THE EARLIEST COMPLEX ANIMALS • SAVING LIVES FROM ANEURYSMS

SCIENTIFIC AMERICAN

NANOBODIES: Tinier Molecules for Better Immunity

AUGUST 2005
WWW.SCIAM.COM

Is the Universe Out of Tune?

BAD HARMONIES IN THE MICROWAVE MUSIC DEFY THEORY

Growing Test-Tube **Teeth**

Chameleon Circuits and **Morphware**

How the Mind Grasps **Symbols**

august 2005

contents

features

SCIENTIFIC AMERICAN Volume 293 Number 2

SCIENTIFIC AMERICAN Volume 293 Number 2

departments


16


96


Mark Kryder, Seagate Technology 32

columns

Cover image by Jean-Francois Podevin; photograph at left: Theo Anderson

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.

Great Teller's Ghost

Take the high ground! The American soldier has been admonished to do so countless times during the nation's hard-fought military history. With strong input from the U.S. Air Force, the Bush administration is reportedly pushing that dictate to the final frontier. The military use of space has proceeded for decades with the deployment of spy, communications and navigation satellites. The new administration directive would move the U.S. from space militarization to "weaponization," opening the way to both offensive and defensive arms in the firmament. The policy makes no sense on technical, geopolitical or economic grounds.

SPACED-BASED LASER obliterates a communications satellite under one scenario for placing weapons in orbit.

A 2001 commission led by incoming Secretary of Defense Donald H. Rumsfeld issued a report that contended that the U.S. should maintain superiority in space by protecting its own satellites, by denying the use of space to adversaries, and by retaining the ability to attack targets from or within space. The directive would codify in policy at least some of those recommendations, which have already served to guide the air force in its planning efforts.

For the past 10 years, the military has disbursed billions in R&D on space weaponry—and more of the same is on the drawing board. The wish list would make Edward Teller's ghost proud. Take two examples: one, a space-based laser, would focus a laser's energy on ground- or space-based enemy targets. "Rods from God"—hypervelocity rods made of tungsten or uranium—would be hurled from an orbiting platform to penetrate underground or hardened targets. Unfortunately, the feasibility of these technologies has not been demonstrated in any conclusive way.

Installation of such weapons could initiate a needless extraterrestrial arms race and leave us less secure. Russia has already indicated that it will retaliate if any country places weapons in space. Sophisticated death-star lasers are vulnerable to low-tech counterresponses, such as $6 smoke grenades that can create an impenetrable haze. Protecting vital communications and surveillance assets can be accomplished more effectively by modest measures, such as having unmanned aerial vehicles at the ready that can supply the communications or espionage capabilities to replace a battle-damaged satellite.

From its current position of strength, the best strategic course for the U.S. is to initiate a process that would affirm the use of space as a global commons where all guns are checked at the border of the mesosphere, 50 miles up. A good first step would be to pledge to abide by that provision of the Anti-Ballistic Missile Treaty, which the administration withdrew from in 2002, that forswears space-based missile defenses. Otherwise, perhaps the greatest hope against realization of yet another episode of Star Wars is the astronomical expenditure—in the hundreds of billions of dollars—that the federal government would confront, at a time of deficits and underfunding of other security needs such as protecting our seaports. We can only hope that the overburdening financial requirements will relegate rods from God to the realm of think tanks and military colleges for a long time to come.

THE EDITORS editors@sciam.com

GEORGE RETSECK

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

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:

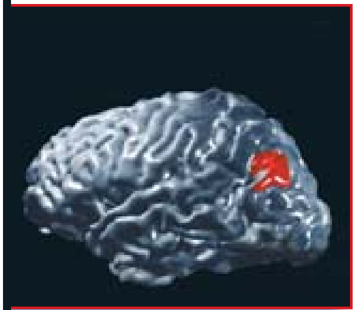
Cosmic "Magnifying Glass" Reveals Distant Planet

Astronomers have discovered an extrasolar planet using a technique known as gravitational microlensing. This is the second such find by the method, which measures an object's effect on light from a distant star. The new planet weighs about three times as much as Jupiter and orbits a star similar to our sun.

Hormone Spray Elicits Trust in Humans

It sounds like the stuff of science fiction, but researchers have determined that a whiff of a certain hormone makes people more willing to trust others with their money.

Brain Region Linked to Metaphor Comprehension



Metaphors make for colorful sayings but can be confusing when taken literally. A study of individuals who are unable to make sense of figures of speech has helped scientists identify a brain region they believe plays a key role in grasping metaphors.

Ask the Experts

Are food cravings the body's way of telling us that we are lacking certain nutrients?

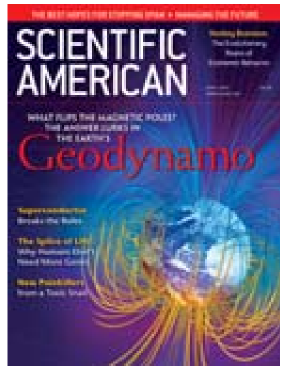
Peter Pressman of Cedars-Sinai Medical Center in Beverly Hills, Calif., and **Roger Clemens** of the University of Southern California School of Pharmacy explain.

SPECIAL OFFER

DAVID A. AGUILAR *Harvard-Smithsonian Center for Astrophysics* (top); CENTER FOR BRAIN AND COGNITION, U.C.S.D. (bottom)

Letters

EDITORS@SCIAM.COM



IN APRIL the editors took delicious advantage of the day that celebrates a uniquely human capacity—humor. In the spirit of the adage "Many truths are spoken in jest," they modestly proposed in SA Perspectives to be fair and balanced—to henceforth cover *all* explanations of natural phenomena. The voluminous and emotional responses ranged from kudos to condemnation. Some admitted they fell for the mock-serious announcement before reading the last sentence that stated this new policy would start on April Fools' Day. Others may not have finished reading before they dashed off frantic pleas for us not to buckle under pressure from gathering antiscience forces. And a few congratulated us for finally seeing the light.

Two e-mails exemplified the range of responses: Susan Scroggins wrote, "If your goal was to rid yourselves of those pesky creationist subscribers, you have succeeded—at least with this one." From Susan Suver came admiration and an extra measure of encouragement—"I am so hot for all of you. If it's wrong to want to marry you en masse and have your smart, snarky little babies, I don't want to be right. Thanks for saying what so desperately needed to be said."

THE EVER WAFFLING EARTH

According to "Probing the Geodynamo," by Gary A. Glatzmaier and Peter Olson, the earth's polarity reverses itself periodically, thereby switching magnetic north and south. This transition could happen over thousands of years. The earth's magnetic field, however, is the one thing that protects our planet from the devastating solar wind. Is there any evidence that suggests that during this polarity reversal the planet becomes vulnerable to the solar wind, and if so, what would be the effects?

Brian Magnani
Manassas, Va.

The authors state that "the core's high temperatures are the result of heat that was trapped at the center of the earth during its formation." The earth formed several billion years ago, so this heat should have diffused as the earth cooled. I always thought that the high temperature of the core was maintained by the heat continuously released by the decomposition of radioactive isotopes.

Bernard Riedl
Laval University, Quebec

GLATZMAIER AND OLSON REPLY: Magnani is correct about the earth's geomagnetic field; it does deflect some of the dangerous radiation in the solar wind to the earth's magnetic poles. More of this radiation presumably reaches the planet's surface (at low latitudes) during a field reversal, because the field intensity typically becomes quite small and the magnetic poles sit near the equator. Greater exposure to cosmic radiation could hypothetically alter the evolution of species, but we are not aware of any evidence for such a correlation.

As for the earth's internal heat referred to by Riedl, radioactive decay in the mantle (the layer of rock that surrounds the core) has slowed cooling in the core but only indirectly. Some scientists have proposed that the core itself contains radioactive heat sources, the most plausible being radioactive potassium. Its actual concentration in the core is basically unknown, however, and two recent computer simulations of the geodynamo—one with no radioactive heating in the core and the other with the best estimate of the amount of potassium that could be there—show very little difference in the structure of the resulting magnetic field.

SURELY WE JEST

Please accept my exultant thanks for the "Okay, We Give Up" April Fools' SA Perspectives. It took guts to level some humorously couched blows at the creationists and others who don't understand the difference between science and faith. I hope to hear more informed citizens espousing science as the basic driver of policy, rather than religion or money. Thanks again for the chuckles and the reminder that we are at a pivotal time for reminding (or teaching) people that science is not a dirty word.

Christina Alba
Fort Collins, Colo.

Fancy fossils, eh? I'd like to see some. And radiocarbon dating? Sure, if it weren't just based on ridiculous assumptions. Peer-reviewed articles? If all the "peers" believe the same thing, why would they go against their own beliefs? That's some fantastic logic. Even though all you do is insult and ridicule without giving any hard evidence, I am still convinced. I am convinced that despite the total lack of indisputable evidence, evolution must be the answer. How do I know? Because the scientists told me!

Ryan D. Thomas III
Liberty University

Although, as a result of your editorial, I choked on my lunch from laughing, it totally made my day.

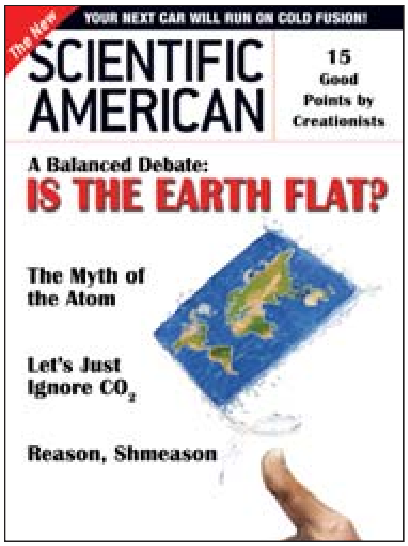
The world's gotten scary. After a threat from consumers to boycott IMAX theaters, several branches decided not to show a video on undersea life because it mentioned evolution. A law is being considered in Florida that would allow students to sue their professors if they don't teach "serious academic theories" (read: intelligent design/creationism) in addition to those theories that they personally agree with (read: evolution).

Amanda L. Smith
via e-mail

Mocking those not sharing your confidence in science's "proven" findings is not only uncalled for but unseemly.

To summarily denounce all those who honestly, diligently toil toward understanding our universe with a flip, "you-flat-earther" dismissal is hubris in its most egregious form. That you may fervently agree that all evidence points to catastrophic global warming or that an antimissile missile defense simply won't work is okay. But a cavalier April Fools' rejection of those who don't agree with you is an embarrassment to the scientific community. Shame on you!

Douglas Ball
Mississauga, Ontario



SCIAM SCAM cover: Not on sale now!!!

CANNING SPAM

I enjoyed "Stopping Spam," by Joshua Goodman, David Heckerman and Robert Rounthwaite. One topic, however—the technical means by which e-mail is authenticated—has an inaccuracy that is worth noting.

The article states: "A new industry standard, the Sender ID Framework, is finally addressing the spoofing problem." Sender ID—used by Microsoft—is not actually an industry "standard"; in fact, the Internet Engineering Task Force, a technical standards-setting organization, disbanded the Working Group that was attempting to standardize this style of authentication.

In addition, Sender ID is just one of the proposed authentication systems under consideration for standardization. The Federal Trade Commission hosted an e-mail authentication conference in September 2004 in which Sender ID, SPF and DomainKeys (a cryptographic signature-based system from Yahoo!) all received wide coverage.

Many in the field are taking a close look at all of the authentication systems, generally recognizing that Internet protocol based systems and cryptographic systems can work together. The article's incorrect reference to the singularity of Sender ID should be corrected so that readers will not be misinformed. Thanks for your consideration.

Mark Delany
Systems architect, Yahoo! Mail
Sunnyvale, Calif.

GOODMAN REPLIES: Delany is correct that Sender ID is not an official standard. Nevertheless, it is a de facto one. Sender ID is a superset of the Sender Policy Framework (SPF), a developing industry practice that publishes the identities of machines sending mail from each domain, thus allowing recipients to ensure that mail comes from the identified source. SPF records are compatible with Sender ID. More than a dozen companies have released or are planning to release products supporting Sender ID or SPF. Independent estimates show that at least 800,000 domains have posted Sender ID/SPF records, whereas fewer than 200 domains currently send mail with DomainKeys.

We do believe that cryptographic signatures will play an important role in e-mail, especially when absolute security is needed, such as for financial transactions. But cryptographic solutions are also more difficult to deploy than IP-based methods. Given the urgency of making an immediate dent in spoofing and phishing, we chose to emphasize the technology that has the most immediate potential to help. Longer term, we agree with Delany that cryptographic approaches are a promising, complementary approach.

ERRATUM In "The Feynman-Tufte Principle," by Michael Shermer [Skeptic], it was incorrectly stated that in the *Challenger* disaster a solid rocket booster (SRB) made by Morton Thiokol blew up. It was the shuttle's external liquid-fuel tank that exploded because of leakage of hot gases through a faulty O-ring seal on one of the adjoining SRBs.

MATT COLLINS

150, 100 & 150 Years Ago

FROM SCIENTIFIC AMERICAN

Prescription Faith ▪ Flying Rebuttal ▪ Medical Fishing

AUGUST 1955

PLACEBO MEDICINE—"Society has never completely divorced the physician from the witch doctor, and for many patients the notion that the doctor lacks omniscience in his domain is extremely disturbing. When the patient expects or demands some tangible evidence of therapeutic capability, it is usually simplest (and wisest) for all concerned to prescribe a harmless pill or liquid. The well-known illegibility of the prescription scripts frequently makes it impossible for a curious patient even to guess at the nature of the medicament. To guard against the possibility that the patient will ask the pharmacist about the contents of the prescription, physicians should steer clear of milk sugar or other well-known ingredients. Names such as ammoniated tincture of valerian can safely be revealed without upsetting the psychological applecart."

AUGUST 1905

RISE OF GALVESTON—"The completion of Galveston's great sea wall marks the successful culmination of one of the most unique and gigantic engineering undertakings in recent years—a solid concrete wall 4½ miles in length and seventeen feet high and the elevation of the city's grade to the level of the top of the wall. The wonderful nerve of Galvestonians is more properly realized after considering that in the great storm of September 8, 1900, over 8,000 lives were wiped out of existence, and more than $20,000,000 worth of property was destroyed, and faith in the rapidly growing city so rudely shaken that five years have not entirely sufficed to restore public confidence. The grade-raising necessitates the lifting up of 2,156 houses."

SCOPOLAMINE—"A vegetable alkaloid may be employed as an anaesthetic, instead of chloroform or any of the ethers. The idea is due to Dr. Schneiderlin, of Baden. What led this surgeon to experiment with this substance? I do not know, for the power to produce general anaesthesia could not be inferred, *à priori,* from its known properties. Scopolamine, the alkaloid recommended, was extracted from the *Scopolia japonica,* a perennial herbaceous plant, popularly known as Japanese belladonna."

THE LANGLEY AEROPLANE—"The experiments undertaken by the Smithsonian Institution upon an aerodrome, or flying machine, capable of carrying a man, have been suspended from lack of funds to repair defects in the launching apparatus. Entirely erroneous impressions have been given by the account of these experiments in the public press, from which they have been judged, even by experts; the impression being that the machine could not sustain itself in flight. It seems proper to emphasize and to reiterate, that the machine has never had a chance to fly at all, but that the failure occurred on its launching ways; and the question of its ability to fly is consequently, as yet, an untried one. —S. P. Langley"

AUGUST 1855

PARASITE TRAP—"The accompanying engravings are views of a patented trap and process for removing tape worms from the human stomach and intestines. The figure shows the hand of the operator fishing for the tape worm. A and B represent two sizes of traps. The patient is put upon a strict course of fasting for several days. The trap, baited with a nutritious food—such as a piece of cheese—is then swallowed. The tape worm, by a little wriggling, displaces a spring under the fine prongs of the bait fork, and is thus transfixed and caught. The worm is abstracted, and great care must be exercised in drawing it up, so as not to break it."

SCIENTIFIC AMERICAN

TAPE-WORM TRAP, baited and ready for use, 1855

GEOPHYSICS

Rebuilding a Volcano

MOUNT ST. HELENS MAY BE PATCHING IN ITS 1980 CRATER **BY KRISTA WEST**

On October 2, 2004, Mount St. Helens let out a noontime burp of steam; nearly an hour's worth of low-frequency tremors soon followed. Magma, it seemed, might be moving underground. Fearing an imminent threat to human life, scientists at the Cascades Volcano Observatory in southwest Washington State issued a volcanic alert level 3, aviation code red. The next day the Federal Aviation Administration restricted air traffic within five and a half miles of the summit, diverting and delaying flights. The alerts seemed justified a couple of days later, when eruptions produced a light dusting of ash as far as 70 miles northeast of the volcano.

But nothing close to Mount St. Helens's catastrophic eruption in 1980 ever happened. No one was killed or injured. No trees toppled. And air traffic soon returned to normal. On October 6, volcanologists downgraded the alert.

Given the excitement and media attention, the average citizen undoubtedly expected a big, explosive blowout. Instead Mount St. Helens erupted slowly and quietly. Although explosive activity has largely subsided, the volcano continues to erupt today—with periodic hiccups of steam, ash and tremors. One explosion in March billowed ash nearly 11 kilometers above sea level.

In retrospect, scientists say the increased activity in late September and early October was not caused by the volcano preparing to blow its top but rather by liquid magma steadily burrowing its way to the surface. The volcano is heaping up a new lava dome in its crater, perhaps even rebuilding itself to a pre-1980 shape.

Mount St. Helens is like "a giant tube of toothpaste, squeezing out lava," says John Eichelberger, a volcanologist at the University of Alaska–Fairbanks. With the current

JOHN PALLISTER/USGS/CASCADES VOLCANO OBSERVATORY

BLOWING SMOKE: Mount St. Helens may not have explosively erupted, but the volcano is slowly spewing lava that may ultimately fill in its crater, which formed after the blowout of May 18, 1980. This aerial view, taken from the west, was photographed on June 8, 2005.

GROWING A MOUNTAIN

Since the renewed activity began in October 2004, Mount St. Helens has been reshaping itself with oozing lava. Volcano data taken during spring 2005:

Height added to lava dome: **150 meters**

Total height of dome: **2,300 meters**

Total volume of dome: **45 million cubic meters**

Lava extrusion per second: **2 cubic meters**

Fatalities from 1980 eruption: **57 humans**, **7,000 big-game animals**, **12 million hatchery salmon**

Human fatalities from 2004–2005 eruption: **0**

Volcano Cam: **www.fs.fed.us/gpnf/volcanocams/msh/**

conformation of the mountain, there is little space for magma to collect. So once the magma established a path to the surface, the volcanic activity decreased and the "toothpaste" began to ooze out. As the lava continues extruding, it is forming a long, narrow, spinelike dome in the volcano's crater.

Known to scientists as the whaleback because of its distinct appearance, the new lava dome is 1,400 feet above the crater floor with a volume of 45 million cubic meters (as of the last measurement in March), enough to fill about 200 oil supertankers. Erupting lava is currently adding to the dome's volume at a rate of two cubic meters per second—slow enough for gases to escape and rocks to settle without exploding.

William E. Scott, a volcanologist at the Cascades observatory, says that the exiting lava could continue for years but that "a repeat of May 18 [1980] events is essentially impossible because the volcano is radically different." He adds: "The concern is about the type of eruption, not whether or not there is an eruption."

In 1980 the catastrophic eruption changed the shape of the mountain quickly. A magnitude 5.1 earthquake preceded the largest landslide in recorded history when the entire northern side of the mountain slid away. This collapse depressurized a giant, building bubble of magma that exploded up and out, killing everything in its path. Today the eruption is seeping instead of exploding, but it could change the contours of the mountain over time, as the lava slowly squeezes out to form the new dome. "The mountain is just rebuilding itself," Eichelberger says.

Eichelberger has studied Bezymianny volcano in Kamchatka, Russia, and says that Mount St. Helens is acting the same way. Bezymianny erupted in 1956, blowing out its side and leaving a Helens-like crater. Fifty years of lava-dome growth and intermittent eruptions have mostly filled Bezymianny's crater and reconstructed the mountain to a pre-1956 configuration.

Scientists are not sure yet if Mount St. Helens is rebuilding itself Bezymianny-style. "Dome-growth episodes come in all sizes and durations," Scott says. "It's impossible to tell if the current eruption will continue." Right now Mount St. Helens has no place to store a 1980-like bubble of magma—that is, unless some mountain reconstruction occurs, enabling the volcano to make room for one. That will be the time to start worrying.

Krista West watched the 1980 eruption of Mount St. Helens from her backyard swing set, safely out of the blast zone.

ORIGINS

Footprints to Fill

FLAT FEET AND DOUBTS ABOUT MAKERS OF THE LAETOLI TRACKS BY KATE WONG

FANCY FOOTWORK: *Australopithecus afarensis* was thought to have made the 3.6-million-year-old Laetoli footprints. But its foot may have been too flat to have permitted the striding gait evident in the trackway.

KENNETH GARRETT *National Geographic Image Collection*

It is one of the most evocative traces of humanity's ancestors ever found, a trail of footprints pressed into new fallen volcanic ash some 3.6 million years ago in what is now Laetoli, Tanzania. Discovered in 1978 by a team headed by Mary Leakey, the Laetoli footprints led to the stunning revelation that humans walked upright well before they made stone tools or evolved large brains. They also engendered controversy: scientists have debated everything from how many individuals made the prints to how best to protect them for posterity. Experts have generally come to agree, however, that the tracks probably belong to members of the species *Australopithecus afarensis*, the hominid most famously represented by the Lucy fossil. Now new research is calling even that conclusion into question.

The case for *A. afarensis* as the Laetoli trailblazer hinges on the fact that fossils of the species are known from the site and that the only available reconstruction of what this hominid's foot looked like is compatible with the morphology evident in the footprints. But in a presentation given at the American Association of Physical Anthropologists meeting in April, William E. H. Harcourt-Smith of the American Museum of Natural History and Charles E. Hilton of Western Michigan University took issue with the latter assertion.

The prints show that whoever made them had a humanlike foot arch, and the reconstructed *A. afarensis* foot exhibits just such an arch. So far, so good. The problem, Harcourt-Smith and Hilton say, is that the reconstruction is actually based on a patchwork of bones from 3.2-million-year-old *afarensis* and 1.8-million-year-old *Homo habilis*. And one of the bones used to determine whether the foot was in fact arched—the so-called navicular—is from *H. habilis*, not *A. afarensis*.

To get a toehold on the Laetoli problem, the researchers first compared the gaits of modern humans walking on sand with two sets of the fossil tracks. This analysis confirmed that the ancient footprints were left by individuals who had a striding bipedal gait very much like that of people today. The team then scrutinized naviculars of *A. afarensis, H. habilis,* chimpanzees and gorillas. The dimensions of the *H. habilis* navicular fell within the modern human range. In contrast, the *A. afarensis* bone resembled that of the flat-footed apes, making it improbable that its foot had an arch like our own. As such, the researchers report, *A. afarensis* almost certainly did not walk like us or, by extension, like the hominids at Laetoli.

But according to bipedalism expert C. Owen Lovejoy of Kent State University, other features of the australopithecine foot, such as a big toe that lines up with, rather than opposes, the other toes, indicate that it did have an arch. Even if it did not, Lovejoy contends, that would not mean *A. afarensis* was incapable of humanlike walking. "Lots of modern humans are flat-footed," he observes. "They are more prone to injury, because they lack the energy-absorptive capacities of the arch, but they walk in a perfectly normal way."

For their part, Harcourt-Smith and Hilton note that a new reconstruction of the *A. afarensis* foot built exclusively from *A. afarensis* remains is needed to confirm these preliminary findings. As for identifying the real culprit, if *A. afarensis* did not make the prints, that would put the poorly known *A. anamensis* in the running. But just as likely, speculates Harcourt-Smith, an as yet undiscovered species left the prints. That is to say, consider the world's oldest whodunit an unsolved mystery.

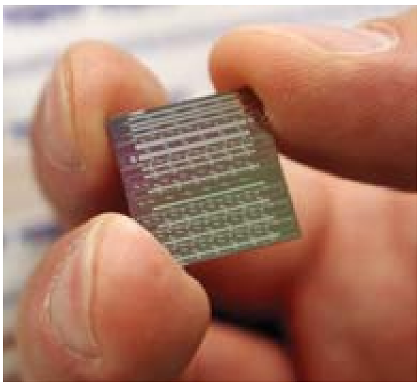
news SCAN

OPTOELECTRONICS

Making Light of Silicon

FINALLY, LASERS FROM THE LOW-COST SEMICONDUCTOR BY STEVEN ASHLEY

Scientists have long sought to build lasers from silicon. Such an advance would enable engineers to incorporate both electronic and optical devices onto cheap silicon chips rather than being compelled to employ costly-to-make lasers based on "exotic" semiconductor materials such as gallium arsenide or indium phosphide. Silicon lasers could lead to affordable light-based systems that harness photons instead of electrons to shuttle huge amounts of data swiftly—at multigigabit-per-second rates. Two research groups, one at the University of California at Los Angeles and the other at Intel Corporation, have recently reported success in making silicon emit continuous laser light.



LIGHT WORK: Eight lasers constructed from inexpensive silicon reside in this chip made by Intel. The coherent light beams could lead to ultrafast computer circuitry that transmits data optically.

This much anticipated feat came despite silicon's dogged resistance to serving as a lasing medium. In a good lasing material, electrons that are pumped up with energy release that energy in the form of coherent photons of light. In silicon, however, excited electrons are more likely to vibrate, thus generating heat instead. "There have been many attempts, but no one had been able to get silicon to lase before now," notes Bahram Jalali, the physicist who led the U.C.L.A. team.

Jalali and his group solved the problem last fall by making clever use of some of the very vibrations that undermined silicon's suitability for lasers in the first place. In particular, they focused on the Raman effect, a process in which the wavelength of light lengthens after it scatters off atomic vibrations. The U.C.L.A. researchers matched the scattered light with the pump energy from another laser in a way that created constructive feedback, resulting in a net amplification of light.

Intel reported its own success in creating a silicon Raman laser several months afterward. The chipmaker's scientists fed light from a separate laser into a waveguide (or light pipe)—basically an S-shaped ridge the engineers sculpted onto a 15-millimeter-square silicon chip—and Raman laser light emerged. Naturally, the task was not that easy. The power of a silicon Raman laser typically hits a limit as photons sporadically collide with silicon atoms and release free electrons. "Unfortunately, the free electron cloud absorbs and scatters light, so you get diminishing returns as you pump the device harder," explains Mario Paniccia, director of Intel's photonics technology laboratory. The team therefore positioned two electrodes on either side of the waveguide, forming a kind of diode. "Placing a voltage across the diode sucks the free electrons away like a vacuum cleaner," he says, and thus keeps the light flowing through the chip.

"This and related research should lead to many useful applications," says Philippe M. Fauchet, an electrical and computer engineer at the University of Rochester. A laser beam generated continuously through silicon could overcome cost and size limitations in lasers that could be used in surgical procedures, for example. The technology could also detect tiny amounts of chemicals in the environment, jam the sensors of heat-seeking missiles or enable high-bandwidth (high-capacity) optical communications.

Looking a bit farther afield, Paniccia believes that the new laser technology could serve as a building block for high-bandwidth photonic devices constructed almost entirely of inexpensive silicon in existing semiconductor foundry and micromachining facilities. "We've already developed the other necessary components of such a system," including fast modulators (optical encoders), light guides and photodetectors, he notes.

Of course, many in the industry hope that this technology will eventually lead to fully optical computers—superspeedy digital systems in which photons rather than electrons serve as 0s and 1s. Paniccia is certainly optimistic about the recent progress: "This work constitutes not only a scientific breakthrough but also a psychological one, because nobody thought it could be done."

EXTENDING LASER'S REACH

Lasers built from silicon represent an important advance, but "the silicon Raman laser is not a replacement for diode lasers," such as those used in DVD players and telecommunications equipment, cautions Bahram Jalali, a physicist at the University of California at Los Angeles. The Raman lasers require pumping with lasers; in contrast, diode lasers are pumped with electricity. Rather the silicon version extends the operating range of lasers to the midinfrared region of the spectrum (from 1.8 to 5 microns wavelength), Jalali points out, "where they cannot today operate."

PAUL SAKUMA AP Photo

MEDICINE

Widening the Window

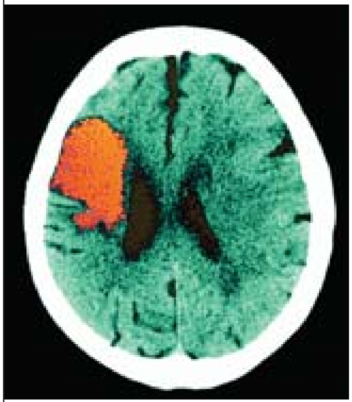
STRATEGIES TO BUY TIME IN TREATING ISCHEMIC STROKE BY CATHRYN M. DELUDE

Nine years ago the Food and Drug Administration approved tissue plasminogen activator (tPA) as the first, and still only, drug for treating ischemic strokes, which are caused by blood clots in the brain that starve neurons of oxygen. Yet only 3 percent of stroke victims receive this clot-busting thrombolytic, largely because they enter the emergency room within three hours of the onset of symptoms. After that, tPA's effectiveness in reducing death and disability sinks, while the relative risk of dangerous hemorrhaging rises. Recently scientists have discovered ways that could extend tPA's window of time, at least for some patients, and have found alternatives that may be both effective and safe beyond three hours.

A key to a bigger tPA window was the realization among researchers that not all neurons deprived of oxygen died after three hours, as was previously assumed. Restoring blood flow can revive enough neurons to significantly improve recovery. The trick is figuring out which patients can still benefit from treatment.

From the beginning, doctors used CT scans to triage patients, separating the many with ischemic stroke, who are candidates for tPA, from the few with hemorrhaging stroke, who are not. (About 80 percent of all strokes are ischemic.) But the images could not show how much of the ischemic tissue was already dead and how much was still salvageable. "We were treating patients blindly," remarks Steven Warach of the National Institutes of Health's Stroke Center. "We didn't know what was going on in the brain."

Today's MRI scans can distinguish between dead and dying cells, and newer CT scans seem able to as well. "It's been a big advance," says Warach, who described MRI's diagnostic effectiveness in the March



ZEPHYR/PHOTO RESEARCHERS, INC.

BLOCKED BLOOD: Brain images, such as this one made by a CT scan, can reveal tissue killed by an ischemic stroke (*orange*), thereby letting physicians know if there's still time to treat a patient.

news
SCAN

2005 *American Journal of Neuroradiology.* "We can push the time window for tPA to eight hours." Selecting patients for treatment on the basis of a "tissue clock" rather than the "ticking clock" gives many more patients a chance for a fuller recovery. As important, MRI and CT imaging may also identify those at risk for bleeding if given thrombolytics, a concern that keeps some physicians from administering tPA even within three hours.

Simple, nondrug measures may keep endangered neurons alive until they can be rescued with tPA, says Aneesh B. Singhal of Massachusetts General Hospital, where a small pilot study gave participants high-flow oxygen through a face mask. "It buys time," explains Singhal, who co-authored the paper on it in the April 2005 *Stroke.* "We can delay the progression of ischemic stroke by several hours. Because oxygen therapy is readily available in ambulances and the ER, it could make logical combination therapy" with tPA.

Meanwhile potentially safer drugs have entered late-stage clinical testing. Desmoteplase, which derives from the saliva of a vampire bat, chews up the fibrin holding a clot in place just as tPA does, but it is more potent and selective. "Even at nine hours, patients had significant long-term clinical benefits, in terms of better recovery after 90 days," says Warach of phase II results. The drug recently entered phase III trials.

Another drug in phase III testing protects neurons by way of a different method. Cerovive (NXY-059) works by scavenging up free radicals that break down the blood-brain barrier and worsen stroke outcome. Preliminary results from a global trial, known as SAINT I, suggest the chemical reduces the amount of disability after a stroke.

Even with the good news, many patients will not qualify for these therapies, because they may still arrive too late or have contraindications. A therapy that encourages brain cells to step up their own repair mechanisms might be the best solution, but that is still a long way off.

Cathryn M. Delude is a freelance writer based in Andover, Mass.

NEED TO KNOW: TOXIC tPA

During an ischemic attack, the lack of oxygen endangers neurons, but the treatment itself—the drug tPA—is toxic to stroke-stressed blood vessels and neurons, according to Berislav Zlokovic of the University of Rochester Medical Center. "We show that tPA breaches the blood-brain barrier and kills neurons directly," Zlokovic says of his team's report last year. The group discovered that activated protein C (APC), a well-known drug used to treat sepsis, can override tPA's toxic effect in mice. As a stand-alone therapy, APC might save neurons that would otherwise die. In combination with tPA, it could increase tPA's effectiveness after three hours by reducing collateral damage.

NATURAL HISTORY

When Extinct Isn't

QUESTIONING THE TERM AFTER A BIRD'S RETURN **BY MARGUERITE HOLLOWAY**

The video images may be tiny, grainy, dark and fleeting, but many looking at them see something glorious: evidence that at least one ivory-billed woodpecker—an 18- to 20-inch-tall bird with a wingspan of some 30 inches, last seen in the U.S. in 1944—is alive in the bottomland forest of eastern Arkansas. After a year of traipsing and canoeing through the Big Woods and its bayous, many inconclusive recordings of ivory-bill-like calls, seven good sightings and one fortuitous videotaping, scientists and conservationists announced in April that the bird was not extinct after all.

If the discovery holds up, the ivory-billed woodpecker will not be the only U.S. species recently returned from oblivion. In May, just a few days after the ivory-bill news, the Nature Conservancy announced the discovery in Alabama of three snails listed as extinct. A few weeks later, botanists at the University of California at Berkeley reported finding the Mount Diablo buckwheat, a tiny pink-flowered plant that had not been seen since 1936. At least 24 species of other presumed or possibly extinct plants, insects and other organisms have been found during natural heritage surveys in North America since 1974, according to Mark Schaefer, president of NatureServe, a nonprofit conservation group based in Arlington, Va. There are examples from elsewhere as well. The Bavarian pine vole, last seen in 1962, scurried back into view in 2000. The New Zealand storm petrel and the Lord Howe Island stick insect are among the other species no longer missing.

With so many "extinct" creatures reappearing, it is reasonable to wonder if the word has lost its meaning—something Ross

MacPhee, curator of mammalogy at the American Museum of Natural History, has been outspoken about. He worries that complacency may set in (if it hasn't already), because many species that people read or hear about are described as either on the verge of extinction, already extinct or formerly extinct. He cites the example of Miss Waldron's red colobus monkey, extinct in 2000, alive in 2004. "The average person has been barraged with the same story over and over again," he says. "People are using the term indiscriminately."

To counteract this trend, MacPhee and his colleagues formed the Committee on Recently Extinct Organisms in the late 1990s, devising criteria to determine extinction reliably—including rigorous taxonomic identification and a 50-year waiting period before declaring something extinct, an idea first put forth by the Convention on International Trade in Endangered Species.

Schaefer and John W. Fitzpatrick of Cornell University, lead author of the ivory-billed woodpecker report, agree that "extinction" should be applied more carefully. "The word 'extinct' is an absolute term, like pregnant or dead," Fitzpatrick remarks, "so we need to describe the probability of that being true." In cases such as the passenger pigeon, which has not been seen by anybody for nearly a century, "we treat it as formally extinct," he notes. For many plants and small vertebrates, "we suspect extinction, but they may still be hiding in some spots. That was the case for the ivory bill." Schaefer says that his organization has a letter scheme to describe species as GX (presumed extinct) or GH (possibly extinct) and then uses several other classifications, such as G1 (critically imperiled)—ranks also used by various federal agencies.

For now, the ivory-billed woodpecker seems to have been successfully upgraded. But "speaking probabilistically, this bird has a very, very slim chance of persisting," Fitzpatrick points out. "The key lies in growing back the old forest. That is exactly what this bird needed. Unlike in other cases of extreme endangerment, like Hawaii, the ivory bill is a case in which the natural habitat of the bird is getting progressively better. And I would add that even if the ivory bill fails, we should accelerate the process."

IVORY-BILLED WOODPECKER, last seen in the 1940s, appears to be alive in the forests of Arkansas. This 1935 photograph has been colorized.

GEORGE M. SUTTON *Cornell Lab of Ornithology*

NOT SO PROOF-POSITIVE?

The fuzzy images of the ivory-billed woodpecker have convinced many scientists of the bird's existence. But doubters remain, says Cornell University's John W. Fitzpatrick: "As we anticipated, there isn't universal acceptance of the video as conclusive hard evidence. The video is blurry, and there are people who dispute our interpretation of the video."

To respond to the skeptics, Fitzpatrick and his colleagues plan to make more evidence available over the next few months. "We have a lot of unpublished data from last year: significant numbers of acoustic signals that do give us hope that there are birds in several places," he says. "But it is all uncertain because the acoustic analysis will never leave us sure unless the bird is yakking right into our microphone." The team will also keep scouring the woods, Fitzpatrick adds, "until we get a beautiful photo."

ASTRONOMY

Cosmic CAT Scan

OBSERVING THE EARLY UNIVERSE—WITH 10,000 TV ANTENNAS **BY W. WAYT GIBBS**

In the beginning, the universe was a void full of energy but without form. And so it remained for many millions of years—exactly how long is still a major mystery of cosmology—until the first stars condensed from the fog of matter and lit up with a blue nuclear glow.

Telescopes are just like time machines: the farther out in space they look, the further back into the past they peer. But even the best optical telescopes cannot make out what the universe was like at an age of less than one billion years. Before that time, a haze of neutral hydrogen gas shrouded

LOW-TECH RADIO TELESCOPE under construction in western China will use thousands of consumer television antennas and hundreds of cheap personal computers to slice through the fog that shrouded the infant universe.

these first beacons in the infant cosmos.

A new radio observatory under construction on the high plateau of Ulastai in remote western China may soon yield images of this formative epoch, however—and for a bargain price, too, because the sprawling instrument is built almost entirely from parts that one could buy at RadioShack. Even though it will cost just $3 million, the Primeval Structure Telescope (PaST) is one of China's largest investments so far in experimental astronomy. The project was launched in 2003 by Xiang-Ping Wu of the Chinese Academy of Sciences in Beijing, Jeffrey B. Peterson of Carnegie Mellon University in Pittsburgh and Ue-Li Pen of the Canadian Institute for Theoretical Astrophysics in Toronto.

Though formally a telescope, PaST is better thought of as an experiment. "We'll get enough data from it to answer our principal questions within a couple weeks of turning it on" next year, Peterson says. (Analyzing those data may take years, however.) That is because the instrument is essentially a giant, incredibly sensitive television receiver.

PaST will combine radio signals picked up by 10,000 high-gain antennas arranged in lines up to three kilometers long. The log-periodic antennas, similar to those sold by the millions for rooftops, cost just $20 each. Household coaxial cable splitters, installed backwards, combine the signals from multiple antennas and feed them into a bank of 320 ordinary Pentium PCs, running free Linux software. The computers merge the data to produce a high-resolution picture of a 10-degree patch of sky centered near the North Star.

Actually "picture" is not the right word, because PaST will record thousands of simultaneous signals within a broad swath of the VHF spectrum. The scientists are writing software to sift out uninteresting signals—such as those from television stations, meteors and black holes at the centers of distant galaxies—to reveal a kind of three-dimensional CAT scan of the early universe that theorists predict lies buried within the noise.

As the first stars flickered on, their ultraviolet light excited neutral hydrogen atoms around them, causing the gas to emit a faint radio signal at 1,420 megahertz. As the starshine intensified, it eventually stripped electrons from the hydrogens, ionizing the atoms.

But over time the expansion of the universe stretched the ancient radio waves, lowering their frequencies by an amount proportional to their age. Astronomers can thus see a particular moment in time and location in space by "tuning" their receiver to the appropriate frequency. "It's a bit like archaeology," says Abraham Loeb of the Harvard-Smithsonian Center for Astrophysics. "We can slice the universe and see more and more ancient layers as we go deeper."

Astronomers expect that PaST will reveal a uniform haze of bright neutral hydrogen at about 200 million years after the big bang that became increasingly punctuated by bubbles of ionized—and thus dark—hydrogen surrounding the first stars. Simulations suggest that these shells then connected, like the voids in a Swiss cheese, to form tunnels. The neutral hydrogen fog gradually dissipated into stray wisps and vanished forever within the first billion years, leaving us with the transparent space we see today. This story will remain fuzzy until PaST or a competing observatory delivers more clarity. Let there be light.

RACE AGAINST THE PaST

The Primeval Structure Telescope (PaST) faces competition from two higher-budget, higher-tech projects. "It's sort of a race right now," says Abraham Loeb, a cosmologist at the Harvard-Smithsonian Center for Astrophysics. "It's not clear who will win. Each experiment has advantages and disadvantages."

- **LOFAR project:** **15,000 antennas in the Netherlands; pilot installation of 100 antennas now in place. Cost: At least $65 million to date. Cons: Noisy radio environment.**
- **Mileura Widefield Array (MWA):** **8,000 antennas in the Australian outback. Cost: $10 million. Cons: Proposed to, but not yet approved by, the U.S. National Science Foundation.**

JEFFREY B. PETERSON

BY THE NUMBERS

Coming to America

IMMIGRATION TODAY RIVALS THE INFLUX OF A CENTURY AGO **BY RODGER DOYLE**

The current wave of immigration, which rivals the massive influx of 1880–1914, started with the Immigration and Nationality Act of 1965. Since then, about 27 million legal immigrants have crossed the border. In addition, an estimated 10.3 million illegal ones live in the U.S. The net result is that, as of 2004, there were 34.2 million foreign-born residents in the country. More than half are from Latin America and about a quarter from Asia, which contrasts with the pre–World War I period, when the foreign-born were overwhelmingly European.

Today's surge, like its predecessor, is profoundly affecting the culture and economics of the U.S., particularly in southern Florida, southern California and the New York metropolitan area. In 2004 the foreign-born accounted for 11.3 percent of the population, and at their present rate of increase, this figure could exceed the record of 14.6 percent in 1890. The foreign-born now account for half the growth of the U.S. population.

One reason that the U.S. draws immigrants is the long-standing shortage of native-born workers. Too few Americans are acquiring scientific and engineering skills: of the foreign-born, 3.3 percent of those 18 years of age and older hold higher degrees, such as Ph.D.s and J.D.s, compared with 2.2 percent of the native-born population. At the same time, the native-born shun many manual jobs. Farm labor, for instance, is largely foreign-born. Several California industries, such as apparel and construction, depend almost exclusively on immigrant workers.

Another stimulus to immigration is U.S. involvement abroad, which has led to waves of migrants from South Korea, Vietnam, Cuba and other countries. Other triggers include civil conflict, as in Colombia, and hard times, as in the former Soviet countries.

Although some immigrants are a burden on the welfare system, as a group they pay far more in taxes than they receive in government benefits, such as public education and social services. A National Academy of Sciences study in 1997 found that immigrants had little impact on the earnings of U.S.-born Americans, except for unskilled jobs, where native-born high school dropouts found their wages going down because of competition from unskilled immigrants. According to the National Foreign Intelligence Board, an advisory body to the Central Intelligence Agency, the more liberal immigration policies of the U.S. have given it a competitive edge over Europe and Japan in industries such as information technology.

Rodger Doyle can be reached at rdoyle2@adelphia.net

FROM AROUND THE WORLD

Two thirds of immigrants to the U.S. have settled in just six states: California, New York, Texas, Florida, Illinois and New Jersey.

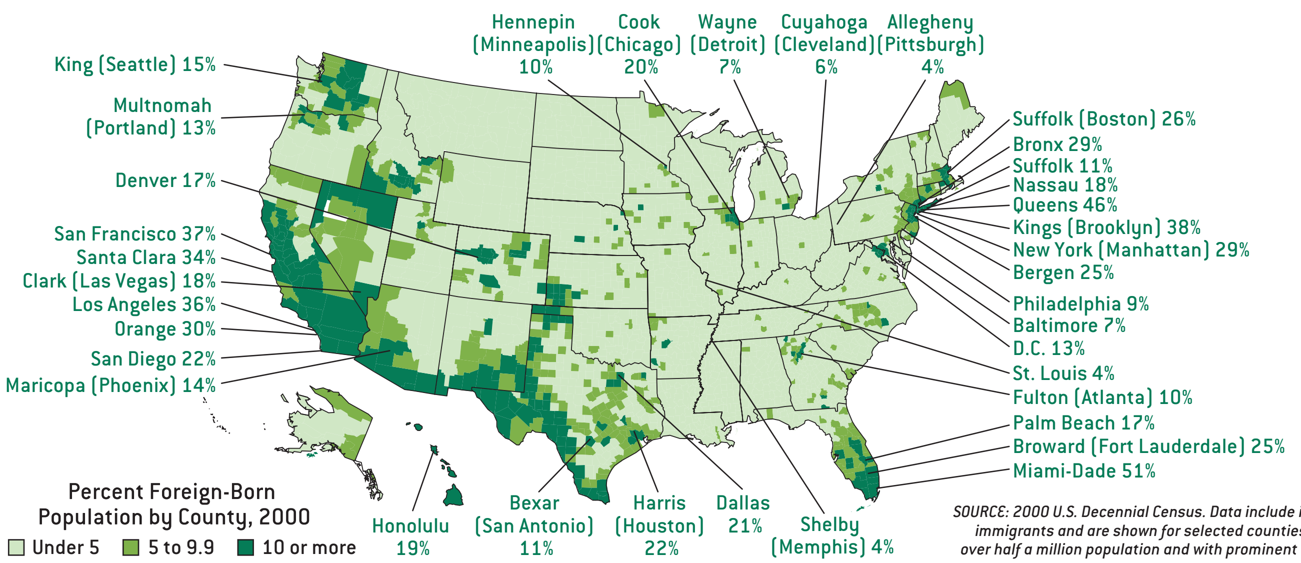
Leading countries from which foreign-born hail (in thousands):

Mexico: 9,177
China: 1,519
Philippines: 1,369
India: 1,023
Vietnam: 988
Cuba: 873
South Korea: 864
Canada: 821
El Salvador: 817
Germany: 707
Dominican Republic: 688

FURTHER READING

Profile of the Foreign-Born Population in the United States: 2000. U.S. Census Bureau, December, 2001.

Migration Policy Institute: www.migrationpolicy.org



SOURCE: 2000 U.S. Decennial Census. Data include illegal immigrants and are shown for selected counties with over half a million population and with prominent cities.

RODGER DOYLE

DATA POINTS: STAYING POWER

Treatment programs against tuberculosis have largely kept the potentially fatal airborne disease under control in the U.S.—the incidence has been dropping steadily since 1992. But the disease's persistence worldwide and multidrug-resistant strains, created perhaps by incomplete courses of antibiotics, still pose challenges. A study examining data between 1994 and 2003 (the latest available) in California, the state with the greatest number of cases, highlights the issues.

Number of U.S. tuberculosis cases, 2003: 14,874

California cases: 3,224

Percent with multidrug-resistant strains: 1.4

Percent dying: 14

Among those who are HIV-positive: 90

Cost to treat someone with multidrug-resistant TB: $28,217 to $1,278,066

Percent of resistant cases born outside the U.S.: 83

Number of foreign-born cases, per 100,000 people: 3.5

Number among U.S.-born cases: 0.2

Percent of patients who reported previous bout of TB: 6

Among those with multidrug-resistant TB: 31

SOURCES: Centers for Disease Control and Prevention; Journal of the American Medical Association, *June 8, 2005*

PSYCHOLOGY

About Face

Hidden messages can drive you to drink. Psychologists sat 39 volunteers in front of computers displaying faces with neutral expressions on them. Unbeknownst to subjects, the computers subliminally flashed smiling or angry faces for $^{1}/_{25}$ of a second for every second of neutral expression. When given cups of lemon-lime drink, volunteers who said they were thirsty before the experiment drank more than double after happy faces than after frowns. In a variation of the experiment, highly thirsty participants exposed to hidden smiles were willing to pay up to triple the price for the drink than others shown grimaces. No volunteers reported consciously seeing the subliminal faces. Being thirsty beforehand was key; subjects who were not apparently remained unaffected. Piotr Winkielman of the University of California at San Diego and his colleagues presented their findings on May 28 at the American Psychological Society meeting in Los Angeles. *—Charles Q. Choi*

HORMONES

Scents of Trust

Oxytocin, a hormone associated with lactation and social bonding, seems to make people more inclined to trust strangers with their cash. In a game in which 58 volunteers could invest money with anonymous trustees, economist Ernst Fehr of the University of Zurich and his colleagues dosed half the subjects with a whiff of oxytocin and the others with an inert placebo nose spray. After sniffing oxytocin, investors were roughly twice as likely to invest all their funds, and on average the oxytocin group put in nearly 20 percent more than placebo subjects. This effect was not seen when the trustee was replaced with a computer, revealing that oxytocin promotes social interaction instead of just making people more likely to take risks. In the June 2 *Nature,* the researchers suggest that this work could help alleviate social phobias, rather than assist mutual-fund agents. *—Charles Q. Choi*

ANIMAL BEHAVIOR

Sponge-Nose Smarty Pants

Dolphins may not be the second smartest species on earth—next to mice—as *The Hitchhiker's Guide to the Galaxy* contends, but they do appear to have culture. Marine biologists have noticed that some wild dolphins in Australia's Shark Bay, most of them females, keep marine sponges on their snouts for protection as they root around the seafloor for food. Researchers suspected it was learned tool use, but they couldn't discount a direct genetic cause for the behavior. Now, based on DNA samples from 13 "spongers" and 172 nonspongers, biologists from the University of New South Wales in Sydney have found that all the spongers are related to one another and probably share a recent ancestor, but they don't fit a pattern in which a gene could be causing the tool use. Hence, they conclude in the June 21 *Proceedings of the National Academy of Sciences USA* that daughters are most likely learning the behavior from their mothers. *—JR Minkel*

AQUA CULTURE: A two-year-old dolphin uses a sponge to protect her snout, a behavior learned from her mother.

JANET MANN/*PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA*, 2005; ILLUSTRATION BY MATT COLLINS

news SCAN

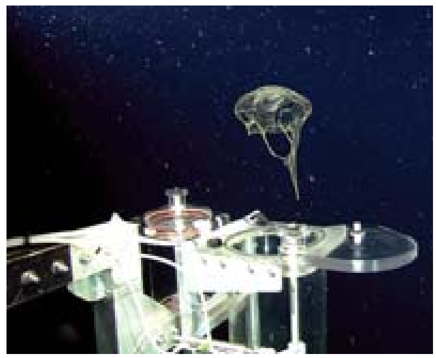
ONCOLOGY

Cancer's John Hancock

One theory of cancer formation holds that some tumors aggressively spread, or metastasize, because they contain a pocket of perpetually renewing cells similar to stem cells. Inspired by this idea, investigators from the Sidney Kimmel Cancer Center in San Diego, Calif., used a mouse model of prostate cancer to identify 11 genes expressed in both neural stem cells and metastatic tumors. They then checked the genes' expression patterns in tissue samples from 1,153 patients treated for one of 11 cancers, including leukemia, lymphoma, and tumors of the breast, prostate and lung. Expression of the signature genes predicted a four- to 20-fold increased likelihood of relapse or death within three to five years, depending on the type of cancer. If confirmed, the signature could help patients and doctors decide whether to adopt aggressive treatment early on, says Gennadi Glinsky, co-author of the June 1 *Journal of Clinical Investigation* report. —*JR Minkel*

DEEP-SEA ECOLOGY

Dinner Slime



MUCUS CATCHER: A sampling chamber from a remote-controlled submersible prepares to collect a sinking larvacean house.

MBARI © 2002

A mystery bedeviling marine biologists for years was how life at the bottom of the ocean gets enough food. The measured amount of sinking organic particles did not seem to be enough to sustain the observed creatures there. The menu secret might be mucus. Tadpole-size creatures called giant larvaceans dwell in delicate houses of slime often more than a meter wide, which they use to snare food. Once a house gets too clogged, the larvacean abandons it and spins a new one. Bruce Robison and his colleagues at the Monterey Bay Aquarium Research Institute in Moss Landing, Calif., spent 10 years collecting larvacean houses and videoed the building and discarding of the mucous abodes. Larvaceans spun themselves a new house every day, with enough old homes descending to account for roughly half as much carbon as the rain of organic particles. Their findings fell into the June 10 *Science*. —*Charles Q. Choi*

PHYSICS

The Next Little Thing

As a step to molecular circuits, researchers would like to control the charge flowing across single organic molecules, thereby turning them into electrical junctions. But the relatively enormous electrodes abutting the molecules typically complicate such attempts. Leaving a single "dangling bond," or unpaired electron, on a silicon electrode may be the key to precise charge control of adjacent molecules, reports a group from the University of Alberta in Canada. They deposited lines of ringed carbon molecules on a silicon surface in such a way that a dangling bond punctuated the end of each line. A scanning tunneling microscope played over the structures felt more charge in carbon molecules closer to dangling bonds, indicating that the bonds could serve to precisely alter a molecule's conductivity. See the June 2 *Nature* for details. —*JR Minkel*

BRIEF POINTS

- **We got the beat: Babies preferred listening to musical rhythms to which they were bounced or rocked, suggesting that the experience of physical movement helps develop the perception of musical beats.**

  Science, June 3

- **The detection of stars extending from the Andromeda galaxy's main disk indicates that the galaxy is 220,000 light-years across, three times bigger than previously thought. The Milky Way spans about 100,000 light-years.**

  Meeting of the American Astronomical Society, May 30

- **A mosquito carrying West Nile virus could infect other mosquitos feeding next to it via the insect's saliva. The finding may explain the virus's rapid spread, which may occur among dead-end hosts.**

  Proceedings of the National Academy of Sciences USA online, June 10

- **Beating binge drinking? A small study has found that ingesting kudzu extract cut subsequent alcohol consumption by almost half, perhaps because the herb sensitized the body's response to alcohol.**

  Alcoholism: Clinical and Experimental Research, May

Full of Holes

The curious case of acupuncture By MICHAEL SHERMER

John Marino was the most driven man I ever met, a monomaniac on a mission to break the U.S. transcontinental cycling record—which he did in 1980, covering the 3,000 miles in 12 days, three hours. I wanted to be like John, so that year I took up serious cycling. In addition to pedaling hundreds of miles a week with him, I followed his training regimen of vegetarian meals, megavitamin dosing, fasting, colonics, mud baths, iridology (iris reading), negative ions, chiropractic, massage and acupuncture.

Although most of the nostrums I tried were useless, I noted with interest (because he beat me) that Jonathan Boyer, the winner of the 1985 Race Across America (co-founded by Marino and me), had a Chinese acupuncturist on his support crew. Given the successes of Marino and Boyer, it seemed possible that there might be a biomedical connection.

Sham acupuncture was as effective as "real" acupuncture.

Traditional Chinese medicine holds that a life energy called Qi ("chee") flows through meridians in the body; each of the 12 main meridians represents a major organ system. On these 12 meridians are 365 acupuncture points, one for each day of the year. When yin and yang are out of balance, Qi can become blocked, leading to illness. Inserting needles at blocked points—now believed to number about 2,000—supposedly stimulates healing and health.

This theory lacks any basis in biological reality, because nothing like Qi has ever been found by science. Nevertheless, a medicinal procedure like acupuncture may work for some other reason not related to the original, erroneous theory. Electroacupuncture—the electrical stimulation of tissues through acupuncture needles—increases the effectiveness of analgesic (pain-relieving) acupuncture by as much as 100 percent over traditional acupuncture. So says George A. Ulett, a practicing physician and acupuncturist (with both an M.D. and Ph.D.) and author of the 1992 *Beyond Yin and Yang: How Acupuncture Really Works* and the 2002 textbook *The Biology of Acupuncture* (both published by Warren H. Green in St. Louis). Ulett posits that electroacupuncture stimulates the release of such neurochemicals as beta-endorphin, enkephalin and dynorphin, leading to pain relief. In fact, he says, the needles are not even needed—electrically stimulating the skin (transcutaneous nerve stimulation) is sufficient. Ulett cites research in which, using this technique, the amount of gas anesthetic in surgery was reduced by 50 percent.

These findings might help explain the results of a study published in the May 4, 2005, issue of the *Journal of the American Medical Association,* in which Klaus Linde and his colleagues at the University of Technology in Munich compared the experiences of 302 people suffering from migraines who received either acupuncture, sham acupuncture (needles inserted at nonacupuncture points) or no acupuncture. During the study, the patients kept headache diaries. Subjects were "blind" to which experimental group they were in; the evaluators also did not know whose diary they were reading. Professional acupuncturists administered the treatments. The results were dramatic: "The proportion of responders (reduction in headache days by at least 50%) was 51% in the acupuncture group, 53% in the sham acupuncture group, and 15% in the waiting list group." The authors concluded that this effect "may be due to nonspecific physiological effects of needling, to a powerful placebo effect, or to a combination of both."

In my experience, "needling" (where the acupuncturist taps and twists the flesh-embedded needle) isn't painful, but it is most definitely noticeable. If acupuncture has effects beyond placebo, it is through the physical stimulation and release of the body's natural painkillers. Finding that sham acupuncture is as effective as "real" acupuncture demonstrates that the Qi theory is full of holes. The effects of being poked by needles, however, cannot be ignored. Understanding the psychology and neurophysiology of acupuncture and pain will lead to a better theory. And for all such alternative medicine claims, testimonials can steer us in the direction of where to conduct research; science is the only tool that can tell us whether they really work or not. SA

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

BRAD HINES

Kryder's Law

The doubling of processor speed every 18 months is a snail's pace compared with rising hard-disk capacity, and Mark Kryder plans to squeeze in even more bits By CHIP WALTER

Over the years there has been a lot of talk about Moore's Law and the way that doubling the power and memory of computer semiconductors every 18 months has driven technological advance. But from where Mark Kryder sits, another force is at least as powerful, perhaps more: the cramming of as many bits as possible onto shrinking magnetic hard drives.

The 61-year-old engineer might be on to something. Since the introduction of the disk drive in 1956, the density of information it can record has swelled from a paltry 2,000 bits to 100 billion bits (gigabits), all crowded in the small space of a square inch. That represents a 50-million-fold increase. Not even Moore's silicon chips can boast that kind of progress.

THEO ANDERSON

MARK KRYDER: STORING DATA

- Leading the way in boosting hard-disk capacity, which has enabled products such as the iPod. His next goal: a terabit in a square inch—equal to about 200 DVD movies—which will most likely spawn new industries.
- Predicts that in a few years the U.S. consumer will have 10 to 20 disk drives in various devices: "He may not realize that he has that many, but he will have them in appliances enabled by disk-drive technology."

Kryder is not denigrating the importance of faster computer processors, but he says, at the very least, both digital elements need each other. Without the continual squeezing of bits onto ever shrinking hard drives, the world of information as we know it today, and tomorrow, will come to a grinding halt.

Information storage has been Kryder's bailiwick most of his career. As founder and director of Carnegie Mellon University's Data Storage Systems Center and now as chief technology officer at hard-drive manufacturer Seagate Technology, he has often spearheaded the breakthroughs that have increased hard-disk densities (and accelerated their corresponding drop in price). But these days, he says, altogether unexpected trends are afoot: smaller, high-capacity drives are spawning not only new products and applications but entirely new industries.

"Who would have predicted the success of handheld digital audio players?" Kryder asks. "We completely missed seeing the iPod coming." Now, he points out, "disk drives are appearing in GPS systems for automobiles and enabling us to record and playback HDTV on TiVo and digital cable systems."

Such devices may relegate Moore's Law to secondary status. "Today the density of information we can get on a hard drive is much more important to enabling new applications than advances in semiconductors," Kryder remarks. Without them, Apple Computer's iTunes Music Store would not have sold hundreds of millions of songs and on-demand TV would still be a pipe dream.

But to Kryder, these new services are just the beginning. Now tiny, capacious hard drives are replacing low-capacity flash memory cards, which use electrically charged transistors rather than moving parts to record information. Soon hard drives will migrate into phones, still cameras, PDAs, cars and everyday appliances. "In a few years the average U.S. consumer will own 10 to 20 disk drives in devices that he uses regularly," Kryder predicts. These advances are forcing manufacturers to become much more nimble as their markets expand. Optimizing a drive for an Xbox or an automobile's diagnostic system is very different from creating a razor-thin, rugged one-inch drive for a flip phone.

Kryder began exploring digital storage in the 1970s as a postdoc at the California Institute of Technology. Later he spent five years at the IBM Thomas J. Watson Research Center, where he researched bubble memory, which records data by magnetizing small circles on gadolinium gallium garnet. When he joined Carnegie Mellon in 1978, Kryder continued his bubble memory work, but it became clear that the technology, used in cruise missiles and other niche applications, faced as obstacle as a mainstream product: gadolinium gallium garnet was expensive. When the fledgling personal computer industry made the hard drive its storage device of choice, Kryder switched gears, assembled a conference of hard-drive industry gurus in 1982, and asked them to name their greatest research needs. Next he persuaded businesses such as IBM and 3M to support an effort to develop those technologies. The result in 1983: the Magnetics Technology Center (MTC), the only operation of its kind in the U.S. For the next five years, the center incubated increasingly efficient hard-drive technologies, while cultivating the field's top thinkers.

TINY HARD DRIVES are creating new industries. This one, introduced by Hitachi in January, can hold 10 gigabytes.

But the MTC tended to react to what the industry wanted, Kryder says, rather than pushing the envelope. So, in 1987, after discussions with the National Science Foundation (NSF), he worked to create an organization that set a technological agenda. In 1990 the MTC became the Data Storage Systems Center (DSSC), one of a handful of NSF-funded engineering research centers. Kryder immediately set an ambitious goal: demonstrate hard drives that could store four gigabits of information in a square inch of disk space. Back then, four billion bits represented an enormous leap. Densities at the time hovered around 100 million bits. But in just four years the DSSC had met the new benchmark, a 40-fold increase.

By 1998, when Kryder joined Seagate to form its advanced research center, the DSSC had set an even loftier target: crowd 100 gigabits into a square inch by the early 21st century. In 2005, just seven years later, Seagate began shipping 110-gigabit drives. Inside of a decade and a half, hard disks had increased their capacity 1,000-fold, a rate that Intel founder Gordon Moore himself has called "flabbergasting."

But now current hard-drive technologies are hitting a new wall. Hard disks typically store bits of information using a tiny head that flies across the surface of the disk and magnetizes billions of discrete areas in horizontal space that represent zero or one, depending on whether they are facing clockwise or counterclockwise. The magnetized areas are becoming so small that it is difficult for them to remain stable.

Kryder and his team are reviving a method called perpendicular recording to fix the problem. It flips the charges north to south, permitting the use of stronger magnetic fields in media that can store smaller bits. Seagate's Pittsburgh lab has already prototyped this approach, which should pack in at least 200 gigabits per square inch within the next two years. Ultimately, Kryder thinks perpendicular drives will record 400 or 500 gigabits within four years. Because that is nothing more than a blink in the world of hard drives, Kryder has already set his next goal: a terabit per square inch, and he has tapped the 100 Ph.D.s at Seagate to work up still more exotic recording systems to make it a reality.

One project tackles a new method called heat-assisted magnetic recording (HAMR), which uses a burst of heat so the drive's head can more easily magnetize even smaller surfaces. When the disk cools, the magnetic field stabilizes. Beyond that he foresees patterned media recording, which would theoretically allow drives to magnetize 10 times more information.

Surprisingly, his team will not be working on holographic storage, considered by many as the ultimate storage technology. Holography uses all three dimensions to store data, and the goal has been to stuff a terabit into a space the size of a sugar cube. But Kryder predicts that in another six years or so, hard drives will reach the terabit benchmark, at which time they will be smaller and cheaper than holographic systems.

Kryder isn't predicting where all these tiny drives lodged everywhere will lead. The big question for him isn't so much how to crowd more bits on drives, but understanding how those drives will shape the industries of the future. SA

Chip Walter, based in Pittsburgh, is working on his third book, Six Traits (That Make Us Human).

More complicated than they look, teeth are actually tiny organs.

Test-Tube Teeth

If tissue engineers can manufacture living replacement teeth, they would blaze a trail for engineering larger organs while leading dentistry into the age of regenerative medicine

By Paul T. Sharpe and Conan S. Young

We take them for granted until they are gone

or require major repairs. And then the options are grim: do without lost teeth or replace them with inert prosthetic versions. In the Western world, an estimated 85 percent of adults have had some form of dental treatment. Seven percent have lost one or more teeth by age 17. After age 50, an average of 12 teeth stand to have been lost.

In theory, a natural tooth made from the patient's own tissue and grown in its intended location would make the best possible replacement, although such bioengineered teeth have for many years been little more than a dream. Recently, however, progress in understanding how teeth first develop has combined with advances in stem cell biology and tissue engineering technology to bring us close to the realization of biological replacement teeth.

Apart from the potential benefit to people who need new teeth, this research also offers two significant advantages for testing the concept of organ replacement: teeth are easily accessible, and whereas our quality of life is greatly improved if we have them, we do not need our teeth to live. These may seem trivial points, but as the first wave of replacement organs start to make their way toward the clinic, teeth will serve as a crucial test of the feasibility of different tissue engineering techniques. With organs essential to life, doctors will have no leeway to make mistakes, but mistakes with teeth would not be life-threatening and could be corrected.

This is not to say that engineering teeth will be simple. Millions of years of evolution went into establishing the complex processes that produce organs, teeth included, during embryonic development. The challenge for tissue engineers is to replicate those processes, which are tightly controlled by the growing embryo's genes. A good way to start learning how to build teeth, therefore, is to observe how nature does it.

Overview/*Cutting-Edge Teeth*

- Tissue engineers working toward creating living replacement teeth take cues from nature as they coax disparate cell types to form a functional organ.
- Alternative methods include building teeth from existing dental cells or growing them from progenitor tissues. Both approaches have already produced structurally correct teeth.
- Remaining challenges include growing roots and identifying ideal raw materials for bioengineered human teeth, but progress has been rapid and test-tube teeth may become the first engineered organs.

Delicate Dialogue

JUST SIX WEEKS after conception, a human embryo is less than an inch long and barely beginning to take recognizable shape. Yet a constant cross talk among its cells is already initiating and guiding the formation of its teeth. The intricacy of such signal exchanges is among the reasons that teeth and other organs cannot as yet be grown entirely in dishes in laboratories. Indeed, scientists may never be able to completely reproduce these conditions artificially. The more we understand these early developmental processes, however, the greater will be our chances of providing engineered tooth tissues with the most important cues for organ building and letting nature do the rest.

Most organs, for example, arise through interactions between two distinct embryonic cell types, epithelial and mesenchymal, and teeth are no exception. In the embryo, oral epithelial cells (which are destined to line oral cavities) send out the first inductive signals to mesenchymal cells (which will produce jawbone and soft tissues), instructing them to begin odontogenesis, or tooth formation. Once the mesenchymal cells have received their initial instructions, they start sending signals back to the epithelial cells. This reciprocal exchange continues throughout embryonic tooth development.

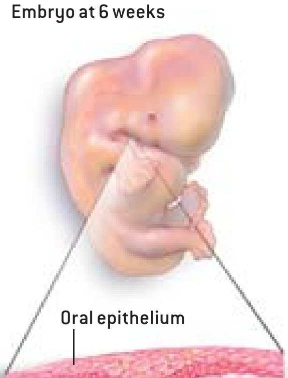
At first, the future tooth is no more than a thickening in the embryonic oral epithelium. As it grows, the epithelium begins to penetrate the underlying mesenchymal tissue, which in turn condenses around the protrusion, forming a tooth bud by the embryo's seventh week [*see box on opposite page*]. As the epithelium penetrates farther, it wraps itself around the condensing mesenchyme, eventually forming a bell-shaped structure, open at its bottom, around 14 weeks. Ultimately, the epithelium will become the visible outer enamel of the tooth that erupts from the baby's gum line some six to twelve months after birth, and the mesenchymal cells will have formed the nonvisible parts of the tooth, such as dentin, dental pulp, cementum, and a periodontal ligament that attaches the tooth to the jawbone.

Even before this tooth begins forming, its shape will be predetermined by its position. Some of the same epithelial signals that trigger initiation of odontogenesis also regulate

CARY WOLINSKY (*photography*); JEN CHRISTIANSEN (*photoillustration*) (*preceding pages*)

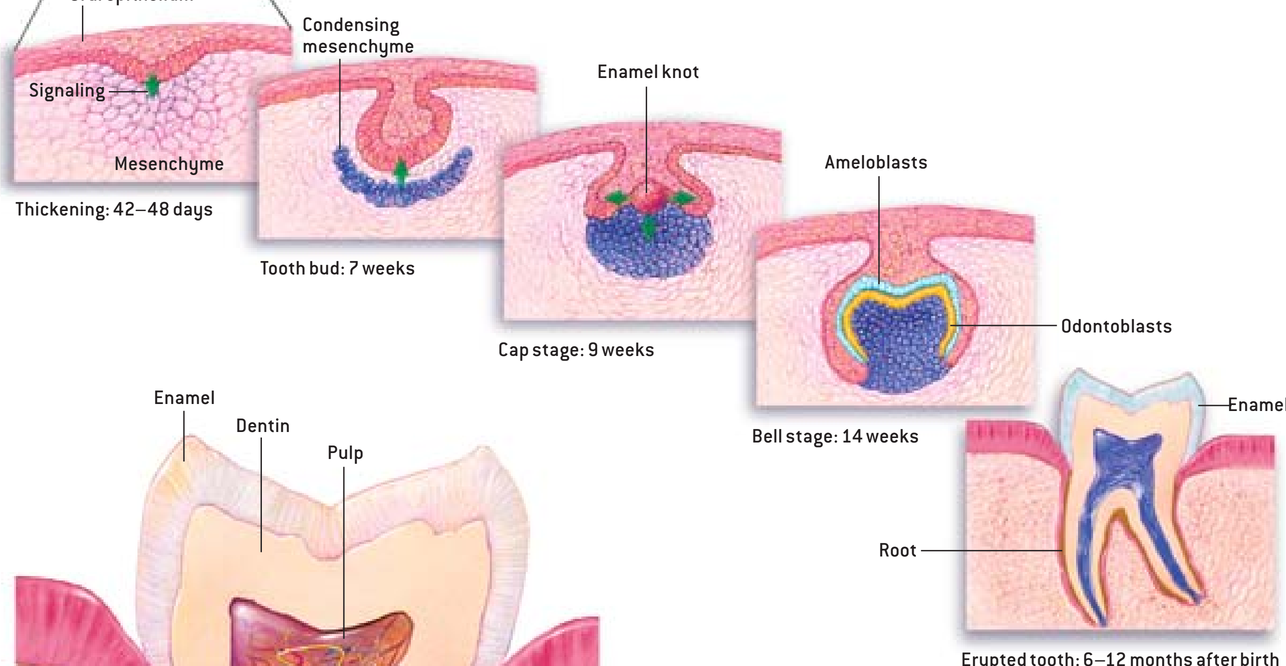
HOW NATURE ENGINEERS A TOOTH

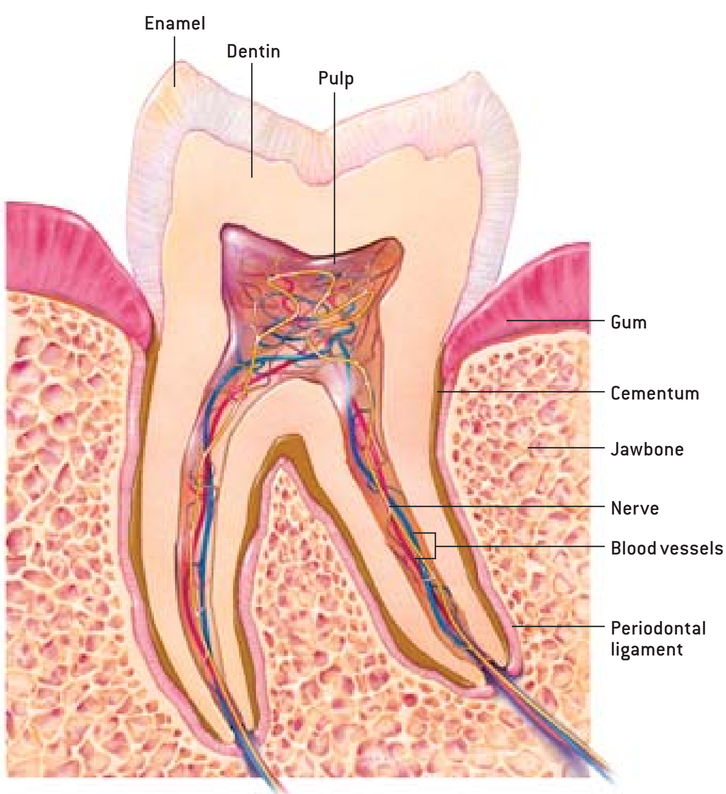
It may look simple from the outside, but on the inside a tooth is a tiny marvel of design and construction that takes about 14 months to complete in a developing human. Two different types of primordial embryonic tissue combine to produce a tooth, and an ongoing molecular dialogue between them directs the process. Tissue engineers are studying these signals and steps to understand the cues they need to replicate as they create living bioengineered replacement teeth.



TOOTH FORMATION

Just six to seven weeks into human embryonic development, when the whole head is still taking shape, teeth are also beginning to form. At the location of a future tooth, oral epithelial tissue thickens slightly and gene activity within its cells causes signals to be sent to underlying mesenchymal tissue. As the epithelium penetrates farther, mesenchymal cells respond by emitting their own signals and condensing around the protrusion to form a tooth bud. By week nine, the epithelium has become a cap atop condensed mesenchyme. A structure at its center called the enamel knot is now a primary source of signals directing the activity of both epithelial and mesenchymal cells. At 14 weeks, the tooth germ has a bell shape comprising differentiating cells called ameloblasts, which will later become enamel, and odontoblasts, which will form dentin. Roots are the last structures to develop, completing their formation as the tooth erupts some six to 12 months after birth.





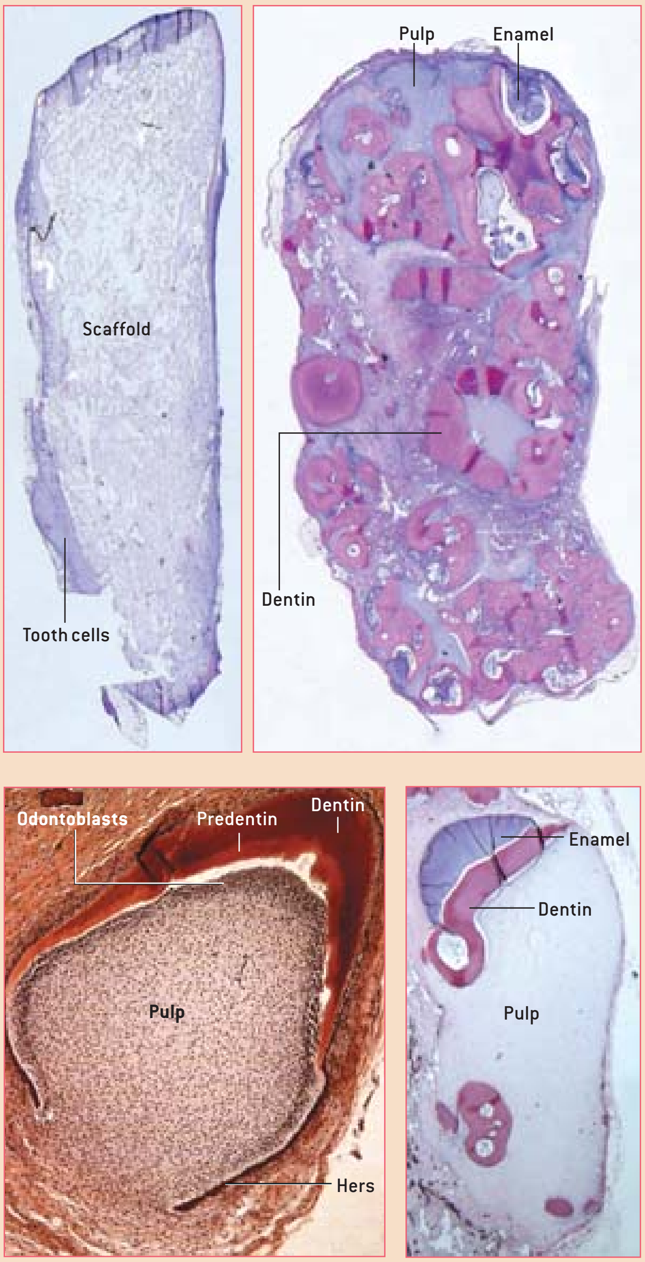
END RESULT

A living tooth is defined as an organ because it comprises multiple tissue types, each with an essential function. Enamel, the body's hardest mineralized surface, seals and protects the interior. Dentin, a bony substance, makes up the bulk of a tooth and serves as a cushion to resist chewing forces. Pulp, in the center, contains nourishing blood vessels and nerves that provide sensory perception. Cementum forms the hard outer surface of a tooth where it is not covered by enamel. Periodontal ligament is a connective tissue that attaches to both the cementum and the jawbone, anchoring the tooth in place yet providing some flexibility.

ANDREW SWIFT

Cells Reunite to Form Teeth

Tooth cells taken from adolescent pigs and seeded onto a biodegradable scaffold are visible in blue along its edges after one week of incubation (*top left*). Following 25 weeks of growth (*top right*), the scaffold has dissolved and new dental pulp, enamel, and dentin have taken its place. In a series of such experiments, tiny toothlike structures grew amid the new tissues. Correct tooth-tissue organization (*bottom left*), including a pre-root structure known as Hertwig's epithelial root sheath (Hers), was observed in 15 to 20 percent of the miniature teeth. In other instances, the tooth structure was incorrect or incomplete (*bottom right*). These bioengineered teeth nonetheless seem to confirm that disaggregated dental cells can reorganize themselves into larger dental tissues.



REPRINTED FROM *DEVELOPMENTAL ANALYSIS AND COMPUTER MODELLING OF BIOENGINEERED TEETH*, BY C. S. YOUNG ET AL. IN *ARCHIVES OF ORAL BIOLOGY*, VOL. 50, PAGES 259–265; 2005. WITH PERMISSION FROM ELSEVIER (*top left* and *right; bottom right*); CONAN S. YOUNG (*bottom left*)

an important category of genes in the jaw mesenchyme. Known as homeobox genes, they participate in determining the shape and location of organs and appendages during embryonic development throughout the body. In a developing human jaw, different homeobox genes are activated in different areas, guiding each tooth bud down a pathway to become a molar, premolar, canine or incisor.

A homeobox gene called *Barx1,* for example, is switched on, or expressed, by mesenchymal cells in the positions where molar teeth will grow. In animal experiments, causing *Barx1* to be misexpressed in mesenchyme that would normally form incisors makes those teeth develop with a molar shape instead. Because the ability to predict and control tooth shape will be essential for the creation of engineered teeth, scientists can use the activity of genes such as *Barx1* as definitive predictive markers of future shape when teeth created in the lab are first growing in culture.

In turn, we must provide the right signals to the developing teeth at the right time. As early as the 1960s, researchers such as Shirley Glasstone of Strangeways Research Laboratory in Cambridge, England, began exploring the possibility of growing teeth by experimenting with mouse tissues. In seminal studies performed over the next three decades, tiny pieces of embryonic mouse dental epithelium and dental mesenchyme were brought together and then either grown in a tissue culture dish or surgically implanted in the body of a host where the recombined tissues would receive a blood supply. These experiments demonstrated that such embryonic tooth primordia could continue to develop as if they were still in the embryo, producing dentin and enamel. Their development arrests early, however, and they do not ultimately yield fully formed teeth. Something is missing from their environment.

The growth factors and other signals required to complete tooth formation in an embryo most likely come from surrounding jaw tissue. Thus, transplanting tooth primordia into the jaw to finish developing would seem to be a simple solution. When replacement teeth are engineered, for instance, they will ideally be grown in their permanent location so that they can create nerve and blood vessel connections and physically attach themselves to the jawbone. The adult jaw is a vastly different environment from the embryonic version, however, and scientists have been unsure whether it would provide the correct signals to a developing tooth.

Moreover, tooth primordia must be constructed from the right combination of cells to reproduce natural tooth material and structure. Being able to use cells from a patient's own body would be preferable to using embryonic cells because the patient's own tissue would not be perceived as foreign and so would not provoke an immune response.

Three key milestones must therefore be reached to establish whether engineering replacement biological teeth is possible. Sources of cells that can form teeth and are easily obtained from patients themselves must be identified. The teeth produced from these cells must be able to develop in the environment of the adult jaw, producing roots that are attached to

Each cell seems to know its place in the larger collective.

the bone by a functional periodontal ligament. And the shape and size of these biological teeth must be predictable and controllable so that they can be made to match the patient's own teeth. These are ambitious goals, but considerable progress toward each is being made by different research groups using somewhat disparate approaches.

Building Bioteeth

IN THE LATE 1980S organ transplant surgeon Joseph P. Vacanti of Harvard Medical School and polymer chemist Robert S. Langer of the Massachusetts Institute of Technology conceived the idea of placing the cells of an organ or tissue on a prefabricated biodegradable scaffold with the goal of generating tissues and organs for transplantation [see "Artificial Organs," by Robert S. Langer and Joseph P. Vacanti; SCIENTIFIC AMERICAN, September 1995]. In simplified terms, their approach was based on the fact that living tissues are made of cells constantly signaling to one another and often moving around within a three-dimensional community of sorts. Each cell seems to know its place and role in the larger collective that forms and maintains a functional tissue. Therefore, if the right mix of dissociated cells is reaggregated within a scaffold that replicates their natural 3-D environment, the cells should instinctively reform the tissue or organ to which they belong.

Vacanti and Langer's early successes regenerating pieces of liver tissue from liver cells using this scaffold-based strategy have since led to widespread experimentation with the technique to produce other complex tissues, such as heart muscle, intestine, mineralized bone and now teeth. Pamela C. Yelick and John D. Bartlett of the Forsyth Institute in Boston began working with Vacanti in 2000 to investigate the feasibility of engineering teeth this way by focusing on pigs, which, like humans, produce two sets of teeth over their lifetime.

One of us (Young) also took part in these experiments for which raw material was derived from the unerupted third molars ("wisdom teeth") of six-month-old pigs. To obtain a heterogeneous random mixture of dental enamel epithelial and pulp mesenchymal cells, the pig teeth were broken into tiny pieces and then further dissolved using enzymes. Tooth-shaped scaffolds were made from biodegradable polyester plastics and coated with a substance that makes the plastic sticky so cells can adhere to it. The cell mixtures were seeded into the scaffolds, and the constructs were surgically implanted into rat hosts, wrapped in omentum, a fatty white material rich in blood vessels that surrounds the intestines. This step is important because the developing tooth tissues require an ample blood supply to provide them with nutrients and oxygen while they grow.

Initially the scaffolds provided support for the cells, but later they dissolved as intended and were replaced by new tissue. When the implants were examined after 20 to 30 weeks, tiny toothlike structures were visible within the confines of the original scaffold. Their shape and the organization of their tissues resembled the crowns of natural teeth [*see box on opposite page*]. They also included most of the tissues that make up a normal tooth, demonstrating for the first time that enamel, dentin, pulp, and features that appeared to be developing tooth roots could be regenerated on scaffolds.

It seemed that mixtures of dental cells could reorganize themselves on scaffolds into arrangements that favor formation of mineralized enamel, dentin and soft tooth tissue. Another possible explanation for these exciting results, of course, was that the random arrangement of cells seeded onto the scaffold favored tooth tissue development only by chance. The Forsyth group therefore tested these possibilities in a new study using dental epithelial and mesenchymal cells isolated from the first, second and third molars of rats. This time, however, the cells were grown and their numbers expanded in tissue culture for six days before their being seeded onto scaffolds and implanted in rat hosts. After 12 weeks' growth, the resulting tissues were extracted and examined. Once again, small tooth structures consisting of enamel, dentin and pulp tissue were observed to have formed within the original scaffold.

These new results were encouraging because they lent some weight to the previous evidence that cells can reorganize them-

THE AUTHORS

PAUL T. SHARPE and *CONAN S. YOUNG* met two years ago at a tooth and bone conference where they discovered a shared fondness for mountain biking and soccer (one calls it "football"), despite their differing approaches to bioengineering teeth. Sharpe established and heads the department of craniofacial development at Guy's Hospital in London and is also Dickinson Professor of Craniofacial Biology at King's College London. In 2002 he founded Odontis Ltd., a biotechnology company devoted to growing human teeth and bone by emulating their formative processes in a developing embryo. Young is an instructor in oral and developmental biology at the Harvard School of Dental Medicine and a staff scientist at the Forsyth Institute in Boston, where he is working toward growing teeth from cells seeded onto biodegradable scaffolds.

No one knew whether the adult jaw would provide signals for teeth to form.

selves into tooth-forming configurations. Moreover, the cells did not appear to have been adversely affected by being expanded in culture—a process that will be essential in engineering human replacement teeth because tissue engineers would probably have to craft a replacement tooth from small samples of the patient's own cells. And, finally, the experiment demonstrated that tooth regeneration is possible in a second mammal, making the success of a similar approach in humans more likely.

Although the Forsyth team was able to generate most of the desired tissue types with cells from an adult source, those tissues organized themselves into the proper arrangement for a natural tooth only 15 to 20 percent of the time. The group is therefore continuing to work on methods of more precisely placing different dental cell types within scaffolds to achieve a more accurate tooth structure.

At the same time, the team is exploring the possibility that the new tooth tissues observed in these experiments might not have been produced solely by reorganization of the dissociated dental cells. Instead the third molar tooth buds that provided cells to seed the scaffolds might have contained hidden stem cells—potent progenitors of other cell types—that were responsible for forming the new tissue. If true, this would mean that new dental stem cells capable of producing nearly all the dental tissue types required for bioengineering teeth might exist within teeth themselves, at least until early adulthood, when wisdom teeth erupt. Such versatile adult dental stem cells would certainly speed efforts to generate teeth on scaffolds, and they might also facilitate the tooth-engineering approach used by the Sharpe group at King's College London.

Teeth from Scratch

RATHER THAN ATTEMPTING to build adult teeth from their constituent cells, one of us (Sharpe) is pursuing a strategy based more closely on reproducing the natural processes of embryonic tooth development described earlier. In essence, the method requires an understanding of the basic principles controlling early tooth formation and a source of cells to play the roles of embryonic oral epithelium and mesenchyme.

To date, the Sharpe group has experimented primarily with mouse cells, using both stem cells and ordinary cells, from embryonic as well as adult sources, to test the potential of various cell types to produce replacement teeth. In most cases, the group began by aggregating mesenchymal cells in a centrifuge until they formed a small solid mass. This pellet was then covered in epithelium and cultured for several days, while the gene activity in its tissues was monitored for indications of early tooth development. Next, these tooth primordia were implanted into the bodies of animal hosts in locations where they could receive a nourishing blood supply, such as the kidney of a mouse, and left to grow for about 26 days.

In the course of these experiments, clear tooth formation was observed but only when the epithelium came from an embryonic source and the mesenchymal cell populations contained at least some stem cells. When stem cells from adult bone marrow took the place of oral mesenchyme, for example, the transplanted constructs produced structurally correct teeth. Thus, it seems embryonic mesenchyme can be replaced with adult stem cells to generate new teeth.

Unfortunately, many years of experiments have established that embryonic epithelium contains a unique set of signals for odontogenesis that disappear from the mouth after birth. The Sharpe group is continuing to seek an effective population of substitute cells that could be derived from an adult source. Still, the results achieved with primordia made from the combination of adult stem cells and embryonic oral epithelium have been extremely encouraging.

Significantly, these teeth were also in the normal size range for mouse teeth, they were surrounded by new bone and connective tissue, and they showed the earliest signs of root formation. The next step was to see whether such explants could also form teeth in the mouth. In the embryonic jaw, soft tissues, teeth and bone are all developing together without external stresses such as chewing and talking, whereas the adult jaw is a hard, busy place. No one knew whether it would provide the necessary signals for teeth to form and integrate themselves into the environment as they would in an embryo.

To find out, the Sharpe group extracted tooth buds from embryonic mice, then transplanted them into the mouths of adult mice. Small incisions were made in the soft tissue of the upper jaw of the host mice, in a region known as the diastema between the molars and incisors where normally there are no teeth. The embryonic tooth primordia were inserted into these pockets and sealed in place with surgical glue. Afterward, the mice were fed a soft diet and the transplants monitored. Just three weeks later teeth could be clearly identified in the diastema. They had formed in the correct orientation, were of appropriate size for the mice, and were attached to underlying bone by soft connective tissue [*see illustration on opposite page*].

Remarkably, it appears that the adult mouth can provide a suitable environment for tooth development. That is just one of the three milestones toward engineering replacement teeth that we identified earlier, however. The road to human bioengineered teeth may yet have a few twists.

On the Cusp

COMPARED WITH EFFORTS to engineer other organs, teeth have made considerable progress in a short time. The overall challenge remains developing methods that are simple yet controlled.

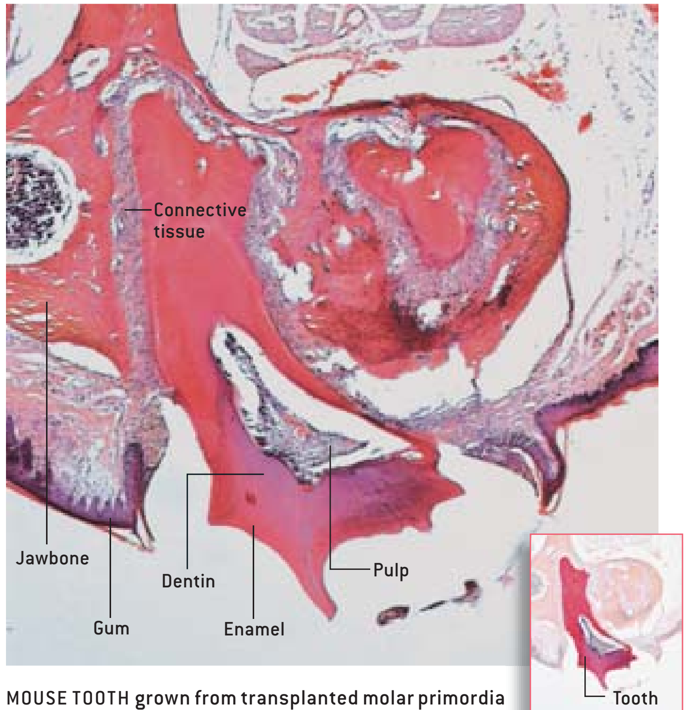
Another of the targets that we established, the ability to predict and control tooth size and shape, is close. In cultured primordia, molar and incisor tooth germs can easily be distinguished by their appearance and their gene activity, although other shapes found in the human mouth, such as premolars and canines, are more difficult.

The teeth grown from embryonic primordia in the mouths of adult mice by the Sharpe group displayed shapes appropriate to their original locations in the embryo—molar primordia grew into molar-shaped teeth, for example. Because shape signals are received at the very start of natural tooth development, the embryonic tooth germs were already programmed. Tissue engineers need to better understand these initial shape signals to induce them in human bioteeth.

To date, the teeth generated by any of the tissue engineering methods we have described have not developed roots. In truth, both root development and the stimuli that initiate tooth eruption are complex and still little understood. Roots are the last part of teeth to form, completing their development during the eruption process, and more research is needed to understand what conditions would best favor their creation in replacement teeth. Another unknown is how long engineered human teeth would take to fully form in an adult mouth. Humans' second set of "adult" teeth also begins developing in the embryo, yet those teeth take six to seven years to finally erupt—or 20 years in the case of wisdom teeth. Our experience with tooth generation in animals suggests that an engineered human tooth would form far more quickly, but we do not know if it might take longer to fully mature and its enamel to completely harden.

Of course, most research into bioengineered tooth production is also working toward finding an effective and easily accessible source of the patient's own cells to use as raw material. Immune rejection would be avoided, and because tooth size, shape and color are genetically determined, the engineered teeth would more closely match the patient's natural teeth. The Sharpe group has found that adult mesenchymal stem cells derived from bone marrow (but also possibly obtainable from fat) can replace embryonic mesenchyme in the tooth formation process. A substitute for embryonic epithelium has yet to be identified, although purported adult stem cells have been discovered in other tissues with epithelial origins, such as skin and hair. These or some other adult cell type may prove effective, perhaps with the aid of gene manipulation to induce the appropriate initiating signals for odontogenesis.

Of the several potential cell sources, teeth themselves may be the most convenient. The Forsyth group's results suggest that stem cells capable of forming tooth tissues, including enamel, could be present within teeth. Researchers elsewhere have also shown that dentin and other tooth tissues experience some natural regeneration after injury, which, too, suggests the presence of progenitor cells capable of generating a variety of tooth tissues. Thus, the possibility exists of someday soon fashioning new teeth from old. SA



MOUSE TOOTH grown from transplanted molar primordia in the upper jaw of a host mouse demonstrates that new teeth can develop in the adult mouth. The tooth at center in this cross section of the jaw's diastema region has broken through the gum line (a second tooth above it and to the right is still forming). Pulp is visible inside the emerged tooth. Red stain colors dental hard tissues, highlighting enamel and dentin. Although lacking roots, the tooth is attached to surrounding jawbone by soft connective tissue.

MATT BRADMAN

MORE TO EXPLORE

Tissue Engineering: The Challenges Ahead. Robert S. Langer and Joseph P. Vacanti in *Scientific American,* Vol. 280, No. 4, pages 86–89; April 1999.

Tissue Engineering of Complex Tooth Structures on Biodegradable Polymer Scaffolds. Conan S. Young, Shinichi Terada, Joseph P. Vacanti, Masaki Honda, John D. Bartlett and Pamela C. Yelick in *Journal of Dental Research,* Vol. 81, No. 10, pages 695–700; October 2002.

Bioengineered Teeth from Cultured Rat Tooth Bud Cells. Monica T. Duailibi, Silvio E. Duailibi, Conan S. Young, John D. Bartlett, Joseph P. Vacanti and Pamela C. Yelick in *Journal of Dental Research,* Vol. 83, No. 7, pages 523–528; July 2004.

Stem Cell Based Tissue Engineering of Murine Teeth. A. Ohazama, S.A.C. Modino, I. Miletich and P. T. Sharpe in *Journal of Dental Research,* Vol. 83, No. 7, pages 518–522; July 2004.

The Cutting Edge of Mammalian Development: How the Embryo Makes Teeth. Abigail S. Tucker and Paul T. Sharpe in *Nature Reviews Genetics,* Vol. 5, No. 7, pages 499–508; July 2004.

The Early Evolution of Animals

Tiny fossils reveal that complex animal life is older than we thought—by at least as much as 50 million years

By David J. Bottjer

"THERE IS A BILATERIAN IN THAT TRUCK," Jun-Yuan Chen said as we watched the vehicle disappear around a bend in the road. Chen, a paleontologist at the Chinese Academy of Sciences in Nanjing, and I, along with Stephen Q. Dornbos, a colleague then at the University of Southern California, had just collected a truckload of black rocks from a 580-million- to 600-million-year-old deposit in Guizhou Province. Chen was sure they held something important.

We had come to Guizhou in 2002 to hunt for microscopic fossils of some of the earliest animals on earth. Specifically, we were hoping to find a bilaterian. The advent of bilateral symmetry—the mirror-image balance of limbs and organs—marks a critical step in the history of life. The first multicelled animals were not bilaterally symmetrical; they were asymmetrical aquatic blobs—sponges—that filtered food particles from currents they generated. Radially symmetrical

OLDEST FOSSIL ANIMAL with a bilateral body plan yet discovered, *Vernanimalcula* lived in the seas some 580 million to 600 million years ago. This reconstruction enlarges the creature to reveal its complexity; in life it was about the size of the period at the end of this sentence.

TWO DEPOSITS IN CHINA have preserved the remains of soft-bodied animals that provide new information about early evolution. In 2004 the author and his colleagues discovered the oldest known bilaterian animal in rocks collected from the 580-million- to 600-million-year-old Doushantuo Formation, near Weng'an. Significantly younger fossils from the approximately 525-million-year-old deposits in the vicinity of Chengjiang have expanded understanding of the Cambrian explosion.

aquatic creatures, the cnidarians, are slightly more complex; they have specialized stinging cells that can immobilize prey. Bilaterians constitute all the rest of us, from worms to human beings. During some stage in their life cycle, all display not only the crucial left-right balance but a multilayered body that typically has a mouth, gut and anus.

Until several years ago, consensus held that bilaterian animals first appeared in the fossil record about 555 million years ago, although the vast majority showed up somewhat later in a burst of innovation known as the Cambrian explosion, which began about 542 million years ago. The dearth of earlier fossils made it impossible to test ideas about what triggered the "explosion" or even to say for sure whether it was real or merely seemed so because earlier animals left few detectable traces of themselves. But research over the past half a dozen years—including ours in Guizhou Province—has changed the long-held view, suggesting that complex animals arose at least 50 million years earlier than the Cambrian explosion.

Overview/*Older Than We Thought*

- The development of bilateral symmetry marks a critical step in the early evolution of animals.
- Genetic analysis has suggested that bilateral symmetry arose 573 million to 656 million years ago, but controversy clouds the date for several reasons. The most telling is that until recently the earliest known bilaterian fossils were dated to only 555 million years ago.
- Now the author and his colleagues have found supporting fossil evidence for the earlier date: microscopic creatures in Chinese deposits 580 million to 600 million years old.
- The minuscule fossils not only support an early date for the beginning of complex animal life but show that internal complexity evolved before large size did.

Molecular Clocks and Lagerstätten

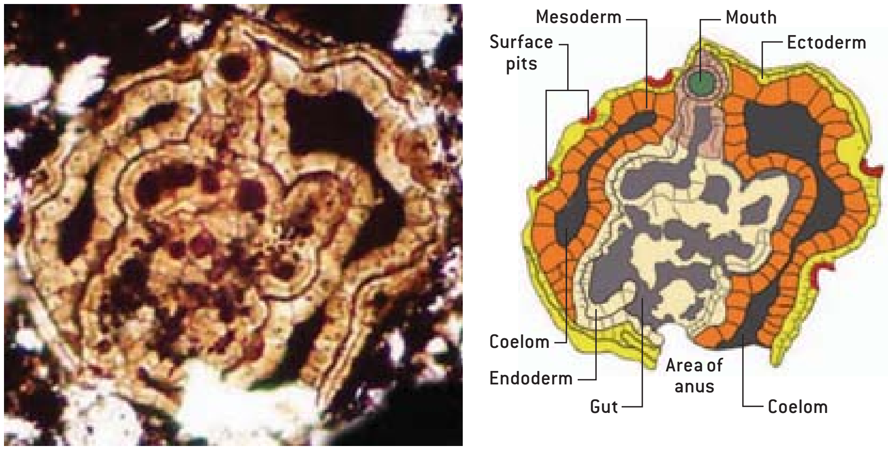
MOLECULAR ANALYSIS, in particular a technique called the molecular clock, has been key in the new thinking about when the earliest animals arose. The clock idea is based on the supposition that some evolutionary changes occur at a regular rate. Over millions of years, for example, mutations may be incorporated in the DNA of genes at a steady rate. Differences in the DNA of organisms, then, can act as a "timepiece" for measuring the date at which two lineages split from a common ancestor, each going its separate way and accumulating its own distinctive mutations.

To estimate the timing of the origin of various major animal groups, Gregory Wray of Duke University and his colleagues used a molecular clock rate based on vertebrates (animals that have a backbone). Their results, published in 1996, postulated that bilaterians diverged from more primitive animals deep into the Precambrian era, as much as 1.2 billion years ago.

Follow-up studies using the molecular clock produced estimates for this split that varied significantly, ranging from as old as one billion years ago to as young as just before the Cambrian period. Such discrepancies naturally generated doubts about the technique, and a more recent study by Kevin Peterson of Dartmouth College and his colleagues addressed some of these concerns. In particular, they used a molecular clock rate derived from invertebrates, which is slower than the one based on vertebrates.

This investigation placed the last common ancestor of bilaterian animals at a much younger date, though still older than the Cambrian explosion, somewhere between 573 million and 656 million years ago. But even this date sparked controversy. It had become clear that only actual fossils would furnish incontrovertible evidence for the time at which bilaterians had emerged. This realization provided a big incentive for paleontologists to get out in the field and find fossils older than the Cambrian. I was among the scientists spurred to search for these elusive specimens.

AMADEO BACHAR (*preceding pages*); ALISON KENDALL (*this page*)



BEST-PRESERVED SPECIMEN of *Vernanimalcula* is shown in the photomicrograph on the left. Diagnostic characteristics of a bilaterian are evident: a multilayered body with paired cavities called coeloms, a mouth and a gut. The drawing on the right highlights key features (the anus is not preserved in this specimen; its location is based on its placement in other specimens).

One huge problem with finding such animals is that they did not have hard skeletons that would mineralize and become fossils. So we must rely on uncovering the rare deposit that, because of the type of rock and the chemical processes involved, preserves intricate details of the remains. These deposits are called *lagerstätten,* a German word that means "lode places" or "mother lode." A lagerstätte that preserves soft tissue is a spectacular rarity; we know of only several dozen scattered over the earth. One of the best known is the Solnhofen Limestone in Germany, where the 150-million-year-old feathered specimens of what is generally considered to be the earliest fossil bird, *Archaeopteryx,* are preserved. In British Columbia, an older deposit, the Burgess Shale, made famous by the writings of Stephen Jay Gould [see, for example, "The Evolution of Life on Earth," SCIENTIFIC AMERICAN; October 1994], reveals a cornucopia of curious soft-bodied organisms from the ancient oceans of the Cambrian period.

A lagerstätte older than the Burgess Shale, in the Chengjiang area of China's Yunnan Province, has yielded many important recent finds of soft-bodied organisms also characteristic of the Cambrian explosion. And, at several spots on the planet, the Ediacaran lagerstätten, named after the Ediacara Hills of Australia where the first example was found, harbor strange Precambrian soft-bodied fossils and animal burrows, including evidence for early bilaterians.

Amazingly, in 1998 two different groups of paleobiologists reported finding fossils with remarkable soft-tissue preservation in another Precambrian lagerstätte—the Doushantuo Formation in Guizhou Province of south China. This deposit contains tiny soft-bodied adult sponges and cnidarians as well as minuscule eggs and embryos. Much of the sediment in which they occur is composed of the mineral calcium phosphate (apatite), which has exquisitely replaced the original soft tissues of these fossils. The latest studies show that these rocks are older than the Ediacara biota, most likely 580 million to 600 million years old, and thus that the microfossils they contain lived 40 million to 55 million years before the Cambrian.

And So to China

THOSE OF US interested in the origin of animals quickly realized that the Doushantuo Formation might be the window through which we would glimpse early bilaterian life. So, in the autumn of 1999, a group of us joined together, at the urging of Eric Davidson, a molecular biologist at the California Institute of Technology, to study the Doushantuo microfossils. The team also included Chen and Chia-Wei Li, who were among the first investigators to report on eggs and embryos in the Doushantuo. Li, a professor at National Tsing Hua University, is an expert on biomineralization, and Chen has extensive experience studying early animal life through his pioneering work on the Lower Cambrian Chengjiang lagerstätte.

Our initial probes suggested that a relatively thin sedimentary layer, which is black in color, would be the most promising for finding a variety of microfossils. Other researchers at the site had applied acid to dissolve the rock matrix in the laboratory, revealing the tiny phosphatized fossils. Unfortunately, the acid dissolution technique was not successful with the layer of black rock that we had targeted. We therefore turned to a different approach: we collected great piles of this black rock and brought it back to Chen's lab at the Early Life Research Center of the Nanjing Institute of Geology and Palaeontology in adjacent Yunnan Province. That is where our dump truck was headed when Chen made his bilaterian prediction.

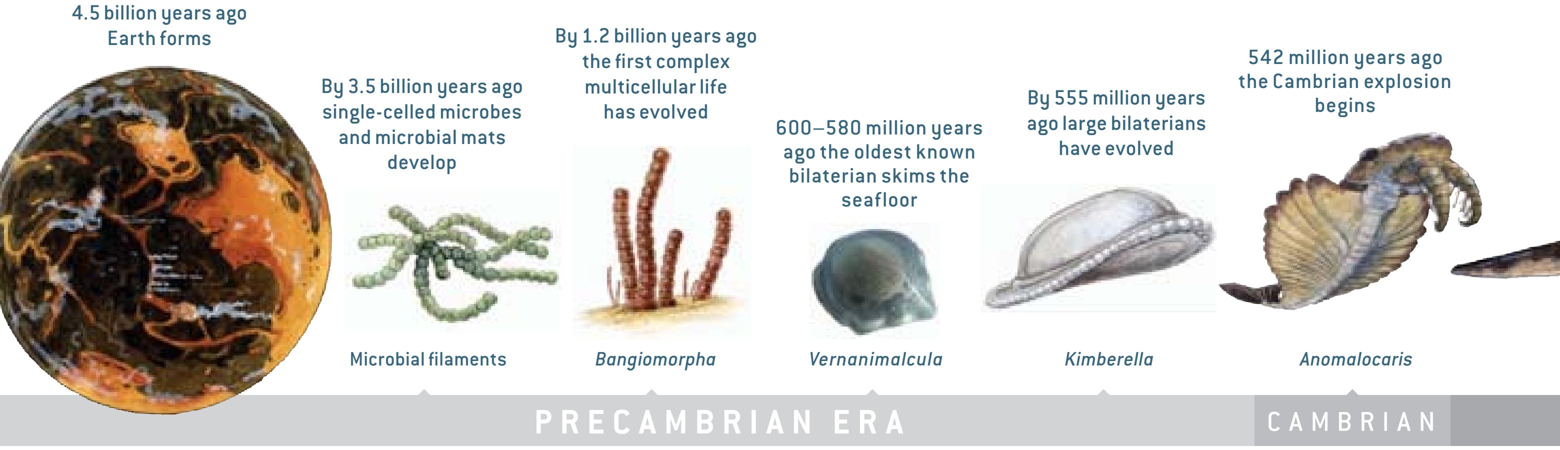
Once back in Yunnan with our rocks, we sliced the samples into thousands of sections, so thin that they were translucent and, when mounted on glass slides, could be examined under a microscope. We made more than 10,000 of these slides, a gargantuan task that Chen and his technicians threw themselves into with optimism and energy. Painstaking analysis of the thousands of slides took several years and revealed myriad eggs and embryos; it confirmed the presence

THE AUTHOR

DAVID J. BOTTJER is a paleobiologist who has focused his research on the origin and subsequent evolutionary history of animals on Earth. He approaches this work in an interdisciplinary fashion, which has led to collaborative ventures with colleagues versed in developmental biology, molecular biology, informatics and geochemistry. He received his Ph.D. in geology from Indiana University and is currently professor of earth and biological sciences at the University of Southern California. He is president of the Paleontological Society (2004–2006) and editor in chief of the journal *Palaeogeography, Palaeoclimatology, Palaeoecology.*

A Tiny Fossil's Place in History

The evolution of complex animal life was formerly thought to have started with a bang during the early Cambrian period, an event often referred to as the Cambrian explosion. The discovery in 2004 of the microscopic *Vernanimalcula* by the author and his colleagues pushes back the origins of complex animal life as much as 50 million years before the Cambrian.



of tiny adult sponges and cnidarians that had been reported previously.

But of course what we were really focused on finding was a bilaterian. Did our catch in the dump truck actually include one of these? In the summer of 2003 we began to zero in on one microfossil type whose complex morphological characteristics particularly intrigued us. Among the 10,000 slides, we were able to locate 10 examples of this type, and, early in 2004, after months of analysis, we came to the conclusion that this tiny organism displayed the basic features of a bilaterian. This was what we were looking for!

Ranging from 100 to 200 microns across, the width of several human hairs, these microscopic fossils are surprisingly complex and constitute almost a textbook example of a bilaterian, including the three major tissue layers (the endoderm, mesoderm and ectoderm familiar from high school biology texts), the

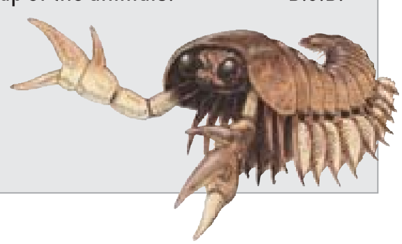
The Real Meaning of the Cambrian Explosion

The Cambrian explosion is generally thought of as a sudden increase in the types of bilaterian animals—those with a right-left balance of limbs and organs. But the story is more complicated, and more interesting, than that. Recent research has shown that a dramatic upsurge in interactions among animals played a large role in this increase in diversity.

First, animals began to alter the environment, and the new conditions created both opportunities and barriers for other denizens of the ancient world. For example, Precambrian animals that lived on the seafloor were adapted to moving about on cushiony microbial mats, which covered most of the ocean floor and had been part of the ecosystem since life originated. At the beginning of the Cambrian (which lasted from 542 million to 488 million years ago), however, evolutionary innovations enabled bilaterian animals to burrow vertically through sediment. The burrowing destroyed the ubiquitous mats and replaced them with a surface that was soupy and unstable. Some organisms, such as the helicoplacoids, small top-shaped animals that lived embedded in the seafloor, most likely became extinct as the sea bottom grew increasingly unstable. In contrast, other organisms reacted to this increase in bioturbation by evolving adaptations for living in the new environments.

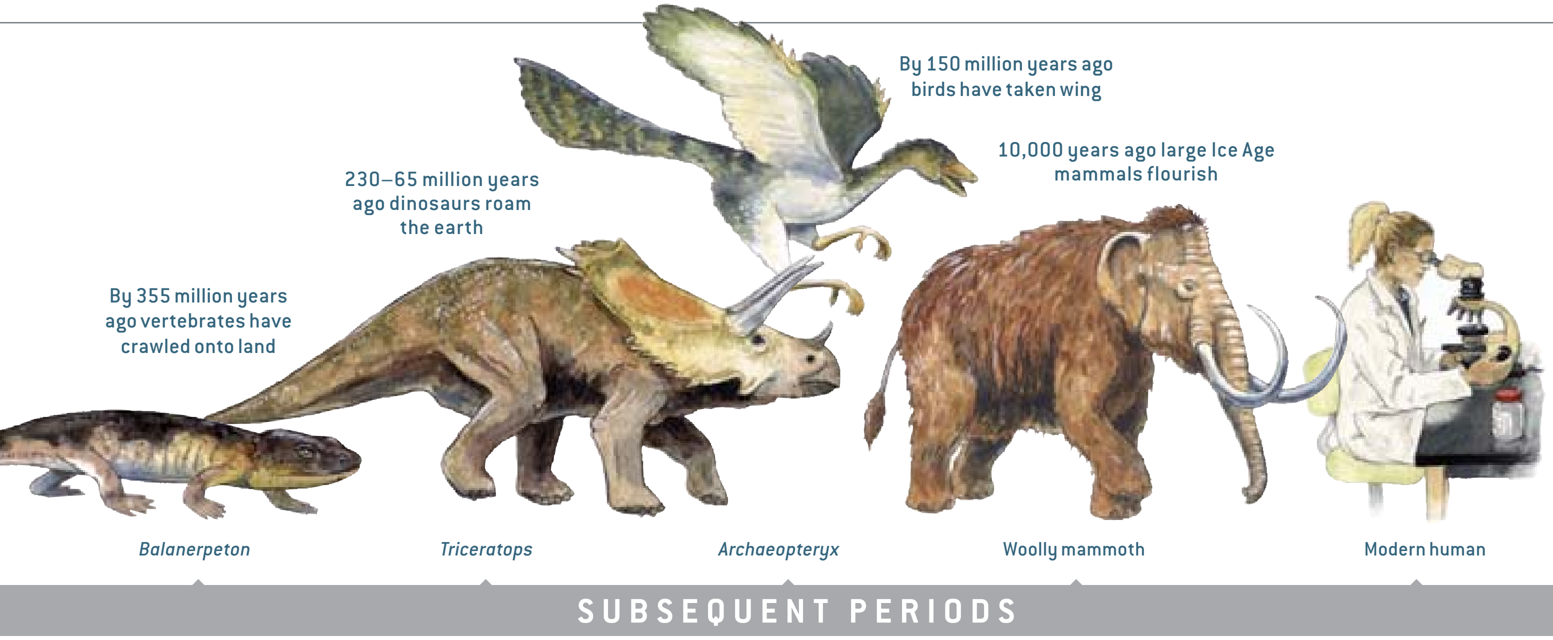
Second, the Early Cambrian marks the time when paleobiologists detect the first presence of bilaterian predators that had evolved to eat other animals. For example, Jun-Yuan Chen and Di-Ying Huang of the Chinese Academy of Sciences in Nanjing and others report several new types of predators from the Chengjiang lagerstätte in China. These include arthropods with strange frontal appendages for capturing prey (*below*), as well as ubiquitous burrowing worms that moved just below the seafloor and fed on other small animals.

These biological interactions played a strong role in the early evolution of animals. Yet as Charles Marshall of Harvard University has argued and as our findings support, the genetic tool kit and pattern-forming mechanisms characteristic of bilaterians had likely evolved by the time of the Cambrian explosion. Thus, the "explosion" of animal types was more accurately the exploitation of newly present conditions by animals that had already evolved the genetic tools to take advantage of these novel habitats rather than a fundamental change in the genetic makeup of the animals. *—D.J.B.*



EARLY PREDATOR, *Haikoucaris* (about four centimeters long)

AMADEO BACHAR



presence of a gut with a mouth and anus, and paired coeloms (body cavities) surrounding the gut. Oval-shaped and looking something like a minute gumdrop, the creature probably scooted along the seafloor to feed. At one end of the oval, the mouth sucked up microbes like a vacuum cleaner. Pits on either side of the mouth may have been sense organs.

We named our find *Vernanimalcula,* which means "small spring animal." The name refers to the long winter of "snowball Earth," when glaciers covered the planet [see "Snowball Earth," by Paul F. Hoffman and Daniel P. Schrag; Scientific American, January 2000]; the rocks holding *Vernanimalcula* are slightly above those marking the final glacial episode.

Legacy of a Small Spring Animal

BIOLOGICAL COMPLEXITY of the kind seen in *Vernanimalcula* implies a period of evolution that transpired long before the 580-million- to 600-million-year-old world in which the tiny animal lived. After all, it could not have gained that degree of symmetry and complexity all at once. We now need to find older lagerstätten that might hold clues to its ancestors.

We also need to move forward in time to try to puzzle out what happened to its descendants. What we know about life during the gap between *Vernanimalcula* and the creatures of the Cambrian explosion 40 million to 55 million years later comes primarily from studies of lagerstätten that contain the Ediacara biota—impressions and casts of soft-bodied organisms that were considerably larger than *Vernanimalcula,* ranging in size from centimeters to as much as a meter. New discoveries by Guy Narbonne of Queen's University in Ontario and his colleagues have confirmed the existence of these animals 575 million years ago; however, only in examples 555 million years old and younger do we find fossils that appear to represent bilaterians. Unlike the minuscule *Vernanimalcula,* these Ediacara bilaterians were macroscopic organisms, such as *Kimberella,* a soft-bodied sea dweller some 10 centimeters long that may have been an ancestor to the mollusks, animals that in today's seas include clams, snails and squid. Unfortunately, no Ediacaran deposits that we have located so far evince the unusual mineral setting essential for preserving microscopic creatures. To learn whether microscopic bilaterians existed alongside the larger Ediacara creatures we must find a fossil deposit of Ediacaran age that has preservation similar to that in the older Doushantuo Formation.

Although we cannot yet track the ancestors and descendants of *Vernanimalcula,* these tiny fossils have revealed a critical step in evolution: they show that bilaterians had the ability to make complex bodies before they could make large ones. Scientists are now speculating on what might have led to the eventual scaling-up of bodies. The most likely explanation is that a drastic rise in the amount of dissolved oxygen in seawater provided the impetus: more oxygen for respiration reduces constraints on size.

Vernanimalcula certainly gives paleontologists new inducements to go out and hunt for fossils of soft-bodied animals. We have a good deal left to learn, but the work so far has given substance to our earlier suspicion that complex animals have a much deeper root in time, suggesting that the Cambrian was less of an explosion and more of a flowering of animal life. SA

MORE TO EXPLORE

Cradle of Life: The Discovery of Earth's Earliest Fossils. J. William Schopf. Princeton University Press, 2001.

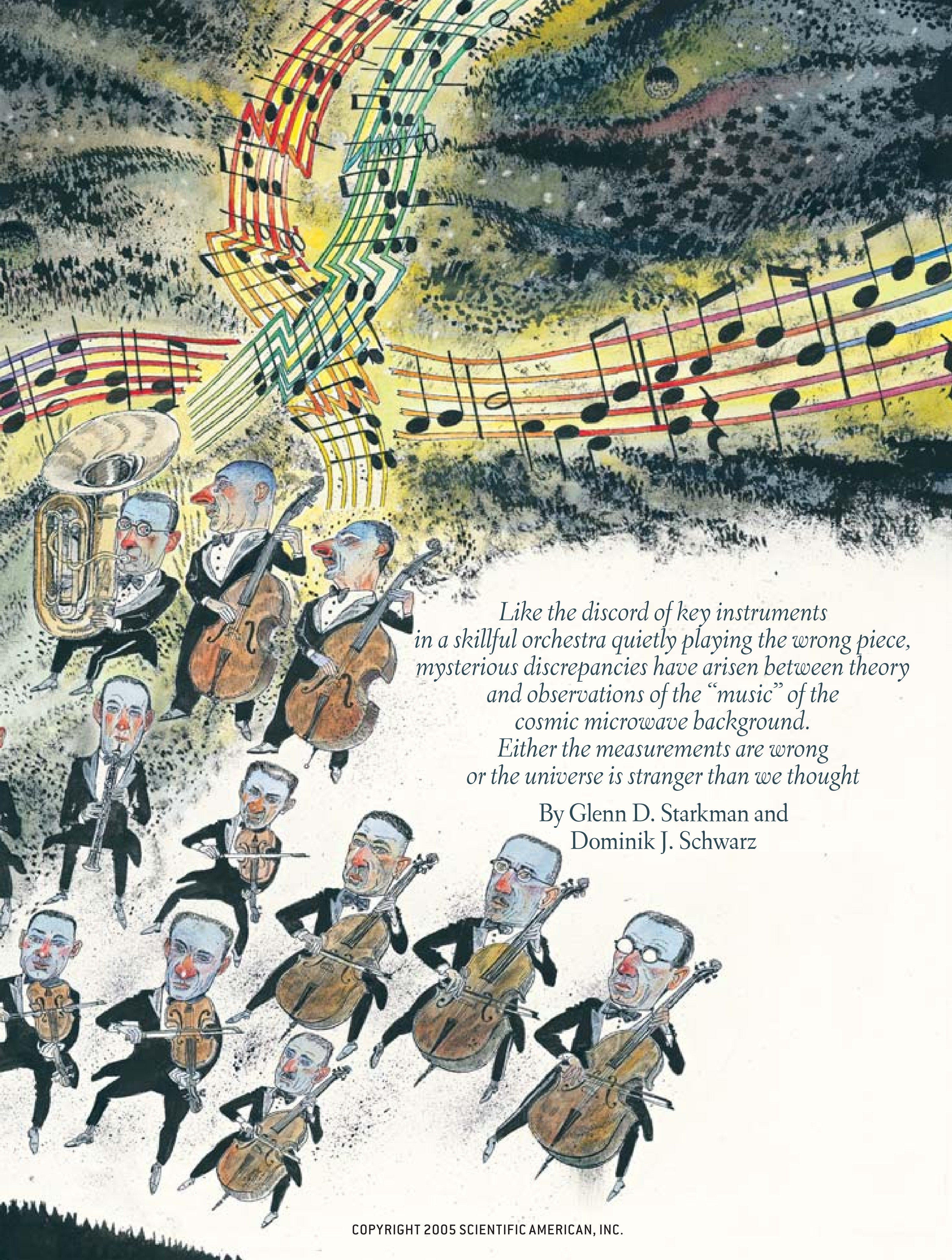
Evolution: The Triumph of an Idea. Carl Zimmer. Perennial (HarperCollins), 2002.

Life on a Young Planet: The First Three Billion Years of Evolution on Earth. Andrew H. Knoll. Princeton University Press, 2003.

On the Origin of Phyla. James W. Valentine. University of Chicago Press, 2004.

University of California, Berkeley, Museum of Paleontology Web site: **www.ucmp.berkeley.edu**

Is the Universe OUT OF TUNE?



Like the discord of key instruments in a skillful orchestra quietly playing the wrong piece, mysterious discrepancies have arisen between theory and observations of the "music" of the cosmic microwave background. Either the measurements are wrong or the universe is stranger than we thought

By Glenn D. Starkman and Dominik J. Schwarz

IMAGINE a fantastically large orchestra

playing expansively for 14 billion years. At first, the strains sound harmonious. But listen more carefully: something is off key. Puzzlingly, the tuba and bass are softly playing a different song.

So it is when scientists "listen" to the music of the cosmos played in the cosmic microwave background (CMB) radiation, our largest-scale window into the conditions of the early universe. Shortly after the big bang, random fluctuations—probably thanks to the actions of quantum mechanics—apparently arose in the energy density of the universe. They ballooned in size and ultimately became the galaxy clusters of today. The fluctuations were a lot like sound waves (ordinary sound waves are oscillations in the density of air), and the "sound" ringing throughout the cosmos 14 billion years ago was imprinted on the CMB. Now we see a map of that sound drawn on the sky in the form of CMB temperature variations.

As with a sound wave, the CMB fluctuations can be analyzed by splitting them into their component harmonics—like a collection of pure tones of different frequencies or, more picturesquely, different instruments in an orchestra. Certain of those harmonics are playing more quietly than they should be. In addition, the harmonics are aligned in strange ways—they are playing the wrong tune. These bum notes mean that the otherwise very successful standard model of cosmology is flawed—or that something is amiss with the data.

Scientists have constructed and corroborated the standard model of cosmology over the past few decades. It accounts for an impressive array of the universe's characteristics. The model explains the abundances of the lightest elements (various isotopes of hydrogen, helium and lithium) and gives an age for the universe (14 billion years) that is consistent with the estimated ages of the oldest known stars. It predicts the existence and the near homogeneity of the CMB and explains how many other properties of the universe came to be just the way they are.

Called the inflationary lambda cold dark matter model, its name derives from its three most significant components: the process of inflation, a quantity called the cosmological constant symbolized by the Greek letter lambda, and invisible particles known as cold dark matter.

According to this model, inflation was a period of tremendously accelerated growth that started in the first fraction of a second after the universe began and ended with a burst of radiation. Inflation explains why the universe is so big, so full of stuff and so close to being homogeneous. It also explains why the universe is not precisely homogeneous: because random quantum fluctuations in the energy density were inflated up to the size of galaxy clusters and larger.

The model predicts that after inflation terminated, gravity caused the regions of extra density to collapse in on themselves, ultimately forming the galaxies and clusters we see today. That process had to have been helped along by cold dark matter, which is made up of huge clouds of particles that are detectable only through their gravitational effects. The cosmological constant (lambda) is a strange form of antigravity responsible for the present speedup of the cosmic expansion [see "A Cosmic Conundrum," by Lawrence M. Krauss and Michael S. Turner; SCIENTIFIC AMERICAN, September 2004].

Overview/*Heavenly Discord*

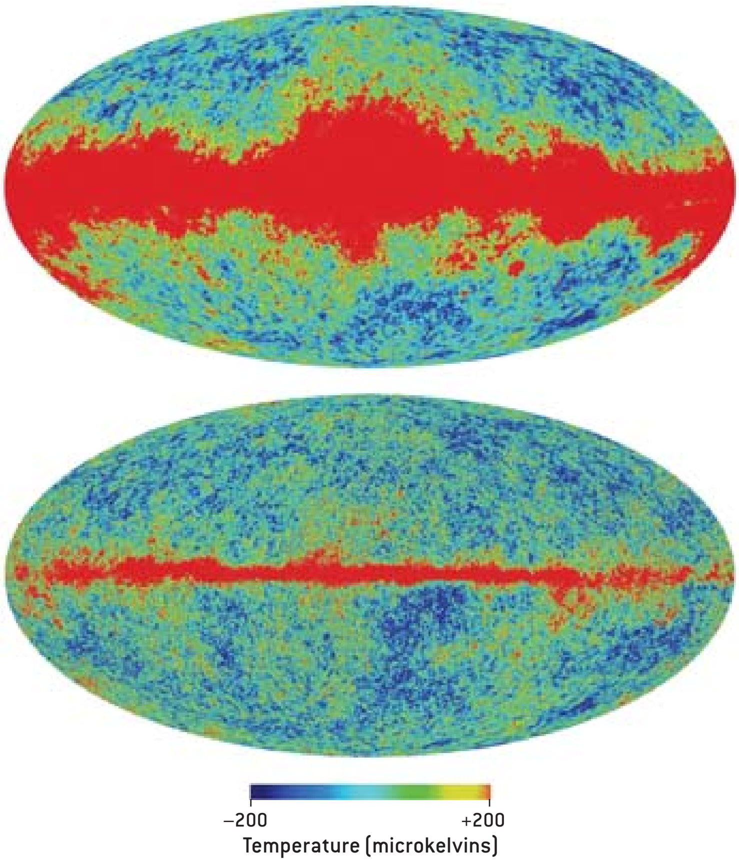
- A theory known as the inflationary lambda cold dark matter model explains many properties of the universe very well. When certain data are analyzed, however, a few key discrepancies arise.
- The puzzling data come from studies of the cosmic microwave background (CMB) radiation. Astronomers divide the CMB's fluctuations into "modes," similar to splitting an orchestra into individual instruments. By that analogy, the bass and tuba are out of step, playing the wrong tune at an unusually low volume.
- The data may be contaminated, such as by gas in the outer reaches of the solar system, but even so, the otherwise highly successful model of inflation is seriously challenged.

The Most Ancient Light

DESPITE THE MODEL'S great success at explaining all those features of the universe, problems show up when astronomers measure the CMB's temperature fluctuations. The CMB is cosmologists' most important probe of the largest-scale properties of the universe. It is the most ancient of all light, originating only a few hundred thousand years after the big bang, when the rapidly expanding and cooling universe made the transition from dense opaque plasma to transparent gas. In transit for 14 billion years, the CMB thus reveals a picture of the early universe. Coming from the farthest reaches, that picture is also a snapshot of the universe at its largest size scale.

Arno Penzias and Robert Wilson of Bell Laboratories first detected the CMB and measured its temperature in 1965. More recently, the cutting edge of research has been studies of fluctuations in the temperature as seen when viewing different areas of the sky. (Technically, these fluctuations are called temperature anisotropies.) The differences in temperature across the sky reflect the universe's early density fluctuations. In 1992 the COBE (Cosmic Background Explorer) sat-

JACK UNRUH (*preceding pages*)



MICROWAVE SKY is measured in the K-band (23 gigahertz, *top*), the W-band (94 gigahertz, *bottom*) and three other bands (*not shown*) by the WMAP satellite. The entire sphere of the sky is projected onto the oval shape, like a map of the earth. The horizontal red band is radiation from the Milky Way. Such "foreground" radiation changes with wave band, allowing it to be identified and subtracted from the data, whereas the cosmic microwave background does not.

ellite first observed those fluctuations; later, the WMAP (Wilkinson Microwave Anisotropy Probe) satellite has made high-resolution maps of them.

Models such as the lambda cold dark matter model cannot calculate the exact pattern of the fluctuations. Yet they can predict their statistical properties, similar to predicting their average size and the range of sizes they span. Some of these statistical features are predicted not only by the lambda cold dark matter model but also by numerous other simple inflationary models that physicists have considered at one time or another as possible alternatives. Because such properties arise in many different inflationary models, they are considered "generic" predictions of inflation; if inflation is true at all, these predictions hold irrespective of the finer details of the model. To falsify one of them would be to challenge the scenario of inflation in the most serious way a scientific theory can be challenged. That is what the anomalous CMB measurements may do.

The predictions are best expressed by first breaking down the temperature fluctuations into a spectrum of modes called spherical harmonics, much as sound can be separated into a spectrum of notes [*see box on page 53*]. As mentioned earlier, we can consider the density fluctuations, before they grow into galaxies, to be sound waves in the universe. If this breakdown into modes seems mysterious, recall the orchestra analogy: each mode is a particular instrument, and the whole map of temperatures across the sphere of the sky is the complete sound produced by the orchestra.

The first of inflation's generic predictions about the fluctuations is "statistical isotropy." That is, the CMB fluctuations neither align with any preexisting preferred directions (for example, the earth's axis) nor themselves collectively define a preferred direction.

Inflation further predicts that the amplitude of each of the modes (the volume at which each instrument is playing, if we think about an orchestra) is random, from among a range of possibilities. In particular, the distribution of probabilities follows the shape of a bell curve, known as a Gaussian. The most likely amplitude, the peak of the curve, is at zero, but in general nonzero values occur, with decreasing probability the more the amplitude deviates from zero. Each mode has its own Gaussian curve, and the width of its Gaussian distribution (the wider the base of the "bell") determines how much power (how much sound) is in that mode.

Inflation tells us that the amplitudes of all the modes should have Gaussian distributions of very nearly the same width. This property comes about because inflation, by stretching the universe exponentially, erases, like a pervasive cosmic iron, all traces of any characteristic scales. The resulting power spectrum is called flat because of its lack of distinguishing features. Significant deviations from flatness should occur only in those modes produced at either the end or the beginning of inflation.

Missing Notes

SPHERICAL HARMONICS represent progressively more complicated ways that a sphere can vibrate in and out. As we look closer at the harmonics, we begin to see where the observations run into troubling conflicts with the model. These modes are convenient to use, because all our information about the distant universe is projected onto a single sphere—the sky. The lowest note (labeled l=0) is the monopole—the entire sphere pulses as one. The monopole of the CMB is its average temperature—just 2.725 degrees above absolute zero [*see box on page 53*].

The next lowest note (labeled l=1) is the dipole, in which the temperature goes up in one hemisphere and down in the other. The dipole is dominated by the Doppler shift of the solar system's motion relative to the CMB; the sky appears slightly hotter in the direction the sun is traveling.

In general, the oscillation for each value of l (0, 1, 2 ...) is called a multipole. Any map drawn on a sphere, whether it be the CMB's temperature or the topography of the earth, can be broken down into multipoles. The lowest multipoles are the largest-area, continent- and ocean-size undulations on our temperature map. Higher multipoles are like successively smaller-area plateaus, mountains and hills (and trenches and valleys) inserted in orderly patterns on top of the larger features. The entire complicated topography is the sum of the individual multipoles.

For the CMB, each multipole l has a total intensity, C_l—roughly speaking, the average heights and depths of the mountains and valleys corresponding to that multipole, or the average volume of that instrument in the orchestra. The collection of intensities for all different values of l is called the angular power spectrum, which cosmologists plot as a graph.

The graph begins at C_2 because the real information about cosmic fluctuations begins with l=2. The illustration on page 54 shows both the measured angular power spectrum from WMAP and the prediction from the inflationary lambda cold dark matter model that most closely matches all the measurements. The measured intensities of the two lowest-l multipoles, C_2 and C_3, the so-called quadrupole and octopole, are considerably lower than the predictions. The COBE team first noticed this deficiency in the low-l power, and WMAP recently confirmed the finding. In terms of topography, the largest continents and oceans are mysteriously low and shallow. In terms of music, we are missing bass and tuba.

The effect is even more dramatic if instead of looking at the total intensities (the C_l's) one looks at the so-called angular correlation function, $C(\theta)$. To understand this function, imagine we look at two points in the sky separated by an angle θ and examine whether they are both hotter (or both colder) than average, or one is hotter and one colder. $C(\theta)$ measures the extent to which the two points are correlated in their temperature fluctuations, averaged over all the points in the sky. Experimentally we find that the $C(\theta)$ for our universe is nearly zero at angles greater than about 60 degrees, which means that the fluctuations in directions separated by more than about 60 degrees are completely uncorrelated. This result is another sign that the low notes of the universe that inflation promised are missing.

This lack of large-angle correlations was first revealed by COBE, and WMAP has now confirmed it. The smallness of $C(\theta)$ at large angles means not only that C_2 and C_3 are small but that the ratio of the values of the first few total intensities—up to at least C_4—are also unusual. The absence of large-angle power is in striking disagreement with *all* generic inflationary models.

This mystery has three potential solutions. First, the unusual results may be just a meaningless statistical fluke. In particular, uncertainties in the data may be larger than have been estimated, which would make the observed results less improbable. Second, the correlations may be an observational artifact—an unexpected physical effect that has not been compensated for in the WMAP team's analysis of its data. Finally, they may indicate a deeper problem with the theory.

Several authors have championed the first option. George Efstathiou of the University of Cambridge was first, in 2003, to raise questions about the statistical methods used to extract the quadrupole strength and its uncertainty, and he claimed that the data implied a much larger uncertainty. Since then, many others have looked at the methods by which the WMAP team extracted the low-l C_l and concluded that uncertainties caused by the emissions of our own Milky Way galaxy are larger than what researchers originally inferred.

The absence of large-angle power is in striking disagreement with most inflationary theories.

Mysterious Alignments

TO ASSESS THESE DOUBTS about the significance of the discrepancy, several groups have looked beyond the information contained in the C_l's, which represent the total intensity of a mode. In addition to C_l, each multipole holds directional information. The dipole, for instance, has the direction of the hottest half of the sky. Higher multipoles have even more directional information. If the intensity discrepancy is indeed just a fluke, then the directional information from the same data would be expected to show the correct generic behavior. That does not happen, however.

The first odd result came in 2003, when Angelica de Oliveira-Costa, Max Tegmark, both then at the University of Pennsylvania, Matias Zaldarriaga of Harvard University and Andrew Hamilton of the University of Colorado at Boulder noticed that the preferred axes of the quadrupole modes, on the one hand, and of the octopole modes, on the other, were remarkably closely aligned. These modes are the same ones that seemed to be deficient in power. The generic inflationary model predicts that each of these modes should be completely independent—one would not expect any alignments.

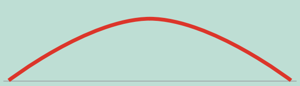
Also in 2003 Hans Kristian Eriksen of the University of Oslo and his co-workers presented more results that hinted at alignments. They divided the sky into all possible pairs of hemispheres and looked at the relative intensity of the fluctuations on the opposite halves of the sky. What they found contradicted the standard inflationary cosmology—the hemispheres often had very different amounts of power. But what was most surprising was that the pair of hemispheres that were the most different were the ones lying above and below the ecliptic, the plane of the earth's orbit around the sun. This result was the first sign that the CMB fluctuations, which were supposed to be cosmological in origin, with some contamination by emission in our own galaxy, have a solar system signal in them—that is, a type of observational artifact.

Detecting Harmonics in the Heavenly Music

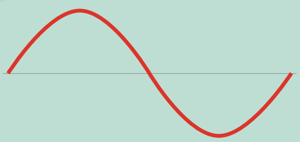
When scientists say that certain instruments in the cosmic microwave background (CMB) seem to be quietly playing off key, what do they mean—and how do they know that?

CMB researchers study fluctuations in temperature measured in all directions in the sky. They analyze the fluctuations in terms of mathematical functions called spherical harmonics. Imagine a violin string. It can sound an infinite number of possible notes, even without a finger pressing it to shorten it. These notes can be labeled *n*, the number of spots (called nodes) on the string other than its ends that do not move when the note is sounded.

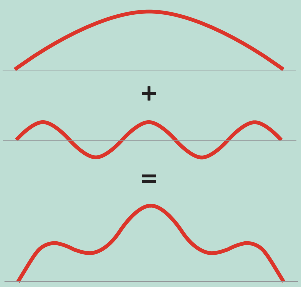
The lowest note, that is, no node ($n=0$), is called the fundamental tone. The entire string, except for the ends, moves back and forth in unison (*below*).



The note with a single node in the middle ($n=1$) is the first harmonic oscillation. In this case, half of the string moves one way while the other half moves the other (*below*). If you sing *do-re-mi-fa-so-la-ti-do,* the final *do* is the first harmonic to the fundamental tone of the first *do*. The note with two equally spaced nodes is the second harmonic, and so on.

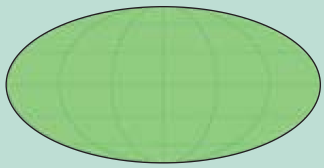


Any complicated way that the string vibrates can be broken down into its component harmonics. For example, we can consider the vibration below as the sum of the fundamental tone ($n=0$) and the fourth harmonic ($n=4$). Note that the fourth harmonic has a lower amplitude (its waves are shallower) in the sum than the fundamental tone. In the orchestra analogy, instrument number four is playing more softly than instrument number zero. In general, the more irregular the vibration of the string, the more harmonics are needed in the sum.

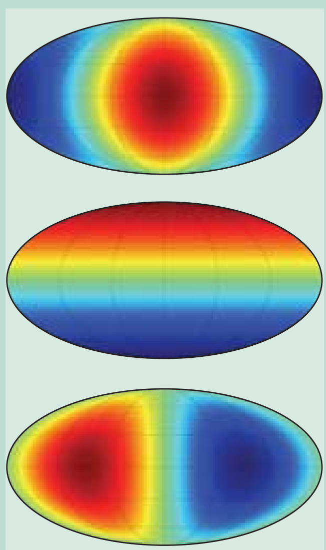


Now let us examine spherical harmonics—denoted Y_{lm}—in which the modes occur around a spherical "drum." Because the surface of the sphere is two-dimensional, we now need two numbers, *l* and *m,* to describe the modes. For each value of *l* (which can be 0, 1, 2, ...), *m* can be any whole number between $-l$ and *l*. The combination of all the different notes with the same value of *l* and different values of *m,* each with its respective amplitude (or in audio terms, the volume), is called a multipole.

We cannot easily draw the spherical harmonics as we drew the violin string. Instead we present a map of the sphere colored according to whether a given region is at a higher or lower temperature than the average. (The map's shape comes from being stretched flat, just like maps of the earth hung in schoolrooms.) The monopole, or $l=0$, is the entire spherical drum pulsing as one (*below*).

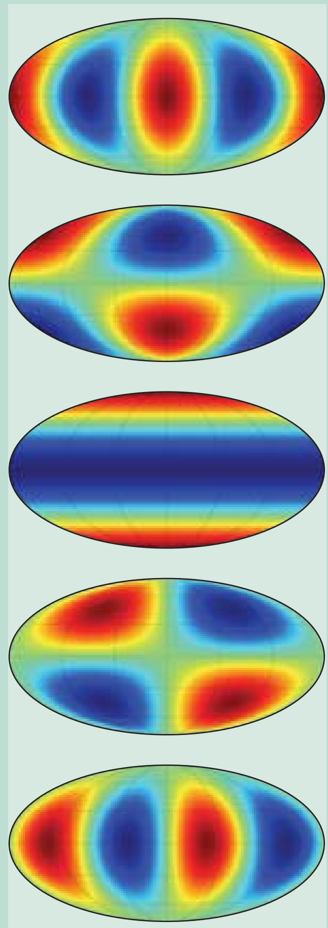


The dipole ($l=1$) has half the drum pulsing outward (*red*) and half pulsing in (*blue*). There are three dipole modes ($m=-1$, 0, 1) in the three perpendicular directions of space (in and out of the page, up and down, and left and right).



The regions of green color are at the average temperature; these nodal lines are the analogues of nodes on the violin string. As *l* increases, so does the number of nodal lines.

The quadrupole ($l=2$) has five modes, each with a more complicated pattern of oscillations or temperature variations on the sphere (*below*).

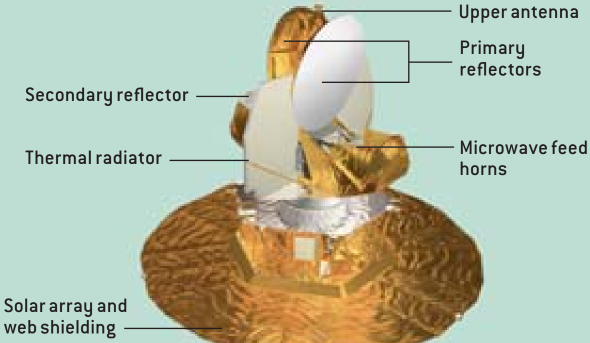


We can break down any pattern of temperature distributions on a spherical surface into a sum of these spherical harmonics, just as any vibration of the violin string can be broken down into a sum of harmonic oscillations. In the sum, each spherical harmonic has a particular amplitude, in essence representing the amount of that harmonic that is present or how loudly that cosmic "instrument of the orchestra" is playing. *—G.D.S. and D.J.S.*

CRAIG COPI *Case Western Reserve University (spherical harmonics images)*; ALISON KENDALL *(strings)*

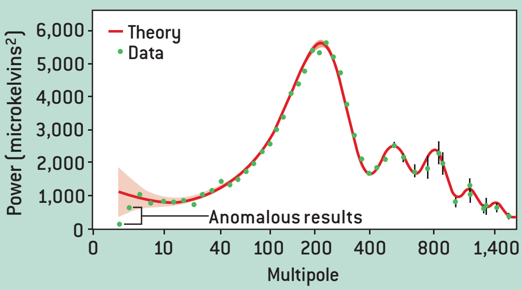
MYSTERIES FROM WMAP

WMAP SATELLITE produces data that are mysterious in three ways.



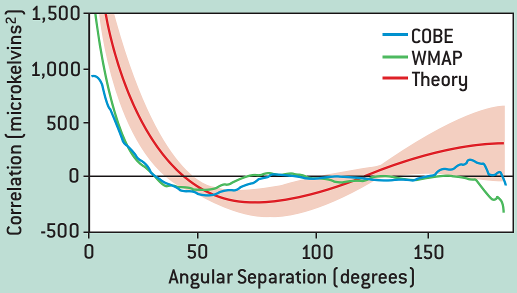
1 ANGULAR POWER SPECTRUM

Most of the WMAP measurements, like those from earlier experiments, are in excellent agreement with values predicted from the inflationary lambda cold dark matter model. But the first two data points (multipoles)—the quadrupole and octopole—are anomalously low in power.



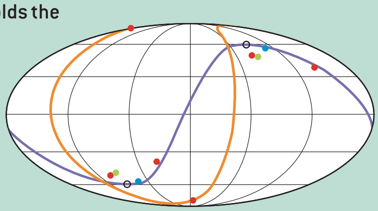
2 ANGULAR CORRELATION FUNCTION

This function relates data from points in the sky separated by a given angle. The data curves from COBE and WMAP should follow the theoretical curve. Instead they are virtually zero beyond about 60 degrees.



3 ALIGNMENT OF THE FIRST TWO MULTIPOLES

The quadrupole (*blue*) and octopole (*red*) should be randomly scattered, but instead they clump close to the equinoxes (*open circles*) and the direction of the solar system's motion (dipole, *green*). They also lie mostly on the ecliptic plane (*purple*). Two are on the supergalactic plane that holds the Milky Way and most of its neighboring galaxies and galactic clusters (*orange*). The probability of these alignments occurring by chance is less than one in 10,000.



Meanwhile one of us (Starkman), together with Craig Copi and Dragan Huterer, then both at Case Western Reserve University, had developed a new way to represent the CMB fluctuations in terms of vectors (a mathematical term for arrows). This alternative allowed us (Schwarz, Starkman, Copi and Huterer) to test the expectation that the fluctuations in the CMB will not single out special directions in the universe. In addition to confirming the results of de Oliveira-Costa and company, we revealed some unexpected correlations in 2004. Several of the vectors lie surprisingly close to the ecliptic plane. Within that plane, they sit unexpectedly close to the equinoxes—the two points on the sky where the projection of the earth's equator onto the sky crosses the ecliptic. These same vectors also happen to be suspiciously close to the direction of the sun's motion through the universe. Another vector lies very near the plane defined by the local supercluster of galaxies, termed the supergalactic plane.

Each of these correlations has less than a one in 300 chance of happening by accident, even using conservative statistical estimates. Although they are not completely independent of one another, their combined chance probability is certainly less than one in 10,000, and that reckoning does not include all the odd properties of the low multipoles.

Some researchers have expressed concern that all these results have been derived from maps of the full CMB sky. Using the full-sky map might seem like an advantage, but in a band around the sky centered on our own galaxy the reported CMB temperatures may be unreliable. To infer the CMB temperature in this galactic band, one must first strip away the contributions of the galaxy. Perhaps the techniques that the WMAP team or other groups have used to remove the galactic thumbprints are not reliable enough. Indeed, the WMAP team cautions other researchers against using its full-sky map; for its own analysis, it uses only those parts of the sky outside the galaxy. When Uros Seljak of Princeton University and Anze Slosar of the University of Ljubljana excluded the galactic band, they found that the statistical significance of some of these alignments declined at some wavelengths. Yet they also found that the correlations increased at other wavelengths. Our own follow-up work suggests that the effects of the galaxy cannot explain the observed correlations. Indeed, it would be very surprising if a misunderstanding of the galaxy caused the CMB to be aligned with the solar system.

The case for these connections between the microwave

THE AUTHORS

GLENN D. STARKMAN and *DOMINIK J. SCHWARZ* first worked together in 2003, when they were at CERN near Geneva. Starkman is Armington Professor at the Center for Education and Research in Cosmology and Astrophysics in the departments of physics and astronomy at Case Western Reserve University. Schwarz has done research on cosmology since he graduated from the Vienna University of Technology in Austria. He recently accepted a faculty position at the University of Bielefeld in Germany. His main scientific interests are the substance of the universe and its early moments.

NASA/WMAP SCIENCE TEAM (*image*); ALISON KENDALL (*graphs*) background and the solar system being real is strengthened when we look more closely at the angular power spectrum. Aside from the lack of power at low *l*, there are three other points—*l*=22, *l*=40 and *l*=210—at which the observed power spectrum differs significantly from the spectrum predicted by the best-fit lambda cold dark matter model. Whereas this set of differences has been widely noticed, what has escaped most cosmologists' attention is that these three deviations are correlated with the ecliptic, too.

Two explanations stand out as the most likely for the correlation between the low-*l* CMB signal and features of the solar system. The first is an error in the construction or understanding of the WMAP instruments or in the analysis of the WMAP data (so-called systematics). Yet the WMAP team has been exceedingly careful and has done numerous cross-checks of its instruments and its analysis procedure. It is difficult to see how spurious correlations could accidentally be introduced. Moreover, we have found similar correlations in the map produced by the COBE satellite, which used different instruments and analysis and so would have had mostly independent systematics.

The results could send us back to the drawing board about the early universe.

A more probable explanation is that an unexpected source or absorber of microwave photons is contaminating the data. This new source should somehow be associated with the solar system. Perhaps it is some unknown cloud of dust on the outskirts of our solar system. But this explanation is itself not without problems: How does one get a solar system source to glow at approximately the wavelength of the CMB brightly enough to be seen by CMB instruments, or to absorb at CMB wavelengths, yet remain sufficiently invisible in all other wavelengths not to have yet been discovered? We hope we will be able eventually to study such a foreground source well enough to decontaminate the CMB data.

Back to the Drawing Board?

AT FIRST GLANCE, the discovery of a solar system contaminant in the CMB data might appear to solve the conundrum of weak large-scale fluctuations. Actually, however, it makes the problem even worse. When we remove the part that comes from the hypothetical foreground, the remaining cosmological contribution is likely to be even smaller than previously believed. (Any other conclusion would require an accidental cancellation between the cosmic contribution and our supposed foreground source.) It would then be harder to claim that the absence of low *l* power is just a statistical accident. It looks like inflation is getting into a major jam.

A statistically robust conclusion that less power than expected exists on large scales could send us back to the drawing board about the early universe. The current alternatives to generic inflation are not terribly attractive: a carefully designed inflationary model could produce a glitch in the power spectrum at just the right scale to give us the observed absence of large-scale power, but this "designer inflation" stretches the limits of what we look for in a compelling scientific theory—an exercise akin to Ptolemy's addition of hypothetical epicycles to the orbits of heavenly bodies so that they would conform to an Earth-centered cosmology.

One possibility is that the universe has an unexpectedly complex cosmic topology [see "Is Space Finite?" by Jean-Pierre Luminet, Glenn D. Starkman and Jeffrey R. Weeks; SCIENTIFIC AMERICAN, April 1999]. If the universe is finite and wrapped around itself in interesting ways, like a doughnut or pretzel, then the vibrational modes it allows will be modified in very distinctive ways. We might be able to hear the shape of the universe, much as one can hear the difference between, say, church bells and wind chimes. For this purpose, the lowest notes—the largest-scale fluctuations—are the ones that would most clearly echo the shape (and the size) of the universe. The universe could have an interesting topology but have been inflated precisely enough to take that topology just over the horizon, making it not just hard to see but very difficult to test.

Is there hope to resolve these questions? Yes, we expect more data from the WMAP satellite, not only on the temperature fluctuations of the sky but also on the polarization of the received light, which may help reveal foreground sources. In 2007 the European Space Agency will launch the Planck mission, which will measure the CMB at more frequency bands and at higher angular resolution than WMAP did. The higher angular resolution is not expected to help solve the low-*l* puzzle, but observing the sky in many more microwave "colors" will give us much better control over systematics and foregrounds. Cosmological research continues to bring surprises—stay tuned. SA

MORE TO EXPLORE

First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. C. L. Bennett et al. in *Astrophysical Journal Supplemental*, Vol. 148, page 1; 2003.

The Cosmic Symphony. Wayne Hu and Martin White in *Scientific American*, Vol. 290, No. 2, pages 44–53; February 2004.

The WMAP Web page is at **http://wmap.gsfc.nasa.gov/**

MORPHWARE

Magnetic logic may usher in an era in which computing devices can change instantly from one type of hardware to another

BY REINHOLD KOCH

Flexibility or performance?

That choice is a constant trade-off for microprocessor designers. General-purpose processors in personal computers execute a broad set of software commands that can cope with any task from graphics to complex calculations. But their flexibility comes at the expense of speed. In contrast, application-specific integrated circuits (ASICs), optimized for a given task, such as the computing required in graphics or sound cards, are very fast but lack adaptability.

Some processors fit a niche between these two types of hardware. Called morphware, they can be reconfigured and optimized for any task. One example—the commercially available field-programmable gate array (FPGA)—consists of large blocks of transistors that perform logic operations and that can be "rewired" by the software. Customization enables FPGAs to accelerate data encryption, automatic military target recognition or data compression by a factor of 10 to 100—enabling, for instance, dramatically enhanced security or faster target acquisition times as compared with a general-purpose CPU (central processing unit).

Overview/*Chameleon Chip*

- The brains of most contemporary computers—the central processing units—are either built to execute a broad set of commands or are optimized for speed.
- A class of processor, known as morphware, offers the best of both worlds. It executes one application quickly and then, based on software commands, changes the processor's wiring to optimize itself for a new application.
- A new type of morphware, called magnetologic, is built from layered metallic materials, each layer of which is magnetized to represent a digital bit. This chameleon of processors can alter its functionality by changing wiring, or gates, many times per second, allowing it to switch rapidly from, say, a cell phone to an MP3 player.
- Because magnetologic stores the results of each processing operation, it can function as both a processor and a memory device.

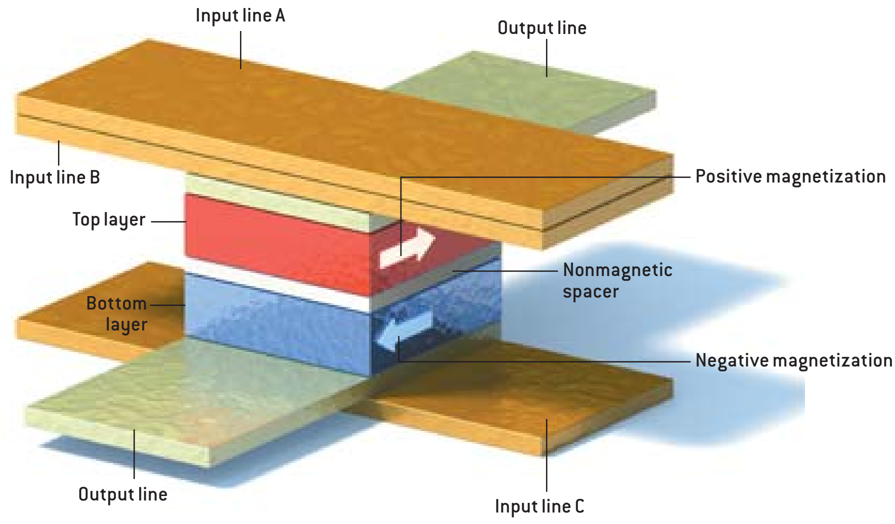
FPGAs rely on the ubiquitous transistor-based technology called complementary metal oxide semiconductor (CMOS). They have limitations, however. Changing operations on the fly—converting, say, a calculation of a matrix of numbers to a parallel-processing computation—requires the relatively slow rewiring of connections between large blocks of transistors, not the individual elements (gates) that perform a processor's logic operations. FPGAs generally take up a large amount of space, resulting in a very low density of circuitry and limiting the number and speed of processing operations.

In the past few years, a number of groups have begun to explore a new type of morphware processor that uses layers of magnetic materials to create reconfigurable logic elements. The advantages of these magnetologic elements are that the information stored does not disappear when external power is shut off and that they do not have to be refreshed while the device is in operation. Unlike CMOS-based systems, the logic is nonvolatile. This stability of magnetic bits explains the key role of magnetic materials in data storage, such as hard disks. In a magnetologic device, nonvolatility of information would also reduce power consumption, and a single element

BRYAN CHRISTIE DESIGN (*preceding pages*)

HOW A BASIC MAGNETOLOGIC GATE WORKS

Two ferromagnetic layers, separated by a nonmagnetic spacer, can switch polarity (*arrows*) to output a digital "1" or "0." Electric currents applied through slablike input lines produce the magnetic field needed for changing magnetization. A current through the two top lines switches the magnetization of the top layer. A current through three input lines shifts both layers. A positive magnetization (*right-pointing arrow*) represents a digital "1," and a negative magnetization (*left-pointing arrow*) a digital "0." After the input currents switch the two layers' parallel or antiparallel states, the result can be read as a bit through the output line.



would be capable of performing different logic functions that typically require multiple transistors.

From Cell Phone to MP3 Player

MAGNETOLOGIC COULD BRING electronic multitasking to a new level, letting a designer create a cell phone that could later morph into a music player, thereby reducing the need for separate microprocessors in electronic equipment. Because the switching speed of magnetologic gates is fast, switching at billions of cycles per second (gigahertz), this chameleon of processors can alter its functionality many times within the space of even one second.

The operation of magnetologic builds on a technology for storing digital bits known as magnetic random-access memory (MRAM), which is now nearing commercialization. Each unit of MRAM consists of two ferromagnetic metallic alloys separated by a nonmagnetic spacer that ensures that the magnetization of one layer does not affect the other and that the polarity (direction of magnetization) can be shifted independently [*see box above*]. The memory element represents the value of a digital bit, which depends on whether the magnetization of the upper and lower layers are aligned in parallel or oppose each other. Lower resistance to the flow of electric current occurs when the magnetization of both layers is in parallel—a state that represents, say, a digital "1." When the polarity of both layers is opposite, the so-called magnetoresistance increases (a "0" state).

To switch the resistance of the MRAM element from low (1) to high (0), or vice versa, an electric current must flow through inputs connected to the memory element. Besides the simple 0 or 1 that it stores in memory, a single MRAM element can be used to represent basic logic functions, such as AND or OR.

Elementary magnetologic gates date back to the early 1960s but were quickly supplanted by silicon microchips. In 2000 William C. Black, Jr., and Bodhisattva Das of Iowa State University published a seminal report on magnetologic based on magnetoresistance. Two years later Siemens Research in Erlangen, Germany, demonstrated experimentally a reconfigurable magnetologic element. Then, in 2003, our group at the Paul Drude Institute in Berlin published a paper that proposed using a simpler implementation for changing the logic states of the various computational elements.

Making a Logic Gate

A MAGNETOLOGIC GATE is very similar to an MRAM cell. It also consists of two magnetic layers separated by a nonmagnetic spacer in which the parallel and antiparallel magnetizations exhibit low and high resistance and provide the logic outputs "1" and "0," respectively. In general, the magnetoresistance of layered systems is significantly higher than that of systems not built in layers, easing the reading and writing of bits. This property is known as giant magnetoresistance or tunneling magnetoresistance, depending on which type of spacer

THE AUTHOR

REINHOLD KOCH is the leader of the nanoacoustics group at the Paul Drude Institute for Solid State Electronics in Berlin. Born in Tirol, Austria, he received his doctorate in chemistry at the University of Innsbruck in 1981. He was granted a Max Kade Fellowship during 1985 and 1986 at the Stanford-NASA Joint Institute for Surface and Microstructure Research in California. From 1988 to 1998 he was assistant teacher in the physics department of the Free University in Berlin and received the Karl Scheel Award in 1994 for his contributions on the evolution of stress in thin films and structural investigations with a scanning tunneling microscope.

BRYAN CHRISTIE DESIGN

Magnetologic stores a digital bit without the need for continuous refreshing by an external current.

material is used. Both effects depend on the electrons' spins (their angular momenta), which are all aligned in the same direction, almost as if the electrons were tiny balls spinning on their axes. These effects are used to "read" the value of a bit.

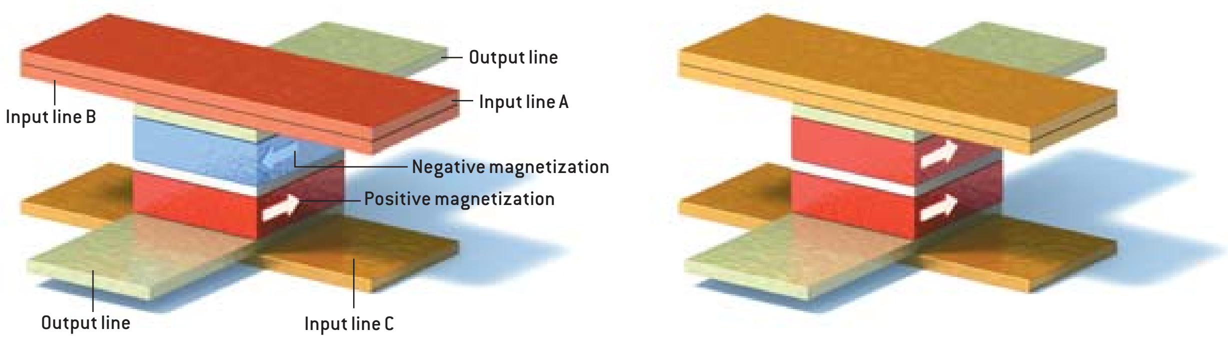
Changing the orientation of spin is used to "write" a bit—in other words, to change the magnetization from one direction to another. The direction of magnetization of either layer can be reversed by the magnetic field of a current flowing through the input lines. But a number of investigators are examining another method, in which spin exerts a torque that can switch a layer's magnetization from one direction to another [see "Spintronics," by David D. Awschalom, Michael E. Flatté and Nitin Samarth; SCIENTIFIC AMERICAN, June 2002].

In the design we put forward at Paul Drude, the magnetologic gate contains three inputs—A, B and C—each of which is addressed by a current of equal magnitude. Our concept makes use of the fact that a magnetoresistive element, though providing only two output values (a 0 and a 1), can be in four different initial states, two of them parallel and two of them antiparallel, allowing the configuration of distinct logic states. Previous magnetologic designs required more complex circuitry that would employ, for example, input currents of different intensity.

In our design, a logic operation begins by setting the gate polarity in one of these four states by addressing two or three of the input lines. Then, in a second step, the logic operation is performed by activating only the upper two input lines, A and B. A chosen initial state can only be reversed when two or three of the input lines are addressed with the same polarity magnetic field, changing the output value from 1 to 0, or the converse [*see box below*]. This process has the advantage that the logic state can be reprogrammed with each new operation.

CHANGING BITS

Logic operations are performed in a magnetologic gate in two steps. An AND gate, for instance, can be created by first setting the device to a 0 state (*left*): the top layer holds a negative magnetization (*left-pointing arrow*); the bottom assumes a positive magnetization (*right-pointing arrow*). Currents that emit a positive magnetic field (*red*) then flow through the two upper input lines, labeled A and B. Magnetization of the top layer becomes positive (*right*). By so doing, the bit changes from a 0 to a 1 and is stored in its new state. The output line is used only to read whether the bit is in a 0 or 1 state.



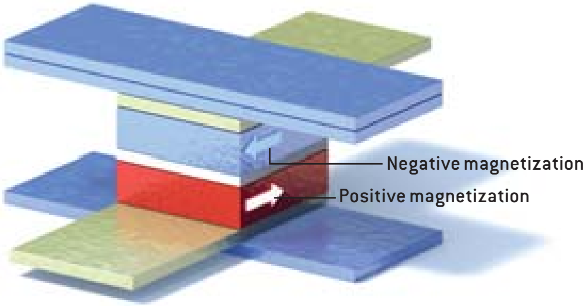
AND GATE

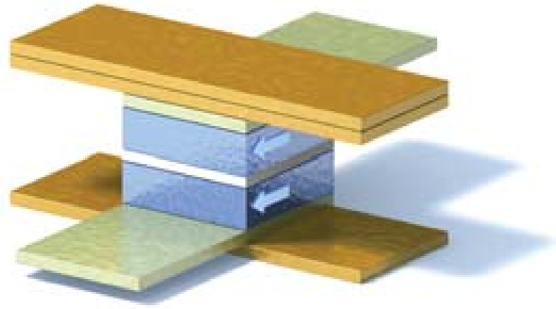
CURRENT	A	B	OUTPUT
⇇	0	0	0
⇆	0	1	0
⇄	1	0	0
⇉	1	1	1

BRYAN CHRISTIE DESIGN

CHANGING GATES

Addressing all three input lines—A, B and C—changes an AND gate to its opposite, a NOT AND, or NAND, gate. Currents that emit a negative magnetic field (*blue*) are applied to all three input lines (*left*). The bottom layer, which exhibits a positive magnetization, then switches to a negative magnetization (*right*). The output of a NAND gate (*purple*) is the opposite of an AND gate (*green*). The C line does not bear on the logic operations shown in the truth tables.





AND GATE

CURRENT	A	B	OUTPUT
⇇	0	0	0
⇆	0	1	0
⇄	1	0	0
⇉	1	1	1

NAND GATE

CURRENT	A	B	OUTPUT
⇇	0	0	1
⇆	0	1	1
⇄	1	0	1
⇉	1	1	0

Because the magnetologic gate maintains its assigned polarity in the absence of an external current, a bit is stored without continuous refreshing and can be read out without deleting the information. Thus, the combined logic and storage capability saves not only energy but also time, compared with information processed by conventional CMOS circuitry.

For the AND function, for example, we start from an antiparallel state with an output of 0. Viewed in cross section, the polarity of the top layer points to the left, whereas the bottom layer points right. Only positive currents that are applied to both inputs A and B—currents that generate a positive magnetic field—can switch the direction of magnetization of the top layer from left to right. The OR gate operates using an analogous method, but the magnetizations of both layers point to the right at the beginning of the procedure. The other two basic logic functions are obtained by switching the bottom layer. All three inputs—A, B and C—are applied to switch the lower layer. The magnetic field needed to switch the polarity of the top layer is less than that for the bottom layer, so the two can be addressed independently. Switching the bottom layer transmutes the output of an AND and OR function into its opposite: NOT AND (NAND) or NOT OR (NOR) [*see box above*].

The OR and AND functions correspond to Boolean addition and multiplication, respectively. Together with NAND and NOR, they represent a powerful basis for describing even the most complex circuits. By changing the procedure of addressing the inputs, magnetoresistive logic gates can produce even more advanced logic functions. For instance, the XOR gate—key to a critical logic unit called a full adder—differentiates between the same and opposite inputs, yielding an output 1 for any two of the same inputs (0/0 or 1/1) and 0 for opposite inputs (0/1 or 1/0). Two magnetoresistive elements can create the XOR gate as compared with eight to 14 transistors in CMOS technology.

The magnetologic gates can also be employed to construct an entire full adder—the most widely used logic unit in a processor. A full adder sums binary inputs A and B plus a carry digit brought forward from a previous calculation. The addition of the three digits produces a new sum as well as a new carry digit. The nonvolatility and the programmability of the magnetologic gates mean that a full adder can be fashioned with only three gates, rather than the 16 transistors with CMOS. The magnetic full adder might become competitive in speed even with the fastest CMOS full adders and boasts superior power efficiency.

Looking Ahead

THE FATE OF MORPHWARE could closely resemble that of commercially announced MRAM cells. The input lines A and B would be arranged in the form of a rectangular grid, a so-called crossbar geometry, similar to that in an MRAM. The magnetoresistive gate elements would sit in the crossing points and be switched only when both input lines are addressed simultaneously. The gates would have to be stacked on top of a template of CMOS transistors that would relay signals indicating when each gate element should begin and stop processing. The transistors in this configuration would also be used to amplify the small currents needed to read a magnetoresistive bit [*see box on page 63*].

The chameleonlike nature of a morphware processor re-

BRYAN CHRISTIE DESIGN

A magnetic chameleon processor constitutes an array of logic gates, each of which is programmable individually by the software.

tains many advantages. Because of the programmability of the logic gates, hardware no longer determines processor capabilities. In CMOS, the logic of a conventional transistor gate is defined by the wiring and is therefore fixed. A magnetic processor constitutes an array of logic gates, each of them programmable individually by the software.

A magnetic chameleon processor therefore needs far fewer logic gates than a conventional processor, in which only a few percent of the hardwired gates are useful for any given task. The programmability also means that newer and better software can easily be implemented, even on older magnetologic processors. Because the switching speed of magnetologic gates is fast, billions of cycles per second, a chameleon processor can alter its functions many times within the course of that second. The nonvolatility of a logic element—the fact that it stores the output of its last operation—also gives the device a benefit in speed. Although magnetologic is fast, its gigahertz switching time is comparable to that of CMOS processors. But nonvolatility means that a clock is not needed to synchronize the extraction of digital bit values from the storage cells in a computer's memory, which simplifies and speeds processing. The bits themselves are stored where they are processed. Unlike CMOS, magnetologic does not necessarily have to reduce component size to increase performance—in other words, it bypasses miniaturization. This advantage may appear increasingly attractive as chip manufacturers struggle to make components ever smaller.

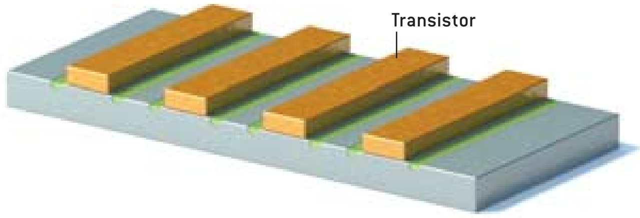
Design of a future chameleon processor is still an academic proposition—for now, no one is considering its development outside of the few laboratories that have published technical

PICK YOUR PROCESSOR

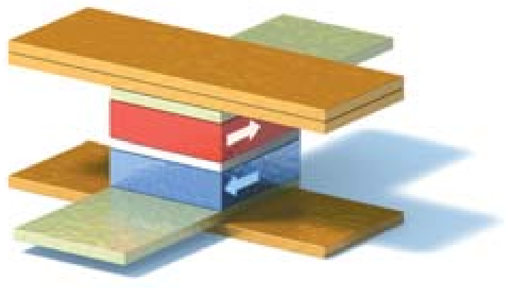
	Programmable	Speed of Reprogramming	Number of Components for Basic Logic	Timing
Magnetologic	Yes	~ 0.1 nanosecond	One magnetologic gate	Executed in parallel without need for a clock to synchronize data transfer
Conventional processor	No	Not applicable	Four transistors	Executed serially; needs a clock, which makes it slower and more complex
Field-programmable gate array	Yes	> 10 nanoseconds	Unable to address individual logic elements	Executed serially; needs a clock

Four transistors are needed (*left*) to perform the functions of a basic AND logic gate, whereas just one magnetologic element can accomplish the same task (*right*).

CONVENTIONAL PROCESSOR (CMOS)



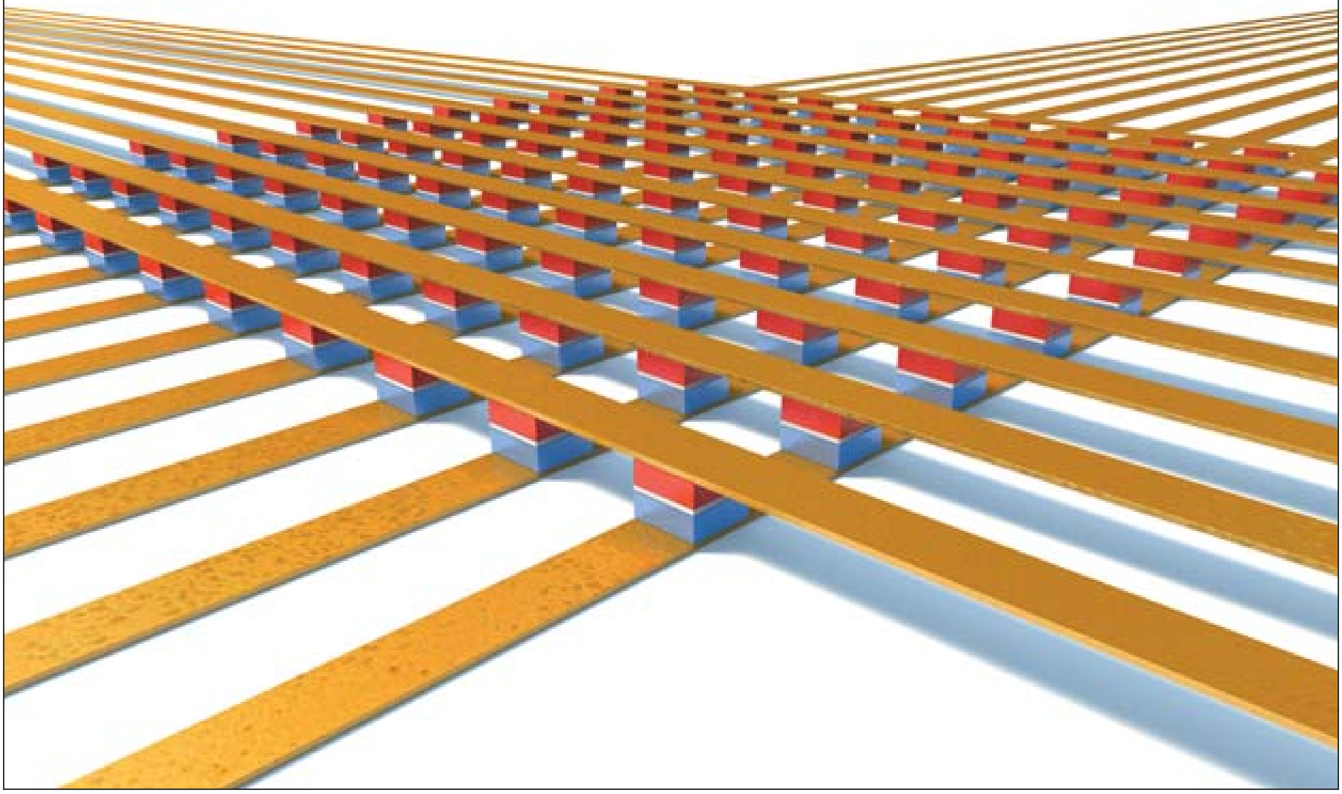
MAGNETOLOGIC



BRYAN CHRISTIE DESIGN

MAGNETOLOGIC PROCESSOR

Arrays of magnetologic gates form the building blocks of a magnetologic processor, which is displayed in an artist's conception of a simplified design. Such a processor would require fewer gates than a conventional one. Transistors (*not shown*) would still be needed to address the various elements—to turn them on and off—and to amplify their small output currents. Software defines the functions that the processor performs at any given time.



papers. Because of its close similarity to MRAM, magnetologic may benefit from engineering work that is addressing problems such as the coupling of magnetic fields between layers in the memories. Similarly, it could suffer if the industry slows development of the technology. Already some companies have hesitated to move ahead with MRAM, estimating that yet another version of random-access memory is unlikely to pull in large revenues. In magnetologic's early implementation, MRAM itself might function as an elementary processor that could be used in early products. But because only one magnetic layer is switched in MRAM, only two programmable functions could be accessed, either AND/OR or NAND/NOR.

To achieve the full potential of a magnetic chameleon processor, many challenging, but ultimately solvable, problems must be surmounted: First, both magnetic layers need to be switched independently, which is still difficult to do in a real working gate. Also, because the processor is working to full capacity most of the time, it generates pockets of heat locally that could compromise the integrity of the data. So reliability requirements for reading and writing operations are much higher. Engineers must show that magnetologic gates can achieve a lifetime as high as 10^{16} to 10^{17} operations, requiring longevity improvements of two or three orders of magnitude. In the meantime, one mitigating factor is that defective gates can be detected and bypassed when a computer boots up. To optimize magnetologic, new magnetic compounds are needed that are compatible with semiconductors and exhibit a giant magnetoresistance [see "Magnetic Field Nanosensors," by Stuart A. Solin; SCIENTIFIC AMERICAN, July 2004].

Perhaps one of the most imposing hurdles is to develop a compiler language and new algorithms that take full advantage of the real-time reprogrammability of the logic gates. To bring a magnetic chameleon processor to market will require an interdisciplinary research effort that uses the combined skills of specialists in materials science and technology, hardware design and electronics, computer sciences, and mathematics. SA

BRYAN CHRISTIE DESIGN

MORE TO EXPLORE

Programmable Logic Using Giant-Magnetoresistance and Spin-Dependent Tunneling Devices. W. C. Black, Jr., and B. J. Das in *Journal of Applied Physics,* Vol. 87, No. 9; pages 6674–6679; May 1, 2000.

Reconfigurable Computing: A Survey of Systems and Software. Katherine Compton and Scott Hauck in *ACM Computing Surveys,* Vol. 34, No. 2, pages 171–210; 2002.

The Future of Universal Memory. Russell P. Cowburn in *Materials Today,* Vol. 6, No. 7–8, pages 32–38; July–August 2003.

Programmable Computing with a Single Magnetoresistive Element. A. Ney et al. in *Nature,* Vol. 425, pages 485–487; October 2, 2003.

Beating a Sudden Killer

When a young woman nearly died from a ruptured aneurysm, the author and the woman's husband began searching for ways to save other aneurysm patients from catastrophe

BY JOHN A. ELEFTERIADES

It was the first beautiful Saturday in spring, and I was in charge of my children. We were out browsing the stores around our local green when the page came in. Lawrence Cohen, a preeminent cardiologist and my colleague at Yale University, was distraught. Normally a man of few words, Cohen was speaking quickly, almost feverishly. "I need you, John. In the ER. Right away. She's dying, John. She's dying right in front of me."

The situation was particularly distressing because Cohen had been following the woman's condition for three years, ever since her husband had come to teach at Yale. So Carmela Kolman was like a member of the family. She was 32 years old and had Marfan's syndrome, a connective tissue disorder that tends to produce thoracic aortic aneurysms: dangerous swellings in the upper part of the large artery that carries blood from the heart, down through the chest and into the abdomen. Left untreated, these aneurysms can grow until they rupture, an event that is often fatal. The only intervention is a preemptive operation to replace the damaged regions with artificial components. But the surgery has its own risks, so physicians hold off on making that call until it seems absolutely necessary. Because Carmela's aorta had been only modestly enlarged, Cohen had not recommended surgery.

Yet this Saturday morning Carmela had come to the emergency room complaining of severe chest pain. A computed tomographic (CT) scan and an echocardiogram showed an aortic dissection: blood had seeped through a tear in the inner part of the arterial wall, causing the inner half of the wall to separate from the outer half, down the entire length of the vessel. Dissection alone can be deadly, because it can result in blocked or diverted blood flow, robbing the heart and other organs of essential oxygen and nutrients. But that was not the worst of the story. The scans indicated that Carmela had blood in her pericardium, the sac that sur-

ANEURYSMS, or bulges, in the aorta pose a silent, but potentially deadly, threat to those who harbor them.

rounds the heart. So the dissection had ruptured. She was drifting in and out of consciousness, her blood pressure was falling, and she was in shock. She needed surgery immediately.

I left the children with a neighbor and rushed to the hospital. There my surgical team and I replaced the weakened part of Carmela's aorta with an artificial vessel made of Dacron, a fabric that is woven into a flexible but sturdy tube. We also traded her damaged aortic valve, which controls the flow of blood as it exits the heart, for a mechanical version. After the surgery, Carmela was very sick. But she clung to life and improved steadily.

Each night on my evening rounds, I spoke with Carmela's husband, John Rizzo, about her condition. As she got better, we found our conversations turning to more scientific topics, particularly issues relating to aortic disease.

Rizzo, it turned out, was an economist working in the epidemiology division of the school of public health and was an expert in data analysis and management. He took a keen interest in my group's work, and in the decade since Carmela's trip to the hospital, Rizzo has helped us compile a database containing the records of all our patients with thoracic aortic aneurysms. As a result, my colleagues and I now have computerized information on more than 3,000 patients with this condition, including some 9,000 images and 9,000 patient-years of follow-up (when our work with all these patients is totaled). We know of no larger organized database on this disorder.

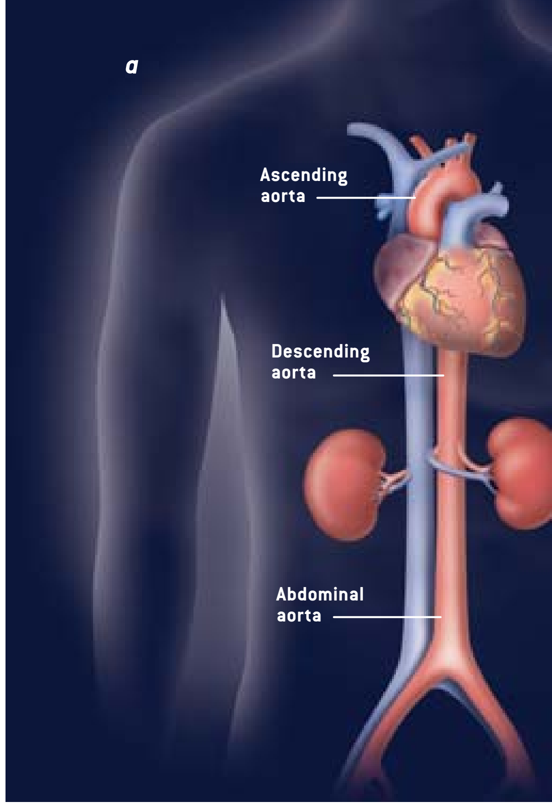
This extensive clinical resource has allowed us to learn more about the behavior of thoracic aortic aneurysms: notably, how fast they grow, how to tell when they are likely to become critical and who is most susceptible to them. These insights have helped physicians decide when to intervene so as to avert the kind of catastrophic event that brought Carmela to the ER that Saturday morning.

Silent Stalker

AS A CARDIAC SURGEON, I focus on disorders that can damage the heart, such as thoracic aneurysms. But aneurysms can arise in any artery. A good number occur in the lower, or abdominal, aorta: the section that runs from the diaphragm to the area above the pelvis, where the artery branches to carry blood to the legs. Research by other investigators has revealed that the mechanisms underlying the growth, dissection and rupture of abdominal aortic aneurysms are similar to those that control the behavior of aneurysms in the chest.

Aneurysms that strike the aorta are the most life-threatening. Every year more than 15,000 people in the U.S. die when an aneurysm in the chest or abdomen bursts or dissects—more people than die from AIDS. Albert Einstein, Olympic volleyball star Flo Hyman, Florida State University basketball player Ronalda Pierce, and actors Lucille Ball, George C. Scott and John Ritter were all killed by thoracic aortic aneurysms. Individuals with Marfan's syndrome are especially susceptible. Medical historians have suggested that Abraham Lincoln might have had this disorder, a condition that, before surgery became feasible, killed most victims by

The normal aorta (*a*) is shaped like a candy cane and is roughly as wide as a garden hose. An aneurysm, which can occur anywhere along the tube, is a ballooning of the wall (*b*), which thins



middle age. Thus, it is possible that our 16th president would have died early even if he had not been assassinated.

Aortic aneurysms are insidious because they are silent stalkers. The vessel can balloon without causing pain. Indeed, most people discover their aneurysms while being tested for something else: a physician will spot the telltale bulge while performing an ultrasound to investigate a heart murmur or a CT scan to evaluate a chronic cough. Pain most often occurs only when an aneurysm ruptures or dissects. And it is severe: the knifelike tearing sensation that accompanies the crisis is described by patients as being orders of magnitude worse than the agony of childbirth or kidney stones.

Survival after such an event tends to be poor. Ruptures usually kill instantly. In some fortunate cases, however, neighboring tissues can press up against the rent in the aorta and hold the structure

TAMI TOLPA (*preceding pages*)

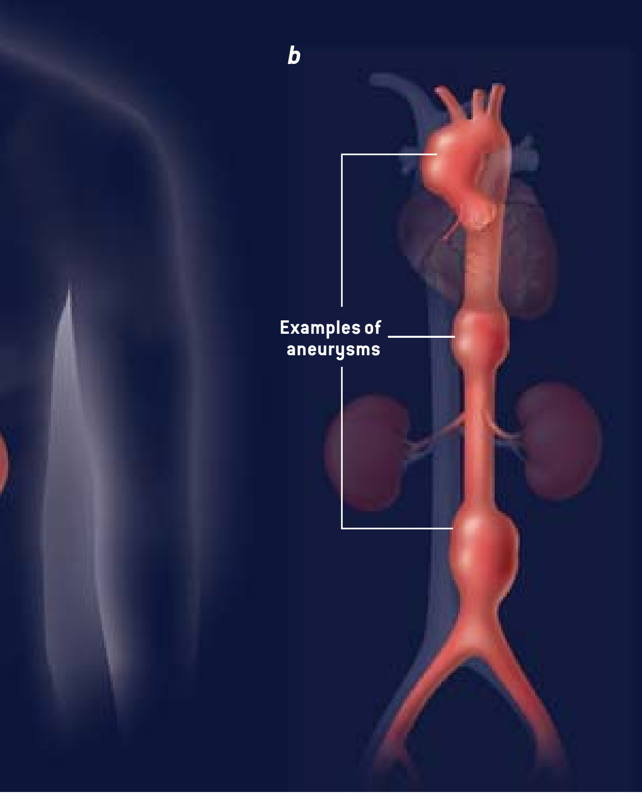
Overview/*Aneurysms*

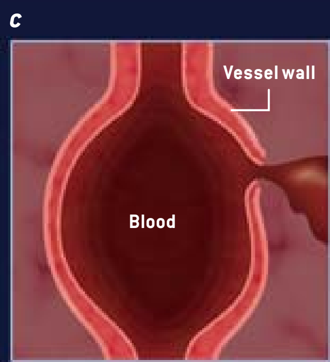
- Clinicians have classically been unsure when to operate on an aortic aneurysm—a swelling in the large artery that carries blood from the heart. Left alone, an aneurysm could fatally rupture or dissect, the inner layer peeling away from the vessel wall. But the only proven intervention, replacement of the damaged aorta with artificial parts, is itself risky.
- Detailed analyses of thousands of patients have now led to guidelines for the best time to operate.
- Weight lifters with aneurysms are in particular danger of sudden death while exercising and should take special precautions.

THE PERILS OF ANEURYSMS

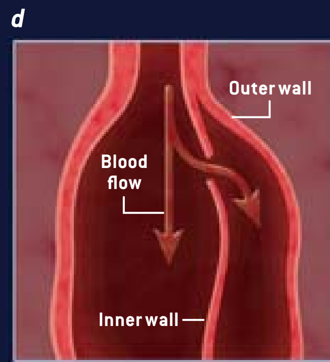
and weakens as it enlarges. The condition can be fatal if the tissue ruptures (*c*) or dissects (*d* and *photograph*), or both. Dissection, the separation of the inner and outer parts of the vessel wall, results when blood seeps into the middle of the wall through a tear of the inner lining.

By analyzing thousands of cases, the author and his colleagues have learned how to predict when an aneurysm is highly likely to rupture or dissect. Such information can help determine when the need for corrective surgery outweighs the substantial risks of the procedure.





RUPTURE



DISSECTION


Cross-sectional view of dissected aneurysm

together long enough for the patient to get to the hospital. For dissections, survival depends on the location. If left untreated, those that begin in the ascending aorta—the segment that emerges from the heart—are fatal within hours or days. Tears in this area can unseat the aortic valve, causing shock, or occlude the coronary arteries, triggering a heart attack. Dissections in the descending aorta, which runs down the back of the chest, are not as threatening: they rupture less frequently than those in the ascending aorta and do not share the same complications.

Surgery can prevent rupture or dissection, but the operation to replace the aorta is very serious and as invasive as procedures get. The operation involves stopping the heart and shunting the blood through a heart-lung machine. In some cases, depending on the location of the aneurysm, surgeons must shut down blood flow entirely and cool the patient from 37 to 18 degrees Celsius to slow metabolism and prevent brain damage while they repair the aorta. Although most people do very well after the surgery, the operation carries risks of stroke, paralysis and death.

To assess whether such a dangerous intervention is warranted, a physician must know how likely it is that an aortic aneurysm will rupture or dissect. In general, a large aneurysm is more dangerous than a small one. But specific data were sorely lacking when Carmela fell ill. Although more than 300 papers had been written on how to operate on the aorta, we could find precious little information on how aortic aneurysms behave before surgery, namely, how fast they expand and how likely they are to burst or tear at any given size. Carmela's aorta, for example, had dissected at 4.8 centimeters in diameter, a relatively modest size—which was why the event was so unexpected. (The normal thoracic aorta is typically about 2.5 to 3.5 centimeters.) Thus, questions regarding aneurysm growth and stability, we reasoned, were a good place to begin our investigations.

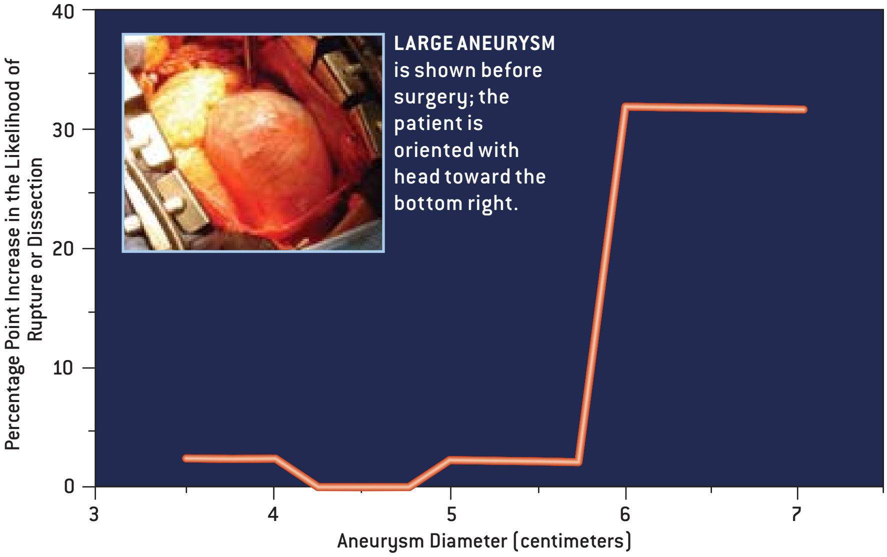
A Threshold Emerges

TO HELP DRAW this information from our clinical database, Rizzo first developed sophisticated statistical techniques that allowed us to determine accurately the growth rate of aneurysms. We found that most grow inexorably and surprisingly slowly: only about 0.12 centimeter per year. Thus, an aneurysm will generally take a decade to grow one centimeter. This finding suggests that aneurysms detected in middle-aged adults probably began growing when the patients were young adults or earlier.

A statistical method devised by Rizzo also permitted us to assess the probabil-

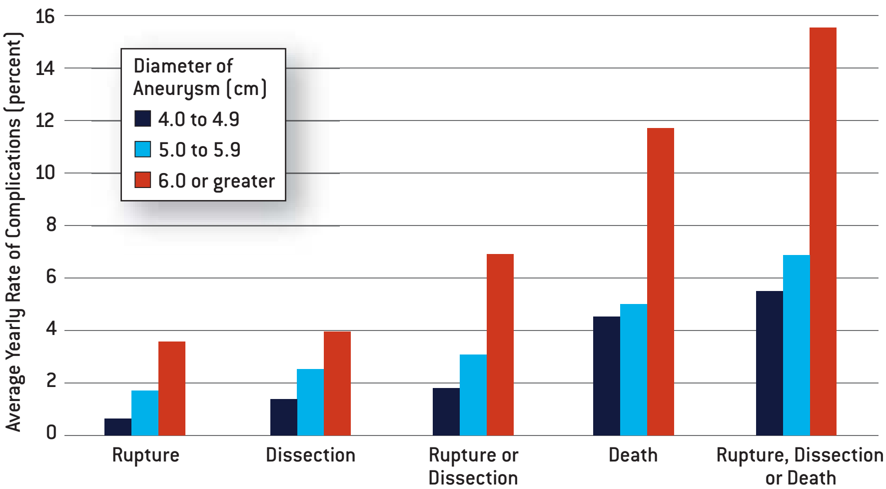
SARA CHEN (*illustration*); CNRI/PHOTO RESEARCHERS, INC. (*photograph*)

RELATION OF SIZE TO RISK OF COMPLICATIONS



LARGE ANEURYSM is shown before surgery; the patient is oriented with head toward the bottom right.

YEARLY RISK OF COMPLICATIONS



PROBABILITIES that rupture or dissection will occur have been calculated for aortic aneurysms in the chest. In one study, the author and his colleagues plotted risk relative to that posed by small aneurysms of 4.0 to 4.9 centimeters. They found a dramatic surge in danger when aneurysms reach six centimeters in the ascending aorta (*top graph*) or seven centimeters in the descending aorta (*not shown*). Another study (*bottom graph*) revealed that the likelihood of rupture, dissection or death within the coming year also jumps sharply for aneurysms that reach six centimeters or higher. (The rates indicated for "Rupture or Dissection" and for "Rupture, Dissection or Death" are lower than the sum of the rates in individual categories because patients with multiple complications were counted only once in the combined categories.) Based on such information, the investigators have determined that many patients with aneurysms in the ascending aorta need corrective surgery when the artery balloons to 5.5 centimeters.

ity of rupture or dissection for thoracic aneurysms of different sizes. We were amazed by the definitiveness of the results. According to our data, the probability of rupture or dissection skyrockets when an aneurysm in the ascending aorta reaches a diameter of about six centimeters, roughly that of a soft drink can. More than 30 percent of patients whose aneurysms reached that size experienced a devastating complication, either rupture or dissection. In the descending aorta, the risk shoots up most dramatically at about seven centimeters.

These numbers represent a lifetime risk of complication: the chance that an aneurysm of a given size will rupture or dissect, although the figures do not indicate when the crisis will happen. But patients who discover that they have an aneurysm are more interested in numbers that predict the yearly rate of complication: in other words, whether their aneurysm might harm them in the near future.

Determining such probabilities requires a large number of cases, and we have recently amassed enough data to begin the appropriate statistical analyses; this data set combines information from patients with aneurysms anywhere in the thoracic aorta, although about two thirds of the patients were affected in the ascending region. We see a trend of gradual increase in the probability of adverse events within the coming year as the aneurysm grows from 4.0 to 5.9 centimeters and then a sharp jump in risk once the aorta reaches six centimeters [*see bottom graph in illustration at left*]. For instance, we find that for a thoracic aneurysm of six centimeters or greater, the risk of rupture, dissection or death within a year soars to a staggering 15.6 percent. Many forms of cancer do not carry as great an annual probability of mortality.

Based on these observations, we recommend that aneurysms in the ascending aorta be surgically removed well before the defect grows to six centimeters. For most people with no family history of aneurysms, we suggest operating at 5.5 centimeters. For the descending aorta, we might perform surgery at six centimeters if a patient is healthy enough to withstand it, but we sometimes delay until about 6.5 centimeters if the patient is frail. We operate at smaller sizes than those listed above for patients with Marfan's or a family history of aneurysm-related disorders, as aneurysms in these people tend to become life-threatening earlier. Using these criteria, we believe, should prevent the vast majority of ruptures and dissections without exposing patients unduly or prematurely to the dangers of aortic surgery. Before an aneurysm warrants going under the knife, doctors may attempt to protect the aorta with medicines that control blood pressure and slow the heart, to limit the stress that is exerted on the ballooned wall.

The abdominal aorta is normally

ALISON KENDALL (*graphs*); JOHN A. ELEFTERIADES (*photograph*); LINE GRAPH SOURCE: M. A. COADY ET AL. IN *JOURNAL OF THORACIC CARDIOVASCULAR SURGERY*, VOL. 113, NO. 3, PAGES 476–491; MARCH 1997; BAR GRAPH SOURCE: R. R. DAVIES ET AL. IN *ANNALS OF THORACIC SURGERY*, VOL. 73, NO. 1, PAGES 17–27; JANUARY 2002.

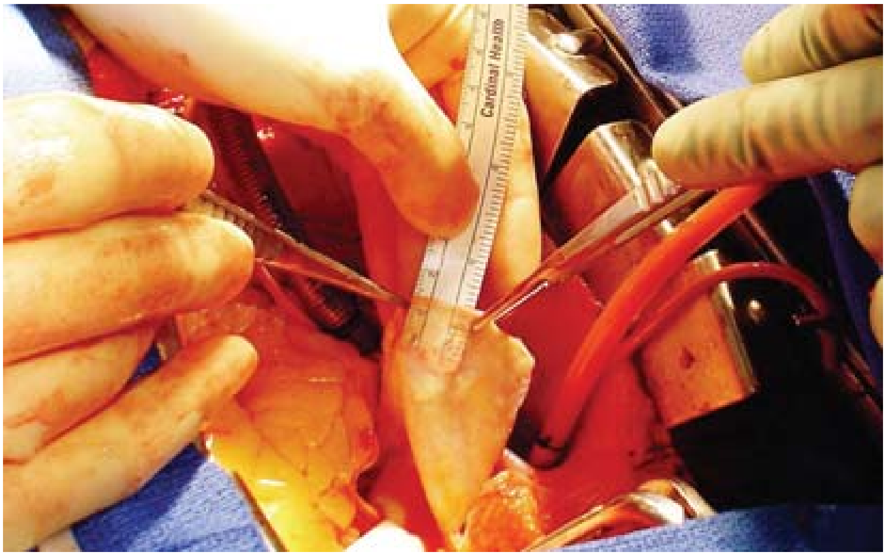
smaller than the thoracic aorta, and rupture of the abdominal aorta usually occurs at smaller sizes than in the thorax. Accordingly, physicians usually intervene surgically at smaller sizes for abdominal aortic aneurysms. Some authorities recommend intervention at four centimeters for women and five centimeters for men, as rough general guidelines.

All in the Family

TO SAVE MORE LIVES, doctors would benefit from knowing which asymptomatic individuals are at risk for aneurysms, so that the condition could be detected early, monitored closely and treated promptly. Marfan's syndrome is a well-known warning; many with the condition wind up having aortic aneurysms. But people with Marfan's account for only 5 percent of all aneurysm patients. The remaining 95 percent of cases are idiopathic—their cause is not yet known.

Physicians once believed that aneurysms were caused by atherosclerosis—the accumulation of plaque (fatty gunk) in the arterial wall. But we have found that patients with aneurysms in the ascending aorta are actually less susceptible to atherosclerosis than the general population, so plaque deposition probably is not causal in their cases. Aneurysms in the descending and abdominal areas, on the other hand, are often accompanied by plaque throughout the aorta and in its branches, which suggests that atherosclerosis probably does contribute to those aneurysms.

Our database has revealed that most thoracic aneurysms have a strong genetic component of some kind—and the same appears to be true for aneurysms in the abdominal aorta and in the brain. Reviewing family histories, we were astounded to discover how often people with aneurysms report having a relative with one or a family member who died suddenly or unexpectedly at a young age. The latter occurrence is often chalked up to cardiac arrest, but in many instances an autopsy would have revealed a ruptured aneurysm. In the 500 families whose pedigrees we have analyzed, approximately 20 percent display some history of aneurysm. In most families, the trait appears to be dominant—in other words, an individual need only inherit an "aneurysm gene" from one parent to be affected; in one of these families, the father passed on aortic disease to all four of his children. Other families showed different patterns of inheritance, suggesting that more than one gene can play a role in susceptibility.



WALL OF AORTA became so thin in a six-centimeter aneurysm that a ruler placed behind it can be seen through the tissue. New findings indicate that aneurysms are caused in part by excessive activity of enzymes known as metalloproteinases (MMPs), which digest proteins that are needed for the elasticity of the artery wall.

JOHN A. ELEFTERIADES

If genetic markers that signify increased susceptibility could be identified, physicians might someday use a simple blood test to pinpoint those who need close monitoring—say, by CT scans or echocardiograms—to catch aneurysms early and determine the best time for surgery. And if the actual genes at fault could be found, researchers might even be able to develop therapies that specifically counteract their ill effects—potentially slowing or preventing the growth of aneurysms by blocking the undesirable activities of the proteins encoded by those genes.

With better detection and, ultimately, improved treatment in mind, we have begun to collaborate with scientists at Celera Diagnostics in Alameda, Calif., to search for genetic markers called SNPs—single nucleotide polymorphisms—that correlate with aortic disease. SNPs are DNA sequences that differ by a single nucleotide, or code letter, between one part of a population and another. James Devlin, Olga Iakoubova and their Celera team are comparing DNA samples obtained from 500 of our patients with thoracic aneurysms and from 500 healthy individuals, in this case the patients' spouses. Then, using automated equipment, they will scan some 16,000 genetic regions for SNPs that appear more often in patients than in the healthy controls.

Our preliminary work has revealed a number of SNPs that might signify increased risk, and we are pursuing these leads in our large patient group. In addition, we are conducting a similar

THE AUTHOR

JOHN A. ELEFTERIADES attended Yale University as an undergraduate and never left. He graduated magna cum laude with a triple major in physics, French and psychology before completing his medical degree and clinical training in general and cardiothoracic surgery. Elefteriades is currently professor and chief of cardiothoracic surgery at Yale and the Yale–New Haven Hospital. He began a weight-training program while on his seventh-grade wrestling team and has been lifting ever since; he benches 75 percent of his weight and has confirmed, by echocardiogram, that he does not harbor an aneurysm.

A Warning for Weight Lifters

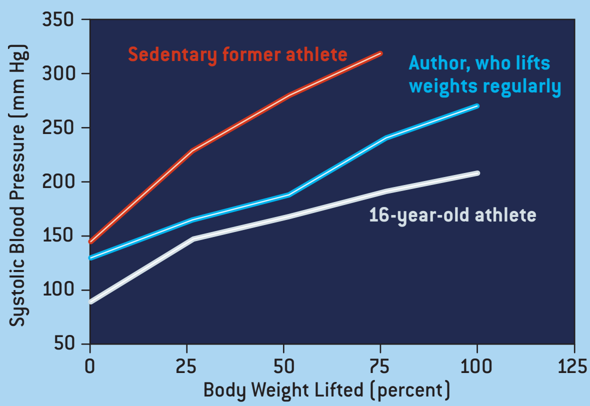
In late 2003 my colleagues and I described in the *Journal of the American Medical Association* the tragic occurrence of a dissection of the aorta in five seemingly healthy individuals who were engaged in strenuous strength training. Each unknowingly harbored a bulge in the part of the aorta emerging from the heart, and the inner half of the distended wall suddenly, and life-threateningly, separated from the outer part. At the time of dissection, two were lifting weights, two were doing push-ups, and the fifth was attempting to lift a heavy piece of granite. Three were saved by surgical intervention. We have since become aware of dozens of additional cases of aortic dissection during weight lifting, suggesting that the phenomenon is not a medical rarity.

What might account for this link? Part of the explanation seems to be that exercise that involves straining against a fixed resistance, as weight lifting does, can push the blood pressure to dangerously high levels. Some studies have recorded a systolic pressure (that in the arteries when the heart contracts) of 380 millimeters of mercury (mm Hg) in competitive weight lifters, as compared with a normal value of 120 or below. We have confirmed such soaring pressures in a small, three-volunteer study of our own. One of our volunteers hit 319 mm Hg when lifting just three quarters of his body weight (*graph*).

Such pressure can be too much for an already stretched artery to bear. From separate studies of the mechanical properties of the distended aorta, we find that at 200 mm Hg, a six-centimeter aneurysm experiences 800 kilopascals of pressure—a value that equals the ultimate tensile strength of the tissue. So it should come as no surprise that an aortic aneurysm subject to a blood pressure that approaches or exceeds 300 mm Hg will not hold.

Because of this pressure elevation, we tell athletes with a personal or family history of aortic aneurysm or any known aortic enlargement to use great discretion in pursuing weight-lifting activities, perhaps limiting bench presses to half their body weight or less. Weight lifting can be highly beneficial for preserving muscle mass and bone strength, but we strongly advise that individuals who intend to embark on heavy weight-training programs get an echocardiogram to check for potential aneurysms before they begin. —*J.A.E.*

YOUNG MAN who suffered an aortic dissection induced by lifting extremely heavy weights bears the scar of the surgery that saved his life. Chances are good that a sharp rise in blood pressure during a workout precipitated the dissection. Even in healthy individuals, blood pressure can soar to astronomical values, exceeding 300 mm Hg, during weight lifting (*graph*).



study on aneurysm patients in Europe, to be sure that the findings will hold up in a different population.

What Goes Wrong

ONCE WE IDENTIFY the genes in which aneurysm-related SNPs occur, we can discern which proteins those genes encode and learn how they contribute to aortic malfunction. But already researchers have a sense of some of the proteins that might be involved. For instance, we know that in most patients with aortic aneurysms, the stretched part of the vessel wall shows a loss of elastic fibers and collagen as compared with healthy tissue. Together these proteins give the artery its strength and flexibility. The defects that contribute to this problem could occur in the genes that code for those proteins or in other genes that regulate the manufacture or maintenance of elastin and collagen.

In Marfan's, the genetic defects that are at fault usually hobble the gene for fibrillin, a protein that combines with elastin to form elastic fibers. As a consequence, the synthesis and deposition of fibrillin are disrupted, a problem that presumably weakens the aortic wall and renders it vulnerable to the formation of an aneurysm. No one yet knows, though, whether mutations in the fibrillin gene are common in patients who do not have Marfan's.

We have recently found evidence that an overabundance of certain enzymes in the aortic wall probably contributes to the formation and growth of aneurysms in many victims. All blood vessels harbor enzymes called metalloproteinases (MMPs) that chew up old proteins to make way for new. The same vessels also possess inhibitory proteins that help hold MMPs at bay. In a healthy

JAKE MESSERE (*photograph*); ALISON KENDALL (*graph*)

Who Should Worry

In some ways, an aortic aneurysm is like a time bomb in the chest. It can sit silently until one day it ruptures or dissects. But certain conditions often signal a susceptibility to aortic aneurysms:

- A family history of aneurysms
- Having someone in the family collapse and die suddenly or unexpectedly
- Marfan's syndrome or its hallmarks. These include long limbs, a tall, thin frame and loose joints (such as are evident in the ability, shown at right, to cross the thumb all the way over the palm while keeping the hand flat)

My colleagues and I tell patients who meet any of these criteria—or who plan on engaging in serious weight training—to have regular CT scans or echocardiograms to check for aneurysms. Weight training does not elevate risk for having an aneurysm, but, as noted in the box on the opposite page, it can increase the likelihood that an existing aneurysm will suddenly become lethal. *—J.A.E.*

JOHN A. ELEFTERIADES

aortic wall the activity of these proteins is balanced, so that protein turnover remains constant. In segments of aorta removed from our aneurysm patients, in contrast, we find an excess of two types of MMP and decreased amounts of one of the inhibitory proteins.

This imbalance could lead to an enhanced degradation of proteins, including elastin and fibrillin, in the aortic wall—a situation that might pave the way for thoracic aortic aneurysms by weakening the vessel wall. In one patient, the aorta had become so thin that the markings of a ruler could be read through its wall. Other scientists have also found evidence of a role for overzealous MMPs. These findings suggest that drugs able to block MMP activity might help retard growth or forestall rupture of aortic aneurysms, but study of this concept is just beginning.

With our Yale colleague George Koullias, we have recently begun to assess the mechanical properties of the dilated aorta to better understand why it becomes more dangerous as it enlarges. Before surgically removing an aneurysm, we measure its diameter, the thickness of its wall, and the blood pressure as the heart contracts and relaxes. From these parameters we can calculate the vessel's mechanical properties.

We have shown that as the aorta grows larger its distensibility, or ability to stretch, falls. We have also demonstrated that by the time an aneurysm in the ascending aorta reaches six centimeters—the same critical value we encountered in our previous studies of aneurysm behavior—the vessel behaves as a rigid tube. This stiffening maximizes the stress that gets absorbed by the wall of the aorta as blood pounds against it with every heartbeat and helps to explain why trouble often ensues when an aneurysm hits the crucial dimension of six centimeters.

Inflexibility sets an aortic aneurysm up for disaster. But what sends it over the edge? We have begun to categorize the specific events that cause dissection to occur at one particular moment in time in a susceptible individual. After interviewing patients in our database, we find that nearly three out of four recall experiencing an intense episode of extreme emotion or physical exertion immediately preceding the dissection. What these activities have in common is that both presumably cause a spike in blood pressure that splits the vulnerable aorta. For one class of athletic activity—weight lifting—we have specific evidence that this is the case; indeed, this activity can put so much stress on an aneurysm that it prompts a crisis even when the swelling has not crossed the six-centimeter mark [*see box on opposite page*]. It seems logical to surmise that pressure spikes from other events could also induce rupture, although we have not looked at that possibility directly yet.

The renowned 19th-century physician Sir William Osler once observed that "there is no disease more conducive to clinical humility than aneurysm of the aorta." Today investigations into the biology and behavior of thoracic aortic aneurysms—from the genetic susceptibility that drives their formation to the physical or emotional events that cause them to blow out or tear—are helping to render the condition a little less humbling.

As for Carmela, she continues to be in good health and has returned to her work as an artist. "I know it sounds clichéd," she says, "but I feel I've been given a second chance to live my life"—a chance her father did not have when he died from an aortic dissection at the age of 34. We hope that our research, inspired by Carmela's crisis that terrifying spring day, will provide many others with the same opportunity. SA

MORE TO EXPLORE

Surgical Intervention Criteria for Thoracic Aortic Aneurysms: A Study of Growth Rates and Complications. Michael A. Coady et al. in *Annals of Thoracic Surgery,* Vol. 67, No. 6, pages 1922–1926; June 1999.

Yearly Rupture or Dissection Rates for Thoracic Aortic Aneurysms: Simple Prediction Based on Size. R. R. Davies, L. J. Goldstein, M. A. Coady, S. L. Tittle, J. A. Rizzo, G. S. Kopf and J. A. Elefteriades in *Annals of Thoracic Surgery,* Vol. 73, No. 1, pages 17–27; January 2002.

Weight Lifting and Rupture of Silent Aortic Aneurysms. John Elefteriades et al. in *Journal of the American Medical Association,* Vol. 290, No. 21, page 2803; December 3, 2003.

Perspectives on Diseases of the Thoracic Aorta. John A. Elefteriades in *Advances in Cardiology,* Vol. 41, pages 75–86; 2004.

Kevin Helliker and Thomas M. Burton's *Wall Street Journal* series on aortic aneurysms: **www.pulitzer.org/year/2004/explanatory-reporting/works/**

PHOTOGRAPHS PROVE TRICKY to many toddlers because they have not mastered dual representation: awareness that a symbolic object is itself (in this case, a highly realistic photograph) as well as a representation of something else (a sneaker). Many try to interact with the objects in photographs, just as this boy is doing by attempting to put his foot in the shoe.

MINDFUL OF *symbols*

On the way to learning that one thing can represent another, young children often conflate the real item and its symbol. These errors show how difficult it is to start thinking symbolically

By Judy S. DeLoache
Photographs by Randy Harris

About 20 years ago I had one of those wonderful moments when research takes an unexpected but fruitful turn. I had been studying toddler memory and was beginning a new experiment with two-and-a-half- and three-year-olds. For the project, I had built a model of a room that was part of my lab. The real space was furnished like a standard living room, albeit a rather shabby one, with an upholstered couch, an armchair, a cabinet and so on. The miniature items were as similar as possible to their larger counterparts: they were the same shape and material, covered with the same fabric and arranged in the same positions. For the study, a child watched as we hid a miniature toy—a plastic dog we dubbed “Little Snoopy”—in the model, which we referred to as “Little Snoopy’s room.” We then encouraged the child to find “Big Snoopy,” a large version of the toy “hiding in the same place in his big room.” We wondered whether children could use their memory of the small room to figure out where to find the toy in the large one.

The three-year-olds were, as we had expected, very successful. After they observed the small toy being placed behind the miniature couch, they ran into the room and found the large toy behind the real couch. But the two-and-a-half-year-olds, much to my and their parents’ surprise, failed abysmally. They cheerfully ran into the room to retrieve the large toy, but most of them had no idea where to look, even though they remembered where the tiny toy was hidden in the miniature room and could readily find it there.

Their failure to use what they knew about the model to draw an inference about the room indicated that they did not appreciate the relation between the model and room. I soon realized that my memory study was instead a study of symbolic understanding and that the younger children’s failure might be telling us something interesting about how and when youngsters acquire the ability to understand that one object can stand for another.

What most distinguishes humans from other creatures is our ability to create and manipulate a wide variety of symbolic representations. This capacity enables us to transmit information from one generation to another, making culture possible, and to learn vast amounts without having direct experience—we all know about dinosaurs despite never having met one. Because of the fundamental role of symbolization in almost everything we do, perhaps no aspect of human development is more important than becoming symbol-minded.

What could be more fascinating, I concluded, than finding out how young children begin to use and understand symbolic objects and how they come to master some of the symbolic items ubiquitous in modern life. As a result of that fortuitous model-room experiment, I shifted my focus from memory to symbolic thinking.

Pictures Come to Life

THE FIRST TYPE of symbolic object infants and young children master is pictures. No symbols seem simpler to adults, but my colleagues and I have discovered that infants initially find pictures perplexing. The problem stems from the duality inherent in all symbolic objects: they are real in and of themselves and, at the same time, representations of something else. To understand them, the viewer must achieve dual representation: he or she must mentally represent the object as well as the relation between it and what it stands for.

A few years ago I became intrigued by anecdotes suggesting that infants do not appreciate the dual nature of pictures. Every now and then, I would hear of a baby who tried to pick up a depicted apple or to fit a foot into a photograph of a shoe. My colleagues—David H. Uttal of Northwestern University, Sophia L. Pierroutsakos of St. Louis Community College and Karl S. Rosengren of the University of Illinois at Urbana-Champaign—and I decided to investigate even though we assumed such behaviors would be rare and therefore difficult to study. Fortunately, we were wrong.

We began testing infants' understanding of pictures in a very simple way. We put a book containing highly realistic color photographs of individual objects in front of nine-month-olds. To our surprise, every child in the initial study, and most in our subsequent studies, reached out to feel, rub, pat or scratch the pictures. Sometimes the infants even grasped at the depicted objects as if trying to pick them up off the page.

We had a unique opportunity to see how universal this response was when anthropologist Alma Gottlieb of the University of Illinois took some of our books and a video camera to a remote Beng village in Ivory Coast. The testing situation there was different: Beng babies sat on the ground or in their mother's lap as chickens and goats wandered around and other children and villagers played, worked, talked and laughed nearby. Yet the Beng babies, who had almost certainly never seen a picture before, manually explored the depicted objects just as the American babies had.

The confusion seems to be conceptual, not perceptual. Infants can perfectly well perceive the difference between objects and pictures. Given a choice between the two, infants choose the real thing. But they do not yet fully understand what pictures are and how they differ from the things depicted (the "referents") and so they explore: some actually lean over and put their lips on the nipple in a photograph of a bottle, for instance. They only do so, however, when the depicted object is highly similar to the object it represents, as in color photographs. The same confusion occurs for video images. Pierroutsakos and her colleague Georgene L. Troseth of Vanderbilt University found that nine-month-olds seated near a television monitor will reach out and grab at objects moving across the screen. But when depicted objects bear relatively little resemblance to the real thing—as in a line drawing—infants rarely explore them.

By 18 months, babies have come to appreciate that a picture merely represents a real thing. Instead of manipulating the paper, they point to pictures and name objects or ask someone else for the name. Melissa A. Preissler of Yale University and Susan Carey of Harvard University recently provided a good example of this development. The two researchers used a simple line drawing of a whisk to teach 18- and 24-month-olds the word for this object that they had not seen before. Most of the children assumed the word referred to the object itself, not just to the picture of it. In other words, they interpreted the picture symbolically—as standing for, not just being similar to, its referent.

One factor we think contributes to the decline of manual exploration of pictures is the development of inhibitory control. Throughout the first years of life, children become increasingly capable of curbing impulses. This general developmental change is supported by changes in the frontal cortex. Increased inhibitory control presumably helps infants restrain their impulse to interact directly with pictures, setting the stage for them to simply look, as adults do.

Experience with pictures must play a role in this development as well. In an image-rich society, most children encounter family photographs and picture books on a daily basis. Such interactions teach children how pictures differ from objects and how they are supposed to be targets of contemplation and conversation, not action.

Nevertheless, it takes several years for the nature of pictures to be completely understood. John H. Flavell of Stanford University and his colleagues have found, for example, that until the age of four, many children think that turning a picture of a bowl of popcorn upside down will result in the depicted popcorn falling out of the bowl.

Pictures are not the only source of symbol confusion for very young children. For many years, my colleagues and

Overview/*Symbols Are Not Intuitive*

- Although symbolic thinking is a hallmark of being human, it is not something infants can do. Instead children learn such thinking over several years.
- During this process, they make many fascinating errors—including mistaking photographed objects for the real thing and treating small toys as if they are much larger.
- Fundamentally, grasping the meaning of a symbol entails dual representation. Only when children can see an object both as itself and as depicting something else can they start to think symbolically.

SCALE ERRORS ARE COMMON among 18- to 30-month-olds, who often interact with small objects as they do with larger versions. Scale errors are another example of failed dual representation. Here the boy couldn't sit on the chair and kept falling off. (In experiments, objects can be even smaller.)

students and I watched toddlers come into the lab and try to sit down on the tiny chair from the scale model—much to the astonishment of all present. At home, Uttal and Rosengren had also observed their own daughters trying to lie down in a doll's bed or get into a miniature toy car. Intrigued by these remarkable behaviors that were not mentioned in any of the scientific literature we examined, we decided to study them.

Gulliver's Errors

WE BROUGHT 18- to 30-month-old children into a room that contained, among other things, three large play objects: an indoor slide, a child-size chair and a car toddlers could get inside and propel around the room with their feet. After a child had played with each of the objects at least twice, he or she was escorted from the room. We then replaced the large items with identical miniature versions. When the child returned, we did not comment on the switch and let him or her play spontaneously. If the toddler ignored the miniature toys for more than three or four minutes, however, we would draw attention to them.

We then examined films of the children's behavior for what we came to call scale errors: earnest attempts to perform actions that are clearly impossible because of extreme differences in the relative size of the child's body and the target object. We were very conservative in what we counted as a scale error.

Almost half the children committed one or more of these mistakes. They attempted with apparent seriousness to perform the same actions with the miniature items that they had with the large ones. Some sat down on the little chair: they walked up to it, turned around, bent their knees and lowered themselves onto it. Some simply perched on top, others sat down so hard that the chair skittered out from under them. Some children sat on the miniature slide and tried to ride down it, usually falling off in the process; others attempted to climb the steps, causing the slide to tip over. (With the chair and slide made of sturdy plastic and only about five inches tall, the toddlers faced no danger of hurting themselves.) A few kids tried to get into the tiny car. Just as they had done with the large version, they opened the door and attempted—often with remarkable persistence—to force a foot inside. One little girl went so far as to take off her shoe in the apparent hope that her foot would then fit!

Interestingly, most of the children showed little or no reaction to their failed attempts with the miniatures. A couple seemed a bit angry, a few looked sheepish, but most simply went on to do something else. We think the lack of reaction probably reflects the fact that toddlers' daily lives are full of unsuccessful attempts to do one thing or another.

Our interpretation is that scale errors originate in a dissociation between the use of visual information for planning an action and for controlling its execution. When a child sees a miniature of a familiar object, visual information—the object's shape, color, texture and so on—activates the child's mental representation of its referent. Associated with that memory is the motor program for interacting with the large object and other similar objects. In half the children we studied, this motor program was presumably activated but then inhibited,

THE AUTHOR

JUDY S. DeLOACHE specializes in early cognitive development—specifically the development of symbolic thinking—at the University of Virginia, where she is William R. Kenan, Jr., Professor of Psychology. DeLoache also holds an appointment in psychology at the University of Illinois at Urbana-Champaign, where she has taught since the late 1970s and where she earned her doctorate. The author thanks the many children and families who have participated in her research.

TWO-YEAR-OLD CANNOT APPRECIATE the symbolic relation between a model of a room and a room itself. This boy can see the toy hidden behind the plant in the model but does not know to look behind the plant in the real room.

and the children did not attempt to interact with the miniature in the same way as they did with the large version.

But in the other half the motor routine was not inhibited. Once the child began to carry out the typical motor sequence, visual information about the actual size of the object was used to accurately perform the actions. Some children, for instance, bent over the tiny chair and looked between their legs to precisely locate it; those trying to get into the miniature car first opened its door and then tried to shove their foot right in. In deciding to interact with the replica, the children relied on visual information linking it to the normal-size object, but in executing their plan, they used visual information about the miniature's actual size to guide their actions. This dissociation in the use of visual information is consistent with influential theories of visual processing—ones positing that different regions of the brain handle object recognition and planning versus the execution and control of actions.

Scale errors involve a failure of dual representation: children cannot maintain the distinction between a symbol and its referent. We know this because the confusion between referent and symbolic object does not happen when the demand for dual representation is eliminated—a discovery I made in 1997 when Rosengren and Kevin F. Miller of the University of Illinois and I convinced two-and-a-half-year-olds—with the full consent of their parents, of course—that we had a device that could miniaturize everyday objects.

The Magical Machine

USING OUR AMAZING shrinking machine, we hoped to see if the need to think of an object in two ways at once was at the heart of young children's inability to appreciate symbols. If a child believes that a machine has shrunk an object or a room, then in the child's mind the miniature is the thing itself. There is no symbolic relation between room and model, so children should be able to apply what they know about the big version to the little one.

We used the powers of our device to turn toys into miniature versions of themselves and to shrink a large tent. In front of the child, we placed a toy—a troll doll with vivid purple hair—in a tent and aimed the shrinking machine at the tent. The child and experimenter then decamped to another room to wait while the machine did its work. When they returned to the lab, a small tent sat where the big one had been. (One of the remarkable things about this study is the fact that the children did not find it at all surprising that a machine could miniaturize objects. Or that it might need privacy to do so.)

When we asked the children to search for the toy, they immediately looked in the small tent. Believing the miniature to actually be the original tent after shrinking, they successfully retrieved the hidden toy. Unlike in our scale model experiment, they had no dual representation to master: the small tent was the same as the large tent, and thus the toy was where it should be, according to the toddlers' view of the world.

Understanding the role of dual representation in how young children use symbols has important practical applications. One has to do with the practice of using dolls to interview young children in cases of suspected sexual abuse. The victims of abuse are often very young children, who are quite difficult to interview. Consequently, many professionals—including police officers, social workers and mental health professionals—employ anatomically detailed dolls, assuming that a young child will have an easier time describing what happened using a doll. Notice that this assumption entails the further assumption that a young child will be able to think of this object as both a doll and a representation of himself or herself.

These assumptions have been called into question by Maggie Bruck of Johns Hopkins University, Stephen J. Ceci of Cornell University, Peter A. Ornstein of the University of North Carolina at Chapel Hill and their many colleagues. In several independent studies, these investigators have asked preschool children to report what they remember about a checkup with their pediatrician, which either had or had not included a genital check. Anatomically detailed dolls were sometimes used to question the children, sometimes not. In general, the children's reports were more accurate when they were questioned without a doll, and they were more likely to falsely report genital touching when a doll was used.

Based on my research documenting young children's difficulty interpreting symbolic objects, I suspected that very young children might not be able to relate their own body to a doll. In a series of studies in my lab using an extremely simple mapping task, Catherine Smith placed a sticker somewhere on a child—on a shoulder or foot, for example—and asked the child to place a smaller version of the sticker in the same place on a doll. Children between three and three-and-a-half usually placed the sticker correctly, but children younger than three were correct less than half the time. The fact that these very young children cannot relate their own body to the doll's in this extremely simple situation with no memory demands and no emotional involvement supports the general case against the use of anatomically detailed dolls in forensic situations with young children. (Because of many demonstrations akin to this one, the use of dolls with children younger than five is viewed less favorably than in the past and has been outlawed in at least one state.)

Educational Ramifications

THE CONCEPT of dual representation has implications for educational practices as well. Teachers in preschool and elementary school classrooms around the world use "manipulatives"—blocks, rods and other objects designed to represent numerical quantity. The idea is that these concrete objects help children appreciate abstract mathematical principles. But if children do not understand the relation between the objects and what they represent, the use of manipulatives could be counterproductive. And some research does suggest that children often have problems understanding and using manipulatives.

Meredith Amaya of Northwestern University, Uttal and I are now testing the effect of experience with symbolic objects on young children's learning about letters and numbers. Using blocks designed to help teach math to young children, we taught six- and seven-year-olds to do subtraction problems that require borrowing (a form of problem that often gives young children difficulty). We taught a comparison group to do the same but using pencil and paper. Both groups learned to solve the problems equally well—but the group using the blocks took three times as long to do so. A girl who used the blocks offered us some advice after the study: "Have you ever thought of teaching kids to do these with paper and pencil? It's a lot easier."

Dual representation also comes into play in many books for young children. A very popular style of book contains a variety of manipulative features designed to encourage children to interact directly with the book itself—flaps that can be lifted to reveal pictures, levers that can be pulled to animate images, and so forth.

Graduate student Cynthia Chiong and I reasoned that these manipulative features might distract children from information presented in the book. Accordingly, we recently used different types of books to teach letters to 30-month-old children. One was a simple, old-fashioned alphabet book, with each letter clearly printed in simple black type accompanied by an appropriate picture—the traditional "A is for apple, B is for boy" type of book. Another book had a variety of manipulative features. The children who had been taught with the plain book subsequently recognized more letters than did those taught with the more complicated book. Presumably, the children could more readily focus their attention with the plain 2-D book, whereas with the other one their attention was drawn to the 3-D activities. Less may be more when it comes to educational books for young children.

As these various studies show, infants and young children are confused by many aspects of symbols that seem intuitively obvious to adults. They have to overcome hurdles on the way to achieving a mature conception of what symbols represent, and today many must master an ever expanding variety of symbols. Perhaps a deeper understanding of the various stages of becoming symbol-minded will enable researchers to identify and address learning problems that might stem from difficulty grasping the meanings of symbols. SA

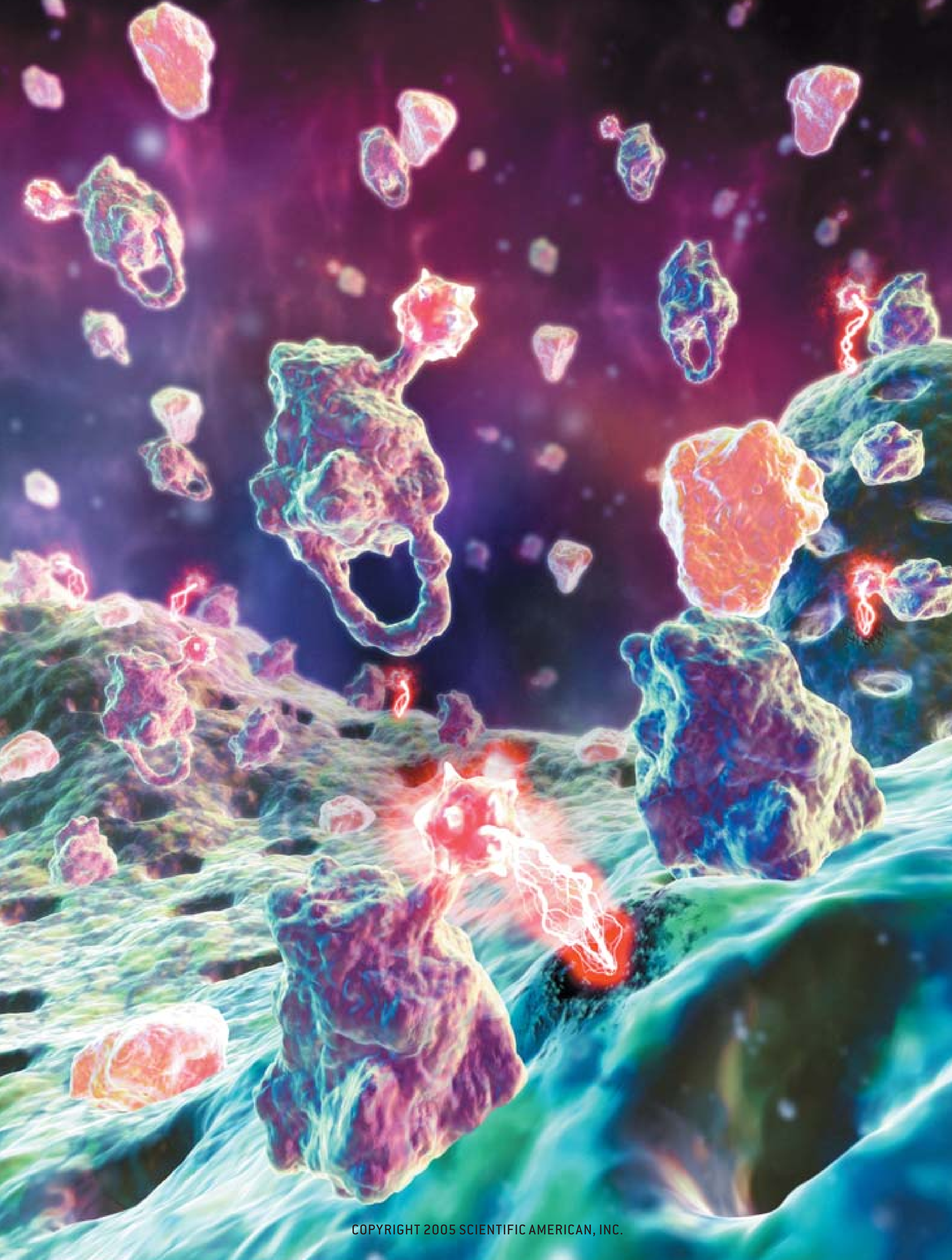
MORE TO EXPLORE

Jeopardy in the Courtroom: The Scientific Analysis of Children's Testimony. S. J. Ceci and M. Bruck. American Psychological Association, Washington, D.C., 1995.

Becoming Symbol-Minded. J. S. DeLoache in *Trends in Cognitive Sciences,* Vol. 8, No. 2, pages 66–70; February 2004.

Scale Errors Offer Evidence for a Perception-Action Dissociation Early in Life. J. S. DeLoache, D. H. Uttal and K. S. Rosengren in *Science,* Vol. 304, pages 1027–1029; May 14, 2004.

Film clips of children making symbolism-related errors can be seen at **www.faculty.virginia.edu/childstudycenter/home.html**



NANOBODIES of several kinds (*purple*) could descend on a cancerous cell (*blue-green*). Some nanobodies might be designed to attach to receptors on the cell, preventing pro-growth signals (*orange*) from reaching the cell. Other nanobodies could deliver radioactive payloads (*clublike appendages*) to cancer-specific targets.

Nanobodies

Antibodies, often described as magic bullets, are actually more like tanks: big, complicated and expensive. Tinier "nanobodies," derived from camels and llamas, may be able to infiltrate a wider range of diseases at lower cost. That is the hope, at least, of one small start-up in Belgium

BY W. WAYT GIBBS

Like many biotech companies, Ablynx emerged from the confluence of a serendipitous discovery, an open window of opportunity and an unreasonable ambition. Housed on two floors in a nondescript gray laboratory on a technology campus outside the university town of Ghent, Belgium, the three-year-old company employs just 45 people, 33 of them scientists and bioengineers. It is a minimal staff with a simply stated mission: find the tiniest sliver of protein that will do the job of a full-size antibody, then turn it into a billion-dollar medicine—or better yet, into the first of a whole new class of "nanobody" drugs against cancer, rheumatoid arthritis, inflammatory bowel disease, perhaps even Alzheimer's disease.

Despite being backed by $40 million of venture capital and partnerships with Genencor, Procter & Gamble and the National Research Council of Canada, Ablynx faces long odds. Its ambitious goal might seem altogether futile were it not for the recent surge in antibody therapies, the problems that still nag these sophisticated drugs, and the insights that Ablynx scientists have into the peculiar biology of the camel family.

Aside from the brain, the most complicated part of the human body is undoubtedly the immune system—and thank goodness. It's a bacteria-eat-man world out there, filled with a nearly endless variety of germs that see us as spawning grounds. Defending against this onslaught are antibodies, which are manufactured by B cells in an equally impressive panoply of models. Antibodies are huge Y-shaped proteins that float about in the blood and the fluid between cells, arms extended, using a chemical sense of touch to interrogate other molecules they encounter. Each model of antibody has its own mission; it patrols for a distinct chemical signature of a certain microbe, allergen or toxin.

Yet despite the sophistication of our immune defenses, we still get sick. No police force is perfect. The immune system is sometimes too slow or complacent in its reaction—for example, to cancers or to infection by respiratory syncytial virus. Other times it overreacts, as happens in organ-transplant rejection and asthma. And when it mistakenly attacks the body's own cells, the immune response can itself cause a degenerative disease such as rheumatoid arthritis.

For years, drugmakers sought to create artificial antibodies that can correct—or at least moderate—these immunological failures. But most early attempts ended in failure and financial disaster. In the two decades following the 1975 invention of a way to produce large batches of antibodies that are identical, or "monoclonal," just two such therapies survived review by the U.S. Food and Drug Administration.

The logjam finally broke in 1997, and by the end of 2004 the FDA had ap-

JEFF JOHNSON *Hybrid Medical Animation*

proved 17 therapeutic antibodies, including promising treatments for all the ailments mentioned above [see "Magic Bullets Fly Again," by Carol Ezzell; Scientific American, October 2001]. Pharmaceutical firms reaped $11.2 billion in sales of these medicines in 2004, the consultancy AS Insights reports.

And the market for monoclonal antibodies (usually abbreviated, idiosyncratically, as MAbs) is still in a formative stage of rapid growth. Dozens more MAbs are now in development or clinical trials, and last year Janice M. Reichert of the Center for the Study of Drug Development at Tufts University projected that 16 of them will gain FDA licenses within the next three years. In 2008, she forecast, MAbs will command roughly $17 billion in worldwide sales.

As Ablynx lines up for its first clinical trials in late 2006, it is aiming for a small slice of that large pie, says Mark Vaeck, the company's chief executive. Nanobodies—relatively simple proteins about a tenth the size of antibodies and just a few nanometers in length—may one day yield new medicines for Alzheimer's and other diseases beyond the reach of current antibodies, but that is not the opening strategy Vaeck chose. Instead he directed his scientists to create nanobodies that do what some of the best-selling antibodies do, only better.

The Trouble with Antibodies

CERTAINLY THERE IS ROOM for improvement. For all their promise, points out Hans de Haard, scientific director at Ablynx, monoclonal antibodies still make pricey and troublesome medicines. According to Medco Health Solutions, treating an asthmatic patient with the antibody Xolair costs about $11,000 a year for the drug alone. Remicade, for rheumatoid arthritis, runs about $4,600 for eight shots. A year's course of Herceptin, an antibody cancer therapy, soars over $38,000.

LLAMA at an Ablynx laboratory makes unusual kinds of antibodies, unique to the camel family, that serve as raw materials for nanobodies.

INA JOHANNA GECKE

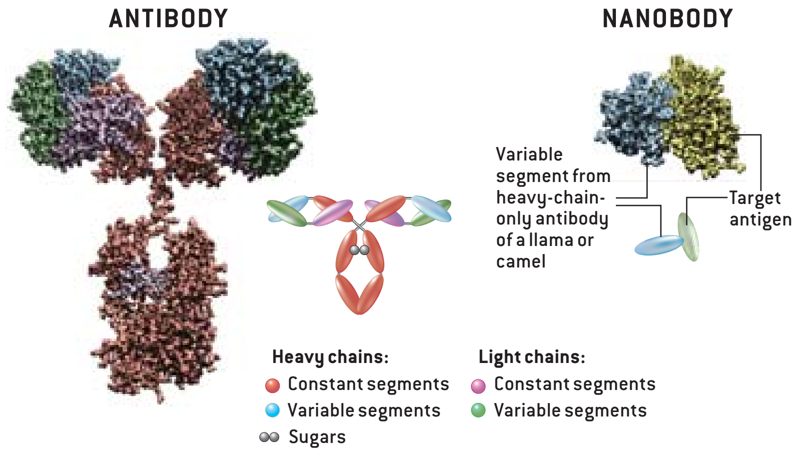
MAbs are so dear in large part because they are so complex. By molecular standards antibodies are giants, each one a conglomerate of two heavy protein chains and two light chains, intricately folded and garnished with elaborate sugars [*see box on opposite page*]. To make a MAb medicine, scientists usually begin with an antibody isolated from mice. They then "humanize" the molecule by tinkering with the genes that encode it to replace some or all of the protein with amino acid sequences copied from human antibodies. (Alternatively, a few companies have genetically engineered mice so that they produce nearly human antibodies directly.)

The humanization reduces the potentially dangerous side effects that antibody therapies often generate when the patient's body perceives the MAbs as foreign invaders and launches an immune attack on the medicine. But the humanization process often consumes many months of highly technical work. And the resulting macromolecules are so complicated that they cannot be synthesized from chemical building blocks, as conventional drugs are. Instead they must be grown in bioreactor vats of mammalian cells that have been engineered to carry the multiple genes needed to make a single antibody.

Cell cultures of this kind do not scale easily for mass production. MAb factories are much more expensive to build and operate than are similarly sized chemical or microbial biosynthesis plants. Drug companies must ensure, for example, that their vats do not take sick with a virus that might ruin the valuable cells or contaminate the antibodies. A recent analysis by Mark C. Via, published by Cambridge Healthtech Advisors, concluded that demand for monoclonal antibodies most likely will far

ANATOMY OF AN ANTIBODY

The millions of kinds of human antibodies all share the same basic structure: two larger (or heavy) protein chains linked with two smaller (or light) chains. The pair of variable segments at the tips of the arms are unique for each model of antibody and determine the target to which it will bind. A nanobody is the variable part of a camel antibody that lacks light chains; it is about one tenth the size of an antibody.



W. WAYT GIBBS AND LUCY READING-IKKANDA

outstrip production capacity for years to come. All these factors conspire to drive up the price of antibody therapies.

The great size of the proteins also imposes practical and medical limitations. High temperatures or extremes of pH make MAbs unravel. They typically expire in weeks unless stored near freezing temperatures. Antibodies are digested quickly in the gut, blocked from entering the brain and held to the periphery of solid tumors. Many illnesses are thus unreachable by monoclonals, and patients who can use MAb therapies must receive them by injection at a clinic.

For certain conditions in which MAbs do not work well, and even for some in which they currently do, simpler, smaller proteins might perform better and be easier to make, easier to handle, easier to take and thus more affordable. This idea predates the invention of nanobodies by many years. In the 1980s protein engineers began experimenting with antibody fragments created by chopping off the stem of the Y, or sometimes the stem and an arm, leaving just one "hand" to do the chemical duty of the antibody.

Like full-size MAbs, these antibody fragments (nicknamed Fabs) can treat illnesses by binding to toxins, pathogens or aberrant cell signals—or alternatively to the cell receptors to which those undesirable molecules dock. But antibody fragments cannot recruit other components of the immune system, such as killer T cells, in the same way that full-size antibodies do, because they lack the protein stem that performs that task.

In their favor, Fabs can be manufactured by bacteria, yeast or fungi, which are less expensive than the mouse or hamster cells needed to synthesize antibodies. Fabs can sneak into the center of tumors, and molecular engineers can rig them to tow toxic payloads—such as radioactive isotopes or chemotherapy drugs—directly to diseased tissue.

On the other hand, Fabs tend to fall apart or filter out of the bloodstream quickly, and so their active half-life typically amounts to mere hours rather than the weeks that full-size antibodies can persist within the body. Fast clearance may be just what is wanted for delivering a toxin, but for many medicines it is a disadvantage. So far only one therapeutic Fab has made it to market in the U.S., and that more than a decade ago.

Some companies, such as Domantis in Cambridge, Mass., have trimmed Fabs further, stripping away all but the tip of one of the two chains. This segment, which is unique to each model of antibody, contains the critical chemical fingers known as complementarity determining regions (CDRs), that determine what target an antibody will recognize—its antigen—and how tightly the two will bind when they meet. The resulting domain antibodies, as Domantis calls its proteins, are similar in size to the nanobodies that Ablynx makes.

But domain proteins evolved as segments of much larger, double-chained antibodies, and that has made them inherently sticky, explains Serge Muyldermans, a protein biologist at the Free University of Brussels. The fragments thus agglomerate together inside the bacteria that make them, as well as inside the patients that take them. The stickiness of the molecules lowers their production yields and hinders them in their work.

From Dromedary to Drug

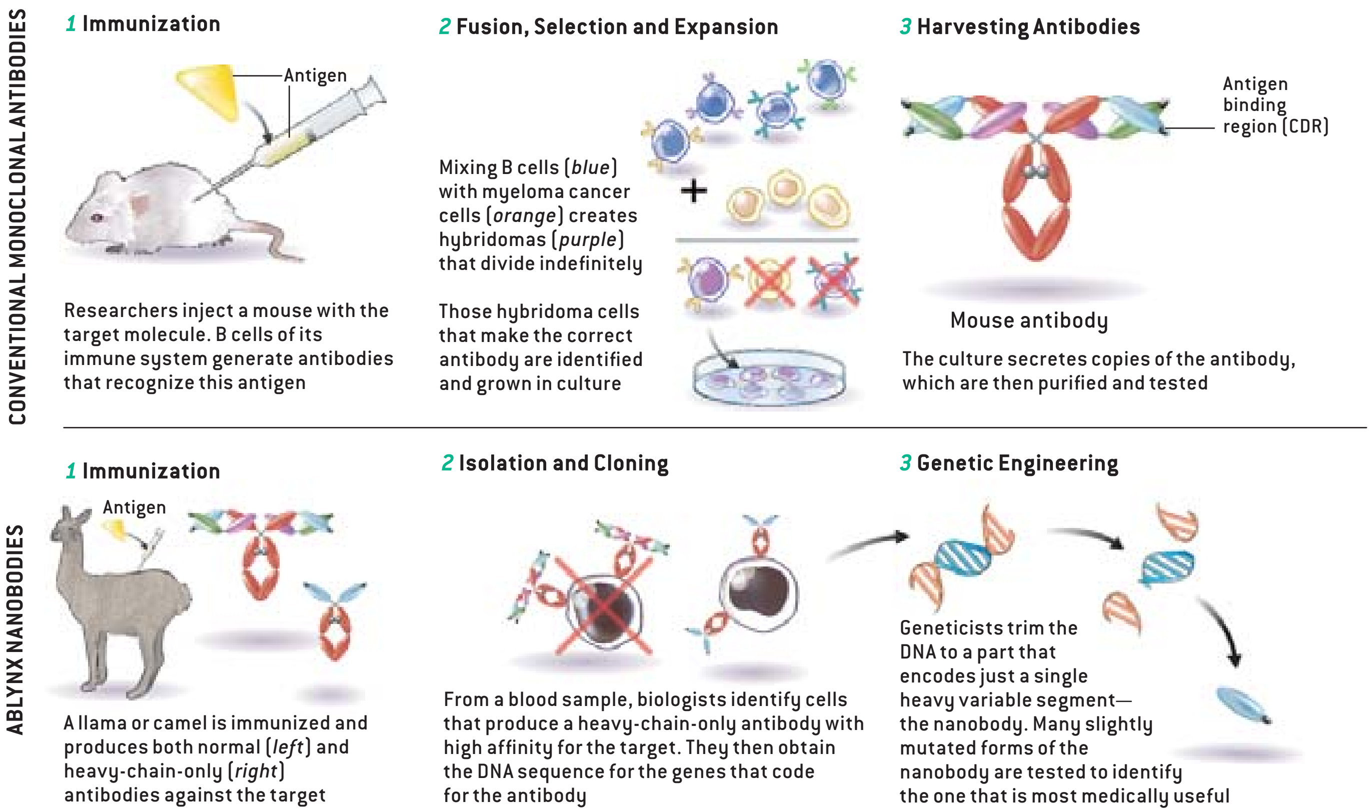
WHILE BIOCHEMISTS continue trying to reengineer antibody fragments to solve these problems, Ablynx is exploiting an alternative offered by nature. In 1989 Muyldermans was among a group of biologists led by Raymond Hamers at the Free University that investigated an odd observation handed in as part of a student project on how dromedary camels (the one-humped, Arabian variety) and water buffalo fight off parasites. One of the tests for antibodies in the dromedary blood seemed to show an error: in addition to normal four-chain antibodies, it indicated the presence of simpler antibodies composed solely of a pair of heavy chains.

After several years of investigation, Hamers, Muyldermans and their colleagues published their serendipitous discovery in *Nature* in 1993. In dromedaries—and also in two-humped Asian camels and South American llamas—about half the antibodies circulating in the blood lack a light chain. Equally surprising, they found, these "incomplete" antibodies are able to grasp their targets just as firmly as normal antibodies do, despite having only half as many CDRs. And unlike Fabs, the heavy-chain-only antibodies do not stick to one another.

Why species in the camel family differ from all other mammals in this respect remains a mystery, but evolution may have handed scientists a work-

CONSTRUCTING ANTIBODIES AND NANOBODIES

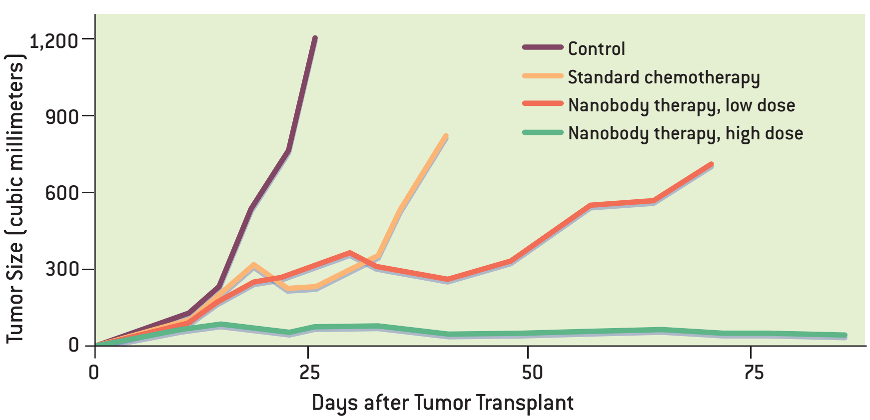
Creating an effective nanobody takes less time and money than a therapeutic antibody requires, according to scientists at Ablynx. In both cases, the immune system of a live animal performs the initial "design" of a protein that can latch onto the target molecule. Geneticists then tweak the DNA encoding that protein to add the properties desired in a medicine.



around for some of the thorniest problems with antibodies and antibody fragments. When Muyldermans's group pared these novel molecules down to just their distinctive, variable segments, the segments retained amazingly strong affinity for their targets, virtually equal to a full antibody 10 times their size. These shortened proteins were also more chemically agile, able to engage targets—including the active sites of enzymes and clefts in cell membranes—too small to admit an antibody. Nanobodies were born, and Ablynx soon followed.

Because nanobodies are so much smaller than antibodies and are not chemical hydrophobes (as are domain antibodies), they are more resistant to heat and pH, Muyldermans says. Pieter Rottiers and Hilde Revets of the Flanders Institute for Biotechnology (VIB) in Belgium have shown that the compounds retain their activity as they pass through the gastrointestinal tract of mice, raising the prospect of nanobody pills to treat inflammatory bowel disease, colon cancer and other disorders of the gut.

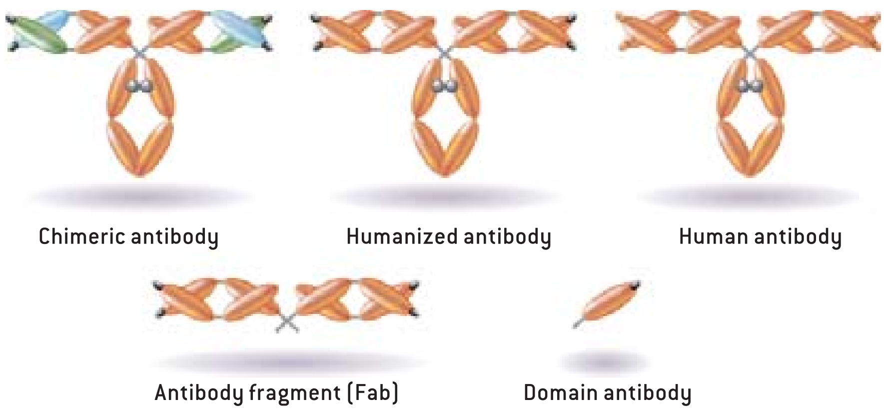
Because nanobodies are so much simpler than antibodies in chemical composition and shape, they can be encoded by a single gene and are easier for microbes to synthesize. In 2002 biologists at Unilever Research in the Netherlands brewed more than a kilogram of nanobodies from a standard 15,000-liter tank of yeast (a yield of 67 milligrams per liter), whereas Ablynx scientists re-



NANOBODIES AGAINST CANCER were tested on mice after they were injected with human tumor cells. Standard chemotherapy merely delayed the growth of the cancers. High doses of nanobody-directed chemotherapy, however, knocked the tumors into remission.

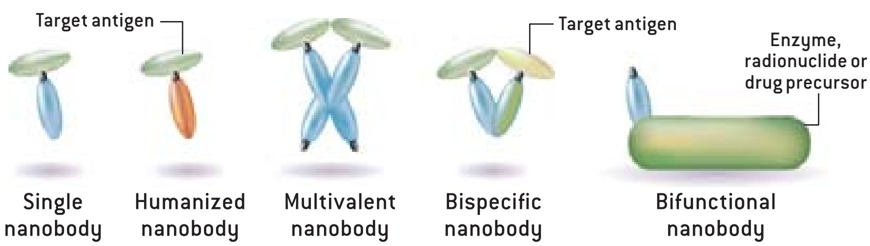
LUCY READING-IKKANDA

4 Humanization



Genetic engineers can replace pieces of the mouse antibody with human segments (*orange*) and can also trim the antibody to create fragments of various sizes

4 Construction of Nanobody Medicine



Nanobody genes can be spliced with genes for other nanobodies or other biochemicals to create medicines that are then produced in bacteria, fungi or yeast cultures

port recent yields exceeding a gram of nanobodies per liter of yeast culture—production rates that far exceed those typical for full-size antibodies. "Plus, our nanobodies are stable at room temperature and have a long shelf life without refrigeration," asserts Tim Van Hauwermeiren, who manages business development for the company.

The creation of new kinds of nanobodies is less difficult—and thus faster and less costly—than it is for antibodies, Van Hauwermeiren claims [*see box above*]. By immunizing llamas with the target antigen and then extracting heavy-chain-only antibodies from their blood, he says, "we can go from isolated target antigen to high-affinity nanobodies within four months." For some conditions, such as rheumatoid arthritis, the nanobody may serve unadorned as a medicine by jamming harmful cellular signals, either by attaching to the signal molecule or by clogging up the receptors for the signal on the surface of cells.

One of the most powerful advantages of nanobodies, however, is the relative ease with which the proteins can be joined to one another or to different kinds of compounds, de Haard says. His team has attached anti-albumin nanobodies to target-specific nanobodies to extend their half-lives in the bloodstream to weeks, he says. They have linked up to four nanobodies to create "multivalent" assemblies that can sop up more antigen per molecule or bind to either, or both, of two different targets.

Recently Revets, Muyldermans and Patrick De Baetselier of VIB published impressive results from an experiment in which they designed nanobodies to bind to a receptor on cancer cells, thus sticking to any tumors the molecules encounter. The researchers tailored a group of such nanobodies to be bifunctional by connecting each protein to an enzyme; the enzyme converts another chemical, called a prodrug, from its normal harmless form into a toxic chemotherapy that kills cells in the immediate vicinity.

The "patients" were mice that the scientists injected with human cancer cells, which soon grew into marble-size tumors. Revets treated some of the mice with the chemotherapy alone; those animals got sick and lost weight, just as happens in all chemotherapies. Their tumors shrank only a little. But the doctors gave another group of mice a high dose of the bifunctional nanobody with its attached enzyme. They waited a bit to give the unbound nanobodies time to filter out of the body, then injected the prodrug. As hoped, the nanobodies focused the chemotherapy on the cancer, sparing healthy tissues while completely driving back the tumors.

Until nanobodies make it through clinical trials, no one knows whether they will work as well in people as they do in mice. But if nanobodies do have an Achilles' heel, it is very likely to be the immune system itself. Ablynx scientists have worked out ways to humanize nanobodies, and studies with baboons have found that they raise no immune response to the tiny llama proteins. But de Haard acknowledges that nanobodies might not be able to evade the more sophisticated web of cellular surveillance that protects humans. The results of next year's clinical safety trials will determine whether nanobodies continue advancing at the recent breakneck pace or get tripped up by the complexities of the human immune system. SA

W. Wayt Gibbs is senior writer.

MORE TO EXPLORE

New Directions in Monoclonal Antibodies. Mark C. Via. Cambridge Healthtech Advisors, October 2004. Available at **www.chadvisors.com**

Nanobodies as Novel Agents for Cancer Therapy. Hilde Revets, Patrick De Baetselier and Serge Muyldermans in *Expert Opinion on Biological Therapy,* Vol. 5, No. 1, pages 111–124; January 2005.

BALL BEARINGS

Ease the Grind

Hidden from sight, ball bearings are key to almost all devices that spin or roll: power plant turbines, steering columns, wheels, skateboards, yo-yos, dentist drills, and the electric motors in everything from refrigerators and can openers to computer hard drives and CD players. In each case, the balls allow efficient, low-friction movement of rotating parts. Each sphere must be perfect, or the motion it facilitates will come to a grinding halt. Yet manufacturers produce them by the millions for pennies apiece. How do they make the balls so incredibly round and smooth?

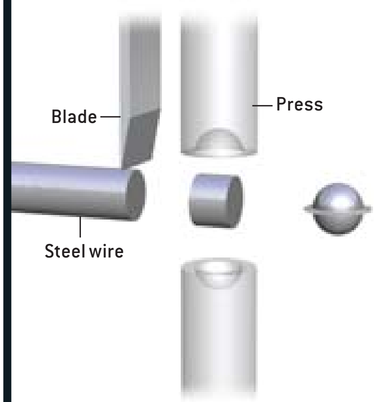
Metal, ceramic and plastic balls ranging from an ounce to hundreds of pounds share a similar genesis. An initial plug is cut from wire or extruded and is then stamped by a press into an approximate sphere. That globe is massaged into greater roundness between heavy plates and honed to exact sphericity and smoothness by rolling through grindstones hundreds, if not thousands, of times [*see main illustration, right*]. The same process creates pellets for air guns and balls for ballpoint pens and roll-on deodorant.

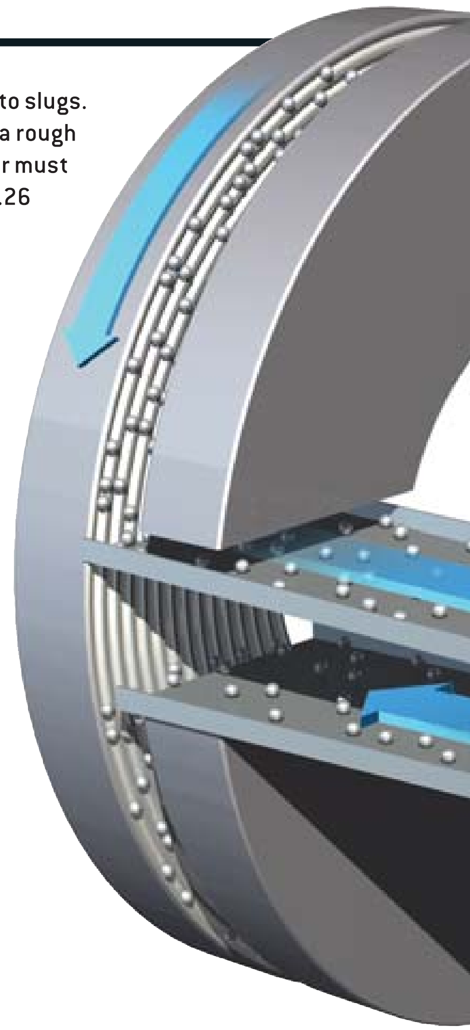
The exactitude is even more impressive given the manufacturing volume. At Hartford Technologies in Rocky Hill, Conn., 500 pounds of wire are processed in a batch, creating 300,000 balls with a 0.25-inch diameter or 2.5 million balls with a 0.125-inch diameter. It takes about a week to finish a batch. The greatest cost variable is wear on the plates and grindstones, which machinists are constantly refacing. "Our machine shop is a very busy place," says Joe Beltrami, engineering director at Hartford. A good shop is important to the bottom line. "Bearings are a high-volume, low-profit-margin business," says Billy Strickland, applications engineer at Noonan Machine Company's BGS division in Charleston, S.C.

Despite their many applications, ball bearings can withstand only so much weight or thrust, because a sphere contacts any given part at a single point. Large loads such as a car axle require roller bearings [*see illustration at far right*] that house cylinders instead of balls; the contact area is a line. And when extremely low friction under light loads is needed, designers may opt for liquid bearings [see "Did You Know ..."].

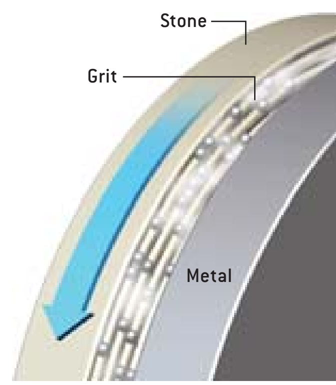
—Mark Fischetti

1 HEADING: Steel wire is cut into slugs. A press squashes a slug into a rough sphere. If the ball's final diameter must be 0.25 inch, it would be about 0.26 inch after this step.





3 SOFT GRIND: A grooved grindstone that is laced with grit applies light pressure as it rotates. Balls make several hundred passes again, taking perhaps six seconds to advance from top to bottom against a stone rotating at 80 revolutions per minute. (Diameter: 0.253 inch.) **HEAT TREAT:** Balls roll around inside a rotary furnace (*not shown*), which tempers them to a specified hardness. **HARD GRIND:** A new grindstone (*not shown*) with a finer grit sands the balls under greater pressure. (Diameter: 0.251 inch.)



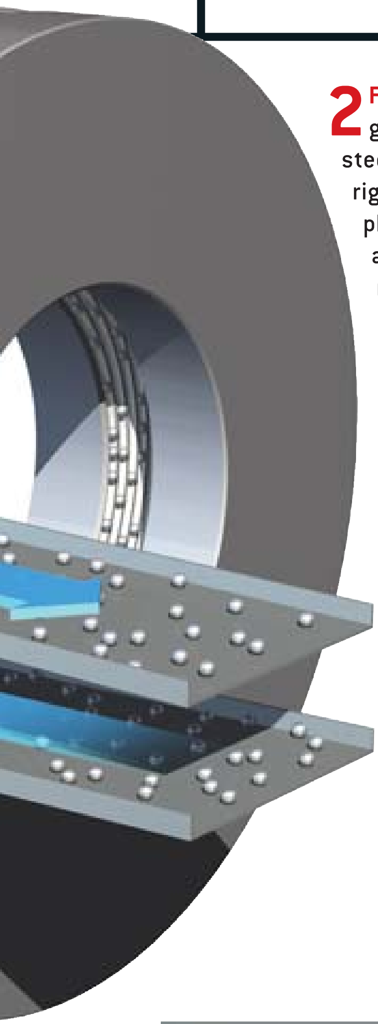
GEORGE RETSECK

DID YOU KNOW. . .

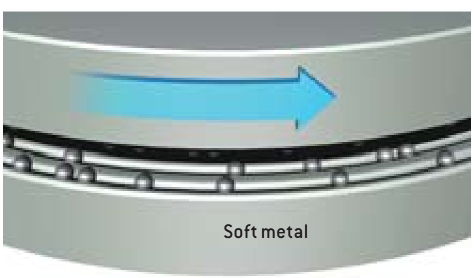
MAKE THE GRADE: A bearing ball's roundness is rated as its grade. A grade 5 ball must be spherical to within five millionths of an inch, and its diameter must be within 50 millionths of an inch. A grade 5, quarter-inch ball diameter must be 0.24995 to 0.25005 inch, and the smallest diameter must be within five millionths of an inch of the largest diameter. Common grades are 1,000 (axle of a farm combine), 100 and 25 (car steering column). "Precision" bearings are grades 10, 5 and 3 (high-speed dentist drill).

IN YOUR COMPUTER: A computer hard drive is spun by a small motor at its center (so is a compact disc in a player). Tiny, precision ball bearings let the motor rotate quickly with little friction. Yet they still generate slight noise and wobble, which limit the density with which data can be packed on the disk and still be reliably located. IBM and others are therefore moving to quieter fluid bearings—essentially, a tiny doughnut filled with liquid instead of balls.

FINE LINE: In 1943 Hungarian journalist Laszlo Biro and his chemist brother, Georg, filed a European patent for a new type of pen with a tiny ball at the tip, which capped the ink reservoir yet allowed ink to flow onto paper. Soon Biro pens were being sold. Today, as then, the rougher the ball the more ink flows; surface finish is the main difference between "fine point" and "medium point" styles.



2 FLASHING: Balls randomly enter grooves formed by two parallel steel plates, which nearly touch. The right plate is stationary. The left plate rotates, massaging the balls as if they were clay, making them rounder as they reach the bottom. There the beads exit and return for another pass. Balls make several hundred passes, ensuring uniform size even if grooves vary slightly. Water cools all parts. (Final diameter: 0.257 inch.)



4 LAPPING: Balls roll through grooved, metal plates filled with a buffing compound that removes slight high points and fine scratches. (Final diameter: 0.2502 inch.) Balls then sift through a washer for cleaning.

LIQUID FLOW in ballpoint pens, roll-on deodorant and pump-spray bottles is controlled by a bearing ball. It is snapped into a cup of the same diameter during manufacturing, but pressure on the ball causes it to turn slightly, drawing a bit of fluid.

BALL BEARING houses spheres between an inner and outer race. Spacers prevent balls from touching, and oil or grease reduces friction (bicycle wheels, inline skates).

BALL THRUST BEARING supports weight while it spins (lazy Susans, bar stools).

ROLLER BEARING uses cylinders instead of balls, providing greater contact area for heavy loads (car and train wheels, conveyor belts).

Topic suggested by reader Tim Moore.
Send ideas to workingknowledge@sciam.com

Heavenly Music in Your Hand

PORTABLE SATELLITE RADIO IS A PALMTOP CORNUCOPIA OF MUSIC BY STEVEN ASHLEY

DIRECT DOWNLOAD of music from the sky to your pocket is now possible anywhere in the continental U.S. with Delphi's XM2Go MyFi device, which receives the XM satellite radio service. A handheld that offers the competing Sirius service is to appear by year's end.

I grew up on classic FM rock radio. (Yes, I'm a dinosaur from the ancient Boomer Age.) The late 1960s and 1970s encompassed a golden era of musical diversity on the airwaves; a multiplicity of bands, styles and vernaculars appeared on eclectic, free-form playlists that were leavened liberally with B sides, live versions, alternative recordings and obscure ditties. And then there was the mystery factor: you never really knew what was cued up next, not until the first bars sounded. But when the DJ struck the right vibe for your mood with an unexpected gem, it was pure magic.

Today FM radio is a wasteland—repetitive, overprogrammed, market-niched to the max, seemingly with more commercials than music. The only strategy that makes it somewhat bearable is to keep switching among several preprogrammed stations, a frustrating enterprise that gets tough on the index finger.

Then there's the portable digital music player—iPods, MP3 players and the like. What's not to love? Nearly distortion-free recordings of your favorite songs (presented randomly if you choose) to listen to anywhere you like. The players require a bit of work to produce your personal soundtrack, but not much. Of course, there is no way around the fact that you are always playing *your favorites*. And as anyone who has owned an extensive collection of LPs, cassette tapes, CDs or iTunes knows, listening to even a large set of well-loved recordings can get boring after a while.

So now comes personal satellite radio, a medium no longer only for the car. It arrived in the form of the Delphi XM2Go MyFi—the first satellite-radio portable music player. I toted around this little marvel of miniaturization for a month to find out whether it's worth the $300 price tag and the $12.95 monthly usage fee. Built by electronics maker Delphi to receive the XM satellite radio service, the device offers 100 channels devoted to whatever musical genre you can think of, each in all its variations and derivations—rock, pop, hip-hop, country, classical, soul, jazz, Latin, world, vintage songs—with little static and commercial-free. It's like old-school FM on steroids. Add to that handy info on track titles and artists' names, streaming sports scores and stock quotes, as well as a wide selection of audio news, sports, weather and traffic reports, talk, comedy and kids' shows.

Satellite radio signals are broadcast to Earth by powerful transmitters circling in geosynchronous orbit rather than from conventional radio towers on the ground. This means one can hear the same programming on the same station anywhere in the continental U.S., a feature that is especially useful for long-distance auto trips or reception in the rural hinterlands. Until recently, however, pay radio units had been designed for use in vehicles or in stationary locations at home or the office. The Delphi MyFi is one of a new generation of handheld satellite radio receivers—a nifty gadget that represents the first real alternative to the iPod in the ongoing battle for the limited space in your pocket.

Whether the MyFi is truly pocketable is debatable, though. At 4.5 inches tall by 2.8 inches wide by 1.2 inches

MATT COLLINS

thick, the seven-ounce unit is a bit bulky to slip easily into your pocket—unless you're wearing painter's pants. It is noticeably heavier and thicker than a 20-gigabyte iPod, for example. In addition, if you put it in your pocket, you would have to attach the clip-on antenna extension wire to get a clear signal. So the MyFi is really much happier strapped to your waist in its belt clip or protective case. And with a five-hour battery lifetime, full-day operation without plugging it into a wall socket is out. In general, however, it did a fine job of delivering quality tunes whenever I took it out jogging or hiking.

Despite the device's chunkiness and limited battery life, its engineers nonetheless did an impressive job of shoehorning all the necessary components into the palmtop gizmo, including a lithium-ion battery, antenna, tuner, memory and display. The monochrome screen is easy to read, and the clickable thumbwheel control permits scrolling through programming and feature menus readily enough. And, like a conventional radio, preprogrammed buttons let you access 30 channels quickly. One cool feature is a built-in transmitter that can be set to relay the satellite radio signal to nearby FM receivers in your house or car.

The MyFi comes well equipped with a full complement of accessories for domestic and automotive use: auxiliary antennas, plug-in power supply, mounting cradles for the home and car console, various utility cables and a useful remote control. The downside to using the attachments is, of course, cable clutter.

The Delphi MyFi delivers clear audio when there is an unobstructed view of the southern sky (where the signal originates) or when it is close to one of the terrestrial repeaters XM has installed in major cities. But reception can be rather spotty when you are mobile and tall buildings or mountains block your line-of-sight downlink. Indoors, more often than not, the signal breaks up when something cuts your invisible tether to the heavens. So inside city buildings, unless you are near a window, reception is intermittent even with the clip-on antenna. At my office, I got good results by attaching the antenna to a venetian-blind slat and tinkering with the skyward aim.

Unfortunately, indoors—at the gym, at work, while shopping, or on planes or subways—is often just where you want to use a portable music machine, and windows are not always accessible. These frequent outages would be a fatal flaw but for the engineers' foresight in equipping the MyFi unit with the ability to digitally record five hours of programming, which you can play back anytime the signal fails.

NEW SETS OF EARBUDS are blossoming from listeners' ears as portable satellite radios arrive.

DELPHI

Use of satellite radio has exploded during the past year. More than 5.4 million subscribers currently pay the $12.95 a month to get service from XM or Sirius, the competing satellite radio providers. Right now XM leads with over four million users, but Sirius is making progress, having signed up 1.4 million customers. Each offers somewhat different (and incompatible) programming, so before you buy I suggest trying out the services through their Web sites (www.xm-radio.com; www.sirius.com). If National Public Radio, shock-jock Howard Stern or all Elvis all the time is your thing, you might want to wait until the end of this year, when the first handheld, Sirius-format radio is expected to arrive. XM features NPR refugee Bob Edwards, Major League Baseball, wild men Opie and Anthony, and a premium Playboy channel. Your choice. Note that Tao and Pioneer plan to roll out portable devices for receiving XM radio by year's end.

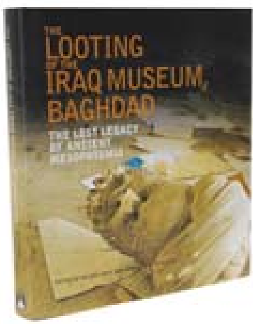
Satellite radio can lead one to develop an insatiable desire for all kinds of unfamiliar music and talk, in the way shortwave radio opened up new worlds for listeners in the past. I ended up educating myself about exotic music styles I had never much listened to before.

One disappointing aspect of the MyFi is its inability to let you produce customized playlists—you can hop between the songs and shows you have recorded, but it will play those automatically only in the order in which you saved them. Nevertheless, I found the MyFi to be a fun gadget, perhaps attractive enough to make me change my tune and buy one. But personal satellite radio is no slam dunk. The cost of admission is still a bit steep, and I might just wait around for some further miniaturization and enhancement of the technology before I think about plunking down the necessary cash.

Truth is, I'm really holding out for a combination iPod/satellite radio—what some geeks call a wireless iPod, iPod Satellite or SkyPod. This version of the "ultimate personal music player" would let you listen to your own tunes or somebody else's whenever and wherever you wanted. The Net has been rife recently with rumors and reports of discussions of this possibility between the two satellite radio companies, Apple, and the makers of MP3 players. Although this marriage would most likely require overcoming some thorny technical challenges, one can only hope that the coming of the SkyPod is in the stars. SA

The Story of the Iraq Museum

PICKING UP THE PIECES OF 40,000 YEARS OF CULTURAL LIFE BY ROGER ATWOOD



THE LOOTING OF THE IRAQ MUSEUM, BAGHDAD: THE LOST LEGACY OF ANCIENT MESOPOTAMIA
Edited by Milbry Polk and Angela M. H. Schuster
Harry N. Abrams, 2005 ($35)

We all know what happened, or think we know. When American troops entered Baghdad in April 2003, hordes of looters rushed into the Iraq Museum, repository of the world's greatest collection of Mesopotamian antiquities, and stripped the place while our GIs were busily pulling down Saddam statues for CNN.

The truth, wouldn't you know it, is a bit more elusive. About 15,000 objects were stolen, not 170,000 as first reported (actually the size of the museum's entire collection), an exaggeration resulting from misunderstandings between the first journalists to reach the shattered museum and distraught Iraqi curators. Some objects were irretrievably damaged, but nearly half those stolen have since been recovered, as museum director Donny George writes in this absorbing book. Its editors aren't interested in raking over old coals or giving a definitive account of how the looting happened. Instead they offer an eloquent, moving and abundantly illustrated history of an institution housing the remains of 40,000 years of Iraqi cultural life, from Neandertals to Ottomans.

Twenty-two writers, including curators and archaeologists, tell the story in essays that evoke the excitement of digging up the world's original civilization and a wistful nostalgia for Iraq's bygone days of field research and camaraderie. The Gulf War, U.N. sanctions and, finally, the explosion of pillage on America's watch all took a devastating toll on museums and archaeology. The only artifacts being found these days in Iraq are those dug up by looters to feed the antiquities trade, and no one in this book ventures a guess as to when, or even if, fieldwork will ever happen again. But slowly, the museum is picking up the pieces.

The Iraq Museum, as it is now known, was created in a room in Baghdad in 1923 by Gertrude Bell, a British amateur archaeologist and Arabist given to rhapsodic gurgles about the objects under her care, writing: "Isn't it fantastic to be selecting pots and things four to six thousand years old!" She died three years later, succumbing in the blazing Iraqi summer at the age of 58, and it took 40 more years for the collection to reach its present location, a hulking box of brick and cement that looks like what it is, a warehouse of history. Because Iraqi law barred the export of antiquities, the place filled with objects excavated by archaeologists all over the country, among them a stunning array of gold jewelry found in 1988 in tombs at the palace of Ashurnasirpal II at Nimrud. Scholars came from around the world, and the museum became one of the Middle East's most prestigious cultural venues.

A string of disasters began in 1991, when, during the first Gulf War, a bomb hit a government building across the street and broke the museum's glass showcases. Curators hurriedly wrapped objects in cotton and rubber padding and closed the museum, and it has remained closed almost uninterruptedly since. The gold of Nimrud was packed off to underground vaults at the Central Bank, which were subsequently flooded, possibly intentionally by Iraqis intent on preventing Saddam diehards from stealing it. (The vaults were final-

WOMAN'S HEAD, from the ninth or eighth century B.C., was thrown down a well during an attack on Nimrud, near present-day Mosul, in 612 B.C. Paradoxically, this traditional way of destroying enemy goods preserved the head. It suffered severely, however, when the storeroom housing it was flooded during the sack of Baghdad in 2003. (Stained ivory, 16.1 centimeters high.)

COURTESY OF HARRY N. ABRAMS

ly drained in 2003 and the gold recovered intact.) By the eve of the current war, the museum was a sad and demoralized place, its employees hunkered down behind steel gates bracing for the next disaster to strike. The museum's story of early promise erased by war and Saddam's megalomania becomes here a kind of metaphor for the recent history of Iraq.

Even after the looting, no institution in the world can tell the story of writing like the Iraq Museum. Cuneiform, the world's first script, was born in southern Iraq, and carbon dating indicates it originated between 3400 and 3300 B.C., writes Robert Biggs in one of the book's finest essays. There must have been quite a burst of innovation, because within a century or two the wheel appeared as well. It was quickly put to use in war, on chariots pulled by recently domesticated donkeys. Cuneiform found its first use in record keeping: receipts for barley bales, notices of gold shipments, more the work of accountants than poets. But before long, people went wild for cuneiform. Clay tablets with its spindly arrangements of flicks and crosses started to appear by the thousands, recording paeans, epics and incantations.

Cuneiform tablets became so common in ancient Iraq that they were used as packing material in building foundations and tossed into trash pits with animal and fish bones. In the 1980s archaeologists found a library of 800 tablets arranged on their shelves at a site called Sippar and sent them to the Iraq Museum, where they were widely and mistakenly reported to have been lost in the 2003 looting. The museum currently holds more than 100,000 tablets, and thousands more circulate elsewhere. Biggs recounts how the Chicago department store Marshall Field's was selling cuneiform tablets from Ur for $10 each as late as the 1960s.

Bad as the theft was, something even worse was happening. Journalists Micah Garen and Marie-Hélène Carleton surveyed sites invaded by bootleg diggers after Saddam fell, and their account in this book suggests not so much looting as industrial-scale leaching. Hundreds of men were digging for treasure, by day and by night with shovels, generators, lightbulbs and trucks. Five Sumerian cities (there are only 18) have had the top nine feet of their surfaces completely sifted by looters, an "unimaginably grim reality, a scene of complete destruction," they write. Just as shockingly, Columbia University's Zainab Bahrani writes that American troops have set up camp atop the ruins of Babylon, removing layers of archaeological material to create a helipad and laying a parking lot on the remains of a Greek theater dating from Alexander's day. Iraqi authorities asked the troops to move, but as of September 2004 they were still there.

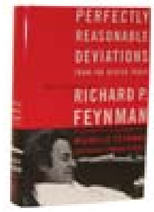
The writers of this book try now and then to sound optimistic but, like grieving widows, keep slipping back into despair. It can get a bit weepy, this "requiem for a departed companion," as one calls it. Still, there is plenty to weep about. Two centuries of research into Mesopotamian civilization have been stopped in their tracks by war, looting and lawlessness. A stone excavated at Nippur carries a long invocation to the goddess Inanna to protect a temple and ends with a humble plea to mortals: "The governor who keeps it permanently in good condition will be my friend." Whoever wrote those words wouldn't have many friends now. SA

Roger Atwood, author of Stealing History: Tomb Raiders, Smugglers, and the Looting of the Ancient World *(St. Martin's Press, 2004), visited Iraq in 2003 and now lives in Venezuela.*

THE EDITORS **RECOMMEND**

PERFECTLY REASONABLE DEVIATIONS FROM THE BEATEN TRACK: THE LETTERS OF RICHARD P. FEYNMAN

edited and with an introduction by Michelle Feynman. Basic Books, 2005 ($26)



This book is more than the sum of its parts, which is to say, much more than a collection of letters. It provides a trajectory of Feynman's life: poignant letters to his first wife, Arline, who died of tuberculosis shortly after they were married; funny, high-spirited letters to Gweneth, whom he married after he was already fairly well known; letters regarding his 1965 Nobel Prize, his resignation from the National Academy of Sciences, his advice to the California State Curriculum Committee.

And it conveys a genuine sense of the man—someone who took his principles very seriously but never himself. "He showed us all how to look at the world. He showed me how to laugh," writes his daughter, Michelle, in the introduction. Throughout she furnishes exactly the right amount of context, in an unself-conscious style that, one imagines, would have pleased her father. She is probably responsible, too, for the inclusion of some wonderful candid photographs.

MARTIN GARDNER'S MATHEMATICAL GAMES: THE ENTIRE COLLECTION OF HIS SCIENTIFIC AMERICAN COLUMNS

Mathematical Association of America, 2005 (compact disc, $54.95)

Gardner's column, which ran in this magazine from 1956 to 1986, introduced thousands of readers to the delights of puzzles and problem solving. His column also broke important mathematical stories—on cryptography, fractals, the game of *Life,* and tilings. Now all these columns are gathered on one, searchable CD, ending frustration for many fans. Gardner, who turns 91 this month, continues to write on a variety of topics.

Kha-Nyou Smell a Rat?

TRACKING DOWN REAL ROCK RODENTS AND IMAGINARY AMAZON APES BY STEVE MIRSKY

On May 20 I was in my office "working," a euphemism for surfing the Web, when I noticed that two articles from very different news sources had a common theme. The stories brought up many questions. Was life imitating art? Was art actually cribbing from life? If I asked my tax adviser's opinion during lunch, was the meal deductible?

In fact, lunch was what connected the two news pieces. The first story, which appeared in the *Australian* and was datelined May 19, carried the title "KEBAB MEAT RODENT A NEW SPECIES." The lead paragraph noted that "an odd-looking rodent, spotted in a food market in Laos where it was going to be turned into a kebab, has turned out to be not only a new species but also the first member of a new family of mammals to be identified in more than three decades." The rodent is known locally as kha-nyou, as well as rock rat. The article also stated that "the discovery was made by Robert Timmins, a member of the New York–based Wildlife Conservation Society (WCS), who spotted a dead rock rat as it was about to be grilled."

I thought of comedian Lewis Black's take on Europeans seeing North American fauna for the first time—"Holy &!@%, look at that! What the @!#% is *that?* [Pause.] Let's eat it."—and continued Web surfing. The next stop was the *Onion,* which calls itself "America's Finest News Source," despite, or because of, the fact that it simply makes its news up. I was thus startled to see the headline "NEW, DELICIOUS SPECIES DISCOVERED." The *Onion* story, datelined May 18 and written in flawless wire-service style, began, "Manaus, Brazil—An international team of scientists conducting research in the Amazon River Basin announced the discovery of a formerly unknown primate species" that was "an amazing biological find" and "tastes wonderful with a currant glaze."

Further research turned up the fact that the WCS had issued a press release about the kha-nyou on May 11 and that the *Scientific American* Web site actually mentioned the find on May 12. (Note to self: Holy &!@%, *Scientific American* has a Web site: www.sciam.com! Let's read it.) Did the *Onion* stoop to relying on real news for its ideas, basing its Brazilian primate, "informally known as the delicacy ape," on the kha-nyou?

Answering this question was going to require true investigative journalism (which stinks because it really *is* work). A call to the *Onion* revealed that they had come up with the delicacy ape piece well before the kha-nyou news broke and that the timing of publication was purely coincidental. "But we could make up a story if you want us to," my source there said, in finest *Onion* fashion.

I next called Robert Timmins at his home in Madison, Wis., and found out that he had in fact beaten the *Onion* by years—he first spotted the kha-nyou in 1996. "It was very early, just after dawn," he recalled. "I was at a fresh food market, where everybody brings in their vegetables from farms, animals from the forest, fish from the river." The kha-nyou was for sale next to some vegetables. "I knew immediately it was something I had never seen before," Timmins said.

It took two years to get specimens out of Laos and another seven for Paulina Jenkins of London's Natural History Museum and C. William Kilpatrick of the University of Vermont to do a complete scientific analysis of the creature and prepare a long, detailed paper for publication in the journal *Systematics and Biodiversity.* (Taxonomy is *really* work.)

They gave the kha-nyou the Latin name *Laonastes aenigmamus,* in the new family Laonastidae. *Laonastes* translates to "inhabitant of the rocks of Laos." *Aenigmamus* means "riddle mouse," which "alludes to the enigmatic taxonomic position of this rodent," the journal article explains. Is it closest to the mole rats, or porcupines, or even chinchillas? That's unclear. Where would the kha-nyou itself prefer to be located? Just not next to the vegetables, thank you. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

ASK THE EXPERTS

What causes headaches?

—Mike A., Wilmington, Del.

Dawn A. Marcus, associate professor at the University of Pittsburgh School of Medicine's department of anesthesiology, offers this answer:

Although they may feel as if they emanate from the brain, headaches actually arise as a result of irritation in nearby structures: skin, joints, muscles, nerves or blood vessels. Brain tissue, encased in the protective coating of the skull, has not evolved the ability to respond to pressure sensations.

Clinicians classify all headaches as either secondary or primary. Secondary headaches, which appear as symptoms of an underlying disorder, have no uniform cause. Anything from a pinched nerve to a sinus infection can lead to secondary head pain.

Most headaches, however, are primary, meaning that the headache is not a symptom of another condition but the problem itself. Research suggests that this type—which includes tension headaches and migraines—may derive from a single, identifiable pathway.

The chain begins when pain centers in the brain are activated, at which point they produce neurotransmitters such as serotonin and norepinephrine. These chemicals call for expansion of meningeal blood vessels enveloping the brain, resulting in increased blood flow. As the vessels swell, they stretch the nerves that surround them. These nerves, in turn, convey signals to the trigeminal system, an area of the brain that relays pain messages for the head and face, and we perceive pain.

Why the pathway is initiated at all is still an open question, although some circumstances seem to make headache onset more likely. These triggers may be internal (for example, hormonal changes during menstruation) or environmental (such as stress or sleep deprivation). So far, however, most evidence for what factors are directly responsible is anecdotal, and the mechanism by which the triggers are converted to chemical signals is little understood.

How can a poll of only 1,004 Americans represent 260 million people?

—C. Bures, Wellesley, Mass.

Andrew Gelman, professor in the departments of statistics and political science at Columbia University, explains:

You can learn a lot about a large population from a smaller cross section—but that does not make the technique flawless.

Mathematically, the margin of error depends inversely on the square root of the number of those sampled; however, the margin of error is an abstraction based on tacit assumptions. In practice, actual errors may be larger than advertised.

One assumption is that the queried group is truly random—that respondents have been chosen one at a time, with everyone in the U.S. equally likely to have been picked. To approximate this ideal, polls use telephone numbers generated randomly by a computer. But if you do not have a phone, you will not be in the survey, and if you have two lines, you have two chances to be included. Another confounding factor is that women, whites, older people and college graduates are more likely to agree to be interviewed. Statistical weighting helps pollsters match the sample to the population, but they can counter only known biases.

Finally, any margin of error is an understatement, because opinions change. For instance, surveying 4,000 people would improve the margin of error to 1.5 percent. Although this sounds appealingly precise, it is generally a waste of time, because public views vary enough day to day that it is meaningless to attempt too exact an estimate. It would be like getting on a scale and measuring your weight as 173.26 pounds; after you drink a glass of water five minutes later, your precise weight would have changed but to an unimportant degree. SA

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

MATT COLLINS