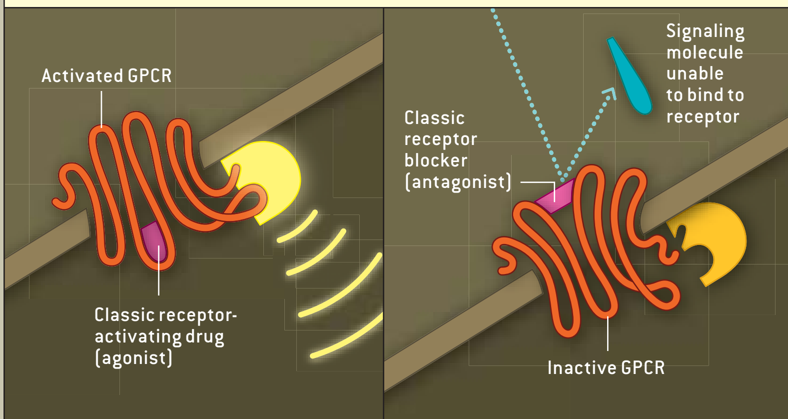
The products listed below are just a sampling of marketed compounds targeting GPCRs; they act on various receptors.

BRAND NAME (GENERIC NAME) AND MAKER	EFFECT
Allegra (fexofenadine) *Aventis*	Blocks histamine action, to control allergic responses
Duragesic (fentanyl) *Janssen*	Relieves pain
Flomax (tamsulosin) *Boehringer Ingelheim*	Eases symptoms of enlarged prostate
Imitrex (sumatriptan) *GlaxoSmithKline*	Eases migraines
Lopressor (metoprolol) *Novartis*	Lowers blood pressure
Oxycontin (oxycodone) *Purdue*	Relieves pain
Pepcid (famotidine) *Merck*	Counteracts stomach acid
Phenergan (promethazine) *Wyeth*	Blocks histamine
Serevent (salmeterol) *GlaxoSmithKline*	Opens airways
Singulair (montelukast) *Merck*	Controls airway inflammation
Sudafed (pseudoephedrine) *Pfizer*	Eases nasal congestion
Zantac (ranitidine) *GlaxoSmithKline*	Counteracts stomach acid
Zyrtec (cetirizine) *Pfizer*	Blocks histamine
Zyprexa (olanzapine) *Eli Lilly*	Eases symptoms of various psychoses

MANY AVENUES OF ATTACK

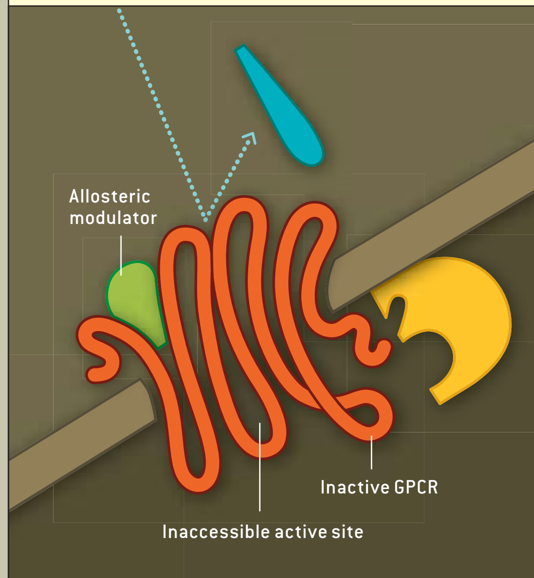
Most drugs on the market target the active site of some cell-surface receptor, and many aim for the active site of a specific GPCR (*below*). Yet molecules acting at regions outside the active site can also influence GPCR activity (*right*). Recent studies encourage hope that small molecules targeted to those additional sites could be administered to activate or quiet GPCRs involved in various diseases.

STANDARD DRUGS



Pharmaceuticals that act on GPCRs today usually fit into the active site and either mimic the action of a natural signaling molecule (*left*) or prevent the native signaler from docking with the receptor and thus from acting on a cell (*right*).

ALLOSTERIC MODULATORS



These agents stabilize the conformation of a GPCR in a way that enhances (*not shown*) or decreases (*above*) the receptor's activity, such as by causing the active site to become inaccessible to a signaling molecule.

depend on both the range of ligands it can detect and the mix of G-protein species it can activate. If, for example, a receptor can detect any of three different signals and can activate any one, two, three or all four of the major G-proteins (as is known to be the case for the GPCR responsive to thyrotropin, the pituitary hormone that stimulates the thyroid gland), the receptor gains the theoretical capacity for dozens of forms of behavior, each seen at one time or another. If it were only a toggle switch, it could have only two.

Research also suggests that drugs can take advantage of this complexity in receptor function. Distinct substances might cause a receptor to hold different biologically active shapes, each of which might interact with a distinct G-protein or G-protein combination, triggering the activity of divergent intracellular paths. Agents that can cause cells to increase or decrease the quantity of receptors at the surface, rather than altering GPCR activity per se, should be valuable as well.

This last strategy could be pursued for combating HIV. One problem that might arise from relying on allosteric modulators to prevent the viral coat protein from finding its docking site on CCR5 is that the virus mutates rapidly. This mutability could lead to the creation of a coat protein that would bind quite well to an allosterically altered CCR5. A plausible way to avert this threat would be to banish the receptor from the cell surface, thereby denying the virus its point of attack.

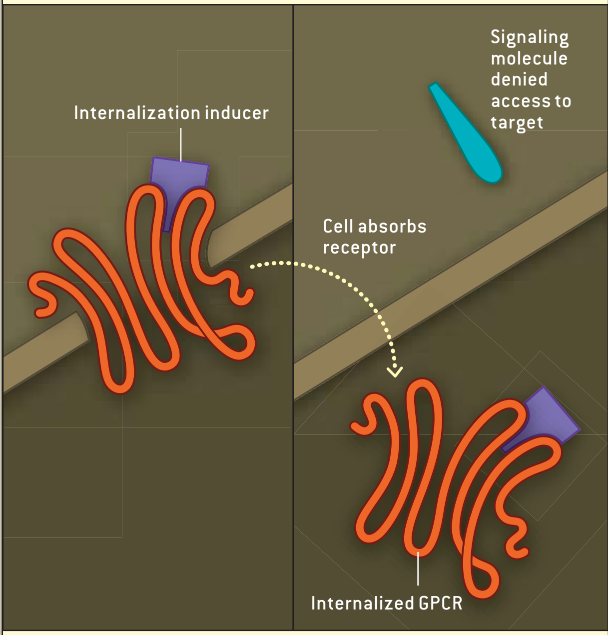
Like all other GPCRs, CCR5 is synthesized endlessly by the cell, stationed at the surface and then drawn back inside for degradation or recycling. And certain chemokines are known to promote CCR5 internalization. This observation raises the possibility of finding pharmacological agents that would not only accelerate the removal of CCR5 from the cell surface but would also serve as therapies to which the virus could not adapt. After all, no change that HIV could undergo would enable it to latch onto CCR5 if that receptor were removed from the cell surface.

Stopping Renegade Signaling

BEYOND BEING controllable by allosteric modulators, GPCRs may exhibit another biologically important behavior, known as constitutive activity—that is, sometimes they activate G-proteins even without being "told" to do so by a bound ligand. As is true in other forms of GPCR functioning, this one arises from a particular shape in the receptor's repertoire. The conformation, however, is one that the receptor rarely takes. Under normal circumstances, the number of molecules that adopt it will therefore be quite small, and so they will have little effect on the cell's overall behavior and will be hard to detect. But if the constitutively active

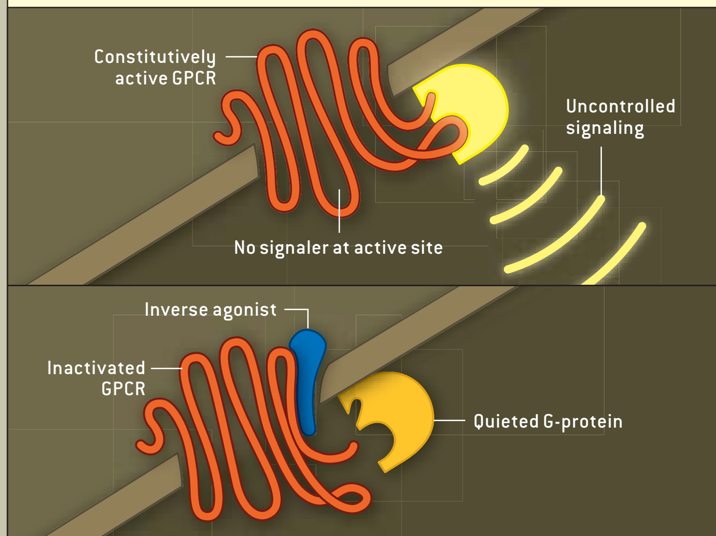
JEN CHRISTIANSEN

INDUCERS OF INTERNALIZATION



Internalization inducers cause receptors to leave the cell surface. They thereby block signal delivery into the cell.

INVERSE AGONISTS



Inverse agonists take aim at "constitutively" active GPCRs. Such receptors improperly behave as if they are bound by a stimulant even when they are not (*top*). Attachment of the inverse agonist quiets the signaling (*bottom*). Cancer cells often have high numbers of constitutively active receptors that theoretically could fuel uncontrolled cell division; inverse agonists, therefore, could potentially become a new form of cancer treatment.

receptors become sufficiently abundant, their combined signaling can exert a powerful influence.

The consequences become especially dramatic in illnesses such as viral infection or cancer, which may advance by inducing one or another receptor to behave in ways that promote the disease. In a form of pancreatic cancer, for example, the receptor for a hormone called vasoactive intestinal peptide (VIP) might be such a bad actor.

In a normal pancreatic cell that displays this GPCR, activation of the receptor by VIP supports cell division. But in people afflicted by this malignancy, the receptor becomes overabundant and the versions that act independently, without need for VIP stimulation, become correspondingly numerous—together they acquire the capacity for driving unconstrained proliferation of tumor cells. Oncologists have long been familiar with destructive constitutive activity in certain non-GPCR receptors, notably one called ras. In those cases, though, mutations in the receptor, rather than an aberrant plentitude of receptors, account for the behavior.

Standard pharmaceuticals cannot quell the cellular misbehavior triggered by constitutively active receptors. A conventional receptor stimulant, or agonist, would only cause more receptors to hold an active shape, to the patient's detriment. A conventional receptor blocker, or antagonist, might prevent natural signals from activating receptors, but such agents will have no effect on receptors that need no outside prompting in order to act. Thus, a new kind of drug is required, one that forces constitutively active GPCRs to maintain an inactive shape.

Such agents, called inverse agonists, might one day constitute an important new form of cancer therapy. They are also being eyed for treating obesity. In this realm, the envisioned targets include the receptor for ghrelin, a recently discovered hormone produced chiefly by the stomach, and the H_3 subtype of histamine receptor; both receptors appear to participate in the brain's regulation of appetite.

Exploiting Phantom Genes

AT LEAST ONE OTHER FORM of GPCR behavior remains to be mined for drug discovery. Cells sometimes mix and match proteins, forming complexes that function as receptors having sensitivities not seen in the individual components. In the most extreme form of this activity, the cell gains a responsiveness to a signal

THE AUTHOR

TERRY KENAKIN has been applying concepts of receptor pharmacology to drug discovery programs for almost three decades, most recently as a principal research investigator at the pharmaceutical firm GlaxoSmithKline. He has written six books on pharmacology and is co-editor in chief of the *Journal of Receptors and Signal Transduction.*

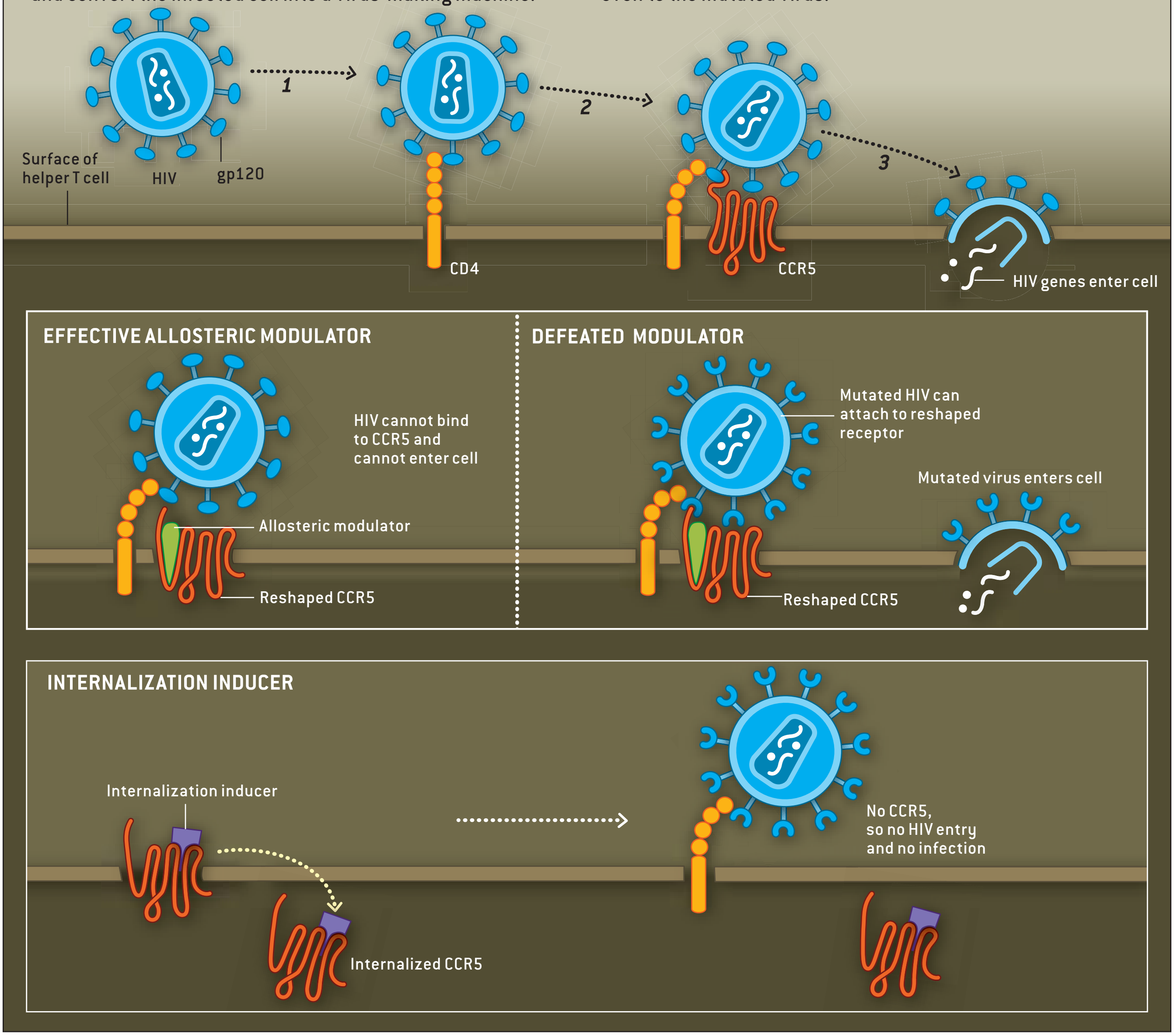
it would otherwise ignore. Individual proteins have their blueprints in specific genes, but these combination receptors have no corresponding single blueprint (from which their behaviors might be predicted), so they might be thought of as products of "phantom" genes.

In some cases, the novel receptor is a complex consisting of two or more GPCRs. In other cases, it consists of a GPCR and a co-protein—one that is not itself a receptor but gives the receptor an altered set of properties. The receptor for a hormone called amylin seems to be of this type. Released by the same pancreatic cells that secrete insulin, amylin modulates the effects of insulin on other cells, but efforts to identify a single protein that serves as its receptor have failed. What is more, analyses of the recently completed human genome se-

A SCENARIO FOR HIV TREATMENT

At least two new kinds of drugs aimed at CCR5, a G-protein coupled receptor, could one day fight HIV. This virus infects helper T cells of the immune system (*top*). To enter the cells, HIV first adheres to a molecule called CD4 (*1*), an attachment that facilitates the binding of a viral protein (gp120) to CCR5 (*2*). This connection induces the cells to take up the attached virus (*3*), enabling the pathogen to release its genes and convert the infected cell into a virus-making machine.

Allosteric modulators that change the shape of CCR5, making it unrecognizable to gp120 (*top left panel*), are already in clinical trials. Unfortunately, though, viral mutations could eventually alter gp120 so that it became able to attach to the allosterically altered CCR5 and enter T cells (*top right panel*). But an internalization inducer could potentially remove CCR5 from the cell surface (*bottom panel*), making it unavailable even to the mutated virus.



JEN CHRISTIANSEN

SOME EARLY PROSPECTS FOR NEW DRUGS

For the most part, investigators are only beginning to devise drugs that influence GPCRs in new ways. But many such agents can be expected to enter pharmaceutical pipelines in the years ahead.

DISORDER	DRUG TYPE	DRUG NAME (MAKER)	TARGET GPCR	STAGE OF DEVELOPMENT
HIV infection	Allosteric modulator	Aplaviroc (GlaxoSmithKline); Vicriviroc (Schering-Plough); UK-427, 857 (Pfizer)	CCR5 (binding by HIV helps the virus enter cells)	All are in phase II or III human trials (early or advanced tests of efficacy)
	Allosteric modulator	AMD3100 (AnorMED)	CXCR4 (this receptor, too, can help HIV enter cells)	In phase III human trials
	Internalization inducer	PSC-RANTES (several institutions)	CCR5	Theoretical
Diabetes	Binder of a receptor formed by two molecules	Symlin (Amylin)	Complex consisting of a protein called RAMP and the GPCR for calcitonin (a thyroid hormone)	Gained U.S. approval in March 2005
Obesity	Inverse agonist	None yet	Constitutively active ghrelin receptor in central nervous system	Theoretical
	Inverse agonist	None yet	Constitutively active histamine H_3 receptor in central nervous system	Theoretical
Cancer	Inverse agonist	None yet	Various constitutively active GPCRs	Theoretical

quence indicate that no gene for such a receptor exists. On the other hand, a complex consisting of the GPCR for the thyroid hormone calcitonin plus a nonreceptor protein called RAMP (receptor activity–modifying protein) responds strongly and selectively to amylin. Apparently RAMP makes the calcitonin receptor "multilingual"—that is, the receptor is reactive to calcitonin if cells lack RAMP, but it is sensitive to amylin if cells contain RAMP.

A different co-protein, called RCP (receptor component protein), induces the calcitonin receptor to obey signals from yet another substance—CGRP (calcitonin-gene-related peptide), a small protein that is the most potent known dilator of blood vessels. This conversion becomes valuable during pregnancy, when blood levels of the dilating peptide soar and RCP levels rise in the uterine wall. As RCP concentrations increase, so do the numbers of calcitonin receptors that become sensitive to the dilator, a change that enhances the blood supply to tissues important for childbirth.

Because co-proteins affect GPCR activity, they might themselves prove valuable as drug targets. One intriguing target is modulin, a co-protein that binds to receptors for serotonin. In the brain, serotonin is most famous as a mood-enhancing neurotransmitter. (Prozac and related antidepressants work by increasing the brain's serotonin levels.) Outside the brain, it acts on the intestines and blood vessels. Perhaps unsurprisingly, serotonin receptors have numerous subtypes, and modulin further tunes the effects of serotonin on particular cells by altering a subtype's sensitivity to it. A drug that mimicked or inhibited modulin, then, could in theory increase or decrease the responsiveness of specific serotonin receptors on specific cell types and might thereby be beneficial in realms ranging from schizophrenia to gastrointestinal function.

Researchers estimate that of the estimated 650 human GPCR genes, about 330 might be blueprints for receptors well worth targeting by drugs. In the past, pharmaceutical scientists would have focused strictly on developing old-fashioned inhibitors or agonists aimed at the receptors' active site. But if many GPCRs offer multiple sites of attack, the opportunities for devising new therapies explode. Because it can take 15 or even 20 years to discover a drug, explore its actions, evaluate its safety, and get it to market, detailed forecasts would be premature. Nevertheless, the new insights into how GPCRs are controlled suggest that these old standbys still have exciting tales to tell. SA

MORE TO EXPLORE

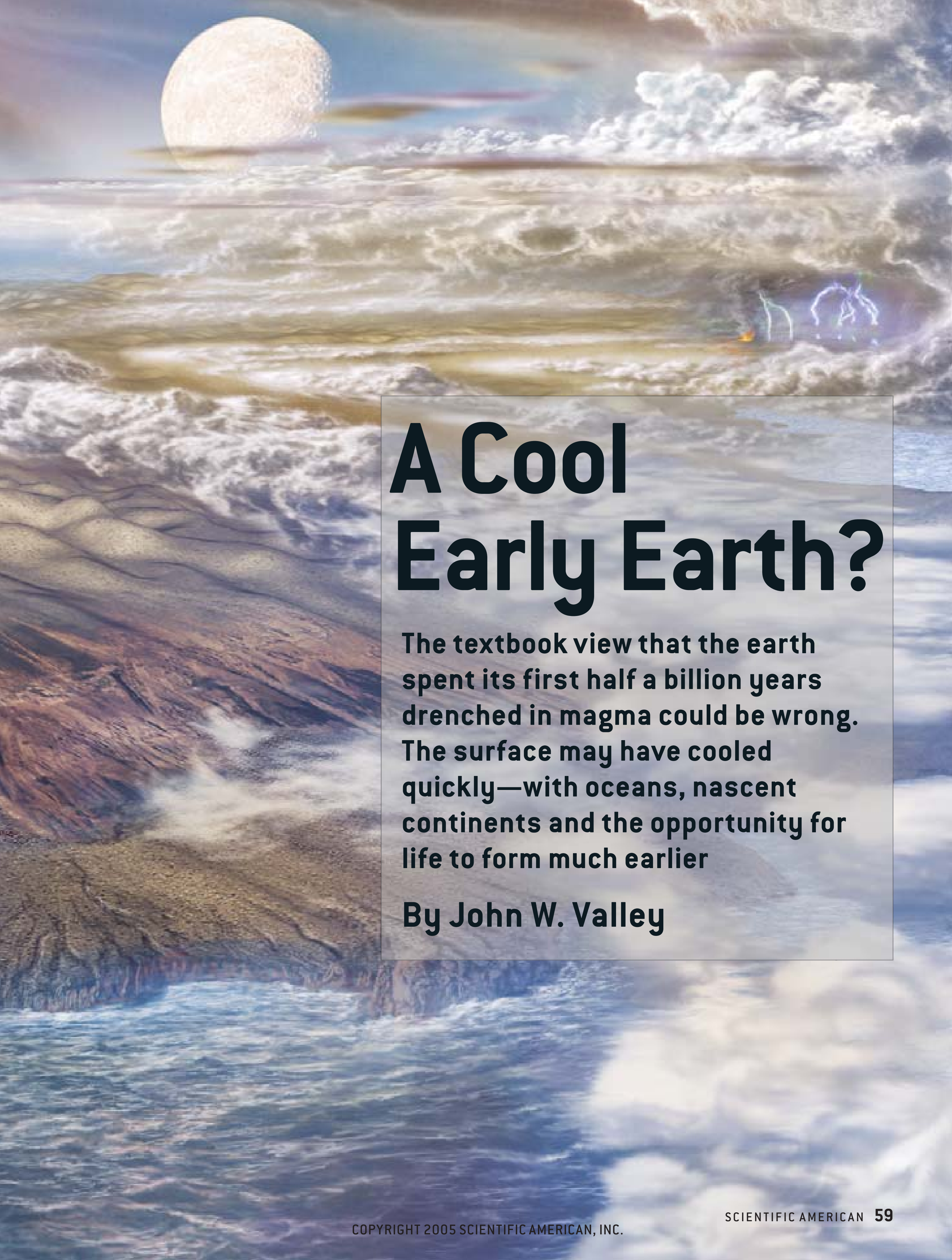
Novel GPCRs and Their Endogenous Ligands: Expanding the Boundaries of Physiology and Pharmacology. A. Marchese, S. R. George, L. F. Kolakowski, K. R. Lynch and B. F. O'Dowd in *Trends in Pharmacological Sciences,* Vol. 20, No. 9, pages 370–375; September 1, 1999.

Drug Discovery: A Historical Perspective. J. Drews in *Science,* Vol. 287, pages 1960–1964; March 17, 2000.

G-Protein-Coupled Receptor Interacting Proteins: Emerging Roles in Localization and Signal Transduction. A. E. Brady and L. E. Limbird in *Cellular Signalling,* Vol. 14, No. 4, pages 297–309; April 2002.

A Pharmacology Primer: Theory, Application, and Methods. Terry Kenakin. Academic Press (Elsevier), 2003.

NEW VIEW of the young earth covered in oceans of liquid water as early as 4.4 billion years ago contrasts sharply with the hot, hostile world typically depicted in textbooks.



A Cool Early Earth?

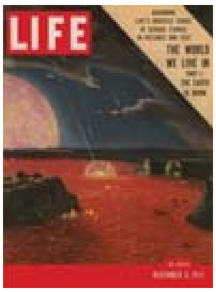
The textbook view that the earth spent its first half a billion years drenched in magma could be wrong. The surface may have cooled quickly—with oceans, nascent continents and the opportunity for life to form much earlier

By John W. Valley

In its infancy, beginning about 4.5 billion years ago,

the earth glowed like a faint star. Incandescent yellow-orange oceans of magma roiled the surface following repeated collisions with immense boulders, some the size of small planets, orbiting the newly formed sun. Averaging 75 times the speed of sound, each impactor scorched the surface—shattering, melting and even vaporizing on contact.

Early on, dense iron sank out of the magma oceans to form the metallic core, liberating enough gravitational energy to melt the entire planet. Massive meteorite strikes continued for hundreds of millions of years, some blasting craters more than 1,000 kilometers in diameter. At the same time, deep underground, the decay of radioactive elements produced heat at rates more than six times greater than they do today.



OLD VIEW of a hot young earth: *Life* magazine, December 8, 1952.

These fiery conditions had to subside before molten rock could harden into a crust, before continents could form, before the dense, steamy atmosphere could pool as liquid water, and before the earth's first primitive life could evolve and survive. But just how quickly did the surface of the earth cool after its luminous birth? Most scientists have assumed that the hellish environment lasted for as long as 500 million years, an era thus named the Hadean. Major support for this view comes from the apparent absence of any intact rocks older than four billion years—and from the first fossilized signs of life, which are much younger still.

In the past five years, however, geologists—including my group at the University of Wisconsin–Madison—have discovered dozens of ancient crystals of the mineral zircon with chemical compositions that are changing our thinking about the earth's beginnings. The unusual properties of these durable minerals—each the size of the period in this sentence—enable the crystals to preserve surprisingly robust clues about what the environment was like when they formed. These tiny time capsules bear evidence that oceans habitable to primitive life and perhaps continents could have appeared 400 million years earlier than generally thought.

Cooling Down

SINCE THE 19TH CENTURY, scientists have attempted to calculate how quickly the earth cooled, but few expected to find solid evidence. Although magma oceans initially glowed at temperatures exceeding 1,000 degrees Celsius, a tantalizing suggestion of a more temperate early earth came from thermodynamic calculations showing that crust could have solidified on the surface within 10 million years. As the planet hardened over, the thickening layer of consolidated rock would have insulated the exterior from the high temperatures deep within the interior. If there were suitably quiescent periods between major meteorite impacts, if the crust was stable, and if the early hothouse atmosphere did not trap too much heat, surface temperatures could have quickly fallen below the boiling point of water. Furthermore, the primitive sun was fainter and contributed less energy.

Still, for most geologists, an undisputed fiery birth and scant clues in the geologic record seemed to point instead to a prolonged ultrahot climate. The oldest known intact rock is the four-billion-year-old Acasta gneiss in Canada's Northwest Territories. This rock formed deep underground and bears no information about conditions on the surface. Most investigators assumed hellish conditions at the planet's surface must have obliterated any rocks that formed earlier. The oldest rocks known to have originated underwater (and thus in relatively cool environs) did not form until 3.8 billion years ago. Those sediments, which are exposed at Isua in southwestern Greenland, also contain the earliest evidence of life [see "Questioning the Oldest Signs of Life," by Sarah Simpson; SCIENTIFIC AMERICAN, April 2003].

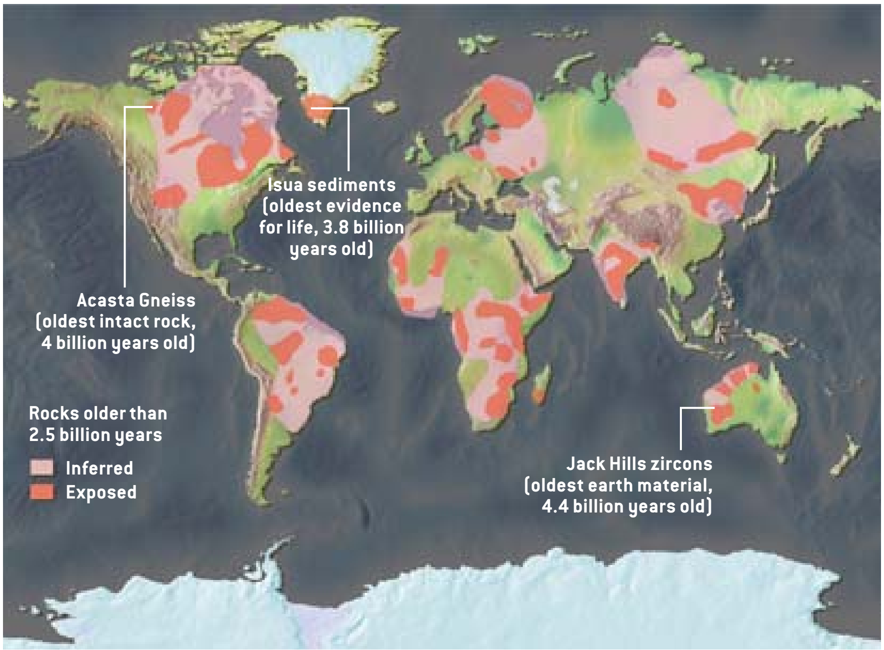
Single crystals of zircon began to add new information about the early earth in the 1980s, when a few rare grains from the Jack Hills and Mount Narryer regions of Western Australia became the most ancient terrestrial material known at that time—the oldest dating back almost 4.3 billion years. But the information these zircons carried seemed ambiguous, in part because geologists were unsure of the identity of their parent rock. Once formed, zircon crystals are so durable that they can persist even if their parent rock is exposed at the surface and destroyed by weathering and erosion. Wind or water can then trans-

Overview/*Zircon Time Capsules*

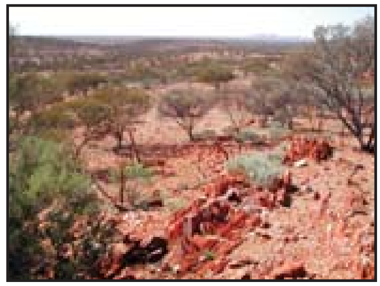
- Geologists have long thought that the fiery conditions of our planet's birth 4.5 billion years ago gave way to a more hospitable climate by about 3.8 billion years ago.
- Now tiny crystals of the mineral zircon, which retain clear evidence of when and how they formed, suggest that the earth cooled far sooner—perhaps as early as 4.4 billion years ago.
- Some ancient zircons even bear chemical compositions inherited from the cooler, wet surroundings necessary for life to evolve.

DON DIXON (*preceding pages*); TIME LIFE PICTURES/GETTY IMAGES (*this page*)

OLDEST PIECES OF THE PLANET



Ancient rocks older than 2.5 billion years crop out or lie just underneath the soil in many spots around the globe (*red*) and are probably hidden below younger rocks across even broader regions (*pink*). Zircon crystals as old as those discovered in the Jack Hills of Western Australia may eventually be discovered at another of these locations.



Fossilized gravel bed in the Jack Hills (*above*) contained the world's oldest zircons yet discovered. Geologists crushed and sorted hundreds of kilograms of this rock (*below*) to find the 20 crystals that bear signs of cool conditions more than four billion years ago.

LUCY READING-IKKANDA; SOURCE: WILLIAM H. PECK *Colgate University* (*map*); JOHN W. VALLEY (*photographs*)

port the surviving grains great distances before they become incorporated into deposits of sand and gravel that may later solidify into sedimentary rock. Indeed, the Jack Hills zircons—separated by perhaps thousands of kilometers from their source—were found embedded in a fossilized gravel bar called the Jack Hills conglomerate.

So, despite the excitement of finding such primeval pieces of the earth, most scientists, including me, continued to accept the view that the climate of our young planet was Hadean. It was not until 1999 that technological advances allowed further study of the ancient zircon crystals from Western Australia—and challenged conventional wisdom about the earth's earliest history.

Digging Deep

THE AUSTRALIAN ZIRCONS did not give up their secrets easily. For one thing, the Jack Hills and their surroundings are dusty barrens at the edge of vast sheep stations, called Berringarra and Mileura, situated some 800 kilometers north of Perth, Australia's most isolated city. The Jack Hills conglomerate was deposited three billion years ago and marks the northwestern edge of a widespread assembly of rock formations that are all older than 2.6 billion years. To recover less than a thimbleful of zircons, my colleagues and I collected hundreds of kilograms of rock from these remote outcrops and hauled them back to the laboratory for crushing and sorting, similar to searching for a few special grains of sand on a beach.

Once extracted from their source rock, individual crystals could be dated because zircons make ideal timekeepers. In addition to their longevity, they contain trace amounts of radioactive uranium, which decays at a known rate to lead. When a zircon forms from a solidifying magma, atoms of zirconium, silicon and oxygen combine in exact proportions ($ZrSiO_4$) to create a crystal structure unique to zircon; uranium occasionally substitutes as a trace impurity. Atoms of lead, on the other hand, are too large to comfortably replace any of the elements in the lattice, so zircons start out virtually lead-free. The uranium-lead clock starts ticking as soon as the zircon crystallizes. Thus, the ratio of lead to uranium increases with the age of the crystal. Scientists can reliably determine the age of an undamaged zircon within 1 percent accuracy, which for the early earth is about plus or minus 40 million years.

THE AUTHOR

JOHN W. VALLEY received his Ph.D. in 1980 from the University of Michigan at Ann Arbor, where he first became interested in the early earth. He and his students have since explored the ancient rock record throughout North America and in Western Australia, Greenland and Scotland. Currently Valley is president of the Mineralogical Society of America and Charles R. Van Hise Professor of Geology at the University of Wisconsin–Madison, where he founded a multimillion-dollar laboratory called WiscSIMS. The cutting-edge capabilities of the lab's new CAMECA IMS 1280 ion microprobe will enable a diverse range of research; besides zircons, Valley and his colleagues will probe many rare or extremely small materials ranging from stardust to cancer cells.

EXTRACTING EVIDENCE

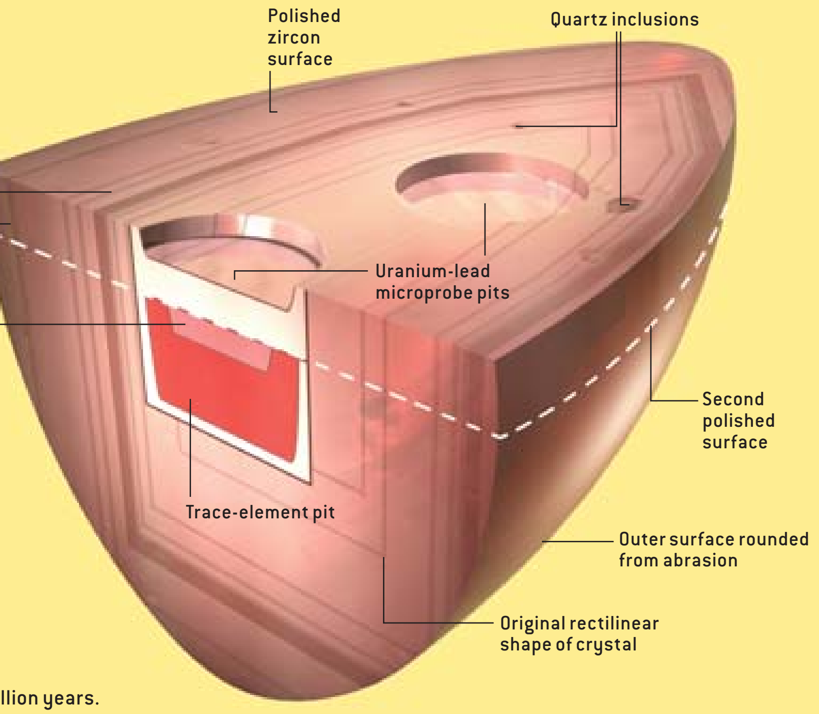
Scientists extract multiple clues about the earth's ancient environment from a single crystal of the mineral zircon (*main cutaway below*). They first embed the zircon in epoxy, then grind and polish the crystal to expose a pristine surface.

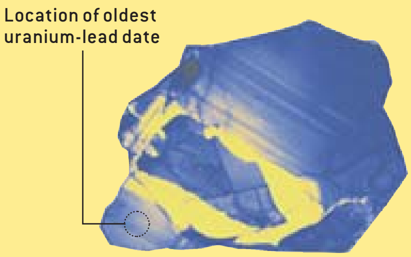
A scanning electron microscope identifies the zircon's growth patterns and any smaller fragments of minerals enclosed as it grew. Inclusions of quartz, for instance, occur most commonly in zircons that came from granite, a type of rock characteristic of continents.

An ion microprobe creates a small pit by sputtering atoms off this polished area using a narrow beam of ions and identifies those atoms by comparing their masses. To determine the age of the crystal, scientists measure atoms of uranium and lead, two impurities trapped within the zircon's atomic structure. Simply put, the constant radioactive decay of uranium to lead means that the more lead present relative to uranium, the older the crystal.

Investigators then grind down the surface to expose a deeper layer of the crystal and make a second microprobe pit in precisely the same location as the first one to measure atoms of oxygen, one of three elements that make up a zircon. The ratio of certain oxygen isotopes—atoms of oxygen with different masses—reveals whether the crystal records hot or cool conditions.

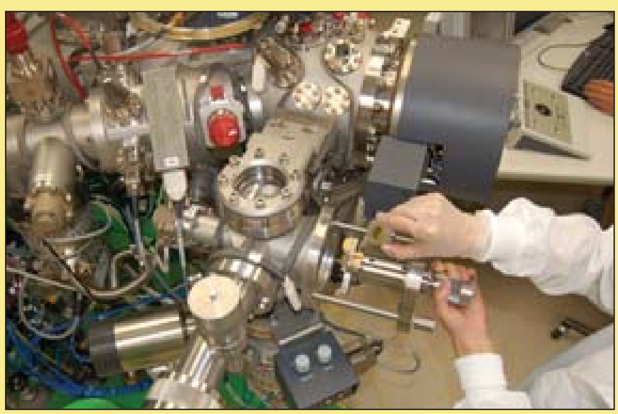
Researchers create a third pit to measure the abundance of certain trace elements—impurities known to make up less than 1 percent of the crystal's structure. Some of these elements are more common in continental crust.





The earth's oldest zircon, viewed as a cathodoluminescent image, dates back 4.4 billion years.

A CLOSER LOOK



Red zircons (*right*), photographed near Roosevelt's nose on a U.S. dime for scale, come from the same rock sample that yielded the earth's oldest crystal. An ion microprobe, such as this one in the author's laboratory at the University of Wisconsin–Madison (*far right*), can analyze isotope ratios or trace elements from spots about 1/15 the diameter of the crystals themselves.

JOHN W. VALLEY (*laboratory*); SIMON A. WILDE *Curtin University of Technology (oldest zircon)*; ALFRED T. KAMAJIAN; SOURCE: AARON J. CAVOSIE *University of Wisconsin–Madison (zircon schematic and red zircons)*

Dating specific parts of a single crystal first became possible in the early 1980s, when William Compston and his colleagues at the Australian National University in Canberra invented a special kind of ion microprobe, a very large instrument they playfully named SHRIMP, short for Sensitive High-Resolution Ion Micro Probe. Although most zircons are nearly invisible to the naked eye, the ion microprobe fires a beam of ions so narrowly focused that it can blast a small number of atoms off any targeted part of a zircon's surface. A mass spectrometer then measures the composition of those atoms by comparing their masses. It was Compston's group—working with Robert T. Pidgeon, Simon A. Wilde and John Baxter, all then at Curtin University of Technology, also in Australia—that first dated the Jack Hills zircons in 1986.

Knowing this history, I approached Wilde. He agreed to reinvestigate the uranium-lead dates of Jack Hills zircons as part of the doctoral thesis of my student William H. Peck, who is now an assistant professor at Colgate University. In May 1999 Wilde analyzed 56 undated crystals using an improved SHRIMP at Curtin and found five that exceeded four billion years. To our great surprise, the oldest dated back to 4.4 billion years ago. Some samples from the moon and Mars have similar ages, and meteorites are generally older, but nothing of this vintage had been found (or expected) from our planet. Almost everyone assumed that if such ancient zircons had ever existed, the dynamic Hadean conditions destroyed them. Little did we know that the most exciting discovery was yet to come.

Evidence of Ancient Oceans

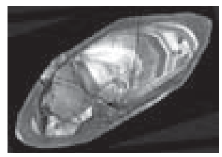
PECK AND I SOUGHT Wilde's zircons from Western Australia because we were looking for a well-preserved sample of the oldest oxygen from the earth. We knew that a zircon could retain evidence not only of when its host rock formed but also of *how*. In particular, we were using the ratios of different isotopes of oxygen to estimate the temperatures of processes leading to the formation of magmas and rocks.

Geochemists measure the ratio of oxygen 18 (^{18}O, a rare isotope with eight protons and 10 neutrons, which represents about 0.2 percent of all oxygen on the earth) to oxygen 16 (^{16}O, the common oxygen isotope with eight protons and eight neutrons, which comprises about 99.8 percent of all oxygen). These atoms are called stable isotopes because they do not undergo radioactive decay and thus do not spontaneously change with time; however, the proportions of ^{18}O and ^{16}O incorporated into a crystal as it forms differ depending on the ambient temperature at the time the crystal formed.

The ^{18}O/^{16}O ratio is well known for the earth's mantle (the 2,800-kilometer-thick layer immediately below the thin, five- to 40-kilometer-thick veneer of continents and ocean crust). Magmas that form in the mantle always have about the same oxygen isotope ratio. For simplicity, geochemists calibrate these ratios relative to that of seawater and express them in what is called delta (δ) notation. The δ^{18}O of the ocean is 0 by definition, and the δ^{18}O of zircon from the mantle is 5.3, meaning that it has a greater ^{18}O/^{16}O ratio than seawater.

That is why Peck and I expected to find a primitive mantle value of around 5.3 when we took Wilde's Jack Hills zircons, including the five oldest, to the University of Edinburgh in Scotland that same summer. There John Craven and Colin Graham helped us use a different kind of ion microprobe specially suited to measure oxygen isotope ratios. We had worked together many times over the preceding decade to perfect the technique and could analyze samples one millionth the size of those then possible in my laboratory in Wisconsin. After 11 days of round-the-clock analysis and little sleep (typical conditions for this difficult procedure), we completed the measurements—and found that our predictions were wrong. The zircons' δ^{18}O values ranged up to 7.4.

The tiny zircons from Western Australia did not GIVE UP THEIR SECRETS easily.



We were stunned. What could these high oxygen isotope ratios mean? In younger rocks the answer would be obvious, because such samples are common. A typical scenario is that rocks at low temperature on the earth's surface can acquire a high oxygen isotope ratio if they chemically interact with rain or ocean water. Those high-δ^{18}O rocks, if buried and melted, form magma that retains the high value, which is then passed on to zircons during crystallization. Thus, liquid water and low temperatures are required on the surface of the earth to form zircons and magmas with high δ^{18}O; no other process is known to do so.

Finding high oxygen isotope ratios in the Jack Hills zircons implied that liquid water must have existed on the surface of the earth at least 400 million years earlier than the oldest known sedimentary rocks, those at Isua, Greenland. If correct, entire oceans probably existed, making the earth's early climate more like a sauna than a Hadean fireball.

Continental Clues

COULD WE REALLY BASE such far-reaching conclusions about the history of the earth on a few tiny crystals? We delayed publishing our findings for more than a year so we could double-check our analyses. Meanwhile other groups were conducting their own research in the Jack Hills. Stephen J. Mojzsis of the University of Colorado and his colleagues at the University of California at Los Angeles confirmed our results, and we published back-to-back technical articles describing our findings in 2001.

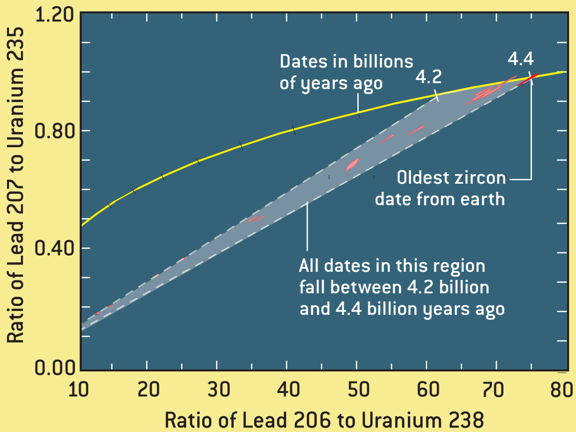
AARON J. CAVOSIE *University of Wisconsin–Madison*

THE TALES THEY TELL

Zircons from the Jack Hills of Western Australia have changed the way scientists think about the early history of the earth. These crystals are the oldest terrestrial materials yet discovered—hundreds of those identified formed more than four billion years ago. Many of these tiny timekeepers also bear clear chemical signs that oceans of liquid water and possibly even continents existed on the earth's surface at a time once thought to be molten and fiery.

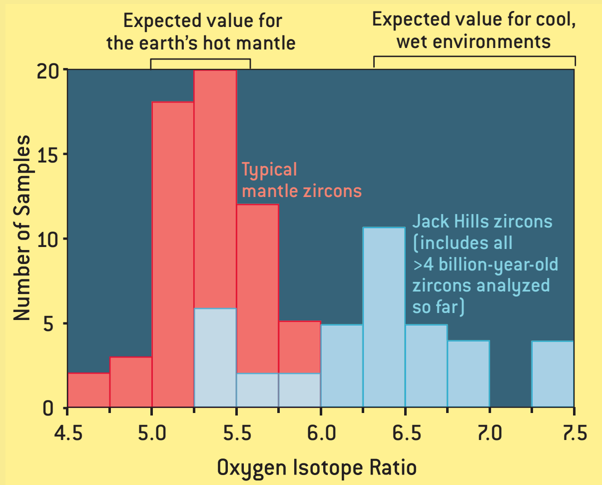
ANCIENT AGE

The oldest age for a Jack Hills zircon—4.4 billion years (*red*)—is an exact match of two geologic "clocks." Two pairs of isotopes—uranium 235–lead 207 (*vertical axis*) and uranium 238–lead 206 (*horizontal axis*)—form two radioactive timekeepers that start ticking when a zircon forms. If they are well preserved, their final ratios plot along a single line (*yellow*). Dates from other parts of the zircon (*pink*) fall off this line because some lead was lost from these areas, but scientists can correct for this damage.



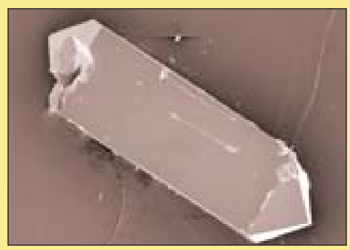
COOL OCEANS

Oxygen isotope ratios in Jack Hills zircon samples (*blue*), which range up to 7.5, are possible only if their source rock formed in a relatively cool, water-rich environment near the earth's surface. Had magma oceans covered the planet when these zircons formed, their values would have clustered near 5.3, as do those of all crystals from hot rock that originates in the planet's deep interior (*red*).



FIRST CONTINENTS?

Rounded surfaces of some Jack Hills zircons under a scanning electron microscope show that wind and possibly running water buffeted these crystals over long distances—possibly across a large continental landmass—before they were finally laid to rest (*right*). Zircons found near their place of origin retain their original sharp edges (*far right*). The large number of ancient, rounded Jack Hills zircons suggests their original source rocks were widespread.



LUCY READING-IKKANDA; SOURCE: SIMON A. WILDE *Curtin University of Technology (uranium-lead)*; LUCY READING-IKKANDA; SOURCE: JOHN W. VALLEY *(oxygen isotope ratios)*; AARON J. CAVOSIE *University of Wisconsin–Madison (zircon grains)*

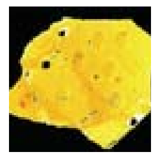
As the possible implications of the zircon discoveries spread through the scientific community, the excitement was palpable. In the superheated violence of a Hadean world, no samples would have survived for geologists to study. But these zircons pointed to a more clement and familiar world and provided a means to unravel its secrets. If the earth's climate was cool enough for oceans of water early on, then maybe zircons could tell us if continents and other features of modern earth also existed. To find out, we had to look more closely into the interiors of single crystals.

Even the smallest zircon contains other materials that were encapsulated as the zircon grew around them. Such zircon inclusions can reveal much about where the crystal came from, as can the crystal's growth patterns and the composition of trace elements. When Peck and I studied the 4.4-billion-year-old zircon, for instance, we found that it contained pieces of other minerals, including quartz. That was surprising because quartz is rare in primitive rocks and was probably absent from the very first crust on the earth. Most quartz comes from granitic rocks, which are common in more evolved continental crust.

If the Jack Hills zircons came from a granitic rock, that evidence would support the hypothesis that they are samples of the world's first continent. But caution is warranted. Quartz can form in the last stages of magma crystallization even if the parent rock is not granitic, although such quartz is much less abundant. For instance, zircons and a few grains of quartz have been found on the moon, which never developed a granitic, continental-style crust. Some scientists have also wondered if the earth's earliest zircons formed an environment more like the early moon or by some other means that is no longer common, perhaps related to giant meteorite impacts or deep-sourced volcanism, but no one has found convincing evidence.

Meanwhile other clues for continental crust came from trace elements (those elements substituting in zircon at levels below 1 percent). Jack Hills zircons have elevated concentrations of these elements as well as patterns of europium and cerium that are most commonly created during the crystallization of crust, which means the zircons formed near the earth's surface rather than in the mantle. Furthermore, the ratios of radioactive isotopes of neodymium and hafnium—two elements used to determine the timing of continental-crust-forming events—suggest that significant amounts of continental crust formed as early as 4.4 billion years ago.

The Jack Hills zircons may be samples of the world's FIRST CONTINENT.



The distribution of ancient zircons provided additional evidence. The proportion of zircons older than four billion years exceeds 10 percent in some samples from Jack Hills. Also, the zircon surfaces are highly abraded and the originally angular crystal faces are rounded, suggesting the crystals were blown long distances from their source rock. How could these zircons travel hundreds or thousands of kilometers as windblown dust and still be concentrated together unless there had once been a lot of them? And how could these zircons escape burial and melting in the hot mantle unless thick continental-type crust was stable so as to preserve them?

These findings imply that the zircons were once plentiful and came from a widespread source region, possibly a continental landmass. If so, it is quite likely that rocks from this earliest time still exist, an exciting prospect because much could be learned from an intact rock of this age.

Furthermore, the age distribution of ancient zircons is uneven. Ages cluster in certain time periods, and no zircons have been found from other eras. My former graduate student Aaron J. Cavosie, now an assistant professor at the University of Puerto Rico, found such evidence even within single zoned zircons where the core formed early, say, 4.3 billion years ago, with a surrounding overgrowth that formed later, between 3.7 billion and 3.3 billion years ago. That the zircons get younger from core to rim is expected because zircon crystals grow concentrically by adding material to their grain boundary. But the great age difference with time gaps between the cores and rims of these particular zircons indicates that two distinct events took place, separated by a major hiatus. In more commonly available, younger zircons, this kind of core-to-rim age relation results from tectonic processes that melt continental crust and recycle the zircons within it. Many scientists are trying to test whether similar conditions produced the ancient Jack Hills zircons.

Most recently, E. Bruce Watson of the Rensselaer Polytechnic Institute and T. Mark Harrison of the Australian National University have reported lower-than-expected levels of titanium in these ancient zircons, suggesting that the temperatures of their parent magmas must have been between 650 and 800 degrees C. Such low temperatures would be possible only if the parent rocks were granitic; most nongranitic rocks melt at higher temperatures, and so their zircons should contain more titanium.

Zircons Are Forever

SINCE MY COLLEAGUES and I analyzed the oxygen isotope ratios in those five Jack Hills zircons in 1999, the data supporting our conclusions have grown rapidly. Investigators in Perth, Canberra, Beijing, Los Angeles, Edinburgh, Stockholm and Nancy, France, have now put tens of thousands of Jack Hills zircons into ion microprobes, searching for the relatively few that are older than four billion years, and other dating techniques have been applied as well.

Hundreds of newly discovered zircons have been reported from several localities with ages from 4.4 billion to four billion years old. David R. Nelson and his colleagues at the Geological Survey of Western Australia have found similarly ancient zircons as far as 300 kilometers south of the Jack Hills. Geochemists are scrutinizing other ancient regions of the earth, hoping to find the first pre-4.1-billion-year-old zircons outside Australia.

And the intensifying search is spurring improved technology. Cavosie has demonstrated much better accuracy of analysis and reported more than 20 Jack Hills zircons with the high oxygen isotope ratios that fingerprint cool surface temperatures and ancient oceans as early as 4.2 billion years ago. My colleagues and I are continuing the search using the first model of the newest generation ion microprobe, called the CAMECA IMS 1280, which was installed in my laboratory this past March.

Many questions will be answered if pieces of the original zircon-forming rock can be identified. But even if we never find that rock, there is still much to learn from the tiny zircon time capsules. SA

MORE TO EXPLORE

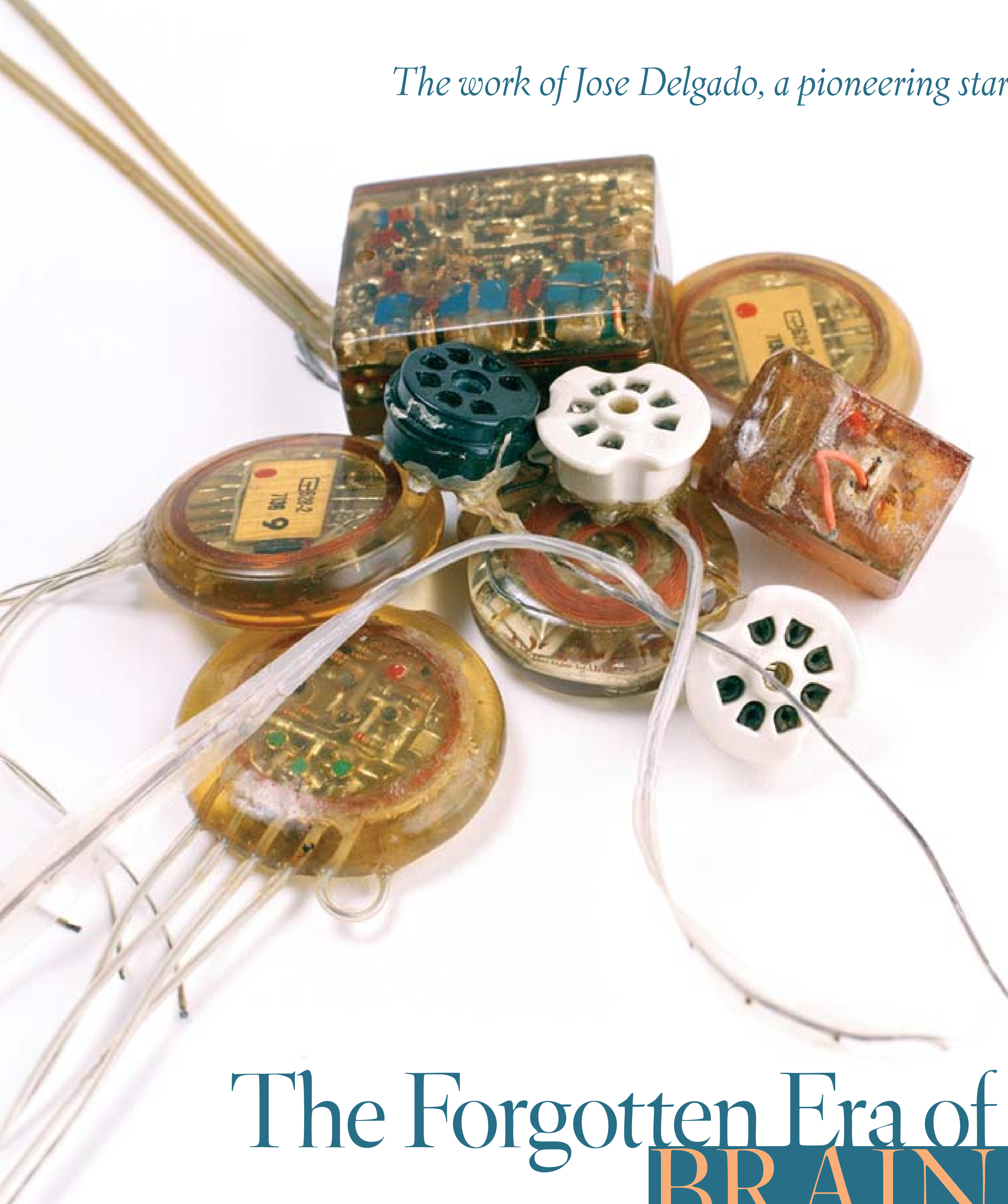
A Cool Early Earth. John W. Valley, William H. Peck, Elizabeth M. King and Simon A. Wilde in *Geology,* Vol. 30, No. 4, pages 351–354; April 2002.

Magmatic $\delta^{18}O$ in 4400–3900 Ma Detrital Zircons: A Record of the Alteration and Recycling of Crust in the Early Archean. Aaron J. Cavosie, John W. Valley, Simon A. Wilde and the Edinburgh Ion Microprobe Facility in *Earth and Planetary Science Letters,* Vol. 235, No. 3, pages 663–681; July 15, 2005.

The author's "Zircons Are Forever" Web site is at **www.geology.wisc.edu/zircon/zircon_home.html**

SIMON A. WILDE *Curtin University of Technology*

The work of Jose Delgado, a pioneering star



The Forgotten Era of BRAIN

By John Horgan

in brain-stimulation research four decades ago, goes largely unacknowledged today. What happened?

In the early 1970s Jose Manuel Rodriguez Delgado, a professor of physiology at Yale University, was among the world's most acclaimed—and controversial—neuroscientists. In 1970 the *New York Times Magazine* hailed him in a cover story as the "impassioned prophet of a new 'psychocivilized society' whose members would influence and alter their own mental functions." The article added, though, that some of Delgado's Yale colleagues saw "frightening potentials" in his work.

Delgado, after all, had pioneered that most unnerving of technologies, the brain chip—an electronic device that can manipulate the mind by receiving signals from and transmitting them to neurons. Long the McGuffins of science fiction, from *The Terminal Man* to *The Matrix,* brain chips are now being used or tested as treatments for epilepsy, Parkinson's disease, paralysis, blindness and other disorders. Decades ago Delgado carried out experiments that were more dramatic in some respects than anything being done today.

CHIPS

ELECTRICAL BRAIN-STIMULATION DEVICES (*opposite page*), invented by Jose Delgado for his research into behavior and motor control, were implanted into apes, monkeys (*shown above*), bulls, cats and humans. Electrodes could remain implanted for more than two years

He implanted radio-equipped electrode arrays, which he called "stimoceivers," in cats, monkeys, chimpanzees, gibbons, bulls and even humans, and he showed that he could control subjects' minds and bodies with the push of a button.

Yet after Delgado moved to Spain in 1974, his reputation in the U.S. faded, not only from public memory but from the minds and citation lists of other scientists. He described his results in more than 500 peer-reviewed papers and in a widely reviewed 1969 book, but these are seldom cited by modern researchers. In fact, some familiar with his early work assume he died. But Delgado, who recently moved with his wife, Caroline, from Spain to San Diego, Calif., is very much alive and well, and he has a unique perspective on modern efforts to treat various disorders by stimulating specific areas of the brain.

When Lobotomies Were the Rage

BORN IN 1915 in Ronda, Spain, Delgado went on to earn a medical degree from the University of Madrid in the 1930s. Although he has long been dogged by rumors that he supported the fascist regime of Francisco Franco, he actually served in the medical corps of the Republican Army (which opposed Franco during Spain's civil war) while he was a medical student. After Franco crushed the Republicans, Delgado was detained in a concentration camp for five months before resuming his studies.

He originally intended to become an eye doctor, like his father. But a stint in a physiology laboratory—plus exposure to the writings of the great Spanish neuroscientist Santiago Ramón y Cajal—left him entranced by "the many mysteries of the brain. How little was known then. How little is known now!" Delgado was particularly intrigued by the experiments of Swiss physiologist Walter Rudolf Hess. Beginning in the 1920s, Hess had demonstrated that he could elicit behaviors such as rage, hunger and sleepiness in cats by electrically stimulating different spots in their brains with wires.

In 1946 Delgado won a yearlong fellowship at Yale. In 1950 he accepted a position in its department of physiology, then headed by John Fulton, who played a crucial role in the history of psychiatry. In a 1935 lecture in London, Fulton had reported that a violent, "neurotic" chimpanzee named Becky had become calm and compliant after surgical destruction of her prefrontal lobes. In the audience was Portuguese psychiatrist Egas Moniz, who started performing lobotomies on psychotic patients and claimed excellent results. After Moniz won a Nobel Prize in 1949, lobotomies became an increasingly popular treatment for mental illness.

Initially disturbed that his method of pacifying a chimpanzee had been applied to humans, Fulton later became a cautious proponent of psychosurgery. Delgado disagreed with his mentor's stance. "I thought Fulton and Moniz's idea of destroying the brain was absolutely horrendous," Delgado recalls. He felt it would be "far more conservative" to treat mental illness by applying the electrical stimulation methods pioneered by Hess—who shared the 1949 prize with Moniz. "My idea was to *avoid* lobotomy," Delgado says, "with the help of electrodes implanted in the brain."

One key to Delgado's scientific success was his skill as an inventor; a Yale colleague once called him a "technological wizard." In his early experiments, wires ran from implanted electrodes out through the skull and skin to bulky electronic devices that recorded data and delivered electrical pulses. This setup restricted subjects' movements and left them prone to infections. Hence, Delgado designed radio-equipped stimoceivers as small as half-dollars that could be fully implanted in subjects. His other inventions included an early version of the cardiac pacemaker and implantable "chemitrodes" that could release precise amounts of drugs directly into specific areas of the brain.

In 1952 Delgado co-authored the first peer-reviewed paper describing long-term implantation of electrodes in humans, narrowly beating a report by Robert Heath of Tulane University. Over the next two decades Delgado implanted electrodes in some 25 human subjects, most of them schizophrenics and epileptics, at a now defunct mental hospital in Rhode Island. He operated, he says, only on desperately ill patients whose disorders had resisted all previous treatments. Early on, his placement of electrodes in humans was guided by animal experiments, studies of brain-damaged people and the work of Canadian neurosurgeon Wilder Penfield; beginning in the 1930s, Penfield stimulated epileptics' brains with electrodes before surgery to determine where he should operate.

Taming a Fighting Bull

DELGADO SHOWED that stimulation of the motor cortex could elicit specific physical reactions, such as movement of the limbs. One patient clenched his fist when stimulated, even when he tried to resist. "I guess, doctor, that your electricity is stronger than my will," the patient commented. Another subject, turning his head from side to side in response to stimulation, insisted he was doing so vol-

Overview/*Brain Implants*

- Jose M. R. Delgado, a pioneer in brain-implant technology, is perhaps most famous for halting a charging bull by merely pressing a button on a device that sent signals to the animal's brain.
- In the early 1970s Delgado went from being acclaimed to being criticized. In 1974 he moved from the U.S. to Spain and then gradually faded from public consciousness and the citation lists of neuroscientists.
- His accomplishments, however, helped to pave the way for modern brain-implant technology, which is enjoying a resurgence today and is improving life for patients with epilepsy and such movement disorders as Parkinson's and dystonia.
- Delgado, now 90, recently returned to the U.S., complete with strong opinions on the promise and perils of the ongoing work.

PRECEDING PAGES: BRAD SWONETZ; COURTESY OF JOSE DELGADO (*x-ray*)

CAROLINE DELGADO, shown monitoring encephalographic readings from a monkey, has assisted her husband since their meeting at Yale University in the 1950s.

COURTESY OF JOSE DELGADO

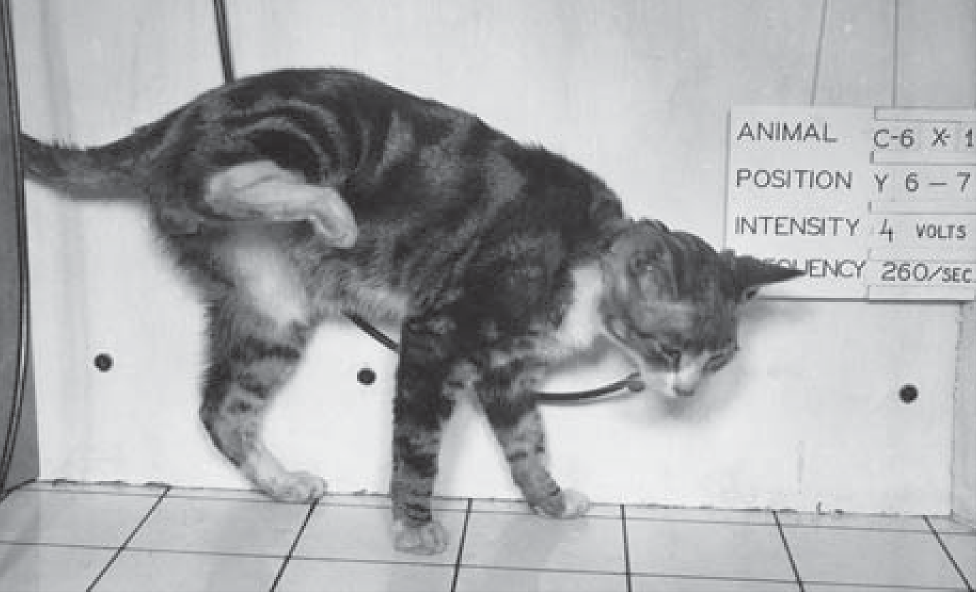
untarily, explaining, “I am looking for my slippers.”

By stimulating different regions of the limbic system, which regulates emotion, Delgado could also induce fear, rage, lust, hilarity, garrulousness and other reactions, some of them startling in their intensity. In one experiment, Delgado and two collaborators at Harvard University stimulated the temporal lobe of a 21-year-old epileptic woman while she was calmly playing a guitar; in response, she flew into a rage and smashed her guitar against a wall, narrowly missing a researcher's head.

Perhaps the most medically promising finding was that stimulation of a limbic region called the septum could trigger euphoria, strong enough in some cases to counteract depression and even physical pain. Delgado limited his human research, however, because the therapeutic benefits of implants were unreliable; results varied widely from patient to patient and could be unpredictable even in the same subject. In fact, Delgado recalls turning away more patients than he treated, including a young woman who was sexually promiscuous and prone to violence and had repeatedly been confined in jails and mental hospitals. Although both the woman and her parents begged Delgado to implant electrodes in her, he refused, feeling that electrical stimulation was too primitive for a case involving no discernible neurological disorder.

Delgado did much more extensive research on monkeys and other animals, often focusing on neural regions that elicit and inhibit aggression. In one demonstration, which explored the effects of stimulation on social hierarchy, he implanted a stimoceiver in a macaque bully. He then installed a lever in the cage that, when pressed, pacified the bully by causing the stimoceiver to stimulate the monkey's caudate nucleus, a brain region involved in controlling voluntary movements. A female in the cage soon discovered the lever's power and yanked it whenever the male threatened her. Delgado, who never shied from anthropomorphic interpretations, wrote, “The old dream of an individual overpowering the strength of a dictator by remote control has been fulfilled, at least in our monkey colonies.”

Delgado's most famous experiment took place in 1963 at a bull-breeding ranch in Cordoba, Spain. After inserting stimoceivers into the brains of several “fighting bulls,” he stood in a bullring with one bull at a time and, by pressing buttons on a handheld transmitter, controlled each animal's actions. In one in-



CAT LIFTED ITS HIND LEG in response to stimulation by an electrode implanted in its brain. The cat, Delgado says, displayed no discomfort in this experiment done in the early 1950s.

stance, captured in a dramatic photograph, Delgado forced a charging bull to skid to a halt only a few feet away from him by stimulating its caudate nucleus. The *New York Times* published a front-page story on the event, calling it "the most spectacular demonstration ever performed of the deliberate modification of animal behavior through external control of the brain." Other articles hailed Delgado's transformation of an aggressive beast into a real-life version of Ferdinand the bull, the gentle hero of a popular children's story.

In terms of scientific significance, Delgado believes his experiment on a female chimpanzee named Paddy deserved more attention. Delgado programmed Paddy's stimoceiver to detect distinctive signals, called spindles, spontaneously emitted by her amygdala. Whenever the stimoceiver detected a spindle, it stimulated the central gray region of Paddy's brain, producing an "aversive reaction"—that is, a painful or unpleasant sensation. After two hours of this negative feedback, Paddy's amygdala produced 50 percent fewer spindles; the frequency dropped by 99 percent within six days. Paddy was not exactly a picture of health: she became "quieter, less attentive and less motivated during behavioral testing," Delgado wrote. He nonetheless speculated that this "automatic learning" technique could be used to quell epileptic seizures, panic attacks or other disorders characterized by specific brain signals.

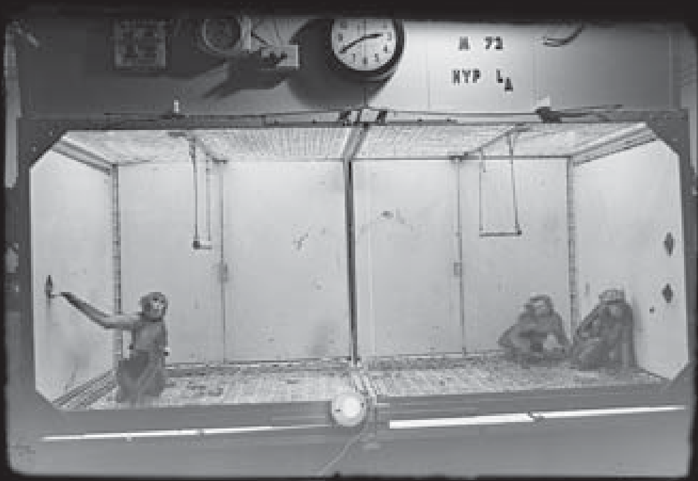
Delgado's research was supported not only by civilian agencies but also by military ones such as the Office of Naval Research (but never, Delgado insists, by the Central Intelligence Agency, as some conspiracy theorists have charged). Delgado, who calls himself a pacifist, says that his Pentagon sponsors viewed his work as basic research and never steered him toward military applications. He has always dismissed speculation that implants could create cyborg soldiers who kill on command, like the brainwashed assassin in the novel and film versions of *The Manchurian Candidate*. (The assassin was controlled by psychological methods in the original 1962 film and by a brain chip in the 2004 remake.) Brain stimulation may "increase or decrease aggressive behavior," he asserts, but it cannot "direct aggressive behavior to any specific target."

Envisioning a "Psychocivilized Society"

IN 1969 DELGADO described brain-stimulation research and discussed its implications in *Physical Control of the Mind: Toward a Psychocivilized Society,* which was illustrated with photographs of monkeys, cats, a bull and two young women whose turbans concealed

FIGHTING BULL with a stimoceiver in its brain (*below*) charged Delgado in a Spanish bullring in 1963 (*middle two photographs*) and then stopped and turned in response to a radio signal from Delgado (*far right*). Critics contended that the stimulation did not quell the bull's aggressive instinct, as Delgado suggested, but rather forced it to turn to the left. Delgado, who grew up in Ronda, Spain, a bastion of bullfighting, admits he felt "frightened" just before his signal made the bull abandon the chase.

COURTESY OF JOSE DELGADO



stimoceivers. (Female patients "have shown their feminine adaptability to circumstance," Delgado remarked, "by wearing attractive hats or wigs to conceal their electrical headgear.") Spelling out the limitations of brain stimulation, Delgado downplayed "Orwellian possibilities" in which evil scientists enslave people by implanting electrodes in their brains.

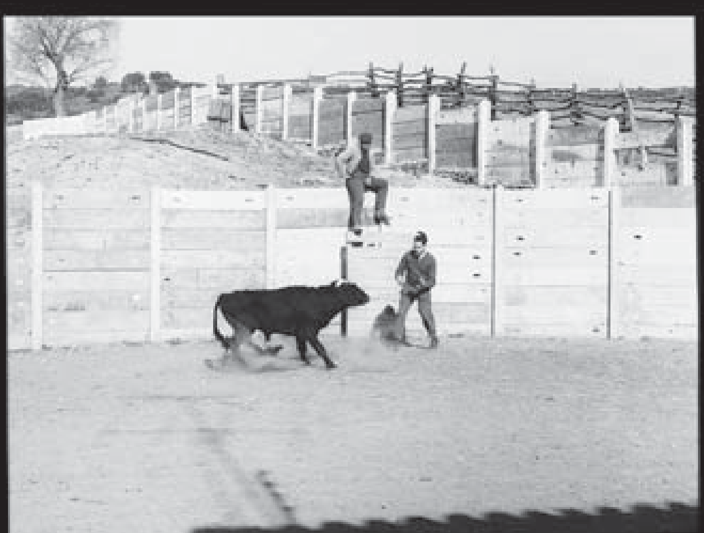
Yet some of his rhetoric had an alarmingly evangelical tone. Neurotechnology, he declared, was on the verge of "conquering the mind" and creating "a less cruel, happier, and better man." In a review in *Scientific American,* the late physicist Philip Morrison called *Physical Control* "a thoughtful, up-to-date account" of electrical stimulation experiments but added that its implications were "somehow ominous."

In 1970 Delgado's field was engulfed in a scandal triggered by Frank Ervin and Vernon Mark, two researchers at Harvard Medical School with whom Delgado briefly collaborated. (One of Ervin's students was Michael Crichton, who wrote *The Terminal Man.* The best-seller, about a bionic experiment gone awry, was inspired by the research of Ervin, Mark and Delgado.) In their book, *Violence and the Brain,* Ervin and Mark suggested that brain stimulation or psychosurgery might quell the violent tendencies of blacks rioting in inner cities. In 1972 Heath, the Tulane psychiatrist, raised more questions about brain-implant research when he reported that he had tried to change the sexual orientation of a male homosexual by stimulating his septal region while he had intercourse with a female prostitute.

The fiercest opponent of brain implants was psychiatrist Peter Breggin (who in recent decades has focused on the dangers of psychiatric drugs). In testimony submitted into the Congressional Record in 1972, Breggin lumped Delgado, Ervin, Mark and Heath together with advocates of lobotomies and

FEMALE MACAQUE (*far left in first photograph*) learned that by pulling a lever in the cage she could escape encounters with an alpha male. The lever sent a signal to a stimoceiver in his brain, pacifying him. The alpha male is in the pacified state at the far right in the left image and has become aggressive in the other shot. Delgado carried out many investigations, such as this one in the early 1960s, into the effects of brain stimulation on social interactions.

accused them of trying to create "a society in which everyone who deviates from the norm" will be "surgically mutilated." Quoting liberally from *Physical Control,* Breggin singled out Delgado as "the great apologist for technologic totalitarianism." In his 1973 book *Brain Control,* Elliot Valenstein, a neurophysiologist at the University of Michigan at Ann Arbor, presented a detailed scientific critique of brain-implant research by Delgado and others, contending that the results of stimulation were much less precise and therapeutically beneficial than proponents often suggested. (Delgado notes that in his own writings he made



THE AUTHOR

JOHN HORGAN, director of the center for science writings at the Stevens Institute of Technology in Hoboken, N.J., was a staff writer for *Scientific American* from 1986 to 1997 and then, until recently, a freelance writer. His books include *The End of Science, The Undiscovered Mind* and *Rational Mysticism*.

many of the same points as Valenstein.)

Meanwhile strangers started accusing Delgado of having secretly implanted stimoceivers in their brains. One woman who made this claim sued Delgado and Yale University for $1 million, although he had never met her. In the midst of this brouhaha, Villar Palasi, the Spanish minister of health, asked Delgado to help organize a new medical school at the Autonomous University in Madrid, and he accepted, moving with his wife and two children to Spain in 1974. He insists that he was not fleeing the disputes surrounding his research; the minister's offer was just too good to refuse. "I said, 'Could I have the facilities I have at Yale?' And he said, 'Oh, no, much better!' "

In Spain, Delgado shifted his focus to noninvasive methods of affecting the brain, which he hoped would be more medically acceptable than implants. Anticipating modern techniques such as transcranial magnetic stimulation, he invented a halolike device and a helmet that could deliver electromagnetic pulses to specific neural regions. Testing the gadgets on both animals and human volunteers—including himself and his daughter, Linda—Delgado discovered that he could induce drowsiness, alertness and other states; he also had some success treating tremors in Parkinson's patients.

Delgado still could not entirely escape controversy. In the mid-1980s an article in the magazine *Omni* and documentaries by the BBC and CNN cited Delgado's work as circumstantial evidence that the U.S. and Soviet Union might have secretly developed methods for remotely modifying people's thoughts. Noting that the power and precision of electromagnetic pulses decline rapidly with distance, Delgado dismisses these mind-control claims as "science fiction."

Except for these flashes of publicity, however, Delgado's work no longer received the attention it once had. Although he continued publishing articles—especially on the effects of electromagnetic radiation on cognition, behavior and embryonic growth—many appeared only in Spanish journals. Moreover, brain-stimulation studies back in the U.S. bogged down in ethical controversies, grants dried up, and researchers drifted to other fields, notably psychopharmacology, which seemed to be a much safer, more effective way to treat brain disorders

Brain Implants Today

When Jose Delgado and a few other intrepid scientists first began exploring the effects of implanting electrodes in the brain half a century ago, they could not foresee how many people would one day benefit from this line of research. By far the most successful form of implant, or "neural prosthesis," is the artificial cochlea. More than 70,000 people have been equipped with these devices, which restore at least rudimentary hearing by feeding signals from an external microphone to the auditory nerve. Brain stimulators have been implanted in more than 30,000 people suffering from Parkinson's disease and other movement disorders (including 17-year-old Kari Weiner, shown at the right). Roughly the same number of epileptics are being treated with devices that stimulate the vagus nerve in the neck.

Work on other prostheses is proceeding more slowly. Clinical trials are now under way to test brain and vagus nerve stimulation for treating disorders such as depression, obsessive-compulsive disorder, panic attacks and chronic pain. Artificial retinas—light-sensitive chips that mimic the eye's signal-processing ability and stimulate the optic nerve or visual cortex—have been tested in a handful of blind subjects, but they usually "see" nothing more than phosphenes, or bright spots.

Several groups have recently shown that monkeys can control computers and robotic arms "merely by thinking," as media accounts invariably put it—not telekinetically but via implanted electrodes picking up neural signals. The potential for empowering the paralyzed is obvious, but so far few experiments with humans have been carried out, with limited success. Chips that might restore the memory of those afflicted with Alzheimer's disease or other disorders are still a year or two away from testing in rats.

The potential market for neural prostheses is enormous. An estimated 10 million Americans grapple with major depression; 4.5 million suffer from memory loss caused by Alzheimer's disease; more than two million have been paralyzed by spinal cord injuries, amyotrophic lateral sclerosis and strokes; and more than a million are legally blind. *—J.H.*

KARI WEINER was confined to a wheelchair (*left*) for seven years by dystonia, a condition that causes uncontrollable muscle spasms. Now (*right*) she walks without assistance, thanks to battery-powered electrodes that were implanted in her brain when she was 13—and to surgeries that then repaired her twisted muscles and lengthened her tendons.

COURTESY OF AUDREY WEINER (*left*); PETER MURPHY (*right*)

DELGADO, holding two of his brain implants in a photograph taken in August, once wrote that humanity should shift its mission from the ancient dictum "Know thyself" to "Construct thyself."

than brain stimulation or surgery. Only in the past decade has brain-implant research revived, spurred by advances in computation, electrodes, microelectronics and brain-scanning technologies—and by a growing recognition of the limits of drugs for treating mental illness.

Delgado, who stopped doing research in the early 1990s but still follows the field of brain stimulation, believes modern investigators fail to cite his studies not because he was so controversial but simply out of ignorance; after all, most modern databases do not include publications from his heyday in the 1950s and 1960s. He is thrilled by the resurgence of research on brain stimulation, because he still believes in its potential to liberate us from psychiatric diseases and our innate aggression. "In the near future," he says, "I think we will be able to help many human beings, especially with the noninvasive methods."

Delgado's successors have faced some of the same questions that he did about possible abuses of neurotechnology. Some pundits have expressed concern that brain chips could allow a "controlling organization" to "hack into the wetware between our ears," as *New York Times* columnist William Safire put it. An editorial in *Nature* recently expressed concern that officials in the Defense Advanced Research Projects Agency, a major funder of brain-implant research, have openly considered implanting brain chips in soldiers to boost their cognitive capacities. Meanwhile some techno-enthusiasts, such as British computer scientist Kevin Warwick, contend that the risks of brain chips are far outweighed by the potential benefits, which will include instantly "downloading" new languages or other skills, controlling computers and other devices with our thoughts, and communicating telepathically with one another.

Delgado predicts that neurotechnologies may never advance as far as many people fear or hope. The applications envisioned by Warwick and others, Delgado points out, require knowing how complex information is encoded in the brain, a goal that neuroscientists are far from achieving. Moreover, learning quantum mechanics or a new language involves "slowly changing connections which are already there," Delgado says. "I don't think you can do that suddenly." Brain stimulation, he adds, can only modify skills and capacities that we already possess.

But Delgado looks askance at the suggestion of the White House Council on Bioethics and others that some scientific goals—particularly those that involve altering human nature—should not even be pursued. To be sure, he says, technology "has two sides, for good and for bad," and we should do what we can to "avoid the adverse consequences." We should try to prevent potentially destructive technologies from being abused by authoritarian governments to gain more power or by terrorists to wreak destruction. But human nature, Delgado asserts, echoing one of the themes of *Physical Control,* is not static but "dynamic," constantly changing as a result of our compulsive self-exploration. "Can you avoid knowledge?" Delgado asks. "You cannot! Can you avoid technology? You cannot! Things are going to go ahead in spite of ethics, in spite of your personal beliefs, in spite of everything." SA

MORE TO EXPLORE

Brain Control: A Critical Examination of Brain Stimulation and Psychosurgery. Elliot S. Valenstein. John Wiley and Sons, 1973. (A contemporaneous scientific critique of the work of Delgado and other neuroscientists.)

Controlling Robots with the Mind. Miguel A. L. Nicolelis and John K. Chapin in *Scientific American,* Vol. 287, No. 4, pages 46–53; October 2002.

Rebuilt: How Becoming Part Computer Made Me More Human. Michael Chorost. Houghton Mifflin, 2005. (A personal story on the pros and cons of brain implants.)

The President's Council on Bioethics Web site is at **www.bioethics.gov**

An overview of modern brain stimulation can be found at **www.bioethics.gov/transcripts/june04/session6.html**

Other Web sites extol the utopian possibilities of brain stimulation, **www.wireheading.com**, or deplore it as a government mind-control plot, **www.mindjustice.org/**

BRAD SWONETZ

SNIFFER DOG patrols after London's suicide bombings this past July.

Better Than a Dog

BY GARY STIX

SERGIO DIONISIO *AP Photo*

The search is on for a sensor that bests a canine at

detecting explosives

The terrorist attack on the London subway system provoked calls from politicians for deployment of new technologies that could warn of the presence of bombs before they go off. But a detector that can discover the presence of multiple types of explosives quickly, accurately and from a far enough distance to protect people and property does not exist. The nearest thing is a sniffer dog, but a canine has a short attention span and needs frequent breaks.

The chemists, materials scientists and electronic engineers who are paid to think about such issues are trying to come up with ideas beyond putting Ritalin in dog chow. Large swaths of the electromagnetic spectrum and the periodic table are fair game. Even the insect kingdom might be recruited to attack the problem. A report by the National Research Council (NRC) last year—*Existing and Potential Standoff Explosives Detection Techniques*—speculated on far-out ideas for finding concealed bombs that use conventional explosives. Bees could be trained, through altered feeding habits, to swarm a vehicle packed with dynamite. Failing that, robotic "insects" with onboard sensors might do the same.

The report, which was commissioned by the Defense Advanced Research Projects Agency (DARPA), recommended that future research be directed at spotting the telltale bodily signs—skin discoloration or increased blood flow, for instance—produced by a suicide bomber or another person handling explosives. Perhaps the wildest notion was using imaging to look for anomalies in the earth's natural ion field, because bomb materials might cause depletion of negatively charged ions around the person carrying the weapon.

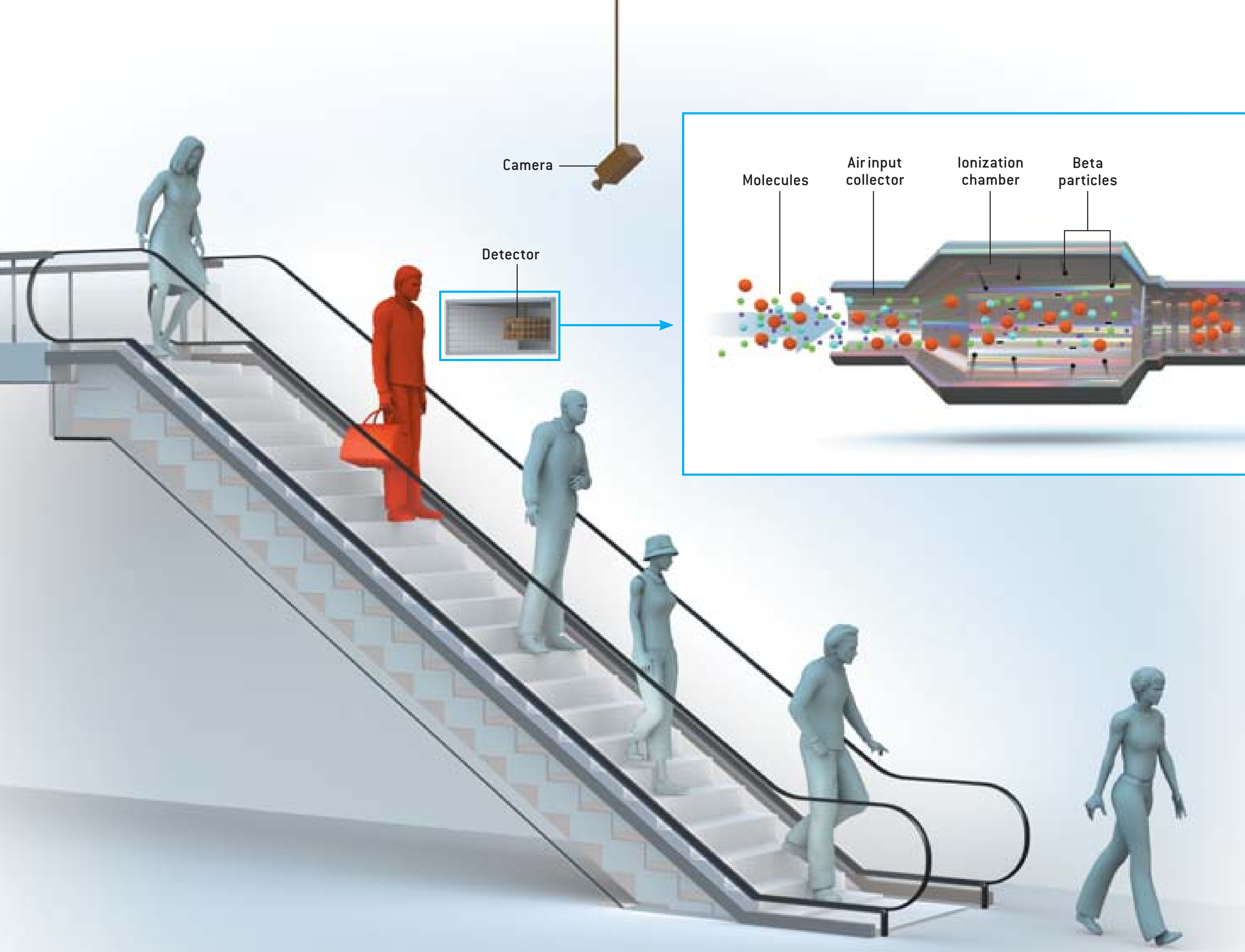
Explosives detectors fall into two categories. Bulk systems—airport x-ray machines are the best example—look for bombs by creating an image of the charge, the detonators and wiring. In some instances, they may even try to discern the chemical composition or other properties. And trace detectors—now coming into widespread use in airports—try to pinpoint vapors or particles from an explosive.

Ion-field fluctuation sensors may not be ready for delivery to mass-transit authorities, although research on methods of both bulk and trace detection has surged, mostly for close-up detection. Imaging machines that deploy millimeter and terahertz wavelengths or x-rays that scatter off a target of interest fulfill all too literally the comic-book adolescent fantasy of peering underneath a dress or a three-piece suit. Privacy issues may be a very real impediment to their deployment. Other techniques that marshal lasers, infrared imagers or even antibodies were also considered by the NRC as prospective avenues for new research. For the moment, standoff detection—the ability to issue an alert from 30 to 50 feet away for a suicide bomber and from 500 to 1,000 feet for a car or truck bomb—is more of a concept than a new class of technology. "I don't think there's a lot out there that is really effective for either London or Iraq," notes Christopher K. Murphy, the program study director for the NRC report.

Standoff could necessitate vast networks of imaging or trace sensors deployed throughout cities. It might require both sensor types to cross-check for false alarms. Powerful software would have to ferret out actual threats from thousands of inputs and issue a warning within a matter of seconds. The challenges are immense. Even with the availability of x-ray vision sensors, it might be simpler just to call Superman.

Deconstructing Fido

YET THE QUEST to unseat the top dog continues. A professor of chemistry from the University of Arizona with an inventor's bent has assembled a souped-up trace detector, a distant cousin of those used in airports, that he believes might one day guard the London subways and other underground transit systems. M. Bonner Denton is an instrumen-



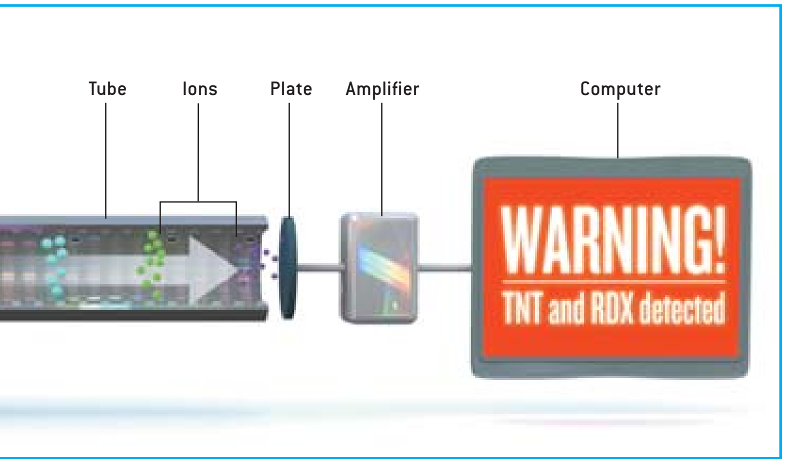
SOMETHING SMELLS: A trace detector positioned alongside an escalator picks up the scent of a bomb-carrying passenger (*in red*) by first sucking air into a collector. The beta particles ionize a gas that ionizes the explosive molecules in the air. The resulting ions are propelled down a tube by an electrostatic field. The smaller ions travel faster and hit the plate first. An amplifier then converts the small current on the plates into a large output voltage, allowing for remote detection. Software looks for the electrical signature of an explosive, and a camera captures the image of the bomber. This representation shows a concept for a system that is at least several years from commercial implementation.

BRYAN CHRISTIE DESIGN

tation jock who has trouble containing his enthusiasm for any type of machine, from spectrometers to race cars. Adjoining Denton's house on the outskirts of Tucson, Ariz., is a machine shop where he builds devices for his university laboratory. But he also uses the workshop to put together sports cars for racing on the Bonneville Salt Flats. In 2001 he garnered the land-speed record for modified sports cars—with an average during two runs of 264.007 miles per hour. And he hopes to break the 300 mark next August, regaining the land record for all classes of sports car, which he yielded in 2004 to a competitor.

Denton has made his mark as more than a car tinkerer, however. Beginning in the late 1970s, several years after joining the faculty at Arizona, he pioneered the marriage of detector arrays—charged-coupled and charge-injection devices—with spectrometers that are staples in chemistry laboratories. He developed instruments for astronomers, environmental scientists and geologists, among others. (Denton holds a joint appointment as professor of geosciences.) A few years ago he adapted a technology that amplifies the signal from an infrared array used by astronomers at the university's nearby Stuart Observatory for use with ions instead of photons.

For explosives detection, he coupled the amplifier to an ion-mobility spectrometer, the technology in airport trace detectors. As an example of its use, vapors or particles of TNT, PETN, RDX or other designations from the pack of nitrogen-based explosives enter a collector and are processed before being ionized and accelerated down a tube toward a metal plate. The smaller ions hit the plate first, the larger ones later. On the other side of the plate, an amplifier converts the current generated by the ions hitting the plate to a fairly large output voltage (16 microvolts for a single ion), while adding relatively little noise. The relevant software then looks for a particular ion signature characteristic of explosives.



The amplifier's sensitivity, according to Denton, is more than 1,000 times that of the amplifiers used in a conventional trace detector. It can detect a few tens of attograms of explosive (an attogram being 10^{-18} of a gram) as it samples the air each 20 milliseconds, Denton says. The sensitivity is such that he estimates it may be able to determine the presence of an explosive at least 15 feet away and perhaps as far away as 50 feet, making it suitable for placement on a subway escalator or staircase. The cost of a unit, Denton says, could conceivably go as low as $2,000.

In March, just before a session at an American Chemical Society meeting at which Denton made a presentation about his work, the university issued a press release in which the tinkerer compared the device to a *Star Trek* tricorder. Some of the researchers at Sandia National Laboratory, which has provided funding for most of Denton's work, went ballistic. "Right now there's no trace technology to do standoff at either 15 or 50 feet," says Kevin Linker, a Sandia expert in explosives detection. Sandia, which has worked on a portable trace detector called MicroHound and which is evaluating Denton's amplifier for incorporation in its system, makes no claims about remote detection of explosives. The low concentrations of particles and vapors involved make tracking from a distance difficult, Linker says. "From our standpoint, the claims he's making haven't been demonstrated," he adds. Denton stands his ground, remarking that the sensitivity that the amplifier lends to the spectrometer will make it the first trace detector capable of identifying explosives at a significant distance.

MILLIMETER-WAVE sensor can detect a hidden gun as well as pectoral muscles.

PHOTOCOMPOSITION BY QINETIQ

Perfect Is Not Good Enough

EVEN IF REMOTE DETECTION becomes a reality in the subway, capable of giving riders precious seconds of advance notice, that solution may not work on the streets above. A subway system has entranceways that can, in theory, be controlled, albeit with a great deal more difficulty than regulating access to airport gates. Stopping suicide bombers in a square, street or other public area may present correspondingly greater challenges. Edward H. Kaplan, a professor of operations research at Yale University, has tried to apply the types of quantitative methods used to improve decision making in business and the military to assess social issues, from needle distributions for AIDS patients to office basketball pools.

Along with Moshe Kress, another operations researcher at the Naval Postgraduate School in Monterey, Calif., Kaplan authored a study commissioned by DARPA that was published in the July 19 issue of *Proceedings of the National Academy of Sciences USA,* around the time of the London attacks. It modeled a scenario in which "perfect sensors" might ensure sufficient forewarning of a suicide bomber. The model showed that sensors might save only a few lives in a crowded plaza or street. In some circumstances, the tip-off might even increase the casualty count. As a crowd grows denser, any given person's probability of exposure to a bomb fragment declines exponentially, because the front line of the crowd acts as a shield. When people flee, the stragglers are still exposed to the blast effects, which can cause more victims than if the initial more closely packed group had remained in place. "The suggestion is clear that last-minute defenses don't seem the way to go when you're talking about random targets," Kaplan says.

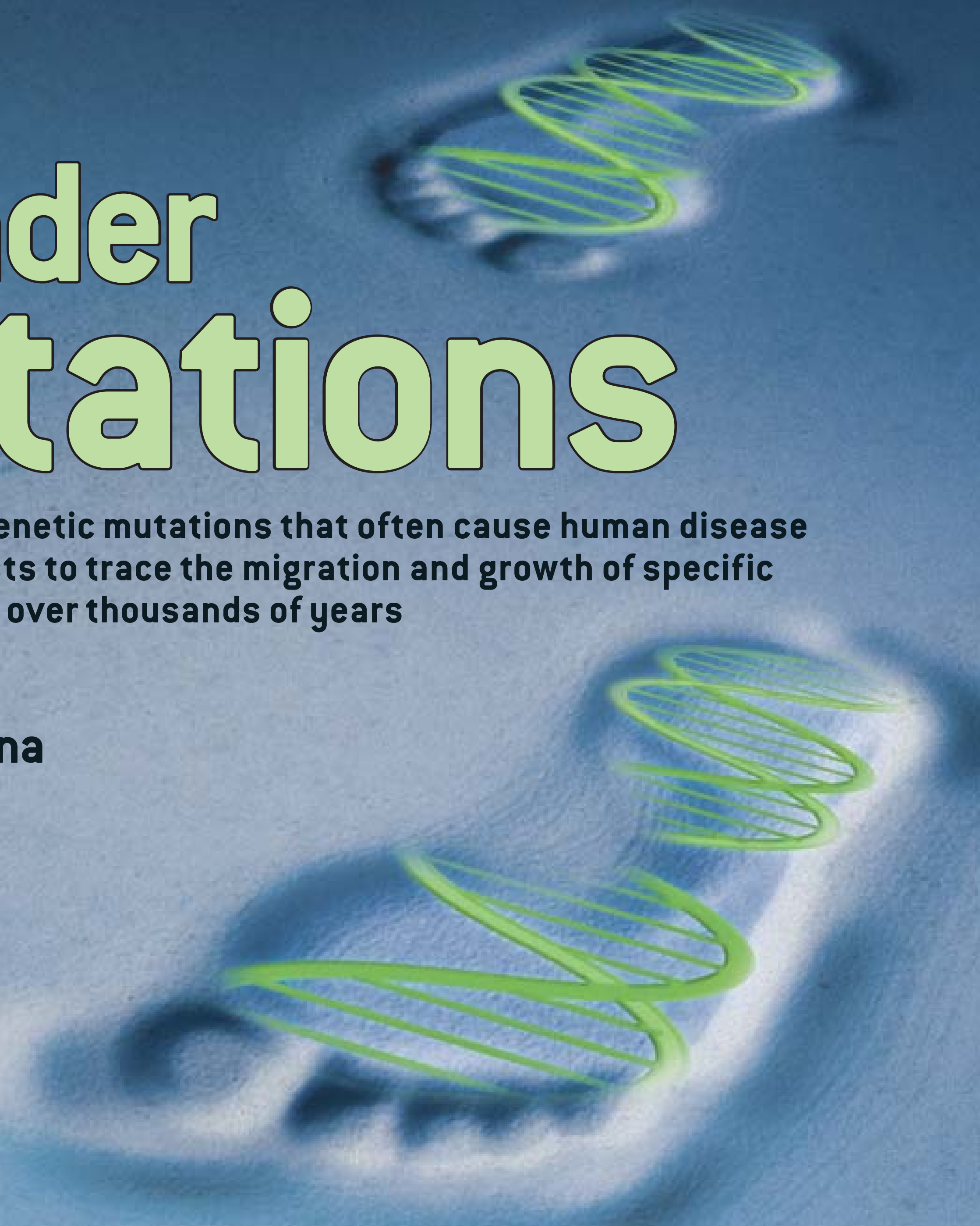
The authors recommend that intelligence resources be brought to bear rather than trying to build a ubiquitous network of detectors. Putting energy into recruiting Pashto, Arabic and Urdu-speaking intelligence officers may accomplish more than a quixotic quest for the tricorder. SA

MORE TO EXPLORE

Existing and Potential Standoff Explosives Detection Techniques. National Academies Press, 2004. Available at **http://books.nap.edu/catalog/10998.html**

Survey of Commercially Available Explosives Detection Technologies and Equipment. NIJ Office of Science and Technology, Washington, D.C., 2004.

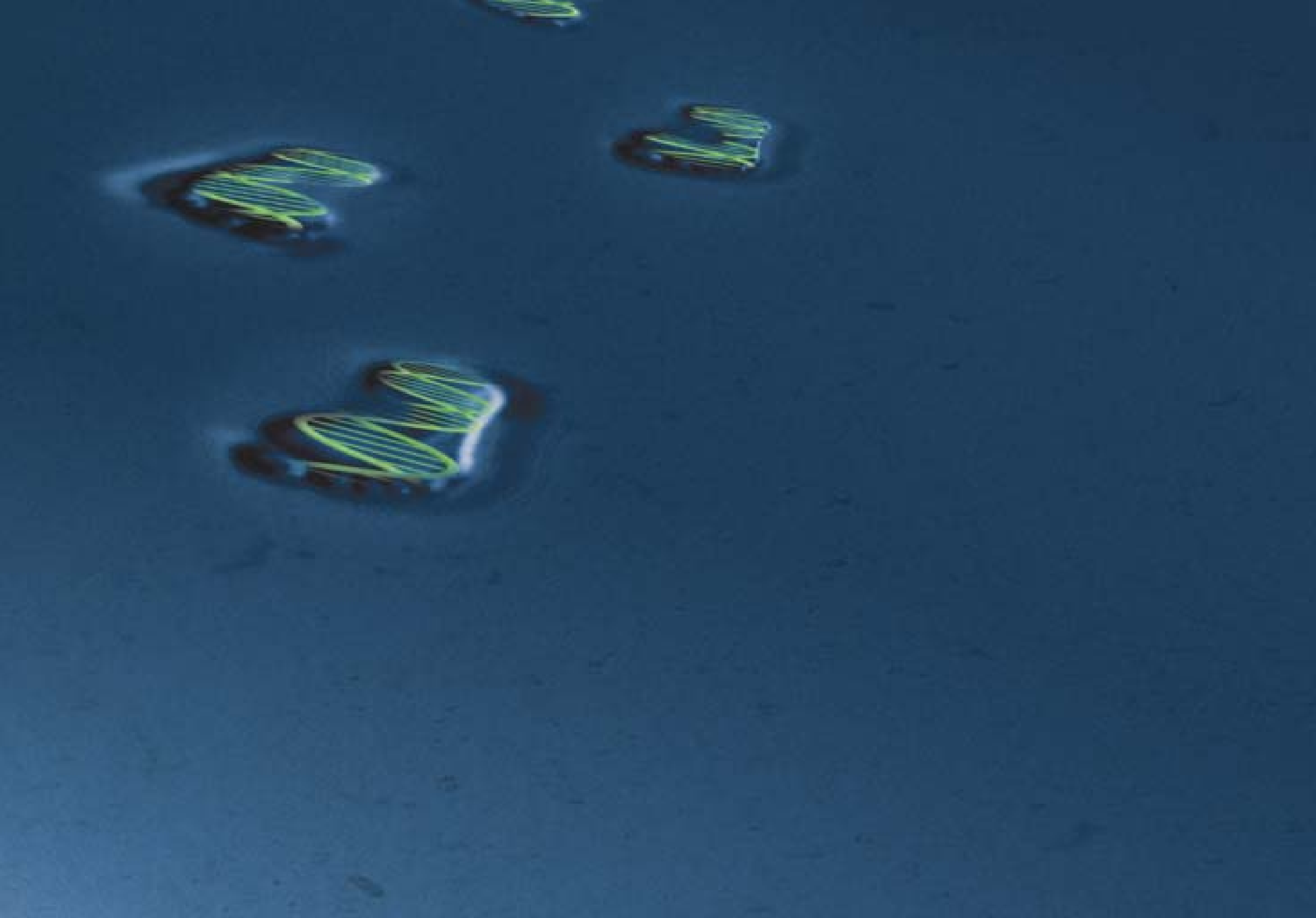
Operational Effectiveness of Suicide-Bomber-Detector Schemes: A Best-Case Analysis. Edward H. Kaplan and Moshe Kress in online *Proceedings of the National Academy of Sciences USA,* Vol. 102, No. 29; July 19, 2005.



Founder Mutations

A special class of genetic mutations that often cause human disease is enabling scientists to trace the migration and growth of specific human populations over thousands of years

By Dennis Drayna



Two middle-aged men who live thousands of miles apart in the U.S. and have never met each other may have a common trait: a propensity to absorb iron so well that this seeming benefit can actually become unhealthy, potentially causing multiple-organ damage and even death. Someone with this condition, called hereditary hemochromatosis, often has it because each of his parents passed on to him the same mutation in a specific gene, an error that originated long ago in a single individual in Europe. The mutation was then carried through time and space in that European's descendants, who now include some 22 million Americans possessing at least one copy of the gene—including the two men, who might be surprised to learn that they are related. The long-gone ancestor is known as the founder of this population, and his or her genetic legacy is called a founder mutation.

Geneticists have discovered thousands of mutations responsible for diseases in humans, but founder mutations stand apart. The victims of many genetic diseases die before reproducing, stopping the mutant genes from reaching future generations. But founder mutations often spare their carriers and therefore can spread from the original founder to his or her descendants. And some of the disorders resulting from these mutations are common, such as the hereditary hemochromatosis caused by the mutation mentioned above, as well as sickle cell anemia and cystic fibrosis. (Why does evolution preserve rather than weed out such seemingly detrimental mutations? Nature's logic will be illustrated presently.)

Medical researchers study disease mutations in the hope of finding simple ways to identify at-risk groups of people, as well as coming up with new ideas for preventing and treating the conditions related to these mutations [*see box on page 83*]. But in a remarkable by-product of such efforts, investigators have discovered that founder mutations can serve as the footprints humanity has left on the trail of time—these mutations provide a powerful way for anthropologists to trace the history of human populations and their migrations around the globe.

The Uniqueness of Founder Mutations

AN APPRECIATION of the unusual status of founder mutations and why they can provide so much information requires a brief examination of mutations in general. Mutations arise by random changes to our DNA. Most of this damage gets repaired or eliminated at birth and thus does not get passed down to subsequent generations. But some mutations, called germ-line mutations, are passed down, often with serious medical consequences to the offspring who inherit them—more than 1,000 different diseases arise from mutations in different human genes.

Founder mutations fit in the germ-line category but are atypical. Inherited diseases ordinarily follow two general rules. First, different mutations in the same gene generally cause the same disease. As a consequence, different families affected by the same disease usually have different mutations responsible for that disease. For example, the bleeding disorder hemophilia is caused by mutations in the gene encoding factor VIII, a component of the blood-clotting system. In general, each new case of hemophilia carries a discrete, single mutation in the factor VIII gene—researchers have spotted mutations at hundreds of locations in the gene.

In a few disorders, however, the same mutation is observed over and over. And there are two ways this identical mutation can arise—as a hot-spot mutation or a founder mutation. A hot spot is a DNA base pair (the individual units of DNA) that is especially prone to mutation. For example, achondroplasia, a common form of dwarfism, usually occurs as a result of a mutation at base pair 1138 in a gene called *FGFR3* on the short arm of human chromosome 4. Individuals who harbor hot-spot mutations are usually not related to one another, and thus the rest of their DNA will vary, as is typical of unrelated people. Founder mutations, which get passed down intact over the generations, are quite distinct from spontaneous hot-spot mutations.

In everyone with a founder mutation, the damaged DNA is embedded in a larger stretch of DNA identical to that of the founder. (Scientists refer to this phenomenon as "identical by descent.") This entire shared region of DNA—a whole cassette of genetic information—is called a haplotype. Share a haplotype, and you share an ancestor, the founder. Furthermore, study of these haplotypes makes it possible to trace the origins of founder mutations and to track human populations.

The age of a founder mutation can be estimated by determining the length of the haplotype—they get shorter over time [*see box on page 82*]. The original founder haplotype is actually the entire chromosome that includes the mutation. The founder passes on that chromosome to offspring, with the founder's mate contributing a clean chromosome. These two chromosomes, one from each parent, randomly exchange sections of DNA, like two sets of cards being crudely cut and mixed.

The mutation will still be embedded in a very long section of the founder's version of DNA after only one recombination, just as a marked card would still be accompanied by many of the same cards that were around it in its original deck after only one rough cut-and-mix. But a marked card will have fewer of its original companions after each new cut-and-mix. And the haplotype that includes the mutated gene will likewise get whittled down with each subsequent recombination.

A young founder mutation—say, only a few hundred years old—should thus be found in the midst of a long haplotype in people who have it today. An ancient founder mutation, perhaps tens of thousands of years old, rests in a short haplotype in current carriers.

The hemochromatosis gene aberration is just one of a rogue's gallery of founder mutations. A number of others are known and well studied in Europeans, and a few are now recognized in

Overview/*History in a Sequence*

- Founder mutations are a special class of genetic mutations embedded in stretches of DNA that are identical in all people who have the mutation. Everyone with a founder mutation has a common ancestor—the founder—in whom the mutation first appeared.
- By measuring the length of the stretch of DNA that includes the founder mutation and by determining who currently carries the founder mutation, scientists can calculate the approximate date at which that mutation first appeared and its route of dispersion. Both pieces of data provide information about the migrations of specific groups of people through history.
- As discrete populations mix, disease-causing mutations now associated with specific ethnic groups will be found more randomly. Future medicine will turn to DNA analysis to determine risks of diseases currently associated with ethnicity.

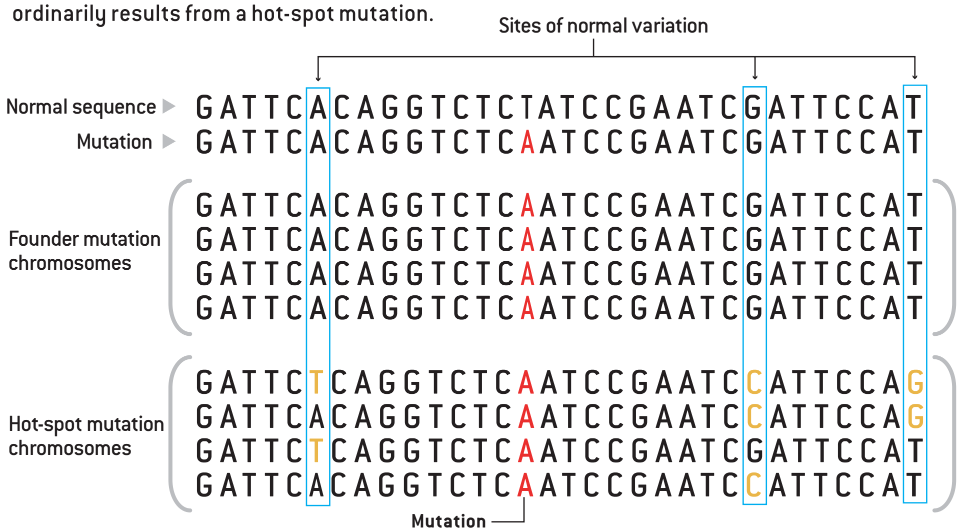
SLIM FILMS (*preceding pages*)

AN OLD ORIGINAL VS. NUMEROUS NEWCOMERS

If a group of patients with the same disease all had the same mutation at a given spot in their DNA, how could physicians know whether they were looking at a hot spot or a founder mutation? They could tell by analyzing the surrounding DNA sequences.

Suppose that in all patients the code at one spot changed from a T to an A (*red, below*). If A were a founder mutation, the surrounding sequences in all patients would be identical—the patients would have inherited the full sequence from the same distant ancestor. But if A were a hot-spot mutation, having occurred spontaneously at a place where DNA is prone to error, the surrounding sequences would also show other differences (*gold*) at sites where DNA codes normally tend to vary without causing disease.

Sickle cell disease, marked by misshapen red blood cells (*top photograph*), is usually caused by a founder mutation. Achondroplasia, a form of human dwarfism (*bottom photograph*), ordinarily results from a hot-spot mutation.



ALISON KENDALL (*illustration*); GOPAL MURTI *Photo Researchers, Inc.* (*top photograph*); WELLCOME PHOTO LIBRARY (*bottom photograph*)

Native American, Asian and African populations [*see box on page 84*]. A striking fact is how common these mutations can be—hundreds or even thousands of times more frequent than typical mutations that cause disease. Most disease mutations exist at a frequency of one in a few thousand to one in a few million. But founder mutations can occur in as much as a few percent of the population.

This anomaly—shouldn't evolution get rid of these harmful genes rather than select for them?—offers an important clue as to why founder mutations persist and spread, over land and sea and across time.

The answer, perhaps not surprisingly, is that under some circumstances founder mutations prove beneficial. Most founder mutations are recessive: only a person with two copies of the affected gene, one from each parent, will suffer from the disease. The much larger percentage of people with only one copy are called carriers. They can pass on the gene to their children and have no symptoms of disease themselves, and the single copy of the founder mutation gives the carrier an advantage in the struggle for survival.

For example, carriers of the hereditary hemochromatosis mutation are thought to be protected from iron-deficiency anemia (a life-threatening condition in the past), because the protein encoded by that mutated gene makes the person absorb iron more effectively than can those who carry two normal copies of the gene. Carriers thus had an edge when dietary iron was scarce.

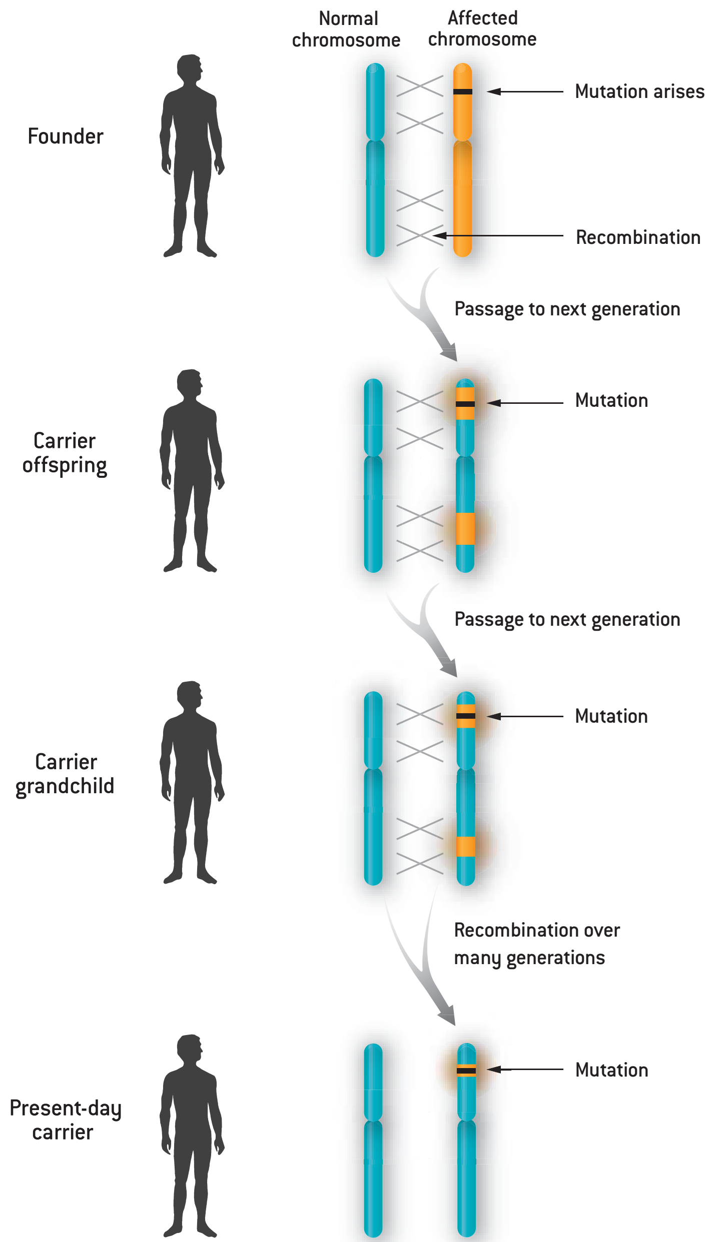
Perhaps the best-known example of a double-edged genetic mutation is the one responsible for sickle cell disease. The sickle cell mutation apparently arose repeatedly in regions riddled with malaria in Africa and the Middle East. A single copy of a sickle cell gene helps the carrier

THE AUTHOR

DENNIS DRAYNA received his bachelor's degree from the University of Wisconsin–Madison in 1975 and his Ph.D. from Harvard University in 1981. He did a postdoctoral fellowship at the Howard Hughes Medical Institute at the University of Utah and then spent 14 years in the biotechnology industry in the San Francisco Bay Area, where he identified a number of different human genes involved in cardiovascular and metabolic disorders. In 1996 he joined the National Institutes of Health, where he currently serves as a section chief in the National Institute on Deafness and Other Communication Disorders. His primary research interests are the genetics of human communication disorders, work that has taken him to eight different countries on four continents in pursuit of families with these disorders. In his spare time he enjoys technical rock and ice climbing in equally far-flung places.

GETTING SHORTER WITH AGE

The uniquely identifiable chromosome region—the haplotype—that surrounds a founder mutation gets shorter over generations as chromosomes mix in a process called recombination. In this example, the yellow chromosome in the founder holds the founder mutation, and the blue chromosome comes from a normal parent. When the founder produces sperm or eggs, the two chromosomes exchange sections. Carrier offspring inherit a newly mixed chromosome that includes the mutation and other parts of the founder haplotype (*yellow region*). Chromosomal mixing over generations inevitably leads to a shortened haplotype.



ALISON KENDALL

survive malarial infection. But two copies doom the bearer to pain and a shortened life span. The sickle cell mutation today can be found in five different haplotypes, leading to the conclusion that the mutation appeared independently five times in five different founders. (Although sickle cell disease usually results from a founder mutation, some cases do arise from other mutations.)

The frequency of a founder mutation in the population is governed by two competing forces—someone who has two copies will probably die before reproducing, but those who have only one copy will survive preferentially over those with no copies. This produces so-called balancing selection, in which the beneficial effects drive the frequency of the mutant gene up while the harmful effects damp down the frequency. Evolution giveth and evolution taketh away, so that over time the gene maintains a relatively steady level in the population.

Researchers still have not found the advantage conferred by some disease-related founder mutations, although a gene's continuing presence does point to such a benefit. For example, a recent discovery may explain the persistence of factor V Leiden, a mutation in the factor V gene, which is responsible for another blood-clotting component. This founder mutation, present in 4 percent of Europeans, leads to thrombosis, a condition of pathological blood clots. In 2003 Bryce A. Kerlin and his colleagues at the Blood Center of Southeast Wisconsin and the Medical College of Wisconsin demonstrated that carriers of this mutation are resistant to the lethal effects of bacterial infections in the bloodstream, a huge threat to survival in the preantibiotics past and still a cause of death today.

A Gene Spread Round the World

LONG BEFORE modern transportation, founder mutations migrated great distances, journeys that in many cases took dozens or even hundreds of generations. The sickle cell trait migrated from Africa west to America on slave ships and north to Europe. A common founder mutation in a gene called *GJB2* causes deafness; this mutation has been traced from its ancient origins in the Middle East along two routes, one along the Mediterranean coast to Italy and Spain and the other along the Rhine and Danube River valleys to northern Europe. A founder mutation in a gene called *ABCA4* that causes blindness appears to have arisen in Sweden about 2,700 years ago and spread to the south and west across Europe.

The most extreme example of migration, however, is probably provided by a genetic variability in our sense of taste. About 75 percent of everyone on earth perceives a substance called phenylthiocarbamide (PTC) as very bitter. The remaining 25 percent do not experience PTC as bitter at all. My colleagues and I at the National Institutes of Health and other institutions recently discovered that the combination of three different changes brings about the form of the gene that codes for the nontaster PTC receptor. Virtually all nontasters worldwide are descended from a founder individual who had these specific alterations in this gene. (Our sense of bitter taste exists to protect us from ingesting toxic substances in plants, but what might be the advantage of the nontaster variant of the gene? We suspect that the nontaster form codes for a version of the PTC detector that has switched to sensing some other toxic substance not yet identified.)

The nontaster mutation is embedded in an exceedingly short stretch of ancestral DNA, only 30,000 base pairs in some carriers, which tells us that the founder mutation is extremely ancient—probably more than 100,000 years old. In the past year, worldwide studies have shown that seven different forms of the PTC gene exist in sub-Saharan Africa. But only the major taster and the major nontaster forms have been found at significant frequency outside of African populations. Of the five remaining forms, one is found only occasionally in non-African populations (and never in New World natives), whereas the other four are exclusively African.

The PTC nontaster mutation provides a remarkable amount of information about early human migration. Its current distribution and frequency confirms anthropological and archaeological evidence that the original population of modern humans lived in Africa and that a small subgroup of those Africans emerged about 75,000 years ago and spread across five other continents—the Out of Africa hypothesis. All existing non-African populations descend from them. But in addition to confirming previous findings, the nontaster form helps to answer one of modern anthropology's most controversial questions: As our *Homo sapiens* ancestors spread across the world, did they interbreed with the more archaic hominids they met in Europe and Asia?

These archaic hominids would al-

Yesterday's Genes, Tomorrow's Medicine

The ability to identify founder mutations has profound implications for the practice of medicine. Knowledge of such mutations can, for instance, help physicians identify patients who should be tested for certain diseases. Currently physicians may rely on an individual's ethnicity to assign some disease risks and perform further tests. For example, most sickle cell disease occurs in those of African ancestry. But as the world's peoples become more genetically mixed, it will become increasingly difficult to assign an ancestral geographic origin or specific ethnicity to any person. With ethnic background disappearing as a diagnostic clue, physicians will therefore rely on testing individuals' DNA more as they try to identify disease risks or the cause of patients' symptoms. And finding founder mutations now, while human populations remain genetically distinct, will help identify the specific genes responsible for numerous conditions.

In fact, known founder mutations may be viewed as special cases of a much larger group of disease-causing variants in our DNA. Although we do not yet know what many of these are, such variants are most likely to be ancient in origin. As the accompanying article notes, such disease-related variants were probably beneficial to humans in their ancestral homes and therefore became common in the population. But the meeting of our old genes from far-flung places with modern environments and behaviors can lead to illnesses, which have become major disorders.

Genetic evaluation will be important in the broad practice of medicine because these numerous variants probably predispose us to many common disorders, not just to rare inherited diseases. Examples of such genetic variants might be those that help us make cholesterol but now contribute to high cholesterol or those that help conserve salt but now lead to salt-sensitive high blood pressure. The recognition of specific genetic profiles tied to common deleterious conditions will mean that genetics will go from being a subspecialty of medicine, concerned with rare and obscure ailments, to center stage in the prevention, diagnosis and management of human disease. *—D.D.*

LARRY WILLIAMS *Corbis*

OBSERVING ETHNICITY is currently a quick way for physicians to estimate the risk of certain disorders. As humanity's DNA becomes ever more mixed, the DNA itself will inform doctors of an individual's predisposition for those diseases.

most certainly have had their own forms of the PTC gene, selected for as a response to natural toxins in the local flora. If other hominids produced offspring with *H. sapiens* partners, we would then expect to find different forms of the PTC gene in European, East Asian or Southeast Asian populations. But there is a conspicuous absence of such variation. We therefore believe that the examination of founder mutations in humans alive today shows that no successful interbreeding between *H. sapiens* and other human groups took place during this great out-migration tens of thousands of years ago.

Finding a Founder

A CLOSER LOOK at the haplotype at the root of hereditary hemochromatosis shows how the conjunction of historical records and genetic analysis of current populations can provide new insights into the causes and history of a particular condition. In the 1980s, before the gene for this disease was identified, medical geneticists found that almost everyone with the condition had a virtually identical stretch of DNA on one part of chromosome 6. This finding was stunning because most of these patients were apparently unrelated to one another and would thus have been expected to have random differences at any place in the sequence. Because of this unique stretch of DNA, researchers realized that patients with hereditary hemochromatosis most likely were all descendants of a common, long-lost ancestor and that the gene responsible for the condition probably sat within the shared area.

BALANCING SELECTION keeps a potentially deleterious gene circulating. In regions with malaria, spread by mosquitoes, having a single copy of a mutation in the hemoglobin gene is protective. Individuals with that mutation have higher survival rates. But those who inherit two copies of the mutation suffer from sickle cell disease and have lower survival rates. The competing forces lead to a stable level of the sickle cell mutation in the population.

Operating on this hypothesis, our research group in the 1990s performed a detailed analysis in 101 patients of the genes we could find in the relevant region of chromosome 6. We also looked at the DNA of 64 control subjects who did not have hemochromatosis. Most patients shared a long region of several million base pairs. A few, however, matched in only a smaller fraction of this region. When we compared the part of chromosome 6 that matched in *all* the patients, we found that this region contained 16 genes. Thirteen of the genes coded for proteins known as histones, which bind to and wind up DNA into sausage-shaped structures visible under the microscope during cell divisions. Histones, and the genes for them, are virtually identical throughout living things, so we thought it was unlikely that they were involved in hemochromatosis. That left three genes of interest.

Two of the genes were the same in the hemochromatosis patients and the healthy control subjects. But in one of those genes, now designated *HFE*, we discovered a mutation that was present in people who had the disease but conspicuously absent from those who did not have an iron problem. This gene thus had to be the one containing the founder mutation that causes hereditary hemochromatosis.

Our discovery of the hemochromatosis founder mutation immediately led to several questions, including, Who was this founder? When and where did this person live?

Chasing the answer to these questions led medical geneticists to join forces with anthropologists and historians, producing answers that have only recently become clear. Surveys showed that hereditary hemochromatosis occurs

Noteworthy Founder Mutations

Affected gene	Condition	Mutation origin	Migration	Possible advantage of one copy
HFE	Iron overload	Far northwestern Europe	South and east across Europe	Protection from anemia
CFTR	Cystic fibrosis	Southeast Europe/Middle East	West and north across Europe	Protection from diarrhea
HbS	Sickle cell disease	Africa/Middle East	To New World	Protection from malaria
FV Leiden	Blood clots	Western Europe	Worldwide	Protection from sepsis
ALDH2	Alcohol toxicity	Far East Asia	North and west across Asia	Protection from alcoholism, possibly hepatitis B
LCT	Lactose tolerance	Asia	West and north across Eurasia	Allows consumption of milk from domesticated animals
GJB2	Deafness	Middle East	West and north across Europe	Unknown

ALISON KENDALL

all across Europe but is somewhat more common in northern Europe. In addition, the founder mutation was present in virtually all patients in the north but appeared in less than two thirds of the eastern and southern European patients. That result meant that the other third had some other mutation in the *HFE* gene or perhaps actually had a different iron disorder altogether.

Focusing in on northwestern Europe, more detailed genetic surveys revealed that the highest frequency of the founder mutation occurs in Ireland, western Great Britain and across the English Channel in the French province of Brittany. This pattern almost perfectly overlaps the current distribution of a particular group of people: the Celts.

The Celts rose to power in central Europe more than 2,000 years ago. Some were displaced northward and westward by the expanding Roman Empire, whereas others intermixed with southern Europeans and remained in their original location. Did the hemochromatosis founder mutation arise in central Europe and move north with its migrating carriers? Or did it originate in the north? Additional studies of the surrounding DNA on chromosome 6 led to the probable answer.

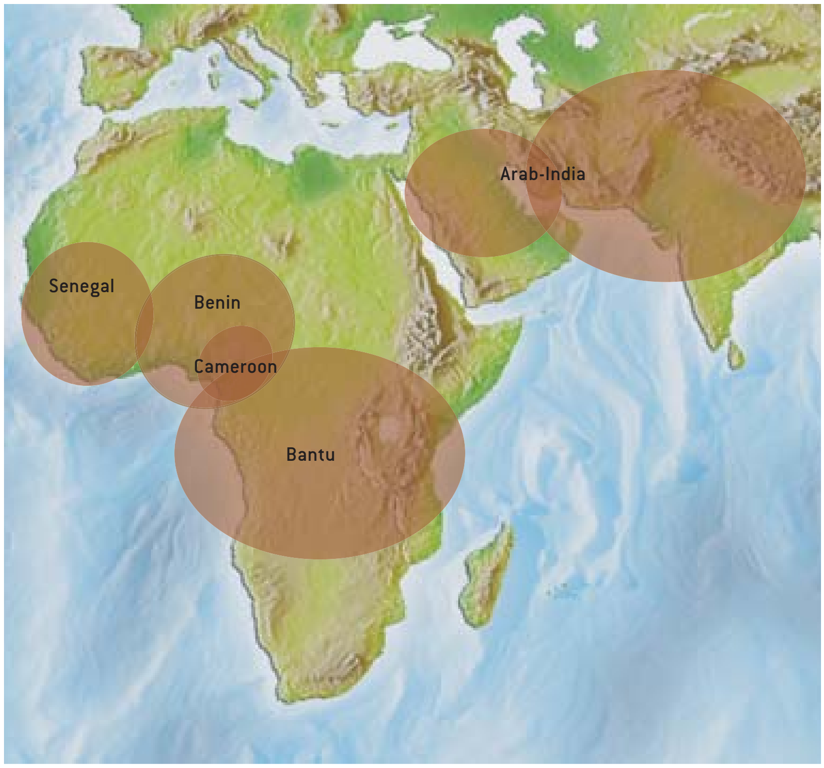
The extensive length of the modern haplotype indicates that the founder mutation is quite young, having come into being probably only between 60 and 70 generations ago, around A.D. 800. An earlier date might have led us to the conclusion that the founder lived in central Europe and that the mutation spread north and west as his descendants were driven out by an expansionist Rome. But the Roman Empire had fallen by 800, so our founder mutation most likely originated in northwestern Europe. It was then spread to the south and east by the founder's descendants.

Anthropologists, notably Luigi Cavalli-Sforza, have previously studied other types of DNA variants to trace populations. Founder mutations now add a new dimension to DNA studies: calibrating the haplotype length dates the mutation, and calculating the frequency of the haplotype in the population measures the geographic spread of the founder's descendants.

Each of us bears biochemical witness to the fact that all humans are indeed members of a single family, bound together by the shared inheritance of our genome. In addition to confirming the Out of Africa hypothesis, analyses of founder mutations have revealed the common ancestry of various other seemingly unrelated groups—recent research by David B. Goldstein of Duke University, for instance, has revealed an unexpected genetic connection between the Celts and the Basques. Further investigations of founder mutations and their haplotypes will no doubt reveal more of the genetic relationships that give us new insights into where we came from and how we arrived at our modern locations. Such study also reveals surprising kinships that may inspire a deeper appreciation for the shared roots of humanity's family tree. SA

UNCOMMON ORIGINS

People with sickle cell disease all have the same mutation. But that mutation can occur within five distinct haplotypes, indicating that the mutation arose independently five different times in human history, as indicated in areas on the map. Patients can have the Senegal, Benin, Bantu, Arab-India or recently discovered Cameroon haplotype. Eight percent of African-Americans carry at least one copy of the sickle cell mutation.



ALISON KENDALL

MORE TO EXPLORE

The Great Human Diasporas: The History of Diversity and Evolution. Luigi Cavalli-Sforza. Addison-Wesley, 1995.

Out of Africa Again ... and Again? Ian Tattersall in *Scientific American,* Vol. 276, No. 4, pages 60–67; April 1997.

Natural Selection and Molecular Evolution in PTC, a Bitter-Taste Receptor Gene. S. Wooding, U. K. Kim, M. J. Bamshad, J. Larsen, L. B. Jorde and D. Drayna in *American Journal of Human Genetics,* Vol. 74, No. 4, pages 637–646; 2004.

The National Human Genome Research Institute's overview of its International Haplotype Map Project can be found at **www.genome.gov/10001688**

Smart Wi-Fi

Wireless access to the Internet via Wi-Fi is increasingly popular, so the technology is being upgraded to ensure that users get prompt, reliable service

By Alex Hills

People love Wi-Fi access to the Internet. More and more, they are using the wireless connection technology at Starbucks cafés, in airport lounges and at home. Wi-Fi seems irresistible because it makes the Net available to users anytime, anywhere. It provides fast communications links that allow e-mail messages to appear almost instantly and Web pages to paint computer screens quickly—all with the mobility and freedom that has made cell phones nearly ubiquitous.

Pyramid Research, a communications industry research firm, predicts the global number of Wi-Fi users could top 271 million by 2008, with 177 million of them in the U.S. Today's Wi-Fi community already supports a vibrant international business in Wi-Fi equipment, estimated at about $3 billion annually, according to extrapolations of figures produced by In-Stat, another market research company. But the very popularity of Wi-Fi also brings problems. As Wi-Fi networks become ever more heavily used, they may be unable to handle the expanded traffic, causing clients' devices to become bogged down with slow service and long delays.

Even when the technology is working properly, wireless access is not as swift as that provided by high-speed wired connections to the Internet, such as digital subscriber line (DSL) or cable modem links, for example. Radio signals cannot hope to match the transmission speeds that copper wires or fiber-optic cables make possible. Nor can Wi-Fi, or other wireless technologies that rely on radio, supply the same degree of security; the transmissions can be intercepted by nearby radio receivers.

Many of these problems were evident even in 1993, when I led a team at Carnegie Mellon University to build Wireless Andrew, the first large-scale wireless local-area network (LAN) and a precursor to today's Wi-Fi networks. Completed in 1999, Wireless Andrew now connects the entire campus [see "Terrestrial Wireless Networks," by Alex Hills; Scientific American, April 1998].

A dozen years since the inception of our wireless network at Carnegie Mellon, much has happened in the world of wireless. Some difficult problems have arisen because of markedly increased Wi-Fi use, but substantial progress has also been made in solving them. Before considering these developments, however, we must discuss how Wi-Fi operates.

ENHANCED WI-FI ACCESS will result as engineers give wireless networks the automated "smarts" they need to deal with the growing number of mobile users trying to get on the Internet.

PETER HOEY


INTERNET
SMART WI-FI

Wi-Fi Workings

WI-FI NETWORKS comprise Wi-Fi-equipped mobile computers (laptops or handhelds) or special Wi-Fi telephone handsets, as well as access points (APs). APs are base stations that communicate by radio and by wire with both mobile systems and the networks that ultimately provide entrée to the Internet. Each AP can send and receive signals within a limited range, typically 20 to 50 meters inside a building. The coverage area of an AP forms a three-dimensional spherelike cell (analogous to a mobile telephone cell but much smaller) that can serve many mobile devices within it simultaneously [*see illustration at right*].

Wi-Fi networks were originally called wireless LANs. Before 1997 wireless LAN equipment did not interoperate—systems made by one manufacturer did not talk to those produced by other companies. But in 1997 the Institute of Electrical and Electronics Engineers adopted the IEEE 802.11 standard, which eliminated that incompatibility. Today most wireless LAN equipment conforms to this standard, which is popularly called Wi-Fi (for wireless fidelity). Although it does not dictate all aspects of network operation, the standard does assure that different equipment types can work together.

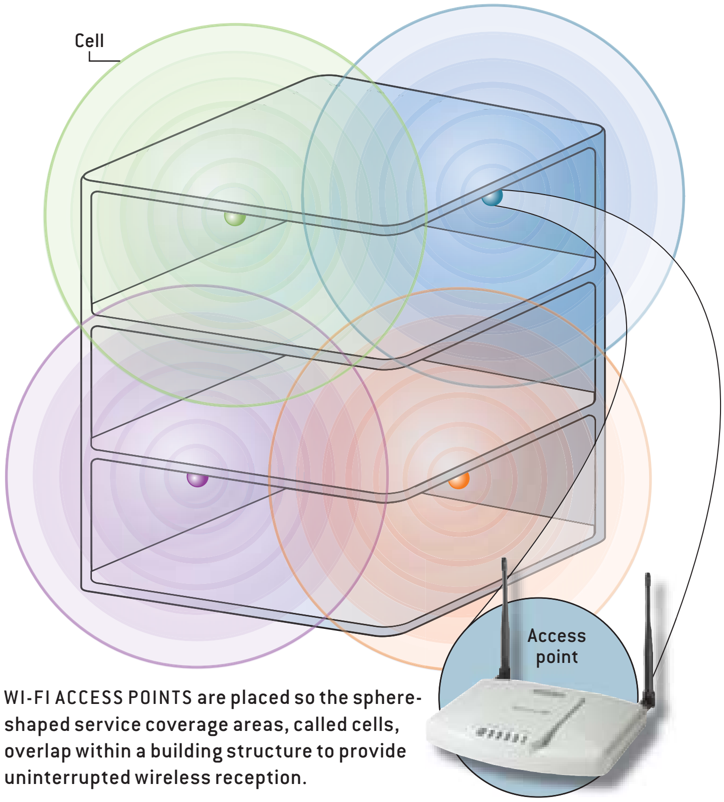
Four major concerns face the designers of Wi-Fi networks: ensuring reliability by making certain that service is not disrupted by poor-quality radio transmissions; maintaining performance by avoiding slow link speeds and overlong delays; designing AP networks that can completely blanket the coverage area; and providing security against unfriendly wireless eavesdroppers or unauthorized users.

The main reason that wireless LANs are subject to these problems is that the technology relies on radio transmission, which has its own specific operational drawbacks [*see box on opposite page*]. A signal received by a client or an AP can be degraded in various ways:

- A wireless transmission is attenuated—that is, weakened by distance even when there are no obstructions (which can cause additional reductions in radio signal strength).
- A radio wave can suffer multipath distortion by reflecting off walls and building structures, furniture, equipment or other objects in the near environs. Signals may then follow multiple paths from transmitter to receiver, which causes numerous copies of the same transmission to arrive at the receiver, each at a slightly different time. The delayed duplicates can corrupt the direct (line-of-sight) signal, creating reception problems.
- A third kind of signal degradation results from interference and noise effects. Interference is generated by conflicting radio transmissions. One common source of Wi-Fi network interference is the microwave oven, which can release stray radio signals. Fortunately, modern microwave ovens are well shielded to keep these emissions to a minimum. Radio noise occurs in nature but also comes from man-made sources such as electrical machinery, automobile engines and fluorescent lighting.

Overview/*Wireless LANs*

- As Wi-Fi technology—wireless access to the Internet—grows ever more popular, increased traffic threatens to overwhelm the radio-based local-area networks (LANs) people use to link with the Net, potentially causing unacceptable delays and service disruptions. A series of technical enhancements, comprising second-generation or smart Wi-Fi technology, will go a long way toward solving these problems.
- Designers of Wi-Fi networks worry about four issues: avoiding poor-quality radio transmissions; preventing slow link speeds and long delays; blanketing user areas with coverage; and providing sufficient security. Smart Wi-Fi, which is just starting to come into operation, will accomplish all these tasks and more.



WI-FI ACCESS POINTS are placed so the sphere-shaped service coverage areas, called cells, overlap within a building structure to provide uninterrupted wireless reception.

Communications engineers are accustomed to overcoming these difficulties, but unfortunately their methods can slow transmission speeds. Whereas wired Ethernet networks provide service at speeds from 100 to 1,000 megabits per second (Mbps), many wireless LANs employ the IEEE 802.11b standard and so operate at rates up to 11 Mbps. Newer IEEE 802.11a and 802.11g equipment can run at speeds up to 54 Mbps—still a bit sluggish compared with Ethernet operations. A soon-to-be-introduced version of IEEE 802.11 will allow communications as fast as 108 Mbps, however.

These numbers in fact overstate Wi-Fi transmission rates. Wi-Fi automatically drops from the maximum speed (11 or 54 Mbps) to a lower rate to cope with radio signal attenuation, multipath, interference and noise conditions. Hence, an

LUCY READING-IKKANDA (*illustration*); SYMBOL TECHNOLOGIES, INC. (*wireless router*)

IEEE 802.11b link may step down from a data rate of 11 Mbps to 5.5, 2 or even 1 Mbps. In addition, overhead bits—extra digital bits that are added to each transmission to control network operation and reduce errors—further reduce the effective data rate.

Since the introduction of the initial Wi-Fi technology, my colleagues and I at Carnegie Mellon and Airespace (now a part of Cisco Systems), as well as engineers at other universities and companies, have worked to solve its shortcomings in the areas of reliability, performance, design and security. The resulting second-generation Wi-Fi equipment (called smart Wi-Fi technology here) embodies various new capabilities intended to overcome existing problems. These enhancements rely on greater intelligence in the Wi-Fi systems.

> Wireless LANs are subject to problems because the technology relies on radio, which has drawbacks.

Avoiding Congestion

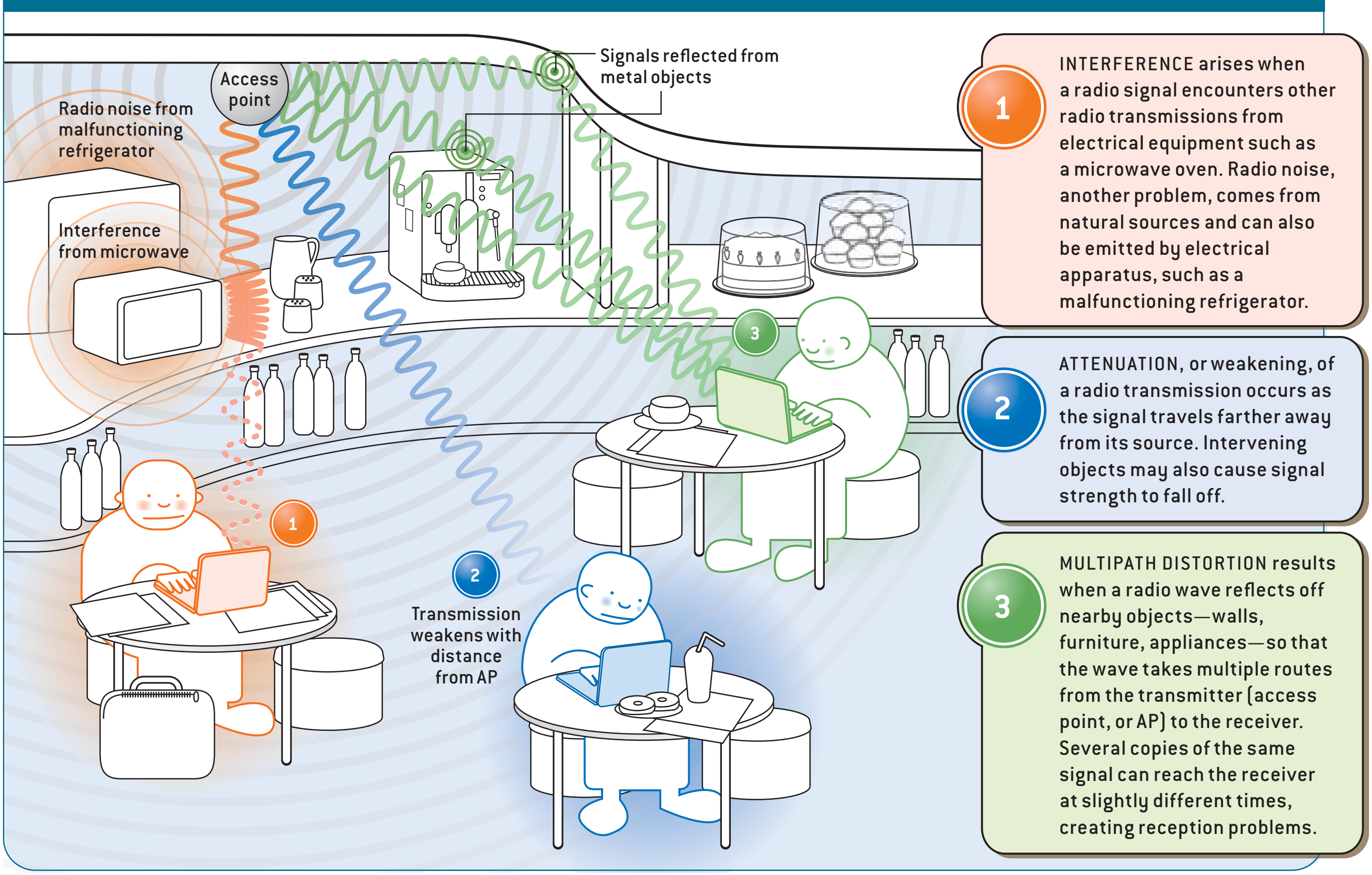
SMART WI-FI TECHNOLOGY will improve a user's experience with a wireless network by dealing with the issues of congestion, the changing radio environment and security in several ways.

Network congestion—when an AP is called on to serve many users and thus becomes overloaded—is likely to cause delays and degrade service significantly. Because a cell's AP and the clients using it must share a single radio channel (segment of the radio spectrum) and only one station (an AP or a client) can successfully transmit at a time, conflicts can occur. Wi-Fi networks currently resolve clashes between competing stations within a cell by using a technique called carrier sense multiple access with collision avoidance: the CSMA/CA protocol.

Under CSMA/CA, each station listens before sending a signal. If a station hears another one in the process of sending, it defers and waits until the channel is free. If two stations attempt to send at about the same time, neither will hear the other and their transmissions will collide. When this happens, neither transmission is received correctly, and repeat transmissions must be made. When many computers are using a single AP, collisions occur frequently, so multiple repeat transmissions are needed and all the users face delays [*see box on next page*].

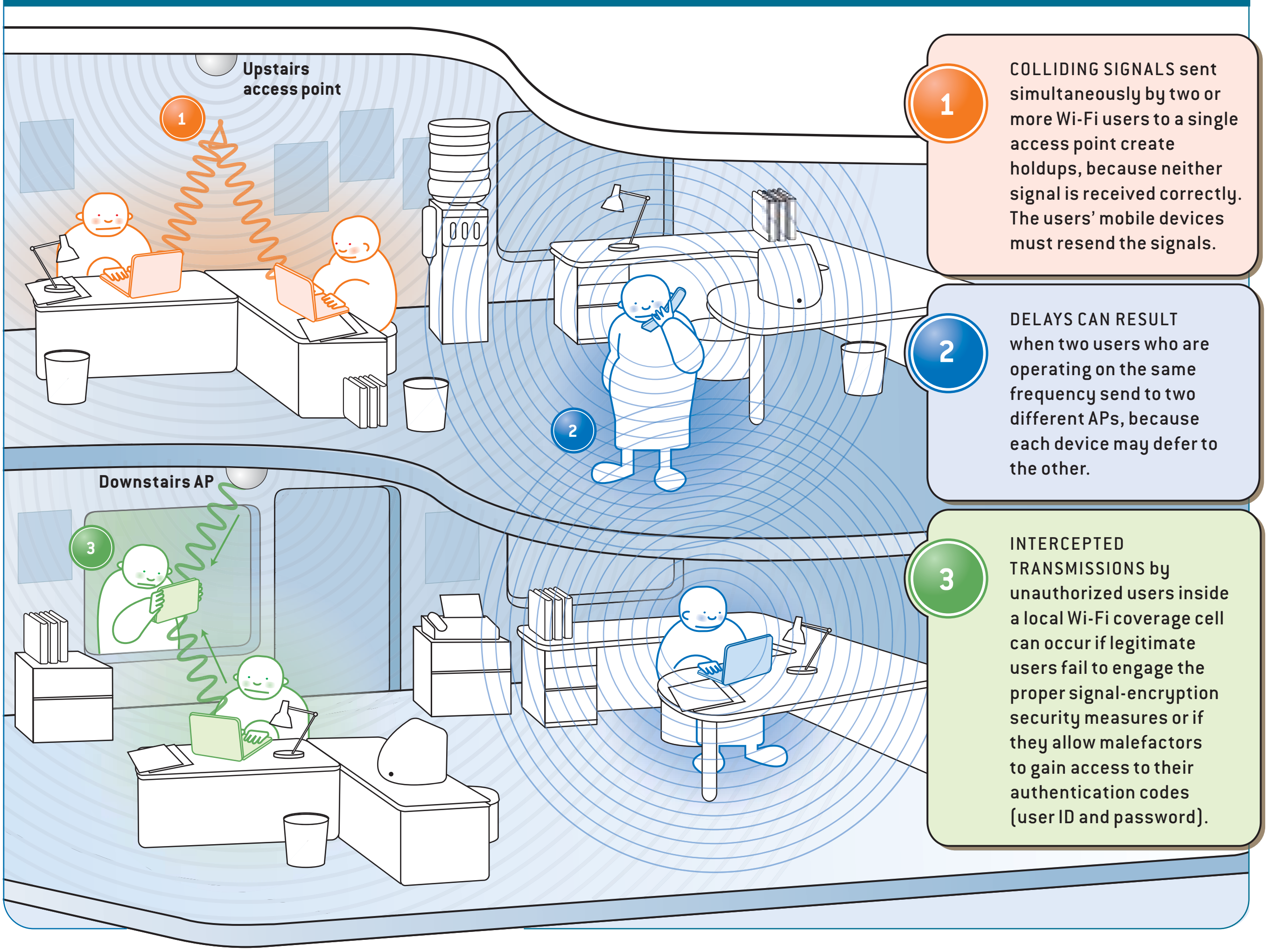
The problem of overloaded APs can be quite severe in areas of high user density. At Carnegie Mellon we first experienced this problem in our large lecture halls and classrooms. My team quickly realized that we could not even approach wired network performance in these crowded spaces,

TRANSMISSION GLITCHES: HOW WI-FI CAN FAIL



LUCY READING-IKKANDA

WOES OF WIRELESS NETWORKS



which can, at times, hold hundreds of mobile computer users.

The CSMA/CA protocol can also cause special difficulties among distant APs and mobile devices that are operating on the same radio channel. If one AP or mobile device can hear a far-off (co-channel) AP or client, it will defer, just as it would to a station transmitting within its own cell. This co-channel overlap produces another kind of performance degradation [*see box above*].

Suppose, for instance, that Jane and Joe are using devices operating on the same radio channel but located in different parts of a building and associated with different APs. If Joe's system can hear Jane's, it will defer every time Jane's system transmits, delaying messages waiting to be sent by Joe's. Similarly, if Jane's system can hear Joe's, it will be unable to send whenever Joe's is transmitting, degrading her communications service. This problem would be particularly noticeable if either is using a voice handset.

Designers can mitigate these situations by assigning channels carefully and by using a new feature called load balancing, which helps to reduce the chance of overtaxing an AP. Load balancing relies on the fact that clients may be within radio range of two or more APs. Smart Wi-Fi networks attempt to relieve congestion by distributing clients among APs more or less uniformly so that no one AP gets swamped, streamlining performance considerably.

A link between a client and an AP is called an association. It begins when a client initiates an association request. When an AP receives a request, it can either accept or deny it. Although the IEEE 802.11 standard does not specify a software algorithm for making this kind of determination, a second-generation AP (or the intelligent switch that controls it) considers the AP's current load and also those being carried by near-

THE AUTHOR

ALEX HILLS is Distinguished Service Professor of engineering and public policy and electrical and computer engineering at Carnegie Mellon University. He has also served as vice provost and chief information officer there. His teaching and research efforts focus on wireless technology and telecommunications policy. Hills developed Wireless Andrew, an early wireless local-area network, and invented Rollabout, a wireless design tool that has been commercialized by Helium Networks. He is an avid runner, backpacker and cross-country skier.

LUCY READING-IKKANDA

When many computers use an access point, collisions occur and all the users face delays.

by APs to help make a decision. A heavily loaded AP might not be the best one to associate with a new client. If such a request is received and the system knows that a lightly loaded AP is also within radio range of the requesting client, the AP may deny the association request, thus boosting the overall performance of the network [*see box below*]. Along with other techniques, load balancing will allow future Wi-Fi networks to perform well, even in high-density locations.

Changing Radio Environments

THE AFOREMENTIONED radio-related problems of attenuation, multipath, interference and noise can be alleviated substantially by good network design. A Wi-Fi network designer has to decide where to place APs within a target space to provide for adequate coverage and performance. He or she must also choose which radio channels to assign to which APs. A designer needs to consider the characteristics of the radio environment and the geometry of the building in which the wireless LAN will be deployed to implement what is actually a three-dimensional radio network.

In selecting AP locations, a network designer aims to avoid coverage gaps, but he or she must simultaneously space the APs as far apart as possible to minimize the cost of equipment and installation. Another reason to separate the APs is that coverage overlap between APs operating on the same radio channel (known as co-channel overlap) degrades performance. Channel assignment, which is the second part of the design process, is typically carried out so as to minimize co-channel overlap, which reduces interaction between stations in different co-channel cells.

Another new smart Wi-Fi feature, automatic cell-size control, allows cell sizes to expand and contract to adjust for changing radio conditions. The technique can also compensate for a less than careful design or for AP failures.

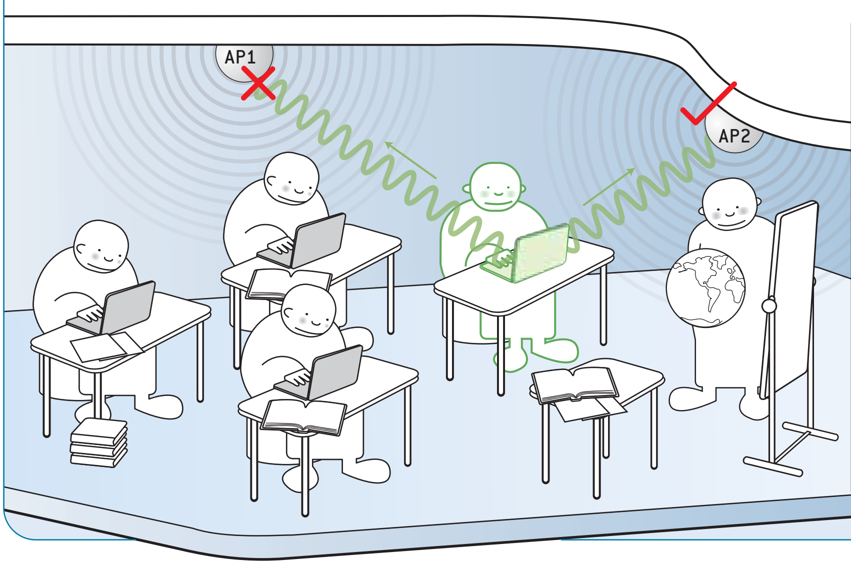
Even for a very carefully configured network, the local radio environment can change from time to time. Thus, the original conditions may no longer exist. When metallic equipment is moved in a factory, for example, a shift in the electromagnetic conditions can lead to coverage gaps. In this case, it is appropriate to expand or contract cell sizes to compensate. Cell sizes can be altered by adjusting the transmitter power output of Wi-Fi APs. If the modifications accurately reflect the new radio environment, continuous network coverage can be maintained throughout the target space without undue cell overlap. (Currently APs can modify only their own transmit power levels, but pending additions to the IEEE 802.11 standard will permit APs to instruct clients to increase or decrease their transmit power as well.)

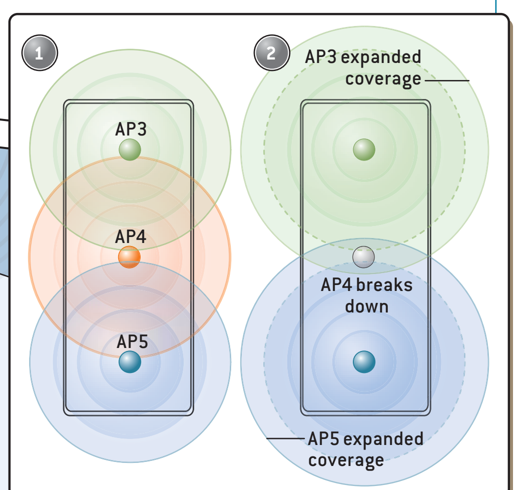
Automatic cell-size control also has the potential to reduce the effort required to design a wireless LAN. This feature makes possible an abbreviated design process that places APs in reasonable, if not optimal, locations. Moreover, APs break

INTELLIGENT WI-FI NETWORKS ADAPT

A smart Wi-Fi network can relieve congestion by distributing user connections evenly among the available Wi-Fi access points. This load-balancing feature (*below, left*) engages when a user attempts to link to a heavily loaded access point, such as AP1. If the system knows that a second, more lightly occupied AP is within radio range of the user, it will deny access to the first AP and connect to the second AP, which will enhance the network's overall operational performance.

When radio conditions change, a smart Wi-Fi system can alter its cell sizes to compensate. In this example (*below, right*), APs 3, 4 and 5 provide service to an interior space (*1*). When the center AP4 cell fails unexpectedly, it leaves a breach in wireless coverage. The neighboring cells—AP3 and AP5—expand to extend coverage across the gap (*2*).





LUCY READING-IKKANDA

> Smart Wi-Fi networks are beginning to behave more like their wired counterparts, and users are starting to notice.

down from time to time. Depending on the particular positions of the APs and the antenna types employed, automatic cell-size control can temporarily fill in coverage holes caused by AP breakdowns [*see box on preceding page*].

Dynamic Channel Assignment

APS MAY ALSO USE dynamic channel assignment in smart Wi-Fi networks to change radio channels automatically. Designers traditionally do channel assignment so as to minimize co-channel overlap based on the radio propagation environment. After the channel assignments have been made, they are typically static. The environment can change, however, so there is no guarantee that these assignments will remain valid.

A second-generation Wi-Fi network senses the radio environment at intervals and then dynamically reassigns channels accordingly. This capability eliminates the need to execute channel assignment during the original design process. If furniture is removed from an office area, for example, it might cause a cell's coverage region to enlarge. Should this expansion result in coverage overlap with another cell operating on the same channel, performance could drop off. It may be appropriate in this case to switch the second cell to a different channel. Channel-switching algorithms assure that co-channel coverage overlap is minimized across the entire network.

Smart Wi-Fi systems usually activate a channel-switching algorithm periodically to ensure that channel assignments reflect the current radio environment. Dynamic channel as-

Wi-Fi vs. WiMAX

Many readers may have heard of a new wireless access technology called WiMAX. What is it, and how is it related to Wi-Fi?

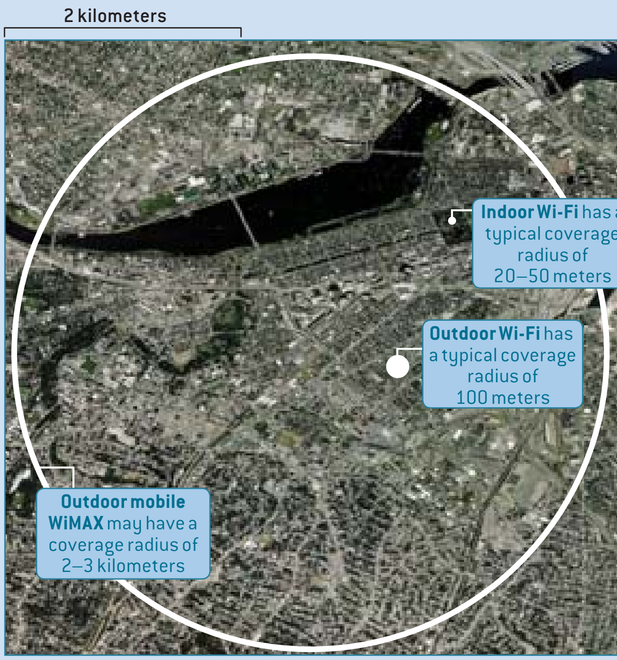
Whereas Wi-Fi is used widely by mobile devices, WiMAX was originally aimed at stationary connections to the Internet. (The term "WiMAX" is a creation of an industry group called the WiMAX Forum.)

Just as Wi-Fi is based on the IEEE 802.11 standard, WiMAX stems from the IEEE 802.16d standard, which was adopted in 2004 to define high-speed wireless service to stationary sites over distances of up to 50 kilometers. The maximum range for Wi-Fi is a few hundred meters. One reason for WiMAX's greater range is that it can transmit at higher power levels, depending on the radio band. WiMAX will have the capability to run at 75 megabits per second—many times the speed of digital subscriber lines (DSL)—but this capacity will usually be divided among many users.

WiMAX was conceived to provide the same kind of fast Internet service that is afforded by DSL, cable modem and even fiber-optic systems. For this reason, it has also been called Wireless MAN (MAN stands for "metropolitan-area network").

Despite the technology's origins, the WiMAX community more recently began work on a mobile version of the standard called IEEE 802.16e, which is popularly known as mobile WiMAX. Like Wi-Fi, it is intended to provide service to laptop computers and other mobile devices, but it will have greater range, probably up to a few kilometers.

Currently there is a good deal of computer industry buzz about mobile WiMAX, but the standard has not yet been adopted. Further, whether the technology will gain a foothold in the market still remains to be seen. It may turn out that WiMAX will not compete directly with Wi-Fi. Because of its higher power and longer range, it will most likely vie with third-generation (3G) cellular service in delivering mobile Internet service, first in urban areas and later over larger regions. Like WiMAX, 3G operates at higher power levels than Wi-Fi, and its base stations cover larger areas than Wi-Fi's do.



It is possible that eventually all three systems—Wi-Fi, WiMAX and 3G—will coexist, each filling a specialized niche. Because 3G and WiMAX run at higher power levels and employ different access schemes than Wi-Fi does, they will neither experience all the same problems nor need the same solutions described in this article.

Increasingly in the future, laptop computers and personal digital assistants (PDAs) will be equipped to work with multiple wireless networks. A laptop might connect with Wi-Fi in the home and office environments but use WiMAX or 3G elsewhere. Thus, Wi-Fi/WiMAX or Wi-Fi/3G combinations may become commonplace at some point, with Wi-Fi/WiMax/3G computers capable of linking to all three networks arriving thereafter. *—A.H.*

TERRASERVER.COM/GLOBEXPLORER; PHOTOILLUSTRATION BY LUCY READING-IKKANDA

SMART WI-FI EQUIPMENT MAKERS

COMPANY	LOCATION	WEB SITE
Aruba Networks	Sunnyvale, Calif.	www.arubanetworks.com
Cisco Systems/ Airespace*	San Jose, Calif.	www.airespace.com
Cisco Systems/ Aironet*	San Jose, Calif.	www.cisco.com
Colubris Networks	Waltham, Mass.	www.colubris.com
Extreme Networks	Santa Clara, Calif.	www.extremenetworks.com
Symbol Technologies	Holtsville, N.Y.	www.symbol.com
Trapeze Networks	Pleasanton, Calif.	www.trapezenetworks.com

**Cisco Systems recently acquired Airespace, a smart Wi-Fi company. Cisco's existing product, called Aironet, incorporates smart Wi-Fi features.*

signment techniques can also help improve performance by allowing APs to choose channels that are not experiencing local noise or interference.

Wireless Security

PROBABLY THE MOST WIDELY DISCUSSED Wi-Fi problem is security. No users want strangers to monitor their e-mail exchanges or gain unauthorized access to their system [*see box on page 90*]. The original IEEE 802.11 standard provided for transmission encryption through a feature called Wired Equivalent Privacy (WEP). Encryption is a way of converting one bit stream to another (encrypted) bit stream such that the original bit stream can be reproduced only with the use of a key, the special cipher that was originally used to do the coding. But many wireless users never bother to activate the encryption feature and so send their transmissions "in the clear," which permits easy interception.

Even with WEP in use, clever people seeking to point out its vulnerabilities found ways to discover the keys and then decrypt messages. It became widely known in 2001 that WEP was flawed, and since then developers have worked to bolster Wi-Fi network security.

Authorized access is also an issue for Wi-Fi networks. Users can identify themselves through an authentication process involving a user ID and password, but if malicious people can easily eavesdrop on others' transmissions, they can readily "snoop" a user ID and password and thereby gain access to the network.

In 2003 and 2004 the IEEE 802.11 working group and the Wi-Fi Alliance (the industry group that coined the term "Wi-Fi") completed work on their related standards, IEEE 802.11i and Wi-Fi Protected Access (WPA), which make much stronger security measures available. These include enhanced encryption techniques and substantially more secure methods for APs and clients to gain access to the keys needed to encrypt and decrypt transmissions.

WPA (which uses another standard, IEEE 802.1X) also provides a considerably stronger authentication process than was previously available. The combination of these new standards dramatically betters overall security for smart Wi-Fi networks.

Some Wi-Fi equipment makers have added other security measures as well. One example is intrusion detection. Wireless networks differ from wired networks in that eavesdropping devices (and even APs) can be anywhere in or near a wireless network's coverage area. (Wired intruders may attack from a distance.) This is why some Wi-Fi equipment uses position location technology to detect the presence of a malicious station. Using this feature, the network can track down the offending station and remove it.

With the development of smart Wi-Fi techniques, wireless networks are beginning to behave more like their wired counterparts, and wireless users are starting to notice the difference. More remains to be done in this regard, though, and research to take Wi-Fi further is continuing. Work, for example, is now under way to automatically find a mobile device in a Wi-Fi network. This capability would allow network operators to quickly locate people (say, physicians in a hospital) or objects (products moving through a factory assembly line) as needed.

Wi-Fi and other wireless communications technologies are growing—and changing—dramatically. More and more people in the U.S. and elsewhere are abandoning landline telephone service in favor of wireless cell phones, and municipal governments such as Philadelphia's are creating citywide Wi-Fi coverage areas. Meanwhile the use of third-generation (3G) cellular telephony is on the upswing, and a new wireless technology called WiMAX [*see box on preceding page*] may soon have a strong presence in the market. Increasingly, we are living in a wireless world. SA

MORE TO EXPLORE

Wireless Andrew. Alex Hills in *IEEE Spectrum*, Vol. 36, No. 6, pages 49–53; June 1999.

Large-Scale Wireless LAN Design. Alex Hills in *IEEE Communications*, Vol. 39, No. 11, pages 98–107; November 2001.

Real 802.11 Security: Wi-Fi Protected Access and 802.11i. John Edney and William Arbaugh. Addison-Wesley Professional, 2003.

Rollabout: A Wireless Design Tool. Alex Hills and Jon Schlegel in *IEEE Communications*, Vol. 42, No. 2, pages 132–138; February 2004.

Radio Resource Management in Wireless LANs. Alex Hills and Bob Friday in *IEEE Communications*, Vol. 42, No. 12, pages S9–S14; December 2004.

Wireless Networks First-Step. Jim Geier. Cisco Press, 2004.

The IEEE 802.11 Handbook: A Designer's Companion. Second edition. Bob O'Hara and Al Petrick. IEEE Press, 2005.

802.11 Wireless Networks: The Definitive Guide. Second edition. Matthew Gast. O'Reilly, 2005.

WORKINGKNOWLEDGE

ROADS

Paving the Way

Lay down some hot black stuff, and you've got a road, right? Not quite. Much more is hidden underneath.

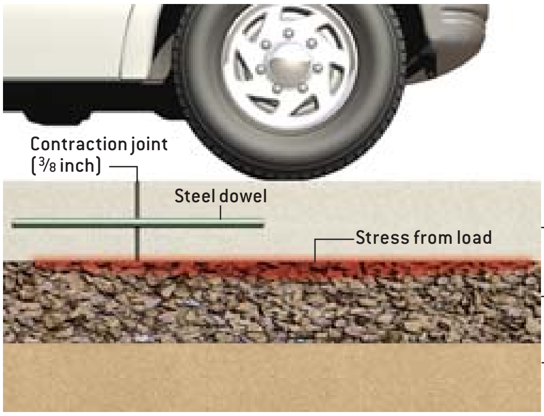
Terminology explains a lot about road construction. Scottish engineer John McAdam is generally credited with designing in the early 19th century the first modern roads made by compressing thick deposits of crushed, angular stones. Builders later poured hot tar to bind the top layer, producing a "tarmacadam" pathway, or tarmac. Although this term lingers, the method has not been used for decades (not even at airports). By the later 1800s asphalt had become the binder of choice. And blacktop? A synonym for asphalt.

Today asphalt roads dominate the landscape, followed by concrete and "unbound aggregate"—gravel. All three formats are built up in layers comprising increasingly finer, denser and harder rock particles [*see illustrations, near right*]. Top levels of asphalt or concrete reduce wear and seal out water that can cause cracking. Most fissures form from the bottom up as underbeds shift or erode. To repair them, road crews typically add a new layer. But if damage is extensive, they may mill (scrape off) several inches and resurface.

Recent innovations include "perpetual pavement," built on a costly but tough base of asphalt that is supposed to resist damage for twice as long as conventional structures. Stone-matrix designs, in which the size and shape of stones within surface layers is carefully controlled, also promise longer life; they are popular in Europe and are becoming widespread in the U.S. Porous pavement, which allows water to percolate through instead of run off, is being tried in parking lots [*see illustration, far right*].

Asphalt and concrete proponents present various arguments about which compositions are better for a given application, but ultimately the choice comes down to economics. "The decision should be made on life-cycle costs," says David E. Newcomb, vice president of research and technology at the National Asphalt Pavement Association in Lanham, Md. "It's a balancing act" of materials and labor, time needed to lay the road (and therefore traffic inconvenience), years of durability, and maintenance and repair. The final analysis, Newcomb allows, "can be a nebulous thing." —*Mark Fischetti*

ASPHALT HIGHWAY is built in layers. Soil is compacted, then covered with stone that is also compressed. Then a base course—a mix of crushed stones (aggregate) about 1.5 inches in diameter and hot liquid asphalt—is laid. Binder and surface courses are added, which have increasingly smaller, harder and more angular aggregates—as fine as three eighths of an inch—that lock tighter when bound by the asphalt. Similar but thinner layers are used for local roads. Asphalt pavement flexes slightly, transferring load directly below a vehicle deep into the base.



CONCRETE HIGHWAY is laid on compacted soil and a compressed aggregate base. A thick layer of concrete is formed in slabs joined by steel dowel bars and kept damp for several days to cure. To control subsequent cracking from thermal expansion and contraction, a joint is cut between slabs and filled with water sealant. Concrete pavement is rigid, transferring load to a wide but shallow region.

GEORGE RETSECK

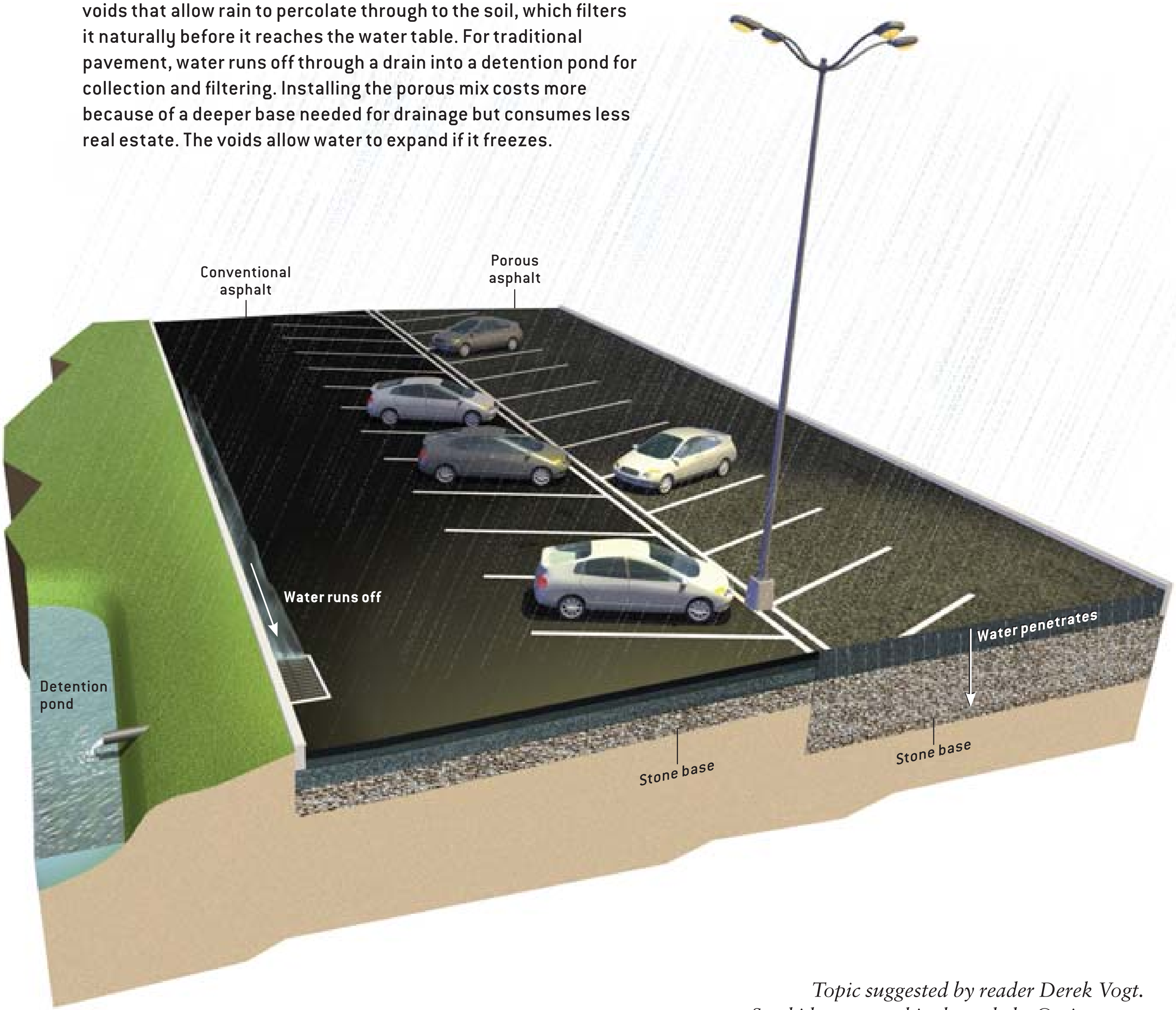
DID YOU KNOW . . .

COLD FACTS: Why the sign "Bridge may freeze before road surface"? Because cold air flows across the pavement's top and bottom, quickly drawing away residual heat as ambient temperature drops below freezing, allowing ice to form. In comparison, only the surface of a road is cooled; warmth from underneath keeps it above freezing, at least for a while. Also, most bridgeways are made of steel and concrete, which cool faster than asphalt, a poor conductor.

HOT STORAGE: "Asphalt" is really 95 percent pulverized rock and 5 percent liquid asphalt, which binds the aggregate. At a plant, the stone is crushed, dried and heated to about 300 degrees Fahrenheit as asphalt is poured in. This hot mix is stored in an insulated silo for up to several days with no further heating. It is then deposited into a dump truck and driven to a paver at a construction site, where it is still a fluid 275 degrees F. References to asphalt road material date to the Babylonians of 600 B.C. Asphalt lakes occur naturally in a few places, such as Trinidad, but most of the commodity is produced as a by-product of petroleum refining.

CHEAP CHIPS: To improve gravel roads at minimal cost, local communities may spray them with a film of oily asphalt, spread small, chipped gravel on top, then run a roller over it, fusing the materials and forming a modest field coat, or chip seal, that reduces rutting and cracking.

POROUS ASPHALT, beginning to be used in parking lots, has small voids that allow rain to percolate through to the soil, which filters it naturally before it reaches the water table. For traditional pavement, water runs off through a drain into a detention pond for collection and filtering. Installing the porous mix costs more because of a deeper base needed for drainage but consumes less real estate. The voids allow water to expand if it freezes.



Topic suggested by reader Derek Vogt.
Send ideas to workingknowledge@sciam.com

Heavy-Metal Sweat

DOES AN INFRARED SAUNA REALLY DETOXIFY THE BODY? BY GARY STIX

POISON PERSPIRATION exudes from the pores when a user is exposed to heat from an infrared sauna, according to claims made for the technology.

I received a call several months ago from a publicist for a company promoting an infrared sauna—a machine that is supposed to heat the body, not the surrounding air, and so produce sweat more efficiently. The company, Sunlight Saunas, contends that users experience the "same healing energy that is released naturally by the sun." On its Web site, it claims, as do abundant other Internet-based sellers of infrared saunas, an amazing list of health benefits: pain relief, weight loss, detoxification, increased circulation, cholesterol removal and a boost for the "immuse [*sic*] system."

The sauna is supposed to emit radiation in the infrared part of the spectrum, which adjoins the microwave spectrum. I started to imagine my body undergoing a gentle, slow cook without sunburn and without any of the icky cold parts that always remain in the middle when I heat up a frozen taco in the microwave. I knew that I could use a little detox, too—and not because I enjoy a glass of wine with dinner. *Scientific American*'s offices on Madison Avenue are less than a block away from environmental monitoring equipment that measures the worst particulate levels in New York City (and some of the worst east of the Mississippi). The diesel fumes from the bus lanes there were an inspiration for an ad from the Natural Resources Defense Council that read "Standing behind this bus could be more dangerous than standing in front of it. . . ."

An x-ray of my lungs could probably illustrate a medical pathology textbook. (My home, by the way, is near the heavily trafficked George Washington Bridge.) What the heck did I have to lose but some excess sulfates and other toxins? I first asked the publicist whether I could go someplace in the New York metropolitan area to sit in a Sunlight Sauna hot box. She said not to worry, the sauna would come to me. A week or two later a box almost big enough to sleep in arrived.

I'm someone who has never contemplated inhabiting any structure that does not come equipped with a superintendent. The thought of standing in a Home Depot aisle trying to distinguish between a ground fault circuit interrupter and a hinged single-pole breaker lockout induces abject terror. But even I could manage the assembly of the Sunlight Sauna Solo.

JOHN FRASER

My spa was a space that stretched nearly underneath the desk in an unused windowless office. No Arizona desert vistas for me. Lying in the sauna, I stared up at a bank of fluorescent lights, my trunk and limbs covered by two cylindrical carbon shells. This personal sauna resembled a minimalist version of a computed tomography machine. I wondered how claustrophobes would do. The carbon shells emitted infrared radiation in the range of four to 50 microns. According to Sunlight Solo literature, the wavelengths coincide largely with those emitted by the body, heating it up at a lower temperature than can be achieved with a conventional sauna. Lying on an infrared heating pad allows for the warming sensation on all sides.

As advised by the publicist, I stayed in the sauna for 30 minutes. I gradually turned up the heat a notch every few minutes until it hit the highest setting of 150 degrees. By the end of the half-hour, my forehead glistened; I had drifted into momentary sleep a few times. When I stood up, I felt lightheaded but not dizzy. The relaxed feeling lasted the rest of the day. I tried the sauna three other times on different days. A colleague who saw me shoeless and in shorts and a *Scientific American* T-shirt inquired about how my “research” was going.

I asked myself: How does this sauna differ from all other saunas? I never came up with a good answer. How did it feel? Well, you could say that it was warm, calescent or maybe even thermogenic—in other words, *caldo, chaud, heiss,* or just hot, hot, hot.

I tried a sauna at the local YMCA to compare. I usually stay in for five minutes after going swimming at lunchtime but decided to stretch it out to half an hour to equal my time in the Solo. It still wasn’t a one-to-one match, because the unadjustable sauna at the Y was more than 180 degrees against 150 for the Solo. Unlike the Solo, I had to leave the Y’s sauna every 10 minutes or so to get a drink of water, and I didn't sweat quite as much. But I found that the same adjectives applied to what is essentially just another form of infrared energy. Any subtleties eluded me. I called Sunlight Sauna to ask what it is I should be feeling. Co-owner Connie Zack said that I might not notice any fundamental physical changes from taking a sauna unless I was actually experiencing some health problem, such as chronic back pain.

The company had sent me two papers from the *Journal of the American College of Cardiology* on small Japanese studies: one showing improvements in vascular function in patients with chronic heart failure, another in patients with risk factors for heart disease. But the most far-reaching assertions for this technology center on detoxification. Claims for the ability of infrared saunas to rid the body of heavy metals and the like populate the Internet like Viagra ads. A press release for Sunlight Saunas mentions Dietrich Klinghardt, a Seattle-area physician who asserts that infrared saunas, but not conventional ones, rid the body of "cholesterol, fat-soluble toxins, toxic heavy metals, sulfuric acid, sodium, ammonia and uric acid."

A trip to Klinghardt's Web site turns up a document that states that sauna therapy can leach toxins from the body. But Klinghardt notes that the poisons can also be displaced from "one body compartment to another." Mercury (beware those old dental fillings) might shift from connective tissue to the brain, according to Klinghardt. That is, unless the patient ingests sufficient quantities of cilantro, garlic and chlorella (green algae) in conjunction with taking saunas. Oooo-kay. Needless to say, I didn't follow up by looking for references to this area of research in the National Library of Medicine. Sunlight Saunas also provided me with testimonials about the Solo's benefits

ENCASED under two carbon shells, the author demonstrates an infrared sauna that provided a pleasant experience, but one not entirely different from sitting in a warm parked car.

from patients with Lyme disease, cancer and "general toxic overload."

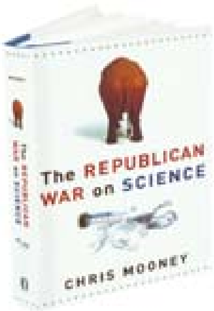
I wasn't ready quite yet to go hunting down chlorella. I wanted to see what the allopathic (nonalternative) world had to say about detoxification. I called Roger Clemens, director of an analytical laboratory at the University of Southern California that evaluates environmental toxins in the food supply. Clemens remarked that the most efficient system for detoxification is not an infrared sauna but rather the kidneys, liver, gastrointestinal tract and immune system. "Except when one of the major organs breaks down, there isn't a medical device or any diet that can accelerate the body's natural process of detoxification," he says.

Hearing this, I decided I would rather rely on the multimillion-year track record of detoxifying my body by just going to the men's room. Shorn of health claims, the sauna was pleasant enough. But a moment of revelation came when I climbed into a car heated by the July sun. Once again, I felt a warmth indistinguishable from what I had experienced at both the Y and when I was wedged into the Sunlight Sauna Solo under the desk at work. Given the almost $2,500 total price for the sauna and the heating pad together, I wagered that I might be able to get the same benefits by spending less time looking for a parking spot in the shade. With a son in college, I don't think I'm going to take the big plunge for a Solo anytime soon. SA

JOHNNY JOHNSON

Science Abuse

SUBVERTING SCIENTIFIC KNOWLEDGE FOR SHORT-TERM GAIN BY BOYCE RENSBERGER



THE REPUBLICAN WAR ON SCIENCE
by Chris Mooney
Basic Books, 2005 ($24.95)

Thomas Jefferson would be appalled. More than two centuries after he helped to shape a government based on the idea that reason and technological advancement would propel the new United States into a glorious future, the political party that now controls that government has largely turned its back on science.

Even as the country and the planet face both scientifically complex threats and remarkable technological opportunities, many Republican officeholders reject the most reliable sources of information and analysis available to guide the nation. As inconceivable as it would have been to Jefferson—and as dismaying as it is to growing legions of today's scientists—large swaths of the government in Washington are now in the hands of people who don't know what science is. More ominously, some of those in power may grasp how research works but nonetheless are willing to subvert science's knowledge and expert opinion for short-term political and economic gains.

That is the thesis of *The Republican War on Science,* by Chris Mooney, one of the few journalists in the country who specialize in the now dangerous intersection of science and politics. His book is a well-researched, closely argued and amply referenced indictment of the right wing's assault on science and scientists. Mooney's chronicle of what he calls "science abuse" begins in the 1970s with Richard Nixon and picks up steam with Ronald Reagan. But both pale in comparison to the current Bush administration, which in four years has:

- Rejected the scientific consensus on global warming and suppressed an EPA report supporting that consensus.
- Stacked numerous advisory committees with industry representatives and members of the religious Right.
- Begun deploying a missile defense system without evidence that it can work.
- Banned funding for embryonic stem cell research except on a claimed 60 cell lines already in existence, most of which turned out not to exist.
- Forced the National Cancer Institute to say that abortion may cause breast cancer, a claim refuted by good studies.
- Ordered the Centers for Disease Control and Prevention to remove information about condom use and efficacy from its Web site.

Mooney explores these and many other examples, including George W. Bush's support for creationism. In almost every instance, Republican leaders have branded the scientific mainstream as purveyors of "junk science" and dubbed an extremist viewpoint—always at the end of the spectrum favoring big business or the religious Right—"sound science." One of the most insidious achievements of the Right, Mooney shows, is the Data Quality Act of 2000—just two sentences, written by an industry lobbyist and quietly inserted into an appropriations bill. It directs the White House's Office of Management and Budget to ensure that all information put out by the federal government is reliable. The law seems sensible, except in practice. It is used mainly by industry and right-wing think tanks to block release of government reports unfavorable to their interests by claiming they do not contain "sound science."

For all its hostility to specific scien-

PRESIDENT GEORGE W. BUSH receives an honorary doctorate of science at Louisiana State University in 2004. Those on the political Right, according to Mooney's book, seldom oppose science directly. Instead they blur the understanding of how science works to give business interests and religious extremists equal footing with mainstream research.

WILLIAM PHILPOTT *Reuters/Corbis*

tific findings, the Right never says it opposes science. It understands the cachet in the word. Perhaps Republicans sense what pollsters have known for decades—that the American public is overwhelmingly positive about science and that there is nothing to be gained by opposing a winner. Instead the Right exploits a misconception about science common among nonscientists—a belief that uncertainty in findings indicates fatally flawed research. Because most cutting-edge science—including most research into currently controversial topics—is uncertain, it is dismissed as junk.

This naive understanding of science hands the Right a time-tested tactic. It does not claim that business interests or moral values trump the scientific consensus. Rather rightists argue that the consensus itself is flawed. Then they encourage a debate between the consensus and the extremist naysayers, giving the two apparently equal weight. Thus, Mooney argues, it seems reasonable to split the difference or simply to argue that there is too much uncertainty to, say, ban a suspect chemical or fund a controversial form of research.

The Republican War on Science details political and regulatory debates that can be arcane and complex, engrossing reading only for dedicated policy wonks. Thankfully, Mooney is both a wonk and a clear writer. He covered many of the battles in real time for publications such as the *Washington Post*, *Washington Monthly*, *Mother Jones* and *American Prospect*.

"When politicians use bad science to justify themselves rather than good science to make up their minds," Mooney writes, "we can safely assume that wrongheaded and even disastrous decisions lie ahead."

Thomas Jefferson would, indeed, be appalled. Writing in 1799 to a young student whom he was mentoring, the patriot advised the man to study science and urged him to reject the "doctrine which the present despots of the earth are inculcating," that there is nothing new to be learned. He concluded by saying opposition to "freedom and science would be such a monstrous phenomenon as I cannot place among possible things in this age and this country." SA

Boyce Rensberger directs the Knight Science Journalism Fellowships at the Massachusetts Institute of Technology and teaches in M.I.T.'s Graduate Program in Science Writing. For many years he was a science reporter and editor at the Washington Post.

THE EDITORS **RECOMMEND**

FACES OF SCIENCE

by Mariana Cook. Introduction by Gerard Piel. W. W. Norton and Company, 2005 ($39.95)

Cook, whose books of photographs include *Fathers and Daughters* and *Couples,* turns her camera on 77 scientists who have answered important questions about the nature of the physical world. Each remarkable image is paired with a brief autobiographical essay.

STATISTICIAN C. R. Rao: "According to statistics, the second born has a lower IQ than the first born, the third a lower IQ than the second, and so on. I am the eighth child!"

BIRDS OF CENTRAL PARK

by Cal Vornberger. Harry N. Abrams, 2005 ($35)

The surprise here is that New York City's Central Park is one of the country's top birding sites. The 843-acre park is a magnet for more than 200 species on their migratory routes. Author/photographer Vornberger has caught the birds in action and interaction; he includes information about how he got some of his spectacular shots and a small pocket guide.

GREAT BLUE HERON catches a crayfish in New York City's Central Park.

VISIONS OF MARS

by Olivier de Goursac. Harry N. Abrams, 2005 ($29.95, paperbound)

Space-imaging specialist de Goursac presents a selection of photographs of the Red Planet: windswept plains, fields of giant dunes, deep canyons, the highest mountains in the solar system, polar ice caps. Drawn from the best of the robotic exploration missions, almost all the images are previously unpublished, and all have been processed by the author to an unusually fine level of detail.

MARTIAN DICHOTOMY: In the north, the great plains and volcanoes; in the south, craters and a small polar cap.

Additions and Corrections

MAKING A CASE FOR A NEW PLANET AND TWO DEAD STARS **BY STEVE MIRSKY**

In news from the far reaches of the solar system, also known as California, astronomers at Caltech are pushing for a distant body to be recognized as the 10th planet orbiting the sun. The object, currently known by the catchy handle 2003UB313, is more than twice as far away as Pluto. Whether Pluto should even be considered a planet has been the subject of recent debate. According to Martin Amis in his 1995 novel *The Information:* "One must never mock the afflicted, of course, but Pluto really *is* an awful little piece of s**t."

Plutotudes aside, if 2003UB313 does pass planetary muster, it'll need a new name. Pluto's moon is dubbed Charon, for the ferryman who brings newcomers across the River Styx to the underworld, run by Pluto. The next-best name for a companion for Pluto is thus Goofy. Which would set the stage for NASA's first *official* Goofy mission.

Speaking of big government, on August 1 President George W. Bush endorsed the teaching of so-called intelligent design as an alternative to evolution. Intelligent design, as noted in this space in February, is "the full-blown intellectual surrender strategy" that proposes that a *scientific* explanation of life's complexity requires the intercession of a supernatural being. But examination of the available data leads only to the conclusion that the biggest beneficiary of the Bush presidency is Warren Harding.

And now for something completely deferent. Twice in this space (August 2003 and June 2005), I noted that British researchers hypothesized that Newton and Einstein might have had Asperger's syndrome, a form of autism. The more recent claim prompted a letter from David Green of the New School in New York City. Here are highlights from that letter, because the whole letter would fill this entire page, and management might wonder why they're paying me and not David Green.

Green began, "You run a fine column." (The phrase "into the ground" seemed to hang in the air.) He then pointed out my reference to the two great physicists possibly having Asperger's and wrote, "You should let this particular piece of foolishness lie. As someone who teaches a course on revolutions in science, I have read umpteen biographies of these men." (I looked up "umpteen," which was defined as the number of incorrect ball and strike calls made in a Major League Baseball game.) Green continued, "The sensible conclusion: Newton could (at least arguably) be afflicted with Asperger's syndrome. Einstein could not."

Green then compared the two physicists. "Newton was emotionally frigid, actively discouraged human contact, was known to laugh only once in his life (when a fellow student asked what use Euclid could be), and died bragging that he was a virgin and thus uncontaminated. Therefore, I'll give the good doctors Newton if they feel they need him, but even Newton is a debatable case. When Einstein died, Bohr eulogized him by saying that the loss to the world was a great physicist, but the loss to those who knew him was his unique warmth and kindness. He loved wine, women and song (perhaps too much of the middle one), had close and deep friendships all his life, and was funny and a fully social being."

In retrospect, diagnosing the mental condition of the dead based on decades- or centuries-old anecdotal evidence does seem somewhat questionable. It would be like a cardiologist practicing teleneurology by watching a few minutes of an old videotape and declaring that the patient shown is not in a persistent vegetative state. One might call such medical practice Mickey Mouse.

Back to geniuses. "The chances that Einstein was a case of even mild autism," Green concluded, "are about as great as the chances that you are or I am Queen Elizabeth."

We are amused. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

What causes shin splints?

—E. Bachman, Austin, Tex.

Claude T. Moorman III, director of sports medicine at Duke University Medical Center, offers an answer:

"Shin splints," the layman's term for the painful sensations felt at the front of the shinbone (tibia) after exercise, occur when the constant pounding and stresses placed on the muscles, bones and joints overwhelm the body's natural ability to repair damage and restore itself. We commonly see shin splints in athletes, military recruits and even in middle-aged weekend warriors, especially at the beginning of milder weather.

Overworked muscles are one major source of the aches. The muscles that connect the tibia to the ankle are held together by fascia—a tough, inelastic covering like a sausage skin. When the muscles naturally expand as a result of exercise, the resulting pressure can cut off blood flow, causing pain. This form of shin splints, known as exertional compartment syndrome, appears in athletes who play field sports such as soccer or who often run on hard surfaces.

Pain can also stem from injuries to the bone, ranging from stress reactions to full-blown fractures. The continual pounding endured during running, for example, can cause many microscopic cracks to develop in leg bones. Normally, with rest, the body easily fixes the tiny fissures. Without sufficient mending time, however, they can coalesce into a stress fracture—a hairline crack—or even a complete fracture, in which the bone breaks all the way through.

People can prevent shin splints by simply adding extra arch support to shoes to redistribute weight or changing to softer running surfaces. Doctors also recommend "active rest," which means that a runner, for instance, should take up swimming or biking for a while. The change of pace gives the affected areas time to heal but maintains the cardiovascular benefits of exercise.

Warming up muscles before exercise to prevent injuries is a controversial subject, with smart people on both sides arguing for and against it. We at Duke, based on research conducted at the university, recommend a slow warm-up period. We believe that about 10 minutes of graduated activity is the best way to prepare the body for working out more strenuously. For shin splints, as with most things in life, moderation appears to be the best medicine.

Why do bees buzz?

—M. O'Malley, Newton, Mass.

Gard W. Otis, a professor of environmental biology at the University of Guelph in Ontario who studies bee behavior, ecology and evolution, explains:

Bees produce their distinctive "zzzz" in two ways. First, their wing beats create wind pulses that people hear as a buzz. This sound is not exclusive to bees—most flying insects produce a similar hum. The pitch of the sound produced is a function of the flapping rate: the faster the wings, the higher the pitch.

Second, some bees, most commonly bumblebees (genus *Bombus*), vibrate their wing muscles and thorax (the middle segment of their body) while visiting flowers. These movements make the pollen fall off the flower's anthers onto the insect's body. Some of that pollen gets deposited when the bee alights on the next bloom, resulting in pollination. The bee also grooms pollen onto basketlike structures on its hind legs, taking it back to the nest to feed to the larvae.

When bumblebees vibrate blossoms to release pollen, the noise is quite loud. Honeybees (genus *Apis*) are incapable of such "buzz pollination" and are usually quiet when foraging. Some plants are adapted to buzz pollination: Tomatoes, green peppers and blueberries all store pollen inside tubular anthers. When the bee shakes the flower, the pollen falls out. Consequently, bumblebees pollinate these crops much more efficiently than honeybees do. SA

For a complete text of these and other answers from scientists in diverse fields, visit www.sciam.com/askexpert

MATT COLLINS