CURING RABIES • EYE MOVIES: WHAT THE RETINA SEES

SCIENTIFIC AMERICAN

Cannibal Galaxies: Tearing Apart the Neighbors

APRIL 2007
WWW.SCIAM.COM

THE DAZZLING FUTURE OF PLASMONICS

New optical technology yields faster computing, brighter LEDs … oh, and **invisibility**

- Storing Hydrogen Fuel
- Genetics of Alcoholism
- Raven Intelligence

april 2007

contents features

SCIENTIFIC AMERICAN Volume 296 Number 4

SCIENTIFIC AMERICAN Volume 296 Number 4

departments

Sonja Lyubomirsky, University of California, Riverside **36**

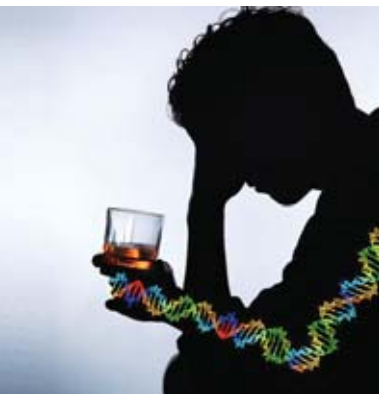
columns

Cover image by Phil Saunders, Space Channel Ltd.; photograph at left by Brad Swonetz.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. 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.

Caution: Sharp Edges

Even as information technology has been driving striking social transformations over the past decade, a new form of information with similar transformative potential has also been coming online: the kind encoded in the DNA sequence of the human genome. Beginning on page 46 of this issue, John I. Nurnberger, Jr., and Laura Jean Bierut describe how the power of genetic analysis is even beginning to reveal the interactions of genes, environment and personal choice that make a person more or less vulnerable to a disorder as complex as alcohol addiction. Their investigations are based on data collected in a massive ongoing national study of alcoholics and their families, all of whom were willing to provide both DNA and detailed personal accounts to help researchers gain insights into this devastating illness.



GENE DATA reveal risks and raise new ones.

Those kinds of volunteers are crucial to the study of how genes work, but fear could make them increasingly hard to find. The National Human Genome Research Institute reports that already about a third of the people approached to participate in genetic studies decline, worried that discovery of a genetic predisposition to illness could later subject them to discrimination from insurance companies or employers. Genetic counselors say that a large proportion of people tested for inherited disease risk factors hide that fact from their insurers for the same reasons.

At press time, the Genetic Information Nondiscrimination Act (GINA), which would bar prejudicial treatment based solely on gene data, was appearing to finally have enough momentum to become law. This bill has been bogged down in the House of Representatives twice in recent years, despite unanimous passage by the Senate both times and support from the White House. Opposition has come largely from business interests, which deride GINA as feel-good legislation that duplicates protections already offered by states and existing federal antidiscrimination laws such as the Americans with Disabilities Act (ADA).

If healthy people with a higher genetic risk for diabetes, for example, could be said to already have the disease, they might well be covered by the ADA. But with rare exceptions, genetic predispositions are no guarantee of illness, and the medical promise of genetic screening is to identify vulnerabilities before sickness arises so as to enable informed choices about risky behaviors, preventive monitoring and medication. If explicit federal protections make people feel better about participating in the genetic information age, it is therefore worthwhile to spell them out.

Passage of GINA will not end controversy over the proper uses of personal genetic data. Increasing access to genetic screening will surely influence debates over universal health care and medical records privacy. Beyond medicine, DNA is contributing to studies of human evolution, the creation of family genealogies, the identification of criminals and the exoneration of the wrongly convicted. Our society is already beginning to enjoy the fruits of this new form of information. But scurrilous entrepreneurs are also, predictably, beginning to offer unreliable DNA tests or bogus interpretations of the data, and more such abuses can be expected. As with other forms of personal information, over which we are willing to trade some measure of control for benefits, such as fast-track passes through airport security, genetic information will offer powerful rewards and some risks. Like any double-edged sword it should be handled carefully.

THE EDITORS editors@sciam.com

SCOTT GIBSON *Corbis (silhouette)*; TIM EVANS *Photo Researchers, Inc. (DNA)*; EMILY HARRISON *(photoillustration)*

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

West/Northwest
818-827-7138
fax: 818-827-7139

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The Netherlands
Insight Publicitas BV
+31-35-539-5111
fax: +31-35-531-0572

Scandinavia and Finland
M&M International Media AB
+46-8-24-5401
fax: +46-8-24-5402

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:

SPECIAL REPORT

16 Mind-Blowing Microphotographs of Living Things

Prize-winning microviews of everything from mouse retinas to slime molds, from the annual Olympus BioScapes International Digital Imaging Competition.

NEWS

Final Report: Humans Caused Global Warming

For the first time, a panel of climate experts has confirmed that global warming is occurring and that it is "very likely"—90 percent certain—man-made.

How Hallucinogens Play Their Mind-Bending Games

Researchers isolate cells affected by LSD and mescaline, potentially leading to more treatments for neurological and psychiatric disorders.

FACT OR FICTION?

Premium Gasoline Delivers Premium Benefits to Your Car

Exploding the myth that premium gasoline delivers better performance in the average automobile.

PUZZLE

Sci-Doku

Try your hand at Sci-Doku, a Sudoku puzzle that uses letters instead of numbers, with an added twist: a science-related clue accompanies each puzzle, and the answer is spelled out in one row or column of the solution.
Only at www.sciam.com

New Podcasts

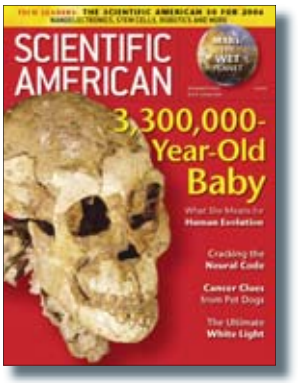
Every weekday: *60-Second Science*
Every Wednesday: *Science Talk*
Listen at www.sciam.com/podcast

2006 OLYMPUS BIOSCAPES INTERNATIONAL DIGITAL IMAGING COMPETITION

SCIENTIFIC AMERICAN®
Established 1845

Letters

EDITORS@SCIAM.COM



OUR SCIENTIFIC AMERICAN 50 award for Policy Leader of the Year to Al Gore for spreading awareness of climate change was the most controversial item of the December 2006 issue. Some correspondents were wholly dismissive of Gore's efforts; others felt he erred in oversimplifying a complicated subject. "The earth's atmosphere is very complex, yet Gore asserts that climate warming is due to one factor: the amount of carbon dioxide in the atmosphere," wrote H. G. Stuebing of Doylestown, Pa. "It seems only prudent to reduce CO_2 emissions, but it is not clear that this will solve the problem."

Questions of scientific consensus regarding other topics in the issue inspired respondents to pontificate on the legitimacy of subliminal persuasion, the definition of an "industrial revolution," and the moral implications of emergent information-scanning technology.

SUBLIMINAL SUBSTANCE

"Voting with the Heart," by Charles Q. Choi [News Scan], details results of experiments based on subliminal messages, conducted by Charles Taber of Stony Brook University. Before you report such results, you must first be able to verify that subliminal messages are valid. But to date, no reputable scientist has found subliminal persuasion to be effective. There is no indication that Taber's experiments have been published in a peer-reviewed journal. What next, a report of an experiment based on voodoo?

Ron Shattil
Oakland, Calif.

TABER REPLIES: Several misperceptions persist in the popular imagination about "subliminal messages." Some say it is voodoo science. Others insist that people can be persuaded to buy products, vote for candidates or imitate chickens against their will. The scientific reality lies between these extremes. Contrary to Shattil's assertion, for decades it has been well established in peer-reviewed studies that stimuli presented below the threshold of awareness influence a variety of cognitive and emotional processes. Recent research uses the subliminal presentation of stimuli to test theories about automatic cognitive processing.

A search using the term "subliminal" on the American Psychological Association's PsycINFO Web site (http://psycinfo.apa.org/psycinfo/), the most prominent research database for peer-reviewed psychology, will return over 1,500 recent articles. The same tool will also turn up some of my own peer-reviewed articles on automaticity in political information processing.

TOTALITARIAN TAGGING

In **"Smart Tags Get Smarter"** [Scientific American 50], Mark Alpert seems to think the main use of new radio-frequency identification (RFID) technology would be to replace bar codes. What he should have noted is that while scanning the contents of your grocery cart, an RFID reader could also scan your purse or wallet. Based on the items you are carrying, store management could discern where you live, what kind of car you drive, your age, your gender, et cetera.

One of the curses of pure science is that it presents itself as not being responsible for what others do with new discoveries. If you thought George Orwell's *1984* was scary, you ain't seen nothin' yet!

John Olsen
Errington, Vancouver, B.C.

CRYSTALLIZING THE KILO

"Weighty Matters," by Ian Robinson, suggests that the kilogram might be defined as a fixed number of silicon 28 atoms. The arrangement of these atoms, however, changes the mass of the sys-

Page Intentionally Blank

tem, according to special relativity. A crystal, for example, has less chemical energy and mass than a cloud of dispersed atoms because the energy (and therefore equivalent mass) used to bind its atoms together must be subtracted from the sum of their masses (and equivalent energy).

Is this difference negligible? Is the kilogram, for all practical purposes, a fixed number of atoms, no matter how they are chemically arranged?

Karl Dahlke
Troy, Mich.

TOWARD A MORE UNIVERSAL DEFINITION: Scientists are currently looking to replace the single metal cylinder used for more than a century as the basis of the kilogram.

COURTESY OF NATIONAL PHYSICAL LABORATORY, U.K., © CROWN, 2005

ROBINSON REPLIES: *The mass of a silicon crystal is less than the sum of the masses of its constituent atoms, but the effect is negligible from the point of view of present efforts to redefine the kilogram, as it is much less than one part in a billion. Any definition of the kilogram made in terms of the mass of an atom would, however, take this effect into account. In a similar vein to the existing International System of Units definition of the mole (the amount of an element that has a mass in grams equal to the element's atomic weight), a kilogram definition would specify that the atoms are unbound, at rest and in their ground state.*

Those making the measurements would then have the problem of relating the masses of the idealized atoms in the definition to those in the crystal artifact that would be created for practical measurements. The issue does not arise if the kilogram is redefined by fixing the value of Planck's constant (one of the principal constants of quantum mechanics) rather than the mass of an atom of a chosen element.

MANUFACTURED REVOLUTION

As evidenced in Rodger Doyle's article, "Not So Revolutionary" [By the Numbers], a conflict seems to exist between what people in manufacturing and others call an "industrial revolution."

To those in manufacturing, the first industrial revolution was the mecha-

nization of various industries by the use of steam-powered machine tools to make goods, and the second one was the addition of computer control to those tools to make more complicated products with greater accuracy and at a lower cost.

The historians, who coin such phrases as the "nth industrial revolution," are unaware of the manufacturing advances necessary to make modern products. Thus, they look at some combination of socioeconomic and technological progress in making their definitions.

In the past 60 years we have seen unremitting progress in making products smaller, more powerful, more efficient, and so on. Underlying most of this progress is the coupling of computers to the manufacturing process to control machine tools.

Steve Jasik
Menlo Park, Calif.

AGING AND CANCER

In "Cancer Clues from Pet Dogs," David J. Waters and Kathleen Wildasin illustrate that both humans and dogs tend to get cancers late in their lives. If mesothelioma takes 30 years to develop in humans but only eight in dogs, is this an indication that the aging process or breakdown of the immune system is a major, even dominant, factor in cancer development?

Paul Allison
Los Alamos, N.M.

WATERS REPLIES: *Most cancers develop in old tissues, implying that the aging process may indeed be a dominant factor in the genesis of tumors. But the precise role aging plays in cancer formation is still unknown, because the intersection between aging and cancer is a surprisingly underexplored territory. The epidemic of cancer and other age-related diseases we share with our pets is an unfortunate, but expected, by-product of our domesticated lifestyles, which favor longevity.*

A silver lining is that domestication's curse offers comparative oncologists a precious field laboratory of naturally occurring cancers that may help explain the aging-cancer connection.

CLARIFICATIONS "Hot Commodity," by Mark Fischetti [Working Knowledge], states that lithium burns when exposed to air. The elemental form of lithium is unlikely to do so, because it forms a nitride directly from the nitrogen in the air. Nevertheless, overheated batteries can ignite materials around them and batteries should never be opened, because most contain hazardous materials.

"Lucy's Baby," by Kate Wong, uses the term "hominin" to describe all the creatures in the human line since it diverged from that of chimpanzees. Although "hominid" was once used for this purpose, the scientific community now favors "hominin."

50, 100 & 150 Years Ago
FROM SCIENTIFIC AMERICAN

Obituary for Parity ▪ Invitation to Flight ▪ Strict Quarantine

APRIL 1957

GRAFTED TISSUE—"We discovered that the power to react against homografts could be prevented from developing if we injected an animal at a very early age with cells from the donor strain—most conveniently cells of the spleen. In adult mice the injection of such cells increases the mouse's resistance to a graft from the donor. But if the spleen cells are injected in a mouse in the fetal stage or very shortly after its birth, the opposite happens: the mouse becomes tolerant of grafts from the strain that provided the spleen cells, though it remains intolerant of homografts from mice of other strains. —P. B. Medawar" [*Editors' note: Medawar shared the 1960 Nobel Prize in Physiology or Medicine for his work on acquired immunological tolerance.*]

END OF PARITY—"There were two mesons, called tau and theta. Tau, in the course of time, disintegrated into three pi mesons; theta, into two pi mesons. What was baffling was that in every property except the mode of decay, tau and theta were identical twins. Could they be one and the same particle? Decay of a particle by two different modes was certainly permitted by theory and precedent, but in this case the principle of conservation of parity stood in the way. Tsung Dao Lee of Columbia University and Chen Ning Yang of the Institute for Advanced Study boldly faced up to an embarrassing but insistent possibility: perhaps the parity-conservation law simply broke down in the realm of particle decays like tau's and theta's! —Philip Morrison" [*Editors' note: Lee and Yang won the 1957 Nobel Prize in Physics for this work.*]

APRIL 1907

PRIZE FOR FLIGHT—"Despite the fact that very many inventors throughout the United States are wrestling with the problem of aerial navigation by means of a true dynamic flying machine, that is, a machine heavier than air, no public flight has been made in this country with such a machine up to the present time. The most advanced knowledge of heavier-than-air navigation seems to be held by two young western experimenters [the Wright brothers], of whom much has been written. As the SCIENTIFIC AMERICAN is the oldest journal in this country treating of Science and the Arts, its proprietors feel that it is fitting that this journal should be the first to encourage the development of the latest great invention and have, therefore, decided to offer a valuable trophy for competition for heavier-than-air flying machines—the Scientific American Aeronautic Trophy."

MERMAID TECHNOLOGY: Special effects for an aquatic opera, 1907

WATERY ILLUSION—"'Neptune's Daughter,' a romantic operatic extravaganza, depends entirely on the great cistern at the New York Hippodrome for the now famous mermaid scene. It is difficult to call this attraction either an illusion or an effect; in truth it is very real, for the mermaids appear at the surface and dive down at will. This very clever act is the invention of H. L. Bowdoin, of New York City, who conceived the idea of utilizing the principle of the diving bell. At the Hippodrome individual diving bells are used. Each bell is provided with an operator, who raises and lowers a little individual lift secured to the outside of the chamber, and who also assists the mermaid in re-entering the air cell after her performance [*see illustration*]. The mermaids are protected from the cold by rubber undergarments."

APRIL 1857

BARRING SICKNESS—"The introduction of the yellow fever into the garrison on Governor's Island, harbor of New York, is attributed to the arrival of invalid soldiers from the forces in Florida. The illness and death of their comrade at Morris' Island is satisfactory on this point, and shows that the rigorous execution of the quarantine law excluded at least one case from the city; and we have every reason to believe, if it be capable of excluding one source of disease, it is capable of excluding many. Merchants, as well as citizens, will believe that there is some virtue in well-regulated and vigorously executed quarantine laws."

GLOBAL WARMING

Conservative Climate

CONSENSUS DOCUMENT MAY UNDERSTATE THE CLIMATE CHANGE PROBLEM **BY DAVID BIELLO**

Paris—The signs of global climate change are clear: melting glaciers, earlier blooms and rising temperatures. In fact, 11 of the past 12 years rank among the hottest ever recorded. After some debate, the scientists and diplomats of the Intergovernmental Panel on Climate Change (IPCC) issued their long-anticipated summary report in February. The summary describes the existence of global warming as "unequivocal" but leaves out a reference to an accelerated trend in this warming. By excluding statements that provoked disagreement and adhering strictly to data published in peer-reviewed journals, the IPCC has generated a conservative document that may underestimate the changes that will result from a warming world, much as its 2001 report did.

ICEBERG from Greenland's Ilulissat Kangerlua glacier shows the influence of climate change.

PAUL SOUDERS *Corbis*

More than 2,000 scientists from 154 countries participated in the IPCC process, which will release three more reports this year. This first report examined only the physical science of climate change. Scientists drafted as lead authors prepared chapters on subjects ranging from a historical overview of climate change science to regional projections. Governments and other reviewers then submitted more than 30,000 comments. Finally, the lead authors and diplomats gathered in Paris to review the final document word by word, changing an emphasis here ("unequivocal" triumphed over "evident") or leaving out a controversial finding there.

For example, after objections by Saudi Arabia and China, the report dropped a sentence stating that the impact of human activity on the earth's heat budget exceeds that of the sun by fivefold. "The difference is really a factor of 10," says lead author Piers Forster of the University of Leeds in England: compared with its historical output, the sun currently contributes an extra 0.12 watt of energy for each square meter of the earth's surface, whereas man-made sources trap an additional 1.6 watts per square meter.

The document's conservatism also reflects the nature of climate change science. Various models running different scenarios predict sea-level rise as little as 18 centimeters (seven inches) or as much as 59 centimeters (23 inches). None of these models, however, completely includes the potentially greater contributions to such a rise from the melting of glaciers in Greenland and Ant-

arctica. Climate modelers do not include effects on land-based ice in these regions because they cannot reduce them to equations, such as *x* amount of extra heat equals *y* amount of melting.

Greenland's glaciers are melting and moving faster on average, but those shifts do not follow a simple, upward linear trend. For example, Kangerdlugssuaq glacier has lost mass from melting and, in its thinner form, has less weight to speed the flow of its ice toward the sea. Additionally, roughly 80 percent of its recent increase in water discharge occurred in just one year before stabilizing, according to Ian Howat of the University of Washington. As glaciologist Richard Alley of Pennsylvania State University notes: "The ice sheet is losing mass, this loss has increased over time, [and] it is not the dominant term in sea-level rise—but it matters." In fact, many variables come into play in Greenland's ice sheet. "You're trying to figure out what is going on with an immense, remote and complex beast, and it isn't easy," Alley adds.

And other important factors, such as the convection that forms thunderstorms, can only be approximated because they occur on too small a scale. "There is no way that the models are able to directly simulate these things," says climate modeler Stephen Zebiak of Columbia University. "So researchers just try to capture the net effect of the processes."

Despite these flaws, global models are increasingly credible: when fed the factors at play in climate over the past 100 years, they accurately match what has been observed to occur. Such precision gives scientists greater confidence in their ability to assign probabilities to the future. And all models agree that the world will warm at least 0.4 degree Celsius in the next 20 years.

This month the IPCC releases its second report, which focuses on global warming's impacts, ranging from intensifying droughts to heavier downpours and other extreme weather events. The third report—due out in May—will discuss options for mitigation, such as alternatives to fossil fuels. In the U.S., the commitment to such alternatives remains precarious—the budget for biofuel and hydrogen research has risen, but funding for other renewable energy sources has declined. And all that has been budgeted for such research represents less investment than the U.S. made in the 1970s.

"There is at least a perception among at least some students," Alley notes, "that the support for the search for solutions to energy and global warming is not yet reliable enough for those students to commit their future to it." Given the conservative IPCC estimates, the need for such solutions seems evident.

news SCAN

THE CLIMATE CRYSTAL BALL

Despite remaining uncertainty, global climate models have become increasingly precise and have enabled researchers to establish with confidence "the range of different scenarios and what their likelihoods are," says climate modeler Stephen Zebiak of Columbia University.

The models agree that the world will warm 0.4 degree Celsius in the next 20 years. And even if the emission of greenhouse gases stopped today, the world would warm 0.6 degree C by the end of the century thanks to heat already absorbed by the ocean and the long-lived effects of carbon dioxide in the atmosphere. If CO_2 concentrations double from preindustrial levels of roughly 280 parts per million (ppm) in the atmosphere, then the world's temperature will rise three degrees C, according to the best estimate of the Intergovernmental Panel on Climate Change. By the end of 2005 the atmosphere bore 379 ppm of CO_2—a level not seen in at least 650,000 years.

ALTERNATIVE MEDICINE

The Autism Diet

CAN AVOIDING BREAD AND MILK EASE THE DISORDER?

BY MARK ALPERT

If you can believe the many testimonials posted on the Web, a diet free of gluten and casein is a miracle treatment for autism. Parents of children suffering from the disorder, which is characterized by impaired social and communication skills, fervently describe astounding improvements that occurred as soon as they removed gluten (a mixture of plant proteins found in wheat, rye and barley) and casein (the main protein in dairy products) from their kids' meals. Surveys indicate that as many as 40 percent of children with autism have been placed on special diets at one time or another. This enthusiasm is grounded more in hope than in science; so far researchers have no good evidence that dietary interventions can alleviate the symptoms of autism. Recently, however, investigators have launched the first rigorous tests of the diets, and the results may be available within a year.

The assumption behind the diets is that people with autism often have gastrointestinal abnormalities that allow unusual

FOOD CHOICE: Families of autistic people want to know if certain diets can alleviate symptoms.

KAT WADE *San Francisco Chronicle*/Corbis

amounts of digestive by-products into the body (the so-called leaky gut syndrome). The by-products of gluten and casein, according to one hypothesis, disrupt brain function by altering opioid activity, which is involved in pain regulation and social bonding. Another theory posits that the gut leakage triggers a harmful immune response. These hypotheses are far from rock-solid; in fact, scientists have not even confirmed that people with autism have a higher-than-normal incidence of gastrointestinal problems. But the causes of autism are so poorly understood and the disorder is so variable that some investigators are willing to consider the possibility that gluten and casein may somehow exacerbate symptoms in some children, perhaps just by producing intestinal discomfort.

Unfortunately, the initial studies of diets that eliminate gluten and casein were badly flawed. Although half a dozen research groups reported improvements in behavior and cognition in autistic children after several months on the elimination diets, nearly all the studies lacked control subjects, individuals who continued to digest the suspect proteins. Because the researchers did not compare the restricted-diet children with a control group, they could not specify whether the behavioral and cognitive gains actually resulted from the diets, from the children's maturation or from other therapies conducted at the same time.

The new studies, in contrast, involve control subjects and have a double-blind design: neither the researchers nor the parents will know whether the autistic children are consuming gluten or casein, so the evaluations of the children's behavior will not be tainted by wishful thinking. In a study led by Robin Hansen of the University of California, Davis, all participants go on a gluten-free diet for two months; then, for the next two months, half the subjects eat daily snacks containing gluten while the other half get indistinguishable gluten-free snacks. Susan Hyman of the University of Rochester is running a similar study testing the behavioral effects of both gluten and casein. An investigation at the University of Pittsburgh Medical Center will monitor the effects of combining the gluten-free, casein-free diet with supplements of omega-3 fatty acids, another popular but unproved therapy for autism.

The researchers have run into some trouble recruiting autistic subjects. Many parents who are committed to the gluten-free, casein-free diet do not want to participate because their children may be included in a control group and receive the offending substances. Hansen says she originally hoped for 60 subjects but now expects 30. Her results should be available in about six months. "It's a hard study to do, but it's worth doing," says Susan E. Levy, director of the Regional Autism Center at Children's Hospital of Philadelphia. "I'm hopeful that we'll know a lot more in a year."

REVERSING THE DAMAGE

Although research on diet and autism is still in its infancy, scientists are making great strides in understanding the genetic roots of the disorder, which strikes about one in 150 U.S. children. Mutations of a gene called *MECP2* cause Rett syndrome, the most physically disabling of the conditions in the autism spectrum.

A team led by Adrian Bird of the University of Edinburgh in Scotland recently showed that reactivating *MECP2* in mice reversed their neurological damage, eradicating tremors and restoring normal breathing in animals that were, in some cases, only days from death. Many genes are involved in causing the more common forms of autism, but the new research holds out the hope that scientists may someday develop therapies that could repair the brains of children suffering from the disorder.

BIOLOGY

Selfless Giving

MOM'S BRAIN CHEMICAL AFFECTS EMBRYONIC DEVELOPMENT **BY GARY STIX**

A mother gives everything to care for her child.

Biologists have recently demonstrated the truth of that greeting-card adage at the molecular and cellular levels. It comes as no surprise that the prenatal environment supplies nourishment, hinders embryonic development in the presence of metabolites from smoke or drink—and can even influence which of a child's genes get turned on or stay silenced. Now French researchers have found that biochemicals from the

mother can get involved directly in the development of the embryo and fetus. Besides blurring the physiological borderline between mother and embryo, the "maternal effect" may help reveal the causes of several diseases.

Jacques Mallet and his co-workers at the French National Scientific Research Center (CNRS) in Paris have shown for the first time in mammals that serotonin is supplied by the mother before the fetus has the capacity to make its own. Serotonin serves as either a neurotransmitter or a hormone that regulates the gastrointestinal system, mood, the sleep-wake cycle, the cardiovascular system, pain perception and appetite, among other functions.

For the experiment, reported in the January 2 *Proceedings of the National Academy of Sciences USA* (*PNAS*), the team used female mice genetically engineered to lack the enzyme tryptophan hydroxylase 1, which produces the form of serotonin that circulates in the blood. Of 43 mouse embryos, 37 were smaller than normal and displayed abnormalities in the brain and other organs. (Some residual serotonin from another maternal gene may explain why all the mice did not share these defects.)

Of 130 or so other pups whose mothers could make serotonin, less than 10 percent experienced such abnormalities in the embryonic stage, even among the subpopulation that had been bred to be incapable of making serotonin by itself.

AL FRANCEKEVICH *Corbis*

MA, CAN I GROW NOW? Mammalian fetuses get biochemical instructions from their mothers on how to proceed with development.

"This experiment," Mallet notes, "proves that the gene of the mother functions in place of that of the pup," which does not produce its own serotonin until the third trimester. Moreover, he adds, "there is no reason to think that this should not happen in humans. It has to be tested now."

Mallet's work builds on other findings of serotonin's importance to the embryo. As early as 1988, Jean Lauder of the University of North Carolina at Chapel Hill had shown in mouse embryonic cultures that serotonin acted as a signaling molecule that plays a critical role in early embryonic development, even hypothesizing that it came from the mother. Her experiments used two yet to be approved drugs, fluoxetine (Prozac) and sertraline (Zoloft). Her report of face, head and heart malformations in mid-gestation delayed FDA approval for Prozac until it could be determined that the drug did not produce these deformities in pregnant mice. In 2005 Michael Levin, director of the Forsyth Center for Regenerative and Developmental Biology in Boston, provided compelling evidence in frog and chick embryos that serotonin acts as a signal for development of the body's left-right axis—placement of the heart on the left side, for instance. Mallet's paper suggests that the maternal serotonin in mice may also function as a developmental signal.

This recent experiment may prompt researchers to look at variations in serotonin levels in the mother to understand the genesis of several diseases in which the chemical has been implicated, including autism, phenylketonuria and irritable bowel syndrome. It also reemphasizes the importance of pregnant women halting antidepressant therapies that modulate serotonin. And absence of the chemical may help explain the developmental problems of very premature infants and suggests the need for careful screening measures before a woman becomes a surrogate mother.

Interest in maternal effects is growing. A few weeks after the report of

Mallet's group, J. Lee Nelson of the Fred Hutchinson Cancer Research Center in Seattle and her colleagues published a study in *PNAS* that showed that cells from the mother can be passed to a fetus in the womb and that they are capable of turning into functioning islet beta cells that make insulin. Their work shows beneficial effects from microchimerism, a process in which a small number of cells from one individual take up residence in another. Most research on microchimerism—either transfer from mother to fetus or fetus to mother—has looked for possible autoimmune effects, although scientists have always considered the possibility of favorable consequences.

For his part, Mallet, whose group was the first to clone the gene for tryptophan hydroxylase 1 in the mid-1980s, continues research on maternal serotonin. Levin notes that further research is needed to determine how the serotonin reaches the embryo and what it does when it gets there, investigations Mallet is now pursuing. His laboratory is also planning to search for other neurotransmitters and hormones from the mother that might contribute to the developing embryo, a task that promises to give new meaning to the notion that a mother's work is never done.

COSMOLOGY

The Not-So-Dark Matter

HOW DARK MATTER MIGHT EMIT DETECTABLE ENERGY **BY GEORGE MUSSER**

GLOW IN THE DARK

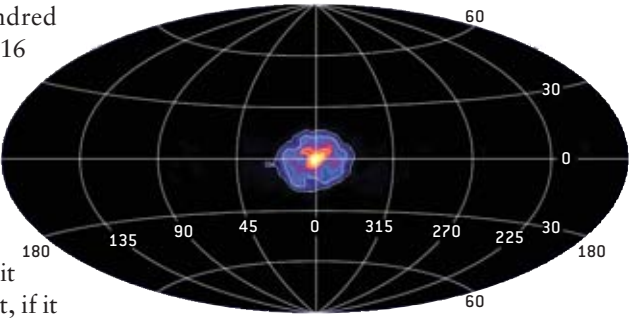
Deep in the core of the galaxy, you might literally be able to read by the light of dark matter. Last year Igor Moskalenko of Stanford University and Larry Wai of the Stanford Linear Accelerator Center revived an earlier idea that a star could sweep up so many dark particles that their annihilations would generate more energy than the star's nuclear fusion. In leaking out from the star, the energy would be stepped down from gamma rays to visible and infrared light. In fact, astronomers have seen preternaturally bright stars in the galactic center. "It is a tantalizing proposal," says astronomer Paolo Gondolo of the University of Utah. Others are skeptical. "The luminous stars in the central parsec do not require any such explanation," says astronomer Mark Morris of the University of California, Los Angeles.

For a blog discussion about another unusual proposal—that dark matter feels a long-range dark force that ordinary matter does not—see www.sciam.com/ontheweb

How is dark matter like the Hundred Years' War? The war lasted 116 years, and dark matter may not be dark. Conventional wisdom holds that darkness is the whole point of dark matter. Something is pulling stars and gas clouds off course; when astronomers go to look for it, they see nothing that fits the bill, so whatever it is must not emit or absorb light. In fact, if it did respond to light, galaxies would not even exist: the sea of radiation that filled the early universe would have buffeted the matter and kept it from clumping.

Nevertheless, astronomers have speculated for years that dark matter could power certain unexplained sources of light in the cosmos. That interpretation has been controversial—not least because each of these unexplained sources would require a different set of dark particle properties. Now a pair of researchers has found a way to make the dark matter explanation more plausible. "It's an idea that actually unifies these things and can explain everything with one new particle," says astronomer Douglas Finkbeiner of the Harvard-Smithsonian Center for Astrophysics.

One way dark matter could produce light would be to meet dark antimatter. The two will annihilate and ultimately yield a burst of gamma rays. Ten years ago the Compton Gamma Ray Observatory (CGRO) detected twice as much gamma radiation as conventional processes should produce. Dark matter annihilation would explain the discrepancy, assuming the dark particles weigh about 100 times as much as a proton.



ANTIMATTER seen near the center of the galaxy might be produced by dark matter.

ESA/J. KNÖDLSEDER *CESR* AND INTEGRAL/SPI TEAM

Another hint dates to the early 1970s, when astronomers detected vast numbers of electrons annihilating with positrons. Lots of things produce positrons, from supernovae to neutron stars, but not in such profusion, and these sources tend to lie in the plane of the galaxy, whereas the emission came from an ellipsoid. In 2003 a team proposed that dark matter annihilation generated the positrons.

The trouble is that positrons are lightweight, so the dark particles would have to be, too—about $^{1}/_{1,000}$ the mass of a proton. Apart from the inconsistency with the CGRO results, such particles would somehow have had to elude detection in particle laboratories.

Finkbeiner and particle physicist Neal Weiner of New York University can explain both sets of observations with the same heavy particle. At the American Astronomical Society meeting in January, Finkbeiner suggested that the particles might careen off one another, transferring some of their kinetic energy into internal energy, whereupon they would spit out an electron and a positron.

It so happens that the kinetic energy of dark particles in the inner galaxy is exactly the amount needed to create an electron-positron pair. To tap it, the particles would interact by a hitherto unknown force of nature that kicks in when they approach within 10^{-14} meter of one another. Such a near miss is 10,000 times more likely than the direct hit required for annihilation. This relative probability would explain both the CGRO results and the observed number of positrons. Particle physicist Jonathan Feng of the University of California, Irvine, says "the idea is more natural than others that have been proposed to date."

Although the need for a novel force may seem ad hoc, the researchers argue that new types of particles generally entail new forces. "The signal for positrons in the center, if we really do take that seriously, is motivating some sort of additional dynamics in the dark sector," Weiner says. Their model might also explain other astrophysical mysteries; for instance, it could be the heat source that keeps intergalactic gas toasty. In any case, NASA's Gamma-ray Large Area Telescope (GLAST) satellite, scheduled for launch this fall, should be able to test it. As scientists gradually uncover what dark matter is and what it does, the particles will take on real names, and the term "dark matter"—which is really just a placeholder for ignorance—will itself fade to blackness.

ONCOLOGY

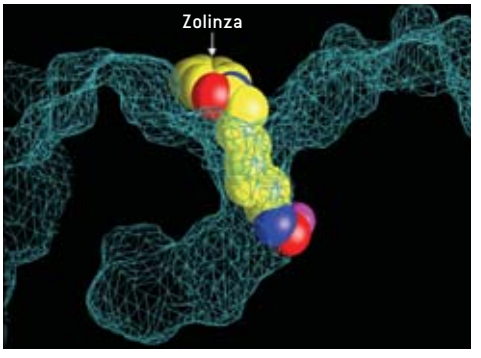
Chemo Control

DRUGS TARGET EPIGENETIC CHANGES IN CANCER CELLS **BY JENEEN INTERLANDI**

Today's cancer chemotherapy consists of little more than a dismal array of toxic drugs that kill healthy cells along with cancerous ones. Physicians must often play a deadly game of trial and error, hoping to find the right dose of the right medicine before time runs out. But a pipeline of new molecular drugs targeting so-called epigenetic phenomena could change that. These treatments could mark a path to a new array of cancer prevention strategies less toxic to the body.

The term "epigenetics" refers to the study of changes in gene expression that do not involve changes in the genetic code; they include an expanding roster of subtle molecular modifications that tell cells which genes to activate (or transcribe) and which to suppress. In cancer cells, these small-molecule regulators can act like broken dimmer switches, turning genes that promote cell growth all the way up and those that suppress tumors all the way down. By comparing the epigenetic patterns of cancerous cells with those of healthy cells, scientists are trying to identify the abnormalities tied to tumor growth. Research has led to the discovery of epigenetic culprits in many cancers, including those found in the colon, prostate, breast and blood.

Like genetic mutations, epigenetic changes can pass from one generation of cells to the next. Unlike genetic mutations, they are reversible, which makes them a drug target. Rather than killing off cells that have become cancerous via epigenetic modifications, therapies that correct such changes might reestablish those cells' normal functions.



EPIGENETIC DRUG Zolinza (vorinostat) blocks a gene-inhibiting enzyme to boost gene activity.

Two molecular switches—methylation and acetylation—have received particularly close attention; in short, methylation turns genes off, whereas acetylation turns them on. The Food and Drug Administration has approved two compounds that inhibit methylation: Vidaza in 2004 and Dacogen in 2006. Both treat myelodysplastic syndrome, a blood disorder that can lead to leukemia. And in October 2006 Zolinza became the first FDA-approved drug that treats cutaneous T cell lymphoma by enhancing acetylation. Dozens of similar drugs are in various stages of clinical trials, and other epigenetic modifications wait to be mined for potential therapeutic value. The goal of epigenetic drugs, however, is not to eliminate tumors permanently. "We aren't looking at a cure for cancer here," cautions Paul Marks, president emeritus of the Memorial Sloan-Kettering Cancer Center in New York City. "But we are looking at a future where some cancers have been reduced to a chronic state."

Researchers could boost the drugs' power by first assessing epigenetic profiles. "There are smokers out there with huge numbers of epigenetic changes, for example," explains oncologist Stephen Baylin of Johns Hopkins University. "The challenge is to identify which of those changes are predictive of lung cancer. Then the question becomes, 'Do we have drugs that we can give to healthy people?' We aren't far away from that." Epigenetic drugs might also increase the efficacy of existing chemotherapies that bind to DNA by helping those drugs gain access to their double-helix targets.

Although much work remains, some say that a paradigm shift is well under way. "The old system feels very Victorian now," remarks Andrew Allen, chief medical officer of Pharmion, the drug company responsible for Vidaza. "We used to say, 'The cancer is in your lungs, so we'll call it lung cancer,' or 'It has small cells, so we'll call it small cell,' et cetera. But we're starting to understand it much more deeply than that."

Jeneen Interlandi is a science writer based in New York City.

GETTING PERSONALIZED

Epigenetic analyses could be a major component of the long-discussed concept of personalized medicine. By diagnosing the epigenetic characteristics of tumors, physicians could better predict which drugs will work in which patients. But those factors do not appear to coincide neatly with traditional clinical diagnoses. One recent study, for example, found 12 different types of biomarkers in prostate cancer alone.

"Classically, you need a piece of the tumor" to conduct a molecular test, explains Marc Ladanyi, chief of the newly formed molecular diagnostics center at the Memorial Sloan-Kettering Cancer Center in New York City. Novel alternatives, however, may soon arise—such as a process that searches the bloodstream for DNA shed by tumors in inaccessible areas of the body. But developing such tests requires funding, and Ladanyi cautions that progress will be slow: "There are all these markers in the literature that never even make it to the clinic."

SOURCE: "DIMETHYL SULFOXIDE TO VORINOSTAT: DEVELOPMENT OF THIS HISTONE DEACETYLASE INHIBITOR AS AN ANTICANCER DRUG," BY P. MARKS ET AL., IN *NATURE BIOTECHNOLOGY*, VOL. 25; JANUARY 2007

SPACE

Natural Gas Rockets

METHANE-FUELED ENGINES GO ON THE ASCENT **BY STEVEN ASHLEY**

Most people know natural gas as home-heating fuel, but methane may soon be powering spaceships into orbit and beyond. Rocket researchers worldwide are now working on engines that burn methane rather than conventional liquid propellants.

During the past half a century, engineers have generally opted to use hydrocarbons, such as kerosene, or hydrogen itself, along with oxygen liquefied at low temperature, as chemical propellants. But kerosene and hydrogen have drawbacks, according to David Riseborough of C&Space in Seongnam, South Korea. "Burning kerosene produces soot, which deposits coking residues on engine surfaces, causing blockages and reusability problems," he says. Hydrogen, meanwhile, requires costly cryogenic storage that can be hazardous to operate and large, insulated tanks that take up space and weight.

Because any failure nearly always spells catastrophe, rocket designers tend to be conservative and stick with well-tested equipment and propellants. In recent years, however, innovators have been investigating alternative fuels, including methane. Methane is relatively benign, which has implications for crew safety and operability in space or on planetary surfaces. When burned, it yields less soot than kerosene and offers superior propulsion performance. In particular, the fuel exhibits a higher specific impulse, a measure of fuel efficiency that describes how much momentum a given weight of propellant can impart.

Although methane's specific impulse lags that of hydrogen, it presents some advantages. Liquid methane is more stable and boils off less rapidly. Further, methane storage vessels need less insulation, a feature that helps to minimize system weight. In addition, at –163 degrees Celsius, liquid methane is closer in temperature to liquid oxygen (–183 degrees C) than liquid hydrogen (–253 degrees C), which eases handling and vehicle design. Another plus for methane is that future Mars astronauts might be able to synthesize it "in situ" from the Red Planet's mostly carbon dioxide–based atmosphere, which would help reduce the size and weight of their spacecraft. In the nearer term, NASA's Orion mission may employ methane to propel the lunar lander's ascent from the moon's surface or the crew capsule's return to Earth.

With funding from NASA, a few companies are amassing engineering experience with methane rockets. Last December, XCOR Aerospace, a rocket maker in Mojave, Calif., ground-tested a 7,500-pound-thrust liquid-methane/liquid-oxygen rocket motor, reports XCOR's Richard Pournelle. A competing company, KT Engineering in Huntsville, Ala., expects to conduct its initial test-firing of a methane engine this year.

In March 2006 engineers at C&Space tested a methane/oxygen motor that delivers 20,000 to 30,000 pounds of thrust. Designed with help from Russian specialists to lift half-ton satellites into low Earth orbit for $2,000 per pound (about 20 percent of today's launch costs), the system could also power suborbital tourist flights or serve as a second-stage booster. "We believe that methane could allow quicker, aircraftlike launch turnarounds," Riseborough explains.

Despite the high hopes, success is by no means assured. The Galaxy Express, a satellite booster with a methane-fueled second stage developed by the Japan Aerospace Exploration Agency and a joint commercial venture, has been delayed by cost overruns in the hundreds of millions of dollars.

LOOKING FOR LESS MESS

Bulletproof reliability is key to rocket scientists. That is why they usually choose so-called hypergolic liquids, such as hydrazine and nitrogen tetroxide, to fuel reaction-control thrusters—small rockets that maintain the attitude of vehicles in space. Hypergolics ignite when the two components come into contact; hence, these rocket motors are dependable and simple to control, needing only two valves. But the toxic fuels are hard to handle and contaminate the vicinity of the burn. Several firms, including Orion Propulsion in Madison, Ala., and Orbital Technologies in Madison, Wis., have developed small methane-fueled rockets (producing 100 pounds of thrust and less) that avoid the safety problems of hypergolics.

SHOCK-WAVE CONES emerge in the rocket plume of XCOR's methane-fueled engine, ground-tested last December. Such engines may power future spaceflight operations.

MIKE MASSEE *XCOR*

BY THE NUMBERS

Darkness on the Water

DEVELOPMENT TAKES ITS TOLL ON ESTUARIES **BY RODGER DOYLE**

Estuaries, where freshwater mixes with saltwater, are dynamic environments of great complexity and a critical habitat for economically important species. Together with coastal waters, which are affected by much the same environmental pressures, they have long suffered degradation. Indeed, last year the United Nations declared the estuaries of China's two greatest rivers, the Yangtze and the Yellow, "dead zones." The damage affects not only the immediate estuarine and coastal environment but also the oceans.

To assess the extent of the harm to estuaries and coastal seas and to reconstruct their ecological history, an international team headed by Heike K. Lotze of Dalhousie University in Halifax, Nova Scotia, has analyzed hundreds of documents. The group focused on 12 temperate-zone estuarine and coastal ecosystems that have been exposed to intense human development for periods ranging from as little as 150 to as many as 2,500 years. They examined 30 to 80 species in each system plus seven water-quality parameters and data on species invasions.

Generally, degradation was minor during the hunter-gatherer and agricultural eras. (Still, even prehistoric populations can do damage: in San Francisco Bay, early hunters depleted the sea otter, large geese, white sturgeon and native oyster populations.) Beginning with the intrusion of market economies about 300 years ago, the ecosystems experienced a rapid decline in species abundance. The past 50 years, however, have witnessed some improvement: populations of birds were up, and populations of reptiles, mammals and vegetation held more or less even. Only invertebrates and fish suffered sharp declines. The size of the ecosystems studied, their richness in terms of number of species, and the density of human population were not related to the degree of degradation. The most degraded systems were those subjected to intense commercial activity for the longest period—the Northern Adriatic, the Western Baltic and the Wadden seas.

By the late 20th century 91 percent of the species studied were depleted (loss of 50 percent of a species or more), 31 percent were rare (loss of 90 percent or more), and 7 percent were extinct. None of this loss could be attributed to invasive species, such as the soft-shelled clam, or to climate change.

During the past 100 years, several systems, including the Outer Bay of Fundy, the Southern Gulf of the St. Lawrence, and Massachusetts Bay, have made modest gains in species abundance, apparently thanks to conservation efforts, whereas in other areas species loss has slowed. These developments suggest to Lotze and his colleagues that degradation has passed a low point and that the systems are on the road to recovery, albeit slowly.

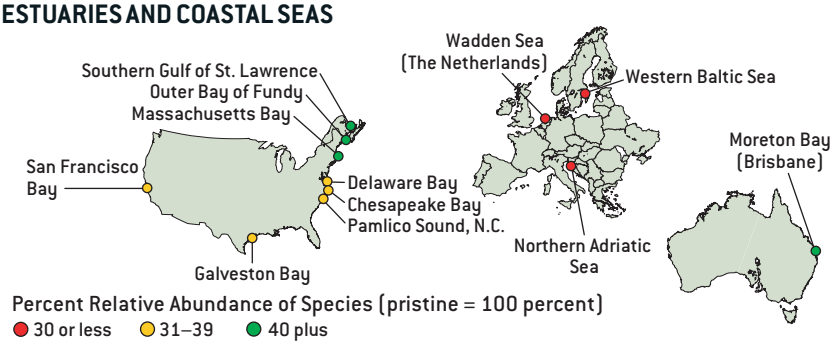
Rodger Doyle can be reached at rodgerpdoyle@verizon.net

FURTHER READING

Depletion, Degradation, and Recovery Potential of Estuaries and Coastal Seas.
Heike K. Lotze, Hunter S. Lenihan, Bruce J. Bourque et al. in *Science*, Vol. 312, pages 1806–1809; June 23, 2006.

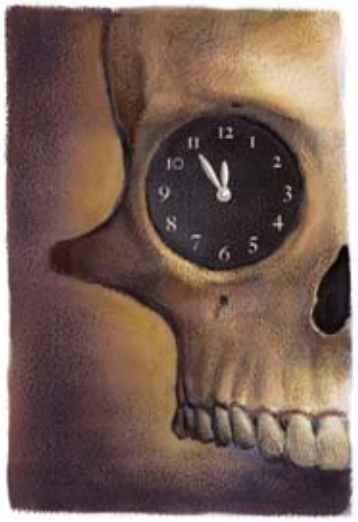
Impacts of Biodiversity Loss on Ocean Ecosystem Services.
Boris Worm, Edward B. Barbier, Nicola Beaumont et al. in *Science*, Vol. 314, pages 787–790; November 3, 2006.

ESTUARIES AND COASTAL SEAS



REGION	PERCENT OF SPECIES REMAINING	YEARS OF HUMAN IMPACT
Northern Adriatic Sea	25	2,500
Wadden Sea	28	1,000
Western Baltic Sea	30	1,000
Pamlico Sound, N.C.	31	300
Delaware Bay	32	240
Chesapeake Bay	35	240
San Francisco Bay	35	180
Galveston Bay	37	180
Moreton Bay	40	150
Massachusetts Bay	42	320
Southern Gulf of St. Lawrence	44	240
Outer Bay of Fundy	47	240

RODGER DOYLE; SOURCE FOR MAP AND TABLE: HEIKE K. LOTZE ET AL.



DATA POINTS: SPRING FORWARD

The "doomsday clock" crept two minutes closer to midnight this past January, a visual cue that represents the increasing likelihood of global catastrophe. Scientists who worked on the Manhattan Project created the symbolic clock in 1947, which is maintained by the board of directors of the *Bulletin of the Atomic Scientists*. For the first time, the board factored into its decision the dangers associated with climate change, in addition to the proliferation of nuclear weapons. —*Thania Benios*

Current doomsday time:
11:55 P.M.

Number of times in the clock's 60-year history that the minute hand has been wound:
Forward: 11
Backward: 7

Doomsday time in 1953 (closest ever to midnight): 11:58 P.M.
Doomsday time in 1991 (furthest from midnight): 11:43 P.M.

Number of nuclear warheads:
In the world: 27,000
In the U.S. and Russia: 26,000
That are launch-ready: 2,000

SOURCE: Bulletin of the Atomic Scientists *statement, January 17*

VACCINES

Artificial Market Power

On February 9, officials from Italy, the U.K., Canada, Norway and Russia met in Rome, where they announced that their governments will commit $1.5 billion to supply developing nations with vaccines against pneumococcus, the leading vaccine-preventable killer of children younger than the age of five (it claims up to one million every year). These nations are creating a so-called advance market commitment to provide incentives for pharmaceutical companies to introduce next-generation pneumococcal injections more rapidly. This strategy requires donors to pay top dollar for a set number of doses, but once funds dry up, participating vaccine makers must supply them at below cost. An existing vaccine protects against as few as 50 percent of cases in parts of the world; however, both Wyeth and GlaxoSmithKline have vaccines in the pipeline that may bump that number closer to 80 percent, a more likely effectiveness requirement that vaccines will have to meet to be eligible for funds. *—JR Minkel*

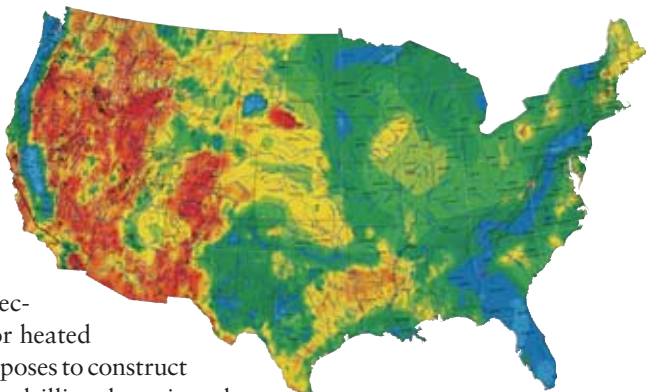
TRANSISTORS

Shrinking by Hafnium

In January, Intel announced that it can further shrink the size of a transistor by replacing oxidized silicon with the metal hafnium in its insulating layer. The insulator regulates the flow of current between a source and sink of electrons and therefore helps to govern how quickly a transistor can switch on or off: the thinner the insulator, the faster the switch. In addition, the tinier the transistor, the more that can fit onto a chip. But as silicon-based insulators have shrunk, they have leaked charge, forcing chips to consume much more power to run. Hafnium has superior insulating properties at such small scales, making it less "leaky" than silicon. With this metal, Intel expects to shrink the smallest dimension of its transistors from today's 65 nanometers to a svelte 45 nanometers, keeping the furious pace of transistor miniaturization on its expected track. IBM says it plans to release a similarly modified transistor next year. *—JR Minkel*

ENERGY

Hot Rocks



HEAT FLOW in the U.S., from 150 (*red*) to 25 (*blue*) milliwatts per square meter, could more than satisfy the nation's electrical needs.

SOURCE: *THE FUTURE OF GEOTHERMAL ENERGY: IMPACT OF ENHANCED GEOTHERMAL SYSTEMS (EGS) ON THE UNITED STATES IN THE 21ST CENTURY*, BY JEFFERSON W. TESTER ET AL., © 2006 MASSACHUSETTS INSTITUTE OF TECHNOLOGY (*U.S. map*); ILLUSTRATION BY MATT COLLINS

Geothermal power plants could supply the energy needs of the U.S. thousands of times over, concludes an 18-member panel led by the Massachusetts Institute of Technology. Geothermal stations create electricity by relying on liquid or vapor heated deep within the earth. The panel proposes to construct many new geothermal power plants, drilling down into the high-temperature bedrock, creating an open reservoir and pumping liquid into it to be heated. The researchers estimate that some 13 trillion trillion (10^{24}) joules lurk deep underneath U.S. soil and that 1.5 percent of that energy is recoverable, without taking cost into account. More than 100 gigawatts of geothermal power (one tenth of the current U.S. electrical generation) could be developed for $1 billion during the next 40 years—at the full cost of one carbon-capturing coal-fired power plant or one-third the cost of a new nuclear generator. The challenge: not to lubricate any faults that could trigger earthquakes as has occurred in Basel, Switzerland. *—David Biello*

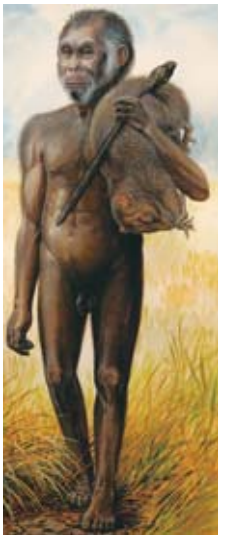
PALEOANTHROPOLOGY

Evidence from the Shire

Three years ago archaeologists found a skull on the Indonesian island of Flores that belonged to a hominin so small that it earned the nickname "Hobbit." A furious debate ensued: the fossil discoverers classify the meter-tall hominin as part of a separate species that lived as recently as 12,000 years ago; others maintain it was a modern human who had microcephaly, in which the brain fails to reach normal size. Fresh evidence has fired up the new-species camp. Paleoanthropologists at Florida State University created virtual skull casts from nine human cases of microcephaly and concluded that the Hobbit skull did not match their dimensions.

The group reports its finding in the February 13 *Proceedings of the National Academy of Sciences USA*. Perhaps more important, officials have reopened the cave where the Hobbit was discovered, along with a new chamber underneath it that may harbor new specimens.

—JR Minkel



"HOBBIT" SPECIES idea gains traction.

PHYSICS

Stop-and-Go Light

In 1999 Lene Hau and her colleagues at Harvard University slowed light traveling at 186,282 miles a second to bicycle speed (38 miles an hour). A few years later the team stopped a beam of light completely. This year Hau's team added another quantum trick: turning light into matter and then back into light again. The matter was a pair of ultracold atomic gases called Bose-Einstein condensates (BECs). In Hau's experiment a laser pulse struck one BEC and imparted its energy to the condensate. Thus imprinted, the BEC's atoms formed a matter wave that traveled 160 microns before hitting the second BEC, which absorbed the atoms. Those atoms radiated a light pulse that was an exact duplicate of the original. Hau speculates that the technique, described February 8 in *Nature*, may someday be used in optical communications or ultraprecise navigation systems.

—JR Minkel

BIOLOGY

Lavender's Hormone Havoc

Lavender and other fragrant oils may cause breast growth in boys. Pediatric endocrinologist Clifford Bloch of the University of Colorado at Denver diagnosed three otherwise healthy boys—ages four, seven and 10—with prepubertal gynecomastia, a rare condition that leads to breast growth in prepubescent males. They all had used lavender-scented soap and skin lotions, or shampoos or styling products that contained lavender oil and tea tree oil. The gynecomastia subsided several months after the boys stopped using the products. Bloch's colleagues at the National Institute of Environmental Health Sciences tested the oils in gene expression studies on lab-grown human breast cancer cells and found that they could mimic estrogens, the primary female sex hormones, and inhibit androgens, the primary male sex hormones. Whether the oils have similar effects on prepubescent girls, adolescents or adults is unknown. The researchers report their findings in the February 1 *New England Journal of Medicine*.

—Charles Q. Choi

BRIEF POINTS

■ **Thirteen of 19 smokers, including a two-pack-a-day fiend, suddenly and easily kicked the habit after they suffered stroke damage to their insulas—a part of the brain linked with emotion and feeling. The results suggest a new target for addiction therapy.**

Science, January 26

■ **A nanomachine version of Maxwell's demon, a 19th-century thought-experiment protagonist who sorts faster gas molecules from slower ones, confirms that the second law of thermodynamics holds up.**

Nature, February 1

■ **Camera shudder: the Hubble Space Telescope's main camera has stopped working and may be damaged beyond repair; about two thirds of astronomers' observing programs requested to use that device.**

NASA announcement, January 29

■ **The mere smell of food undoes the long life attainable from a calorie-restricted diet in fruit flies. Insects genetically engineered to lack a sense of smell did not lose the longevity benefit.**

Science online, February 1

Extended coverage of these News Scan briefs can be found at ***www.sciam.com/ontheweb***

PETER SCHOUTEN *National Geographic Society/Reuters/Corbis (top)*; MICHAEL ROSENFELD *Getty Images (bottom)*

Free to Choose

The neuroscience of choice exposes the power of ideas

By MICHAEL SHERMER

Have you ever watched a white rat choose between an 8 and 32 percent sucrose solution by pressing two different bars on variable-interval schedules of reinforcement? No? Lucky you. I devoted two years of what would otherwise have been a misspent youth to running choice experiments with rats in Skinner boxes for my master's thesis on "Choice in Rats as a Function of Reinforcer Intensity and Quality." Boys gone wild!

Since then, the behaviorists' black box has been penetrated by neuroscientists, most recently by Read Montague of the Baylor College of Medicine with *Why Choose This Book?* (Dutton, 2006). Montague argues that our brains evolved computational programs to evaluate choices in terms of their value and efficiency: "Those that accurately estimate the costs and the long-term benefits of choices will be more efficient than those that don't."

Our brains evolved computational programs to evaluate choices in terms of their value and efficiency.

Life, like the economy, is about the allocation of limited resources that have alternative uses (to paraphrase economist Thomas Sowell). It all boils down to energy efficiency. To a predator, Montague says, prey are batteries of energy: "This doctrine mandates that evolution discover efficient computational systems that know how to capture, process, store, and reuse energy efficiently." Those that do so pass on their genetic programs for efficient computational neural processing to make efficient choices. As a result, our brains consume only about one-fifth the energy of a lightbulb.

Unfortunately, these evolved computational programs can be hijacked. Addictive drugs, for example, rewire the brain's dopamine system—normally used to reward choices that are good for the organism, such as obtaining food, family and friends—to reward choosing the next high instead. Ideas do something similar, in that they take over the role of reward signals that feed into the dopamine neurons. This effect includes *bad* ideas, such as the Heaven's Gate cult members who chose suicide to join the mother ship they believed was awaiting them near Comet Hale-Bopp. The brains of suicide bombers have been similarly commandeered by bad ideas from their religions or politics.

In *The Science of Good and Evil* (Times Books, 2004), I argued that we evolved moral emotions that operate similarly to other emotions, such as hunger and sexual appetite. Thinking of these emotions as proxies for highly efficient computational programs deepens our understanding of the process. When we need energy, we do not compute the relative caloric values of our food choices; we just feel hungry, eat and are rewarded with a sense of satisfaction. Likewise, in choosing a sexual partner, the brain employs a computational program to make you feel attracted to people with good genes, as indicated by such proxies as a symmetrical face and body, clear complexion, and a 0.7 to 1 waist-to-hip ratio in women and an inverted pyramid build in men. Similarly, in making moral choices about whether to be altruistic or selfish, we feel guilt or pride for having done the wrong or right thing. But the moral calculations of what is best for the individual and the social group were made by our Paleolithic ancestors. Emotions such as hunger, lust and pride are stand-ins for such computations.

How can we utilize this theory of choice to our advantage? Montague employed functional magnetic resonance imaging (fMRI) to discover that certain brands, such as Coke, "change dopamine delivery to various brain regions through their effect on reward prediction circuitry." The Coke brand has a "flavor" in the ventromedial prefrontal cortex, a region essential for decision making. Just as Coke is a proxy for flavor, hunger a proxy for caloric need, lust a proxy for reproductive necessity, and guilt and joy proxies for immoral and moral behavior, so, too, can we market moral brands to rewire brains to value and choose good ideas.

In honor of the late economist Milton Friedman, author of the radical book *Free to Choose,* I propose that we begin by marketing this brand—the Principle of Freedom: *all people are free to think, believe and act as they choose, as long as they do not infringe on the equal freedom of others.* SA

Michael Shermer is publisher of Skeptic *(www.skeptic.com). His new book is* Why Darwin Matters.

BRAD HINES

Rapid Victories against Extreme Poverty

Focused steps taken now could rapidly put the poorest poor on a self-sustaining course to productivity and health By JEFFREY D. SACHS

Around one billion people live in extreme poverty, suffering from economic deprivation so severe that they must struggle daily for survival. Extreme poverty is sometimes defined as living on under $1 a day, but more accurately it is the lack of reliable access to basic needs, including adequate food, basic health services, safe drinking water and connectivity with the wider world (via roads, power and telecommunications).

Recent orthodoxy holds that extreme poverty results from corruption, mismanagement and weak institutions. A corollary is that institutional improvements take considerable time, so the escape from extreme poverty is likely to take decades. Without denying the benefit of stronger institutions, I suggest that excessive focus on institutional reforms has gotten the policy sequencing more wrong than right. Often, more direct aid can dramatically reduce extreme poverty in just a few years.

Asia's Green Revolution worked that way.

The proximate cause of extreme poverty is that the poor lack basic tools to achieve adequate productivity. "Tools" should be interpreted broadly to include not only machinery and software for production but also agricultural inputs, clinics, medicines, schools and safe water. Even in poor countries with weak institutions, aid from rich countries can help the poor gain access to the needed tools in a very short time. The result can be a dramatic surge of productivity that raises household incomes and initiates self-sustaining economic growth.

Consider the case of agricultural output. Impoverished farmers in Africa grow around one ton of cereal grains per hectare, roughly a third of the average yield in other developing countries. Their low yields reflect a lack of fertilizers, high-yield seeds and small-scale water management. The farmers are too poor to buy such inputs, and they lack the collateral to borrow. Thus, they plant with what is available, get insufficient yields and remain too poor to buy better inputs or diversify production.

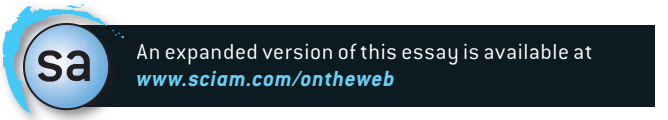
Yet if poor farmers are subsidized for a few years to obtain improved inputs and thereby raise yields and diversify output, and if they reinvest the resulting boost to income on the farm and in the community, two things can happen. First, farm productivity can increase persistently, even after subsidies are removed. Second, the households can accumulate wealth, which can be used as collateral or to self-finance purchases of improved technologies. Temporary assistance can put the farmers on the path of long-term growth. It's not a hunch. Asia's Green Revolution worked that way.

In practice, it is a group of interacting technologies that matter, and that can be provided simultaneously. They include farm inputs, health services, safe water, latrines, computers and training, motor vehicles for village use, on-grid or off-grid electricity and all-weather roads. Because these interventions are low cost and high impact, they can be provided throughout the poorest regions of the world at a cost well within the promised (but unfulfilled) commitment by rich countries to donate 0.7 percent of their gross domestic product as aid.

The Millennium Villages Project, a public-private partnership, is already applying that package of measures fruitfully in 10 African countries. Such interventions can be expanded quickly to the national level. Immunization campaigns, mass distribution of bed nets, countrywide deworming, vitamin supplementation, voucher-based provision of seeds and fertilizers, treadle pumps for irrigation, elimination of user fees for schools and clinics, and mass training of community health workers are all being introduced successfully across Africa.

But this scale-up remains fragile because of the limited and unpredictable flows of donor aid. The U.N. Millennium Development Goals to fight poverty, hunger and disease among the poorest of the poor by 2015 are still achievable. A major part of success will come through mobilizing adequate donor resources and directing them through the grassroots provision of practical tools for ending poverty. SA

Jeffrey D. Sachs is director of the Earth Institute at Columbia University (www.earth.columbia.edu).



BRUCE GILBERT/EARTH INSTITUTE

The Science of Lasting Happiness

Through controlled experiments, Sonja Lyubomirsky explores ways to beat the genetic set point for happiness. Staying in high spirits, she finds, is hard work By MARINA KRAKOVSKY

The day I meet Sonja Lyubomirsky, she keeps getting calls from her Toyota Prius dealer. When she finally picks up, she is excited by the news: she can buy the car she wants in two days. Lyubomirsky wonders if her enthusiasm might come across as materialism, but I understand that she is buying an experience as much as a possession. The hybrid will be gentler on the environment, and a California state law letting some hybrids use the carpool lane promises a faster commute between her coastal Santa Monica home and her job at the University of California, Riverside, some 70 miles inland.

BRAD SWONETZ

SONJA LYUBOMIRSKY: HAPPINESS IS HARD

- Found through experiments that exercises in gratitude, kindness and optimism can make people happier—but only if they keep doing them.
- Sees an American preoccupation with happiness (she left Moscow at age 9).
- Of unrealistic expectations in the U.S. that happiness should be easy: "It's the American quick-fix obsession."
- Skeptical of self-help books offering empirically untested advice, she wrote her own book, *The How of Happiness,* due out in 2008.

Two weeks later, in late January, the 40-year-old Lyubomirsky, who smiles often and seems to approach life with zest and good humor, reports that she is "totally loving the Prius." But will the feeling wear off soon after the new-car smell, or will it last, making a naturally happy person even more so?

An experimental psychologist investigating the possibility of lasting happiness, Lyubomirsky understands far better than most of us the folly of pinning our hopes on a new car—or on any good fortune that comes our way. We tend to adapt, quickly returning to our usual level of happiness. The classic example of such "hedonic adaptation" comes from a 1970s study of lottery winners, who a year after their windfall ended up no happier than nonwinners. Hedonic adaptation helps to explain why even changes in major life circumstances—such as income, marriage, physical health and where we live—do so little to boost our overall happiness. Not only that, but studies of twins and adoptees have shown that about 50 percent of each person's happiness is determined from birth. This "genetic set point" alone makes the happiness glass look half empty, because any upward swing in happiness seems doomed to fall back to near your baseline.

"There's been a tension in the field," explains Lyubomirsky's main collaborator, psychologist Kennon M. Sheldon of the University of Missouri–Columbia. "Some people were assuming you can affect happiness if, for example, you picked the right goals, but there was all this literature that suggested it was impossible, that what goes up must come down."

Lyubomirsky, Sheldon and another psychologist, David A. Schkade of the University of California, San Diego, put the existing findings together into a simple pie chart showing what determines happiness. Half

the pie is the genetic set point. The smallest slice is circumstances, which explain only about 10 percent of people's differences in happiness. So what is the remaining 40 percent? "Because nobody had put it together before, that's unexplained," Lyubomirsky says. But she believes that when you take away genes and circumstances, what is left besides error must be "intentional activity," mental and behavioral strategies to counteract adaptation's downward pull.

Lyubomirsky has been studying these activities in hopes of finding out whether and how people can stay above their set point. In theory, that is possible in much the same way regular diet and exercise can keep athletes' weight below their genetic set points. But before Lyubomirsky began, there was "a huge vacuum of research on how to increase happiness," she says. The lottery study in particular "made people shy away from interventions," explains eminent University of Pennsylvania psychologist Martin E. P. Seligman, the father of positive psychology and a mentor to Lyubomirsky. When science had scrutinized happiness at all, it was mainly through correlational studies, which cannot tell what came first—the happiness or what it is linked to—let alone determine the cause and effect. Finding out that individuals with strong social ties are more satisfied with their lives than loners, for example, begs the question of whether friends make us happier or whether happy people are simply likelier to seek and attract friends.

Count your blessings every day? Not if you want to be really happy.

Lyubomirsky began studying happiness as a graduate student in 1989 after an intriguing conversation with her adviser, Stanford University psychologist Lee D. Ross, who told her about a remarkably happy friend who had lost both parents to the Holocaust. Ross explains it this way: "For this person, the meaning of the Holocaust was that it was indecent or inappropriate to be unhappy about trivial things—and that one should strive to find joy in life and human relationships." Psychologists have long known that different people can see and think about the same events in different ways, but they had done little research on how these interpretations affect well-being.

So Lyubomirsky had to lay some groundwork before she could go into the lab. Back then, happiness was "a fuzzy, unscientific topic," she says, and although no instrument yet exists for giving perfectly valid, reliable and precise readings of someone's happiness from session to session, Lyubomirsky has brought scientific rigor to the emerging field. From her firm belief that it is each person's self-reported happiness that matters, she developed a four-question Subjective Happiness Scale. Lyubomirsky's working definition of happiness—"a joyful, contented life"—gets at both the feelings and judgments necessary for overall happiness. (If a sleep-deprived new mom feels fulfilled but frazzled, and an aimless party girl feels empty despite loads of fun, neither would consider herself truly happy.) To this day, she rarely sees her studies' participants; they do most exercises out in the real world and answer detailed questionnaires on the computer, often from home. To assess subjects' efforts and honesty, she uses several cross-checks, such as timing them as they complete the questionnaires.

The research needed to answer questions about lasting happiness is costly, because studies need to follow a sizable group of people over a long time. Two and a half years ago Lyubomirsky and Sheldon received a five-year, $1-million grant from the National Institute of Mental Health to do just that. Investigators have no shortage of possible strategies to test, with happiness advice coming "from the Buddha to Tony Robbins," as Seligman puts it. So Lyubomirsky started with three promising strategies: kindness, gratitude and optimism—all of which past research had linked with happiness.

Her aim is not merely to confirm the strategies' effectiveness but to gain insights into how happiness works. For example, conventional wisdom suggests keeping a daily gratitude journal. But one study revealed that those who had been assigned to do that ended up less happy than those who had to count their blessings only once a week. Lyubomirsky therefore confirmed her hunch that timing is important. So is variety, it turned out: a kindness intervention found that participants told to vary their good deeds ended up happier than those forced into a kindness rut. Lyubomirsky is also asking about mediators: Why, for example, does acting kind make you happier? "I'm a basic researcher, not an applied researcher, so I'm interested not so much in the strategies but in how they work and what goes on behind the scenes," she explains.

Initial results with the interventions have been promising, but sustaining them is tough. Months after a study is over, the people who have stopped the exercises show a drop in happiness. Like a drug or a diet, the exercises work only if you stick with them. Instilling habits is crucial. Another key: "fit," or how well the exercise matches the person. If sitting down to imagine your best possible self (an optimism exercise) feels contrived, you will be less likely to do it.

The biggest factor may be getting over the idea that happiness is fixed—and realizing that sustained effort can boost it. "A lot of people don't apply the notion of effort to their emotional lives," Lyubomirsky declares, "but the effort it takes is enormous." SA

Marina Krakovsky (marinakrakovsky.com) often writes about experimental psychology.

The Ghosts of Galaxies Past

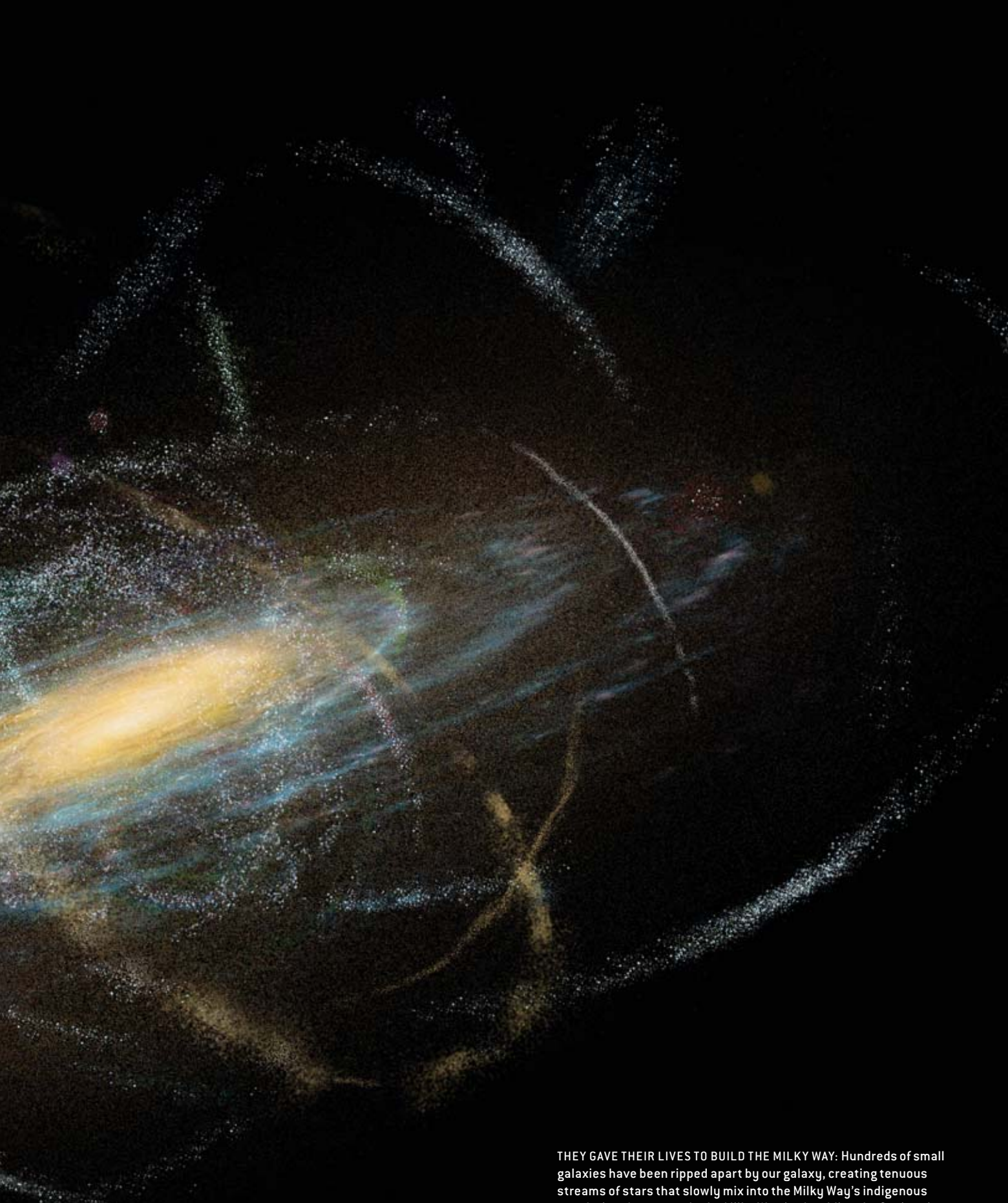
Strangely moving stars may be the remnants of past galaxies devoured by our Milky Way

By Rodrigo Ibata and Brad Gibson

When you look up at the night sky, the stars that you see all reside in our own galaxy, the Milky Way. The nearest large galaxy to us, Andromeda, is more than two million light-years away, a distance 20 times the size of the main disk of our galaxy. With the unaided eye, you cannot make out its stars individually; they blend together into a faint fuzz. As far as our galaxy is concerned, those stars might as well inhabit a separate universe. Conversely, it is natural to think of the stars in our sky as native suns, born and bred within the confines of the Milky Way.

But then what do you make of Arcturus, the second brightest star in the northern sky? Arcturus moves in a subtly different way and has a slightly different chemical composition from that of most stars in the Milky Way; it shares its curious properties with a few other stellar mavericks scattered throughout the galaxy. The origin of these and other atypical stars has been a matter of heated debate since the 1960s. Did the gravity of our galaxy's spiral arms force them into oddball orbits, or are they immigrants, formed in regions beyond the Milky Way from material that was never part of it?

By applying sophisticated forensic techniques akin to those employed elsewhere in the sciences, astronomers have discovered in recent years that the answer is: yes. Some galactic natives were indeed born in or pushed into peculiar orbits, but a surprisingly large number of the anomalies, including Arcturus, are genuine immigrants. A better metaphor than "immigrants" might be "kidnap victims" or "subject peoples," because astronomers think these stars were born into smaller galaxies that the Milky Way then captured, plundered and assimilated. Over time our galaxy may have vanquished hundreds of its neighbors. Their former inhabitants now intermingle with the Milky Way's



THEY GAVE THEIR LIVES TO BUILD THE MILKY WAY: Hundreds of small galaxies have been ripped apart by our galaxy, creating tenuous streams of stars that slowly mix into the Milky Way's indigenous population. As we speak, the Milky Way is shredding the Sagittarius dwarf galaxy, creating a stream (*gold*) that completely surrounds the central disk of the galaxy.

natives and retain only fading memories of home. By finding them out, observers can reconstruct the violent history of our galaxy and probe its darkest secret of all: the nature of the unseen dark matter that governs its existence.

Space Invaders

SPOTTING STELLAR immigrants takes a sharp eye. In principle, they give themselves away by lining up in long streams, like a conga line on a crowded dance floor. Many streams lead back in a starry path to a globular star cluster or one of the Milky Way's satellite galaxies—presumably the original home of the stars in the stream or what is left of it. In practice, though, these streams are hard to find because they barely stand out against the relatively smoothly distributed native star populations. To overcome this problem, many recent discoveries have made use of the matched-filter technique developed during World War II to produce cleaner images of incoming aircraft. The technique filters out the native stars—given an approximate idea of what patterns they and the immigrants form.

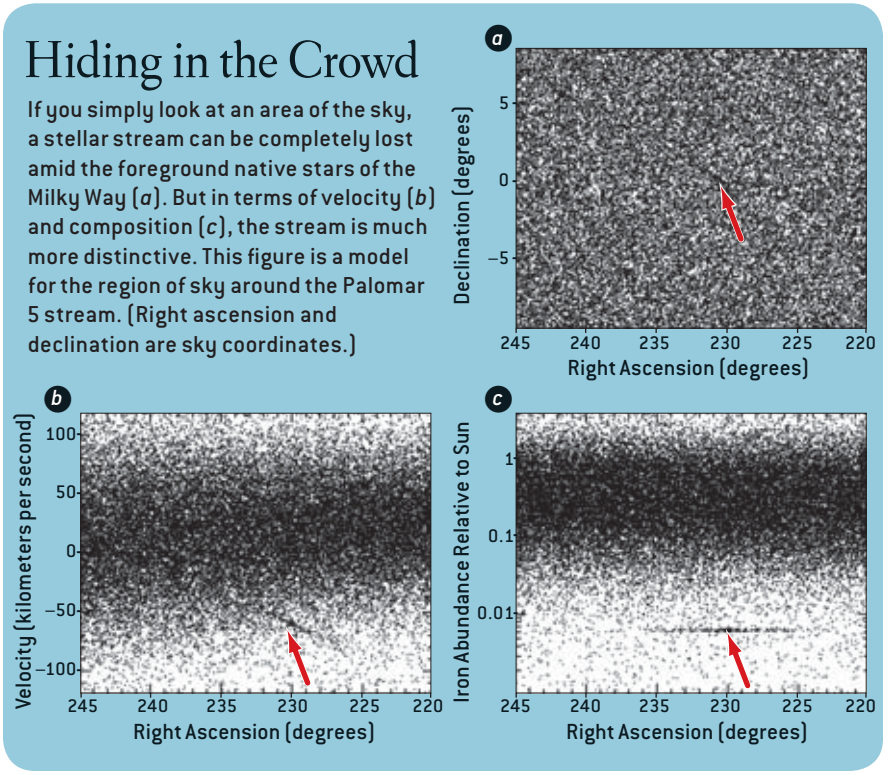
Probably the most impressive of the known star streams is the Sagittarius stream, which one of us (Ibata) and his colleagues discovered in 1994. This stream is an immense necklace of stars encircling our galaxy. It stretches over a million light-years, contains an estimated 100 million stars and connects to the Sagittarius dwarf elliptical galaxy—one of some 15 to 20 mini galaxies that orbit the Milky Way like moons around a planet. These range in size from the Large Magellanic Cloud (about a tenth the mass of our galaxy) to the Sagittarius dwarf (one-hundredth the mass) down to faint ones barely a millionth as massive.

Living close to a giant galaxy is hardly cozy. These small satellites gradually become deformed and eventually are destroyed. Sagittarius has been going through its death throes for several billion years and is now a barely coherent body in its final stages of dissolution. Its stars will disperse throughout our galaxy. The current streamlike coherence of their arrangement will be lost, and future astronomers will be hard put to distinguish them from natives of the Milky Way. Several other small galaxies are also being dismembered, and for some the stream is all that is left [*see table on page 45*]. In a few cases, such as the Large Magellanic Cloud, what is leaking away is not stars but gas [see "Our Growing, Breathing Galaxy," by Bart P. Wakker and Philipp Richter; SCIENTIFIC AMERICAN, January 2004].

The mechanism of destruction is very familiar to us all: gravitational tides, the same process that causes the ebb and flow of the oceans on Earth. Tidal forces arise when different parts of a body experience different gravitational pulls. Earth's moon exerts a larger force on the side of Earth facing the moon than on the side facing away from it—the differential force is too weak to tear our planet apart but enough to cause the sea to bulge out slightly. The motions of the two bodies cause the bulge to sweep around the globe, raising and lowering the sea level in a regular

Hiding in the Crowd

If you simply look at an area of the sky, a stellar stream can be completely lost amid the foreground native stars of the Milky Way (*a*). But in terms of velocity (*b*) and composition (*c*), the stream is much more distinctive. This figure is a model for the region of sky around the Palomar 5 stream. (Right ascension and declination are sky coordinates.)



Overview/*Star Streams*

- The Milky Way galaxy has assembled itself from a kit of hundreds of galactic building blocks—and is still doing so. Whenever a small galaxy or star cluster passes too close for its own good, our galaxy's gravity tears it apart, draws its stars out into a long stream and absorbs them.
- Even recently formed streams are hard to see against the backdrop of the rest of the galaxy, and as they disperse the spatial pattern vanishes altogether. Nevertheless, subtle patterns in stars' motion and composition will forever betray their origin. By undertaking a thorough census of stars, astronomers aim to reveal these galactic immigrants and reconstruct how the Milky Way grew to its present size.
- Streams also provide a novel way to measure the enigmatic dark matter that envelops our galaxy. The shapes of the streams are sensitive to the amount and distribution of this unseen material.

DON DIXON (*preceding pages*); RODRIGO IBATA (*this page*)

cycle. Similarly, the Milky Way deforms a satellite galaxy or star cluster by pulling harder on one side of the stellar system than on the other. In so doing, it can literally lift off some of the stars [*see box below*]. The satellite loses more and more stars, which form a virtual trail of bread crumbs leading the way back to their former home.

A Work in Progress

THE SAGITTARIUS DWARF and other galactic satellites are therefore contributing to the buildup of our own galaxy. These observational findings have revolutionized theoretical understanding of galaxy formation. Astronomers once thought that all galaxies evolved directly from barely perceptible density enhancements in the otherwise smooth primordial universe, then experienced an early period of runaway growth and settled quickly into their present form. Today, based in part on observations of the star streams, researchers think that only dwarf galaxies (up to a billion suns in mass) went through such an abrupt formation period. Large galaxies such as the Milky Way (about one trillion suns) formed later through the accretion or progressive coalescence of dwarfs—a process that continues to this day, albeit at a more leisurely pace than in the past.

Having dramatically caught the Milky Way in the act of assimilating nearby galaxies, astronomers are moving on to the next level of questions: What chemical makeup did these ancient galactic building blocks have? What is the fraction of immigrants to natives in giant galaxies of the present day? How did the chemical elements that these small galaxies brought into the Milky Way alter its early history? Apart from their direct interest as fossil records of the galaxy assembly process, star streams are also of great use as probes of the distribution of dark matter [*see box on next page*].

To tease out the answers, astronomers need to know not only which stars are currently migrating to the galaxy but also which ones migrated in the past. The trouble is that once the immigrant stars and gas mix into the Milky Way,

THE AUTHORS

RODRIGO IBATA and *BRAD GIBSON* started working together in 1997 to use observations of the Sagittarius dwarf elliptical galaxy to understand dark matter. Ibata is an observer who dabbles in theory, and Gibson is a theorist who dabbles in observation. Ibata, who works at the French National Center for Scientific Research, based at Strasbourg Observatory, keeps breaking his own records. He led the team that in 1994 discovered the Sagittarius dwarf, the closest known satellite galaxy of the Milky Way. In 2004 he and his colleagues found an even closer one, the Canis Major dwarf. He traces his love of science to physicist Steven Weinberg's book *The First Three Minutes:* "It certainly gave me a great impetus to learn math, which until then had appeared to be simply an exercise in patience." Gibson is Chair in Theoretical Astrophysics at the University of Central Lancashire in England. He pioneered computer simulations of galaxy formation and has studied mysterious anomalous velocity clouds orbiting the Milky Way.

THAT'S THE WAY THE GALAXY CRUMBLES

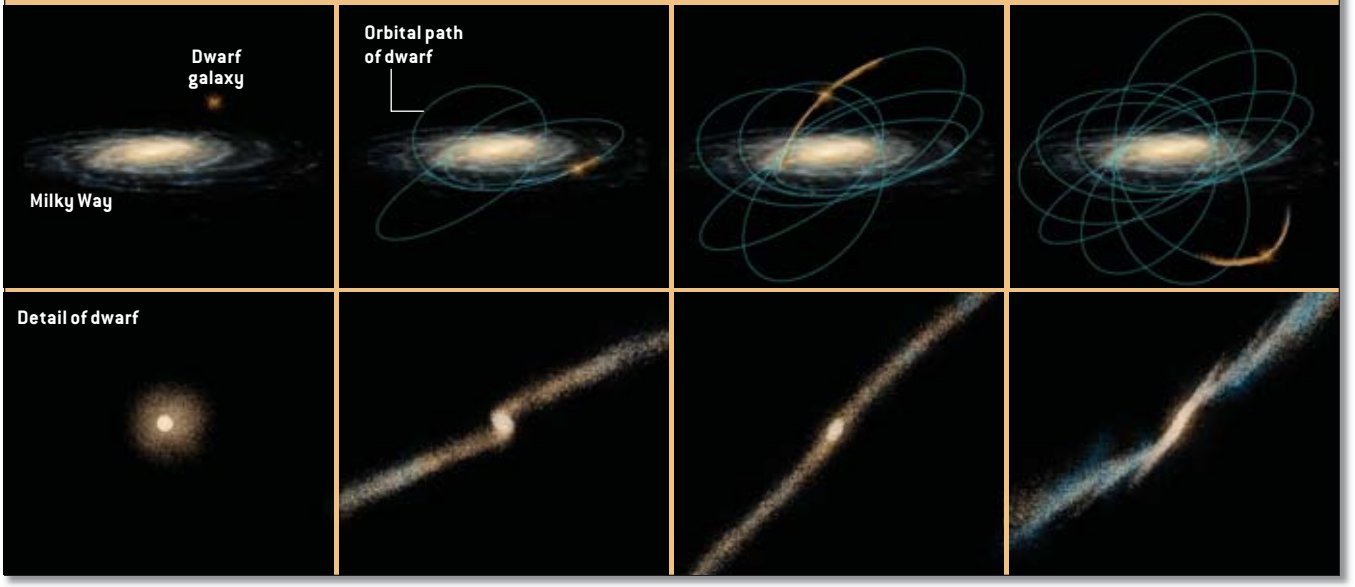
A computer simulation shows how tidal forces exerted by the Milky Way distort and eventually shred a hapless dwarf galaxy to bits.

▼ 3 BILLION YEARS AGO
Having developed in relative isolation, the dwarf galaxy makes its first pass at the Milky Way.

▼ 2 BILLION YEARS AGO
Four close encounters with our galaxy's disk have stretched the dwarf. Escaping stars either slow down and fall behind the dwarf or speed up and run ahead of it.

▼ 1 BILLION YEARS AGO
The dwarf is in serious trouble. Almost all its stars now form a stream for tens of thousands of light-years along its orbit.

▼ TODAY
The dwarf has utterly fallen apart. The tidal stream, meanwhile, is itself spreading out and mingling with the Milky Way's native stars.



DON DIXON

Seeing That Which Cannot Be Seen

In reconstructing the history of the Milky Way, astronomers naturally focus on what they can see: the stars. Yet stars are a very small fraction of galaxies. The bulk resides in a form unknown to science: the enigmatic dark matter. All that we know about it—which is not much—comes from observing its gravitational effect on the stars and gas we do see. Needless to say, it is extremely hard to map out something you cannot see.

What would-be surveyors crave is the ability to follow the motion of a star as it completes a full orbit around the galactic center. The star would slow down and speed up depending on the gravitational field—and hence would reveal the way mass is arranged in our galaxy. Unfortunately, astronomers' careers are much too short to undertake such observations, given that it takes a star several hundred million years to finish one orbit.

Streams of stars provide a way to overcome this handicap. The constituent stars follow similar orbits; the only difference is that they started at different times. Thus, they trace the path a single star would follow over hundreds of millions of years. Our team has already measured the shape of the Sagittarius stream and concluded that the dark matter around the Milky Way is distributed not as an ellipsoid, as computer simulations had predicted, but as a sphere. Interestingly, the behavior of the stream matches the predictions of certain nonstandard theories of gravity, such as modified Newtonian dynamics (MOND) [see "Does Dark Matter Really Exist?" by Mordehai Milgrom; SCIENTIFIC AMERICAN, August 2002].

This result has recently been corroborated by Michael Fellhauer of the University of Cambridge and his collaborators. A caveat is that the Sagittarius stream probes only a small portion of the dark-matter distribution. Measurements of multiple streams are required to determine whether the distribution truly is spherical.

Another question is whether dark matter is distributed smoothly or lumpily, which would bear on the dark matter's composition. If dark matter comprises particles that interact only via the force of gravity, there is nothing to stop them from clumping. If it comprises particles that can also interact by other means (such as the action of nuclear forces), they might resist clumping and smooth themselves out. Observations of star streams are one of the few ways to look for lumpiness. By carefully mapping the positions and velocities of stars in thin streams, the Gaia satellite should be able to confirm or rule out the existence of dark matter lumps as small as 100 light-years across. In this way, star streams larger than an entire galaxy may one day reveal the properties of particles smaller than atoms. *—R.I. and B.G.*

they become essentially impossible to recognize as distinct spatial features. Forensic astronomers must look for more subtle traces of the origins of these stars, such as harder-to-erase aspects of their motion and chemical composition.

We are all conditioned to characterize the movement of bodies in terms of position and velocity. But motion also has other properties, such as energy and angular momentum. Just as position is specified in terms of three-dimensional space, position plus momentum can be specified in terms of an abstract six-dimensional space, known as phase space. The advantage of phase space is that the arrangements of stars within it are more resilient than their arrangements in real space. Although the assimilation process can often destroy the spatial coherence of a stream, it cannot eliminate the stream's phase space coherence because of a key principle in statistical mechanics known as Liouville's theorem.

Thus, by measuring the energy, angular momentum and phase-space density of random samples of stars, researchers can discern stellar groups that are impossible to see directly. These are the ghosts of long-dissolved satellites. Several groups—including those led by Amina Helmi of Kapteyn Astronomical Institute in Groningen, the Netherlands, and Chris B. Brook of the University of Washington—have used the technique to uncover a number of fossils of satellite accretion. All are located in the neighborhood of the solar system, because instruments do not yet have enough precision to measure the full three-dimensional motion of more distant stars.

Which Ones Are Not Like the Others

ANOTHER EXCITING TECHNIQUE, which astronomers are only beginning to test, is so-called chemical fingerprinting. Most stars are not born in isolation but instead in groups of several thousand to tens of thousands, all of which emerge from the same parent cloud of gas. Each cloud has a unique and homogeneous mix of chemical elements and isotopes, which its stellar progeny inherits. Even when the stars disperse, they

DON DIXON

retain their unique chemical tag, allowing forensic astronomers to trace them back to their birthplace. As Kim A. Venn of the University of Victoria in British Columbia and her colleagues have shown, stars born in dwarf galaxies have a very different chemical makeup from that of typical Milky Way natives.

A more difficult question is whether it will be possible to trace individual stars back to their birthplaces. Gayandhi De Silva of the European Southern Observatory and her collaborators recently did a detailed chemical study of one of the Milky Way's best-known star clusters, the Hyades. They found that stars in this cluster all have an essentially identical distribution of chemical elements, suggesting that chemical fingerprinting will indeed allow astronomers to identify stars of common origin. They may one day be able to identify the sun's own siblings—stars formed from the same gas cloud as the sun but now spread far and wide over the galaxy.

Having verified the plausibility of chemical fingerprinting, astronomers have started to put these ideas into practice. RAVE (Radial Velocity Experiment) at the Anglo-Australian Observatory and SEGUE (Sloan Extension for Galactic Understanding and Exploration) at the Sloan telescope have measured the velocities and chemical compositions of more than 100,000 nearby stars. They, in turn, are precursors to the European Space Agency's Gaia mission and the NASA Space Interferometry Mission (SIM). From late 2011 through 2020, Gaia, the most technologically ambitious space telescope ever attempted, will aim to map the three-dimensional structure of our galaxy, obtaining precise positional and velocity measurements for an astonishing one billion stars—nearly 1 percent of the galaxy. The satellite is also intended to measure the chemical compositions of several million of these stars. Working on a comparable timescale, SIM should provide positional information for a subsample of faint stars, thereby probing the most tenuous star streams. These two missions therefore represent the best of both worlds: a sensitive but narrow survey (SIM) and a broad but less sensitive one (Gaia).

A decade ago, when we discovered the stellar stream of the Sagittarius dwarf galaxy, many of our colleagues considered it a mere curiosity with no broader significance. But it soon became the poster child for the Milky Way's tangled history of merging and accreting—processes now thought to be the main drivers of galaxy formation and evolution. Assimilated galaxies have brought new stars, gas and dark matter and have triggered waves of star formation. These immigrants have kept our galaxy a vibrant place. SA

Field of Streams

Astronomers have discovered just under a dozen stellar streams. If current theories of galaxy formation are correct, the Milky Way has devoured hundreds of small stellar systems, most now so thoroughly mixed into the Milky Way that they may never be found.

Stream	Probable Origin	Approximate Mass	Approximate Length	Composition	Year of Discovery
Arcturus stream	A defunct dwarf galaxy	Unknown	Unknown	Old stars (deficient in heavy elements)	1971
Magellanic stream	Large and Small Magellanic clouds	200 million solar masses	1 million light-years	Hydrogen gas	1972
Sagittarius stream	Sagittarius dwarf galaxy	100 million solar masses	1 million light-years	Wide variety of stars	1994
Helmi stream	A defunct dwarf galaxy	10 million to 100 million solar masses	Wrapped several times about the Milky Way's disk	Old stars	1999
Palomar 5 stream	Globular cluster Palomar 5	5,000 solar masses	30,000 light-years	Old stars	2001
Monoceros ring	Canis Major dwarf galaxy	100 million solar masses	200,000 light-years	Intermediate-age stars	2002
Anticenter stream	A defunct dwarf galaxy	Unknown	30,000 light-years	Old stars	2006
NGC 5466 stream	Globular cluster NGC 5466	10,000 solar masses	60,000 light-years	Very old stars	2006
Orphan stream	Ursa Major II dwarf galaxy	100,000 solar masses	20,000 light-years	Old stars	2006

MORE TO EXPLORE

The Ghost of Sagittarius and Lumps in the Halo of the Milky Way. Heidi Jo Newberg et al. in *Astrophysical Journal,* Vol. 569, No. 1, pages 245–274; April 1, 2002. Preprint available at **arxiv.org/abs/astro-ph/0111095**

The New Galaxy: Signatures of Its Formation. Ken Freeman and Joss Bland-Hawthorn in *Annual Reviews of Astronomy and Astrophysics,* Vol. 40, pages 487–537; 2002. **arxiv.org/abs/astro-ph/0208106**

Uncovering Cold Dark Matter Halo Substructure with Tidal Streams. R. A. Ibata, G. F. Lewis, M. J. Irwin and T. Quinn in *Monthly Notices of the Royal Astronomical Society,* Vol. 332, No. 4, pages 915–920; June 2002. **arxiv.org/abs/astro-ph/0110690**

Galactic Chemical Evolution. Brad K. Gibson, Yeshe Fenner, Agostino Renda, Daisuke Kawata and Hyun-chul Lee in *Publications of the Astronomical Society of Australia,* Vol. 20, No. 4, pages 401–415; 2003. **arxiv.org/abs/astro-ph/0312255**

Taking Measure of the Andromeda Halo: A Kinematic Analysis of the Giant Stream Surrounding M31. R. Ibata, S. Chapman, A.M.N. Ferguson, M. Irwin, G. Lewis and A. McConnachie in *Monthly Notices of the Royal Astronomical Society,* Vol. 351, No. 1, pages 117–124; June 2004. **arxiv.org/abs/astro-ph/0403068**

SEEKING THE CONNECTIONS: ALCOHOLISM AND OUR GENES

BY JOHN I. NURNBERGER, JR., AND LAURA JEAN BIERUT

Identifying genetic influences on vulnerability to alcohol addiction can lead to more targeted treatments and help those at risk to make informed choices about their own lives

The tendency to become dependent on alcohol has long been known to run in families, which for some only added to the social stigma attached to this complicated condition. But to scientists, that apparent heritability suggested that some genetic component underlying vulnerability to alcohol problems was being transmitted from generation to generation.

With rapid advances over the past 10 years in technologies for discovering and analyzing the functions of genes, researchers are now increasingly able to get at the biological roots of complex disorders such as substance abuse and addiction. The power to examine patterns of inheritance in large populations, and to survey hundreds of thousands of tiny variations in the genomes of each of those individuals, enables investigators to pinpoint specific genes that exert strong or subtle influences on a person's physiology and his or her resulting risk for disease.

As is true of many other human disorders, alcoholism does not have a single cause, nor is its origin entirely genetic. Genes can play an important role, however, by affecting processes in the body and brain that interact with one another and with an individual's life experiences to produce protection or susceptibility. Teasing these effects apart is challenging, and to date fewer than a dozen genes that influence one's risk for alcoholism have been identified, although more surely exist.

Variants of each of the known genes only modestly alter an individual's vulnerability to alcohol, but many are common in the general population and may have wider effects on drinking habits, on other addictions or problematic behaviors, and on disorders such as depression and anxiety. Finding the genes involved in our responses to alcohol and understanding their effects may thus illuminate a broader array of conditions, too. Revealing the biological processes that can

IAN MCKINNELL *Photonica/Getty Images (man with glass)*; BRIAN BERMAN *Taxi/Getty Images (baby face)*; EMILY HARRISON *(photoillustration)*

build and reinforce alcohol addiction will most certainly help to better target existing treatments and devise new ones to break alcohol's hold.

Clues in Human Variations

GENES POWERFULLY INFLUENCE a person's physiology by giving rise to some 100,000 different types of protein, each of which has a direct role in the daily functioning of the body and brain or in regulating the activity of other genes. The strong connection between variations in basic physiology and an individual's susceptibility to alcohol problems is well illustrated by the very first gene to be identified as affecting the risk of developing alcohol dependence.

Decades ago researchers began investigating the widely observed tendency of persons from Chinese, Japanese or other East Asian backgrounds to become "flushed" when they drank an alcoholic beverage. Blood tests on subjects displaying this effect showed increased levels of acetaldehyde, a breakdown product of alcohol, which resulted in an uncomfortable sensation of warmth in the skin, palpitations and weakness. By the 1980s investigators traced the reaction to an enzyme involved in alcohol metabolism, aldehyde dehydrogenase, and eventually to the gene that encodes it, *ALDH1*. The enzyme breaks down acetaldehyde, but slight variations in the gene's DNA code in these subjects caused the enzyme to work more slowly. When these individuals ingested alcohol, the acetaldehyde—which may be toxic in high doses—was building up in their bodies.

This *ALDH1* gene variant has since been found to be common in Asian populations—seen in 44 percent of Japanese, 53 percent of Vietnamese, 27 percent of Koreans and 30 percent of Chinese (including 45 percent of Han Chinese)—yet it is rare in people of European descent. As might be expected, people with this slow-metabolizing gene variant also have a decreased risk, by up to sixfold, for alcoholism, so it is an example of a genetic variation that can protect against developing the disorder.

Other enzymes that break down alcohol have also been studied for their genetic contribution to alcohol dependence. Alcohol dehydrogenase (ADH), the enzyme responsible for the first step in the conversion of alcohol to acetaldehyde, for example, is actually produced by a family of genes, each of which affects different properties of the enzyme. The genes most important to alcohol metabolism are the *ADH1* group and *ADH4*. Our own studies of a U.S. population of European descent have recently provided strong evidence that variants in the *ADH4* genes in particular enhance the risk of alcoholism in members of that population, although exactly how these *ADH4* variants affect alcohol metabolism remains to be discovered.

Alcoholism is genetically complex, meaning that multiple genes are likely to be involved, and their interactions with one another and with an individual's environment also have to be examined before a complete picture of the processes that can lead to the disorder is assembled. People are also complex and manifest problems with alcohol in diverse ways, especially in the early stages of disease, although cases come to resemble one another clinically in the later stages of illness. Thus, when investigating the biology of alcoholism, researchers must carefully define the problem—for example, distinguishing between true dependence on alcohol and alcohol abuse, which is a less medically severe syndrome.

People who meet criteria for dependence often have multiple cases of alcoholism in their families.

A widely used psychiatric standard for diagnosing dependence, be it on alcohol or another substance, requires that a person have experienced at least three of the following symptoms within the preceding 12 months: tolerance for large doses, withdrawal reactions, loss of control over use of the substance, efforts to stop or cut down, a great deal of time spent on the activity, giving up other activities, and continued use despite resulting physical or psychological problems. People who meet these criteria often have multiple cases of alcoholism in their families. With the willing participation of these subjects, we and other researchers have begun connecting individual symptoms with their physiological origins and ultimately with the responsible genes.

Indeed, an important strategy in the search for genes that affect a person's risk for alcohol dependence has been the examination of endophenotypes, which are physical traits—phenotypes—that are not externally visible but are measurable, and can therefore be studied to see whether certain patterns are more common in people with a complex disorder and may be linked to risk for that condition. The idea is grounded in an assumption that endophenotypes can reveal the biological bases for a disorder better

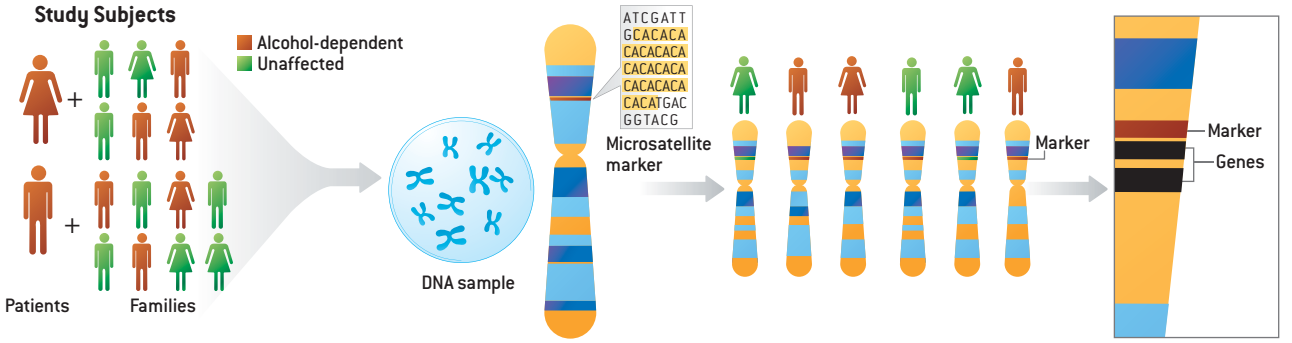
Overview/*Seeking Alcoholism Genes*

- Dependence on alcohol is a complex and controversial disorder, but susceptibility to it shows clear patterns of inheritance, which indicates that genes transmit some biological basis for greater vulnerability.
- Physiological traits, such as distinctive brain activity patterns in alcoholics and their children, help scientists pinpoint variant genes that affect a person's responses to alcohol.
- Finding the genes that influence alcoholism and related disorders provides insight into how the conditions develop, opens the way for better treatments, and allows individuals at high risk to make informed choices about their own health and behavior.

FINDING LINKS THROUGH FAMILIES

Identifying genes that influence a disorder as complex as alcoholism first involves linking the condition's traits to specific regions on chromosomes. This "linkage analysis" is easiest in genetically similar groups, such as families, with multiple members affected to some degree by the disorder.

Chromosomal features known as markers that appear more frequently in the affected relatives can flag potentially significant stretches of DNA. Detailed investigation of those regions can then reveal a gene whose function affects responses to alcohol.



RECRUITMENT
Alcoholics seeking treatment and their willing relatives are interviewed and diagnosed according to psychiatric criteria for alcohol dependence. All subjects provide DNA samples.

CHROMOSOME SURVEY
Researchers scan every person's chromosomes for patterns of repeated DNA known as microsatellite markers. In one individual, an alternating sequence of the bases cytosine and adenine might repeat 17 times, for example, whereas at the same location another relative has only 12 repeats of the sequence.

LINKAGE ANALYSIS
Markers frequently found in people with a specific trait of the disorder, but less often in unaffected kin, flag a chromosomal region linked to that trait.

GENE ASSOCIATION
Closer mapping of the DNA region near a marker reveals specific genes whose role in the disorder can be investigated.

than behavioral symptoms because they represent a fundamental physical trait that is more closely tied to its source in a gene variant. Although this approach to studying complex behaviors was first proposed in the 1970s by psychiatric researchers investigating schizophrenia, it has recently proved even more valuable with modern tools for assessing biologic processes and analyzing genetic data.

The brain's electrical activity patterns, for example, are a form of endophenotype. Using electroencephalography (EEG) to detect such activity through electrodes on the scalp, researchers can record patterns of neural firing. Sophisticated computer algorithms can analyze the data to identify the brain regions where the signals are likely to have originated, offering additional clues to the type of cognitive processing taking place. The overall brain waveforms and spikes in neural activity in response to specific stimuli seen in such EEG readings are distinctive in different individuals and serve as a kind of neurological fingerprint. These patterns can also reflect the general balance between excitatory processes within the brain, which render neurons more responsive to signaling from other neurons, and those that are inhibitory, making neurons less responsive.

Such electrophysiological patterns are highly heritable and they differ in characteristic ways in alcoholics and nonalcoholics, with excitation exceeding and overpowering inhibition in the alcoholic subjects' brains. This imbalance, or "disinhibition," can also be seen in the children of alcoholics and strongly predicts their own subsequent development of heavy drinking and alcohol dependence, which suggests that these patterns are a marker for a biologically inherited predisposition to alcoholism. Moreover, the signature patterns may point to the heritable vulnerability itself: disinhibition is believed to stem from a generalized lack of functioning inhibitory neurons in the brain areas responsible for judgment and decision making, and people lacking these inhibitory circuits may be more prone to acting on impulses originating in lower brain regions, such as the amygdala.

In the 1980s evidence from several laboratories showing that electrical activity in the brain could reveal a person's risk of alcohol dependence helped to stimulate the idea that an intensive search for the genes underpinning alcoholism-associated phenotypes was feasible and worthwhile. With support from the National Institute on Alcohol Abuse and Alcoholism, the Collaborative Study on the Genetics of Alcoholism (COGA), in which we are both participants, started in 1989. The study currently involves eight research centers across the U.S. and thousands of alcoholics and their family members who have agreed to help in this ongoing investigation.

Family Ties

AT COGA'S OUTSET, researchers at sites around the country sought to identify families severely affected by alcoholism. Previous twin, adoption and family studies had indicated that alcohol problems are strongly heritable—indeed, more than 50 percent of the overall risk for alcoholism is attributable to inherited factors, which makes family groups a powerful re-

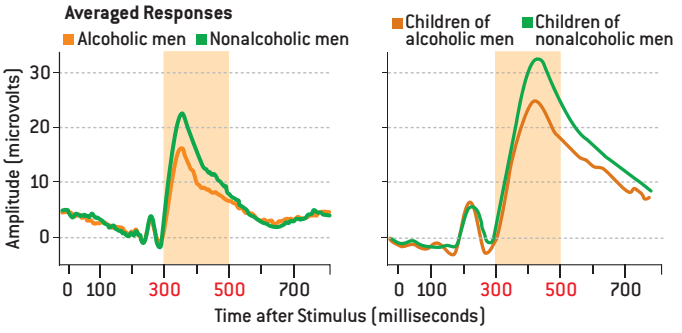
LUCY READING-IKKANDA

SIGNATURES IN THE BRAIN

Certain patterns of brain electrical activity serve as measurable traits, known as endophenotypes, which reveal distinctive physiological characteristics of alcoholics and others at high risk for the disorder. Investigators have used these signature differences in brain function to uncover genes linked to alcoholism and related conditions.

The P300 Response

Measuring brain activity through electrodes on the scalp reveals a spike in signal strength (amplitude) between 300 and 500 milliseconds after a stimulus, such as a flash of light. Known as P300, this distinctive evoked response is significantly weaker in alcoholics, even when abstinent, than in nonalcoholics. A muted P300 is also typical in the children of alcoholic parents, indicating that this functional brain difference predates the onset of heavy drinking and is itself a risk factor for becoming alcoholic.

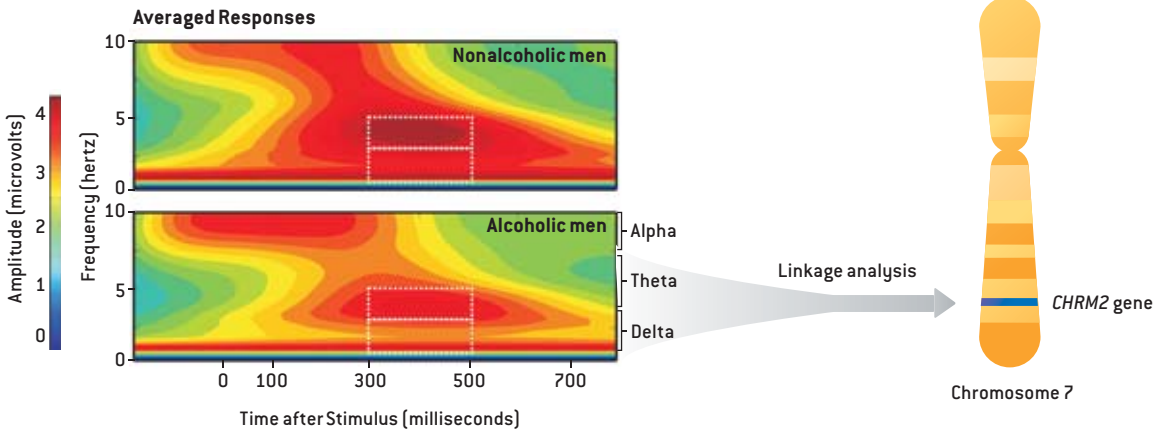


Dissecting the Response

P300 consists largely of neural signaling in low-frequency ranges known as delta and theta, which are associated with awareness and decision making. Mapping the EEG readings of nonalcoholic (*below*) and alcoholic (*bottom*) subjects by frequency reveals weaker signal strength in those bands among the alcoholics after 300 milliseconds. This trait was linked in family studies to both alcoholism and depression.

Linkage to a Gene

Reduced delta- and theta-frequency signal strength in alcoholic subjects was also traced to variants of *CHRM2,* a gene encoding a cellular receptor for the neurotransmitter acetylcholine, which regulates neural excitability.



source for tracking specific traits and linking them to the relevant genes [*see box on preceding page*].

Some 1,200 subjects seeking treatment for alcohol dependence and their relatives—more than 11,000 people in all—were extensively interviewed. Among these, 262 families were found to be "deeply affected," which means that they included two or more first-degree relatives of the patient—such as parents or siblings—who were also diagnosed as alcohol-dependent. The electrophysiological brain endophenotypes of both affected and unaffected members of those families were assessed, and the subjects underwent further interviews to evaluate additional characteristics that are associated with alcoholism risk and believed to be genetically influenced. These traits include "low response," meaning that the person must consume larger-than-average amounts of alcohol before feeling its effects; previous experience of major depression; and certain drinking history patterns, such as a high maximum number of drinks ever consumed in a 24-hour period.

The participants also provided DNA samples, which allowed COGA scientists to examine the chromosomes of each individual and take note of distinctive molecular features, which can serve as markers for a potentially significant region of a chromosome. Markers appearing most frequently in family members exhibiting alcoholism-associated phenotypes would suggest a causal link between that region of a chromosome and the trait. Significant linkages were identified in this manner on chromosomes 1, 2, 4 and 7, and many years of genetic mapping subsequently pinpointed several specific genes in those regions, including *ADH4* and *GABRA2* on chromosome 4, as well as *CHRM2* on chromosome 7. Other research groups studying separate populations have also documented associations between a risk for alcoholism and these chromosomal regions and genes, confirming their likely role in the disorder.

LUCY READING-IKKANDA; FROM "EVENT-RELATED POTENTIALS IN COA'S," BY BERNICE PORJESZ AND HENRI BEGLEITER, IN *ALCOHOL HEALTH & RESEARCH WORLD,* VOL. 21, NO. 3; 1997 [*top right graph*]; FROM "THE UTILITY OF NEUROPHYSIOLOGICAL MARKERS IN THE STUDY OF ALCOHOLISM," BY BERNICE PORJESZ ET AL., IN *CLINICAL NEUROPHYSIOLOGY,* VOL. 116, NO. 5; MAY 2005; USED WITH PERMISSION FROM ELSEVIER [*top left graph* and *bottom left graph*]

For instance, a growing body of research has revealed that some variants of genes that encode cell-surface docking sites for the protein GABA (gamma-aminobutyric acid), which carries signals between certain nerve cells, increase vulnerability to alcoholism. GABA is the most common inhibitory neurotransmitter in the mammalian nervous system. It modulates the activity of neurons by binding to GABA-specific receptors in their cell membranes and literally inhibiting their responsiveness to signaling. One class of these receptors, known as $GABA_A$, is made of protein subunits arrayed around a channel that admits chloride ions into the cell. Variations in the *GABRA2* gene, which encodes one of the $GABA_A$ receptor subunits, have been found to strongly influence an EEG endophenotype, known as the beta frequency, that appears to play a role in mediating neuronal disinhibition.

Neurons that bear GABA receptors are especially abundant in the brain's frontal cortex, where a generalized loss of inhibition can cause seizures, and seizure disorders are commonly treated with medications that boost GABA activity, promoting inhibition. A less generalized loss of GABA-induced inhibition, however, is thought to be involved in behavioral undercontrol or impulsivity, which is a feature of a number of psychiatric disorders, including bipolar affective disorder, substance abuse and chronic conduct problems. Studies by COGA consortium members have demonstrated that variants of the *GABRA2* gene are linked to alcoholism, a finding that has since been confirmed by at least four groups. Interestingly, these variations in *GABRA2* do not change the protein structure of the $GABA_A$ receptor; instead they seem to modify production of the affected protein subunit, perhaps reducing the total number of functioning receptors.

Studies are under way to identify exactly how this GABA receptor gene variant affects disinhibition in the brain, but a connection between GABA activity and alcohol dependence certainly makes sense, because impulsivity is a feature of many cases of alcohol dependence. That trait is particularly associated with an early-onset form of addiction seen primarily in males. People with this kind of addiction are generally prone to "externalizing" disorders that involve problematic behavior, as opposed to "internalizing" disorders such as anxiety and depression. Thus, even without a genetic screen of such a patient, understanding the likely involvement of GABA in that addiction profile can help target therapeutic approaches.

Another neurotransmitter highlighted in the development of alcoholism by the study of endophenotypes is acetylcholine, which, like GABA, affects neurons widely distributed through the central nervous system. Neurons that respond to acetylcholine—described as cholinergic neurons—also have an important role in modulating the overall balance between excitation and inhibition in the brain. Our measures of brain responses in COGA subjects uncovered a connection to the chromosomal region containing the *CHRM2* gene, which encodes a particular type of cholinergic receptor known as the M2 muscarinic acetylcholine receptor (CHRM2).

There are different paths to alcoholism and different pathways underlying them.

Activation of the CHRM2 receptor alters neural signaling in the slow delta and theta frequencies, which are associated with cognitive functions such as decision making and attention [*see box on opposite page*]. We were also able to link variants of the *CHRM2* gene to the clinical conditions of alcohol dependence and major depression. As with *GABRA2*, the *CHRM2* variants that appear to influence the brain's electrical activity, alcoholism and depression do not seem to alter the structure of the receptor protein but rather its manufacture.

This particular association is exciting because it confirms part of a hypothesis articulated in 1976 by psychiatrist David Janowsky and his colleagues at Vanderbilt University regarding the brain's need to maintain a fine balance between different signal-regulating processes to function normally. Janowsky's group proposed that muscarinic supersensitivity—that is, an enhanced effect of acetylcholine on the muscarinic cholinergic receptors—in persons prone to depression and related conditions was an underlying source of imbalance in the brain.

The recently discovered links between *CHRM2*, alcoholism and depression are the first to show a direct connection between a specific gene and such hypersensitivity, and these findings about the cholinergic system provide new targets for the development of more specific pharmacological treatments for alcoholism and depression. They also underscore the need to understand how subtle differences in physiology can contribute to a disorder as complex as addiction.

THE AUTHORS

JOHN I. NURNBERGER, JR., and LAURA JEAN BIERUT are psychiatric geneticists who often collaborate to study the genetic influences on many forms of substance addiction as well as on mental illnesses such as depression and bipolar disorder. Nurnberger is professor of psychiatry at Indiana University School of Medicine, where he is also director of the Institute of Psychiatric Research. Bierut is associate professor of psychiatry at Washington University in St. Louis and, like Nurnberger, an investigator in the Collaborative Study on the Genetics of Alcoholism (COGA). Both authors wish to acknowledge the late Henri Begleiter and Theodore Reich, principal and co-principal investigators of COGA since its inception: "We are indebted to their leadership in the establishment and nurturing of COGA and acknowledge with great admiration their seminal scientific contributions to the field."

Insight, Not Destiny

THE COGA PROJECT has been structured around families, but this type of research has also strengthened understanding of the relative importance of specific gene variants as risk factors in different ethnic groups. This is not to say that certain ethnicities are more prone to alcoholism; instead, like the *ALDH1* gene version that makes many East Asians intolerant of alcohol, certain of the genetic variants that contribute to risk are much more prevalent in some ethnic groups than in others. The knowledge that such genes are likely to be influencing dependence in patients belonging to one of these populations is another tool that can be used to assess the nature of an individual's problem and to tailor treatment accordingly.

Our research group recently discovered, for example, that variation in a gene encoding a receptor involved in taste perception, known as hTAS2R16, is significantly linked to alcoholism in the COGA subjects. The risk variant, which causes decreased sensitivity to many bitter taste compounds, is uncommon in European Americans, whereas 45 percent of African-Americans carry this version, making it a much more significant risk factor in that population.

The genetic contributions to dependence identified so far affect many different aspects of human physiology, from alcohol metabolism to brain activity and taste perception just in the examples we have described. The effect of each of these genes by itself is modest, probably increasing average risk by 20 to 40 percent, and other as yet unidentified genes undoubtedly also contribute to vulnerability to alcohol problems.

An important test to confirm and refine these genetic findings is to see how they influence people early in life, even before the onset of heavy drinking, and whether the variants can predict the later development of alcoholism. COGA added such a prospective arm to the study to follow young members of the high-risk families. Initial results have shown that in adolescents, the *ADH* gene risk variants are indeed associated with early drinking and subsequent development of alcohol problems. Carriers of the *CHRM2* risk variants, however, are more likely to have early symptoms of depression than drinking problems when they are adolescents. Youngsters with the *GABRA2* risk variant more frequently display conduct problems, such as trouble with the police, fighting and expulsion from school, rather than early drinking. In young adults, on the other hand, the GABA receptor gene risk variants do become associated with alcohol dependence.

These findings reinforce the notion that there are different paths to alcohol dependence and different physiological pathways underlying them. The *ADH* risk variants may contribute to the development of alcoholism directly by promoting heavy

RISKY GENES

The genes found so far to influence a person's risk of becoming alcohol-dependent represent a wide variety of physiological processes, including breakdown of alcohol itself, balanced brain function, taste and reward reinforcement. Variations in the genes controlling these traits can protect people from, or predispose them to, responding to alcohol in ways that enhance addiction. Some of the same gene variants have also been linked to other traits or disorders, which suggests that certain problems involving behavior, mood and dependence may have overlapping origins.

GENE; LOCATION	ENCODED PROTEIN; FUNCTION	GENE VARIANT EFFECT	LINKED TO OTHER TRAITS OR DISORDERS
ADH4; chromosome 4	Alcohol dehydrogenase; alcohol-metabolizing enzyme	Increased risk (certain variants)	None
ALDH1; chromosome 4	Aldehyde dehydrogenase; alcohol-metabolizing enzyme	Protective	None
CHRM2; chromosome 7	Muscarinic acetylcholine receptor M2; regulates neural signaling	Increased risk	Major depression; delta- and theta-frequency EEG variations
**DRD2;* chromosome 11	Dopamine D2 receptor; regulates reward reinforcement	Increased risk	Habitual smoking
GABRG3; chromosome 15	$GABA_A$ receptor g3 subunit; regulates neural signaling	Increased risk	None
GABRA2; chromosome 4	$GABA_A$ receptor a2 subunit; regulates neural signaling	Increased risk	Drug dependence; conduct disorders; beta-frequency EEG variations
HTAS2R16; chromosome 7	hTAS2R16 receptor; contributes to bitter taste sensitivity	Increased risk	None
OPRK1; chromosome 8 *PDYN;* chromosome 20	Kappa opioid receptor and prodynorphin, the peptide to which the receptor binds; both participate in regulating aversion and reward	Increased risk	Stress response; may play a role in heroin and cocaine habituation

To date, evidence for DRD2 is contradictory. Further investigation is needed to confirm this gene's role in alcohol or nicotine dependence.

drinking, whereas the *GABRA2* variants predispose a person to conduct problems, which are themselves a risk factor for alcoholism. Meanwhile *CHRM2* may act through depression and other internalizing symptoms to foster drinking.

As more genes are linked to the development of alcohol dependence, these insights will be used to improve tools for gauging an individual's risk for developing alcoholism and identifying those with alcohol problems who may respond better to specific treatments. Doctors commonly consider a person's genetic profile and other family and environmental risk factors when combining medications and behavioral prescriptions for complex conditions such as hypertension, cancer and bipolar affective disorder. Clinicians are in the earliest stages of using genetic variants to shape treatment decisions for alcoholism, and in the future we expect to have molecular guidelines to help develop such individualized strategies.

The recent genetic findings related to alcoholism may also suggest ways to improve the prevention and treatment of smoking and other forms of substance dependence that are frequently seen in people with alcohol problems and tend to cluster in the same families. Mood and anxiety disorders fall into this category as well, and the association between *CHRM2* variations, alcoholism and depression illustrates how these problems may stem in part from a common source. Improved understanding of alcohol dependence should therefore help dissect factors involved in the development of related conditions.

Genetics is never destiny, however. Genes may interact with specific toxic environments, such as abuse or neglect, to result in problems for some gene carriers but not for others. And if half of alcoholism risk is heritable, the other half must derive from other sources. Nobody gets to be alcohol-dependent without making some poor choices, but clearly some people are more sensitive to alcohol than others in the same set of circumstances, and scientists are working to identify the sources of that vulnerability.

Critics have argued that genetic research into alcohol dependence and other forms of addiction, including smoking, is not cost-effective from a public health perspective. For instance, some claim that it would make more sense to direct resources toward reducing the use of potentially addictive substances across the board than to identify—and potentially stigmatize—the individuals who would be most affected by such reductions. Undoubtedly, there is value in limiting the use of alcohol, nicotine and other mood-altering drugs in general. There is also value, however, in supporting individual self-knowledge as it pertains to susceptibility so that people can make informed choices for themselves and in shaping a culture that regards this as a positive goal.

Genetic testing is already providing opportunities for self-assessment that were impossible in the past, and the demand for genetic profiling will increase in the coming years. Microarrays, often called gene chips, can be used to detect a person's gene variants as well as variations in gene activity and to produce a series of medical, psychiatric and behavioral recommendations that the individual may take or leave as he or she wishes. This use of scientific knowledge is surely inevitable, especially in free nations with capitalist economies, where it will be market-driven and competitive. The scientific and academic communities must therefore help guide this process by distinguishing true physiological relations from false claims and by encouraging socially responsible uses for these discoveries.

SCOTT GIBSON *Corbis*

The risks of smoking were first widely publicized by the Surgeon General's Report of 1964, and the combination of that medical information and social pressure has reduced the prevalence of smoking over the subsequent decades. An individual's awareness of personal genetic medical risks may similarly change his or her choices. The broader health and social effects of this new type of information may not be seen quickly, but they could be quite profound over time. SA

MORE TO EXPLORE

The Collaborative Study on the Genetics of Alcoholism: An Update. Howard J. Edenberg in *Alcohol Research & Health,* Vol. 26, No. 3, pages 214–218; 2002.
Available at **www.niaaa.nih.gov/Publications/AlcoholResearch**

Evidence of Common and Specific Genetic Effects: Association of the Muscarinic Acetylcholine Receptor M2 (CHRM2) Gene with Alcohol Dependence and Major Depressive Syndrome. Jen C. Wang et al. in *Human Molecular Genetics,* Vol. 13, No. 17, pages 1903–1911; June 30, 2004.

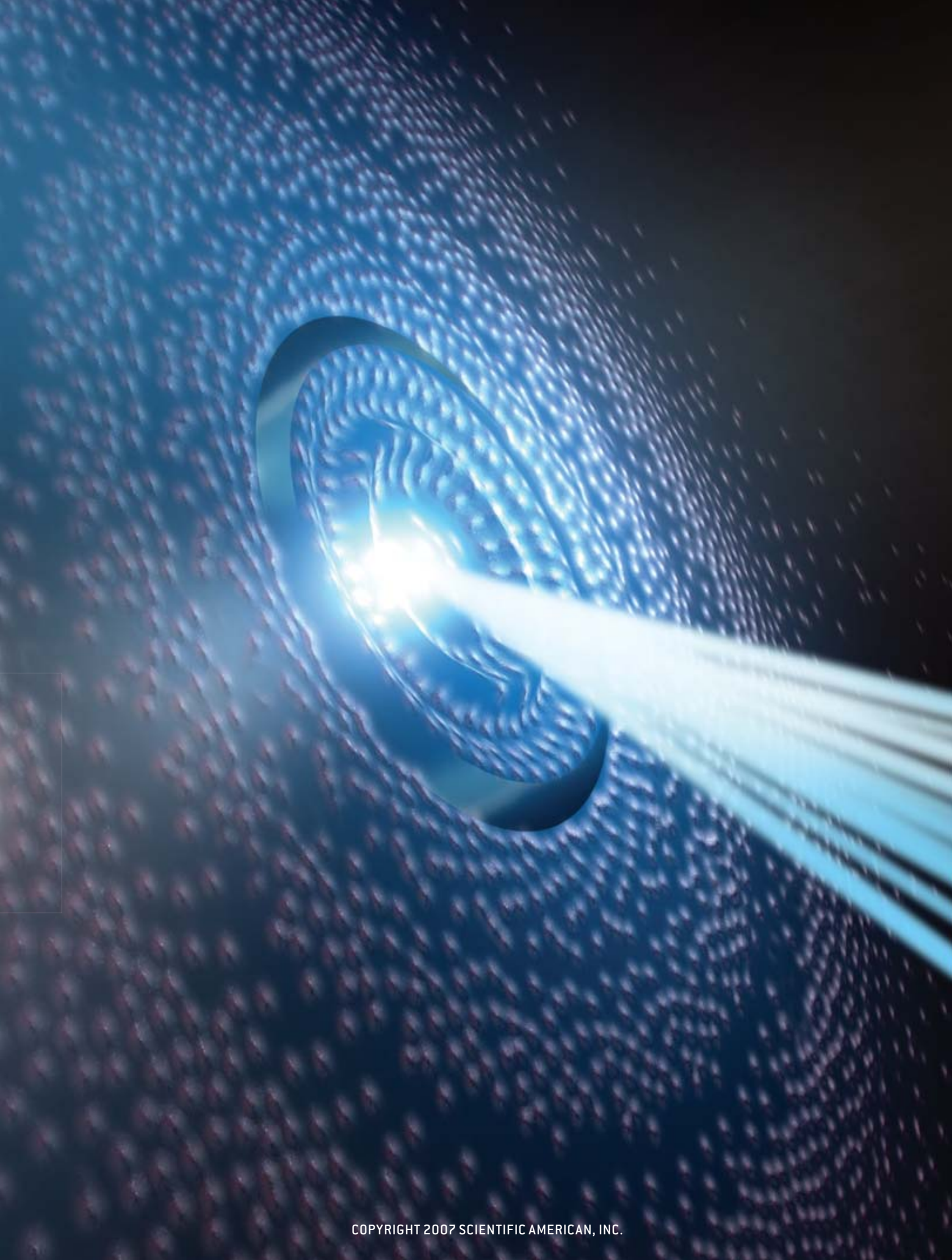
Endophenotypes Successfully Lead to Gene Identification: Results from the Collaborative Study on the Genetics of Alcoholism. Danielle M. Dick et al. in *Behavior Genetics,* Vol. 36, No. 1, pages 112–126; January 2006.

A technology that squeezes electromagnetic waves into minuscule structures may yield a new generation of superfast computer chips and ultrasensitive molecular detectors

The Promise of PLASMONICS

By Harry A. Atwater

LIGHT BEAM striking a metal surface can generate plasmons, electron density waves that can carry huge amounts of data. If focused on a surface etched with a circular groove, as in this artist's rendering, the beam produces concentric waves, organizing the electrons into high- and low-density rings.



Light is a wonderful medium for carrying information.

Optical fibers now span the globe, guiding light signals that convey voluminous streams of voice communications and vast amounts of data. This gargantuan capacity has led some researchers to prophesy that photonic devices—which channel and manipulate visible light and other electromagnetic waves—could someday replace electronic circuits in microprocessors and other computer chips. Unfortunately, the size and performance of photonic devices are constrained by the diffraction limit; because of interference between closely spaced light waves, the width of an optical fiber carrying them must be at least half the light's wavelength inside the material. For chip-based optical signals, which will most likely employ near-infrared wavelengths of about 1,500 nanometers (billionths of a meter), the minimum width is much larger than the smallest electronic devices currently in use; some transistors in silicon integrated circuits, for instance, have features smaller than 100 nanometers.

Recently, however, scientists have been working on a new technique for transmitting optical signals through minuscule nanoscale structures. In the 1980s researchers experimentally confirmed that directing light waves at the interface between a metal and a dielectric (a nonconductive material such as air or glass) can, under the right circumstances, induce a resonant interaction between the waves and the mobile electrons at the surface of the metal. (In a conductive metal, the electrons are not strongly attached to individual atoms or molecules.) In other words, the oscillations of electrons at the surface match those of the electromagnetic field outside the metal. The result is the generation of surface plasmons—density waves of electrons that propagate along the interface like the ripples that spread across the surface of a pond after you throw a stone into the water.

Overview/*Plasmonics*

- Researchers have discovered that they can squeeze optical signals into minuscule wires by using light to produce electron density waves called plasmons.
- Plasmonic circuits could help the designers of computer chips build fast interconnects that could move large amounts of data across a chip. Plasmonic components might also improve the resolution of microscopes, the efficiency of light-emitting diodes, and the sensitivity of chemical and biological detectors.
- Some scientists have even speculated that plasmonic materials could alter the electromagnetic field around an object to such an extent that it would become invisible.

Over the past decade investigators have found that by creatively designing the metal-dielectric interface they can generate surface plasmons with the same frequency as the outside electromagnetic waves but with a much shorter wavelength. This phenomenon could allow the plasmons to travel along nanoscale wires called interconnects, carrying information from one part of a microprocessor to another. Plasmonic interconnects would be a great boon for chip designers, who have been able to develop ever smaller and faster transistors but have had a harder time building minute electronic circuits that can move data quickly across the chip.

In 2000 my group at the California Institute of Technology gave the name "plasmonics" to this emerging discipline, sensing that research in this area could lead to an entirely new class of devices. Ultimately it may be possible to employ plasmonic components in a wide variety of instruments, using them to improve the resolution of microscopes, the efficiency of light-emitting diodes (LEDs) and the sensitivity of chemical and biological detectors. Scientists are also considering medical applications, designing tiny particles that could use plasmon resonance absorption to kill cancerous tissues, for example. And some researchers have even theorized that certain plasmonic materials could alter the electromagnetic field around an object to such an extent that it would become invisible. Although not all these potential applications may prove feasible, investigators are eagerly studying plasmonics because the new field promises to literally shine a light on the mysteries of the nanoworld.

Shrinking Wavelengths

FOR MILLENNIA, alchemists and glassmakers have unwittingly taken advantage of plasmonic effects when they created stained-glass windows and colorful goblets that incorporated small metallic particles in the glass. The most notable example is the Lycurgus cup, a Roman goblet dating from the fourth century A.D. and now held in the British Museum [*see illustration on page 62*]. Because of plasmonic excitation of electrons in the metallic particles suspended within the glass matrix, the cup absorbs and scatters blue and green light—the relatively short wavelengths of the visible spectrum. When viewed in reflected light, the plasmonic scattering gives the cup a greenish hue, but if a white light source is placed within the goblet, the glass appears red because it transmits only the longer wavelengths and absorbs the shorter ones.

Research into surface plasmons began in earnest in the 1980s, as chemists studied the phenomenon using Raman spectroscopy, which involves observing the scattering of laser light off a sample to determine its structure from molecular

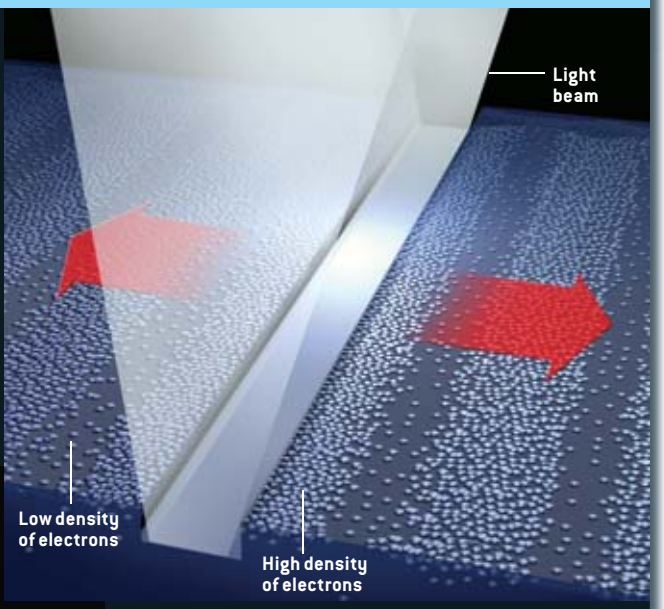
PHIL SAUNDERS *Space Channel Ltd.* [*preceding pages*]

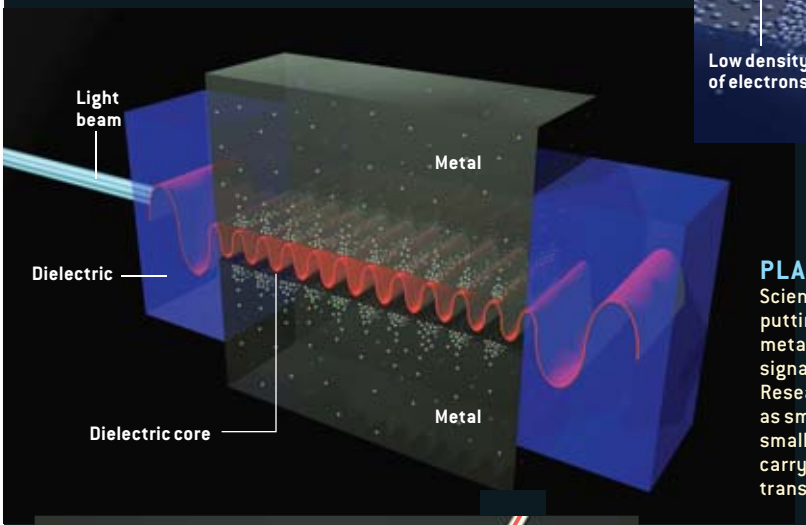
FUNNELING LIGHT INTO TINY WIRES

The study of plasmonics is relatively new, but researchers have already developed prototype devices that demonstrate the promise of the technology.

PLANAR WAVEGUIDE

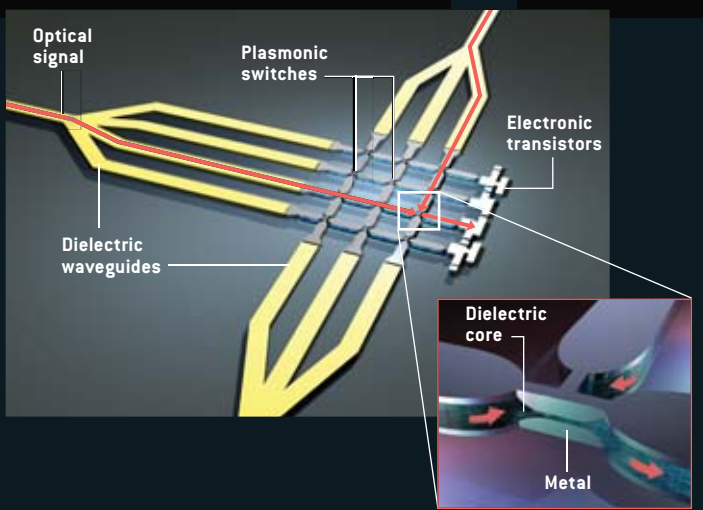
Plasmons always flow along the boundary between a metal and a dielectric (a nonconductive material such as air or glass). For example, light focused on a straight groove in a metal will generate plasmons that propagate in the thin plane at the metal's surface (the boundary between the metal and air). A plasmon could travel as far as several centimeters in this planar waveguide—far enough to convey a signal from one part of a chip to another—but the relatively large wave would interfere with other signals in the nanoscale innards of a processor.





PLASMON SLOT WAVEGUIDE

Scientists have built much smaller plasmonic circuits by putting the dielectric at the core and surrounding it with metal. The plasmon slot waveguide squeezes the optical signal, shrinking its wavelength by a factor of 10 or more. Researchers have constructed slot waveguides with widths as small as 50 nanometers—about the same size as the smallest electronic circuits. The plasmonic structure can carry much more data than an electronic wire, but it cannot transmit a signal farther than 100 microns.



A FASTER CHIP

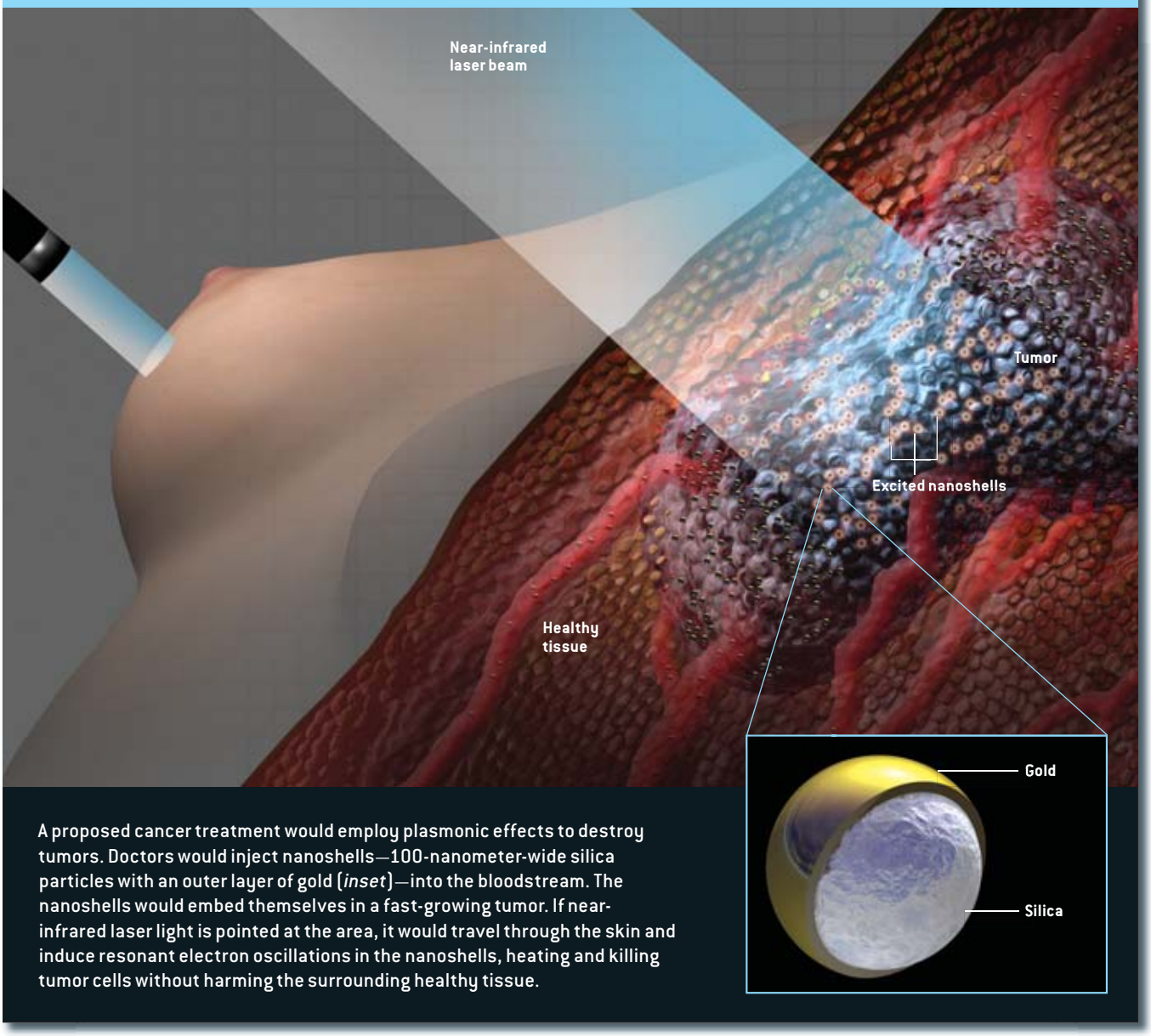
Slot waveguides could significantly boost the speed of computer chips by rapidly funneling large amounts of data to the circuits that perform logical operations. In the rendering at the left, relatively large dielectric waveguides deliver optical signals to an array of plasmonic switches (dubbed "plasmonsters"), which in turn distribute the signals to electronic transistors. The plasmonsters are composed of slot waveguides that measure 100 nanometers across at their broadest points and only 20 nanometers across at the intersections (*inset*).

PHIL SAUNDERS *Space Channel Ltd.*

vibrations. In 1989 Thomas Ebbesen, then at the NEC Research Institute in Japan, found that when he illuminated a thin gold film imprinted with millions of microscopic holes, the foil somehow transmitted more light than was expected from the number and size of the holes. Nine years later Ebbesen and his colleagues concluded that surface plasmons on the film were intensifying the transmission of electromagnetic energy.

The field of plasmonics received another boost with the discovery of novel "metamaterials"—materials in which electron oscillations can result in astounding optical properties [see "The Quest for the Superlens," by John B. Pendry and David R. Smith; Scientific American, July 2006]. Two new classes of tools have also accelerated progress in plasmonics: recent increases in computational power have enabled investi-

PLASMONIC THERAPY FOR CANCER



A proposed cancer treatment would employ plasmonic effects to destroy tumors. Doctors would inject nanoshells—100-nanometer-wide silica particles with an outer layer of gold (*inset*)—into the bloodstream. The nanoshells would embed themselves in a fast-growing tumor. If near-infrared laser light is pointed at the area, it would travel through the skin and induce resonant electron oscillations in the nanoshells, heating and killing tumor cells without harming the surrounding healthy tissue.

gators to accurately simulate the complex electromagnetic fields generated by plasmonic effects, and novel methods for constructing nanoscale structures have made it possible to build and test ultrasmall plasmonic devices and circuits.

At first glance, the use of metallic structures to transmit light signals seems impractical, because metals are known for high optical losses. The electrons oscillating in the electromagnetic field collide with the surrounding lattice of atoms, rapidly dissipating the field's energy. But the plasmon losses are lower at the interface between a thin metal film and a dielectric than inside the bulk of a metal because the field spreads into the nonconductive material, where there are no free electrons to oscillate and hence no energy-dissipating collisions. This property naturally confines plasmons to the metallic surface abutting the dielectric; in a sandwich with dielectric and metal layers, for example, the surface plasmons propagate only in the thin plane at the interface [*see top illustration in box on preceding page*].

Because these planar plasmonic structures act as waveguides, shepherding the electromagnetic waves along the metal-dielectric boundary, they could be useful in routing signals on a chip. Although an optical signal suffers more loss in a metal than in a dielectric such as glass, a plasmon can travel in a thin-film metal waveguide for several centimeters before dying out. The propagation length can be maximized if the waveguide employs an asymmetric mode, which pushes a greater portion of the electromagnetic energy away from the guiding metal film and into the surrounding dielectric, thereby lowering loss. Because the electromagnetic fields at the top and bottom surfaces of the metal film interact with each other,

PHIL SAUNDERS *Space Channel Ltd.*

the frequencies and wavelengths of the plasmons can be adjusted by changing the thickness of the film. In the 1990s research groups led by Sergey Bozhevolnyi of Aalborg University in Denmark and Pierre Berini of the University of Ottawa developed planar plasmonic components that could perform many of the same functions—such as splitting guided waves—usually done by all-dielectric devices. These structures could prove useful in transmitting data from one part of a chip to another, but the electromagnetic fields accompanying the plasmons are too large to convey signals through the nanoscale innards of a processor.

To generate plasmons that can propagate through nanoscale wires, researchers have explored more complex waveguide geometries that can shrink the wavelength of the signal by squeezing it into a narrow space. In the late 1990s my lab group and a team led by Joachim Krenn of the University of Graz in Austria launched parallel efforts to produce these "subwavelength" surface-plasmon waveguides. Working with me at Caltech, Stefan Maier built a structure consisting of linear chains of gold dots, each less than 100 nanometers across. A visible beam with a wavelength of 570 nanometers triggered resonant oscillations in the dots, generating surface plasmons that moved along the chains, confined to a flattened path only 75 nanometers high. The Graz group achieved similar results and imaged the patterns of the plasmons carried along the chains. The absorption losses of these nanowires were relatively high, however, causing the signal to die out after it traveled a few hundred nanometers to a few microns (millionths of a meter). Thus, these waveguides would be suitable only for very short-range interconnections.

Fortunately, the absorption losses can be minimized by turning the plasmonic waveguides inside out, putting the dielectric at the core and surrounding it with metal [*see middle illustration in box on page 59*]. In this device, called a plasmon slot waveguide, adjusting the thickness of the dielectric core changes the wavelength of the plasmons. My lab at Caltech and Mark Brongersma's Stanford University group have shown that plasmon slot waveguides are capable of transmitting a signal as far as tens of microns. Hideki Miyazaki of the National Institute for Materials Science in Japan obtained a striking result by squeezing red light (with a wavelength of 651 nanometers in free space) into a plasmon slot waveguide that was only three nanometers thick and 55 nanometers wide. The researchers found that the wavelength of the surface plasmon propagating through the device was 51 nanometers, or about 8 percent of the free-space wavelength.

Plasmons propagate like the ripples that spread across the surface of a pond after you throw a stone in the water.

Plasmonics can thus generate signals in the soft x-ray range of wavelengths (between 10 and 100 nanometers) by exciting materials with visible light. The wavelength can be reduced by more than a factor of 10 relative to its free-space value, and yet the frequency of the signal remains the same. (The fundamental relation between the two—frequency times wavelength equals the speed of light—is preserved because the electromagnetic waves slow as they travel along the metal-dielectric interface.) This striking ability to shrink the wavelength opens the path to nanoscale plasmonic structures that could replace purely electronic circuits containing wires and transistors.

Just as lithography is now used to imprint circuit patterns on silicon chips, a similar process could mass-produce minuscule plasmonic devices with arrays of narrow dielectric stripes and gaps. These arrays would guide the waves of positive and negative charge on the metal surface; the alternating charge densities would be very much akin to the alternating current traveling along an ordinary wire. But because the frequency of an optical signal is so much higher than that of an electrical one—more than 400,000 gigahertz versus 60 hertz—the plasmonic circuit would be able to carry much more data. Moreover, because electrical charge does not travel from one end of a plasmonic circuit to another—the electrons bunch together and spread apart rather than streaming in a single direction—the device is not subject to resistance and capacitance effects that limit the data-carrying capacity of integrated circuits with electrical interconnects.

Plasmonic circuits would be even faster and more useful if researchers could devise a "plasmonster" switch—a three-terminal plasmonic device with transistorlike properties. My lab at Caltech and other research groups have recently developed low-power versions of such a switch. If scientists can produce plasmonsters with better performance, the devices could serve as the core of an ultrafast signal-processing system, an advance that could revolutionize computing 10 to 20 years from now.

THE AUTHOR

HARRY A. ATWATER is Howard Hughes Professor and Professor of Applied Physics and Materials Science at the California Institute of Technology. His research interests center on subwavelength-scale photonic devices for computing, imaging and renewable energy applications. In addition to devising plasmonic nanostructures, his group is actively exploring the use of new materials for solar power generation (photovoltaics), as well as the solar-driven generation of chemical fuels.

Nanoshells and Invisibility Cloaks

THE POTENTIAL USES of plasmonic devices go far beyond computing, however. Naomi Halas and Peter Nordlander of Rice University have developed structures called nanoshells that consist of a thin layer of gold—typically about 10 nanometers thick—deposited around the entire surface of a silica particle about 100 nanometers across. Exposure to electromagnetic waves generates electron oscillations in the gold shell; because of the coupling interaction between the fields on the shell's inner and outer surfaces, varying the size of the particle and the thickness of the gold layer changes the wavelength at which the particle resonantly absorbs energy. In this way, investigators can design the nanoshells to selectively absorb wavelengths as short as a few hundred nanometers (the blue end of the visible spectrum) or as long as nearly 10 microns (the near infrared).

This phenomenon has turned nanoshells into a promising tool for cancer treatment. In 2004 Halas, working with her Rice colleague Jennifer West, injected plasmonic nanoshells into the bloodstream of mice with cancerous tumors and found that the particles were nontoxic. What is more, the nanoshells tended to embed themselves in the rodents' cancerous tissues rather than the healthy ones because more blood was circulated to the fast-growing tumors. (The nanoshells can also be attached to antibodies to ensure that they target cancers.)

Fortunately, human and animal tissues are transparent to radiation at certain infrared wavelengths. When the researchers directed near-infrared laser light through the mice's skin and at the tumors, the resonant absorption of energy in the embedded nanoshells raised the temperature of the cancerous tissues from about 37 degrees Celsius to about 45 degrees C.

The photothermal heating killed the cancer cells while leaving the surrounding healthy tissue unharmed. In the mice treated with nanoshells, all signs of cancer disappeared within 10 days; in the control groups, the tumors continued to grow rapidly. Houston-based Nanospectra Biosciences is currently seeking permission from the Food and Drug Administration to conduct clinical trials of nanoshell therapy in patients with head and neck cancer.

Plasmonic materials may also revolutionize the lighting industry by making LEDs bright enough to compete with incandescent bulbs. Beginning in the 1980s, researchers recognized that the plasmonic enhancement of the electric field at the metal-dielectric boundary could increase the emission rate of luminescent dyes placed near the metal's surface. More recently, it has become evident that this type of field enhancement can also dramatically raise the emission rates of quantum dots and quantum wells—tiny semiconductor structures that absorb and emit light—thus increasing the efficiency and brightness of solid-state LEDs. In 2004 my Caltech colleague Axel Scherer, together with co-workers at Japan's Nichia Corporation, demonstrated that coating the surface of a gallium nitride LED with dense arrays of plasmonic nanoparticles (made of silver, gold or aluminum) could increase the intensity of the emitted light 14-fold.

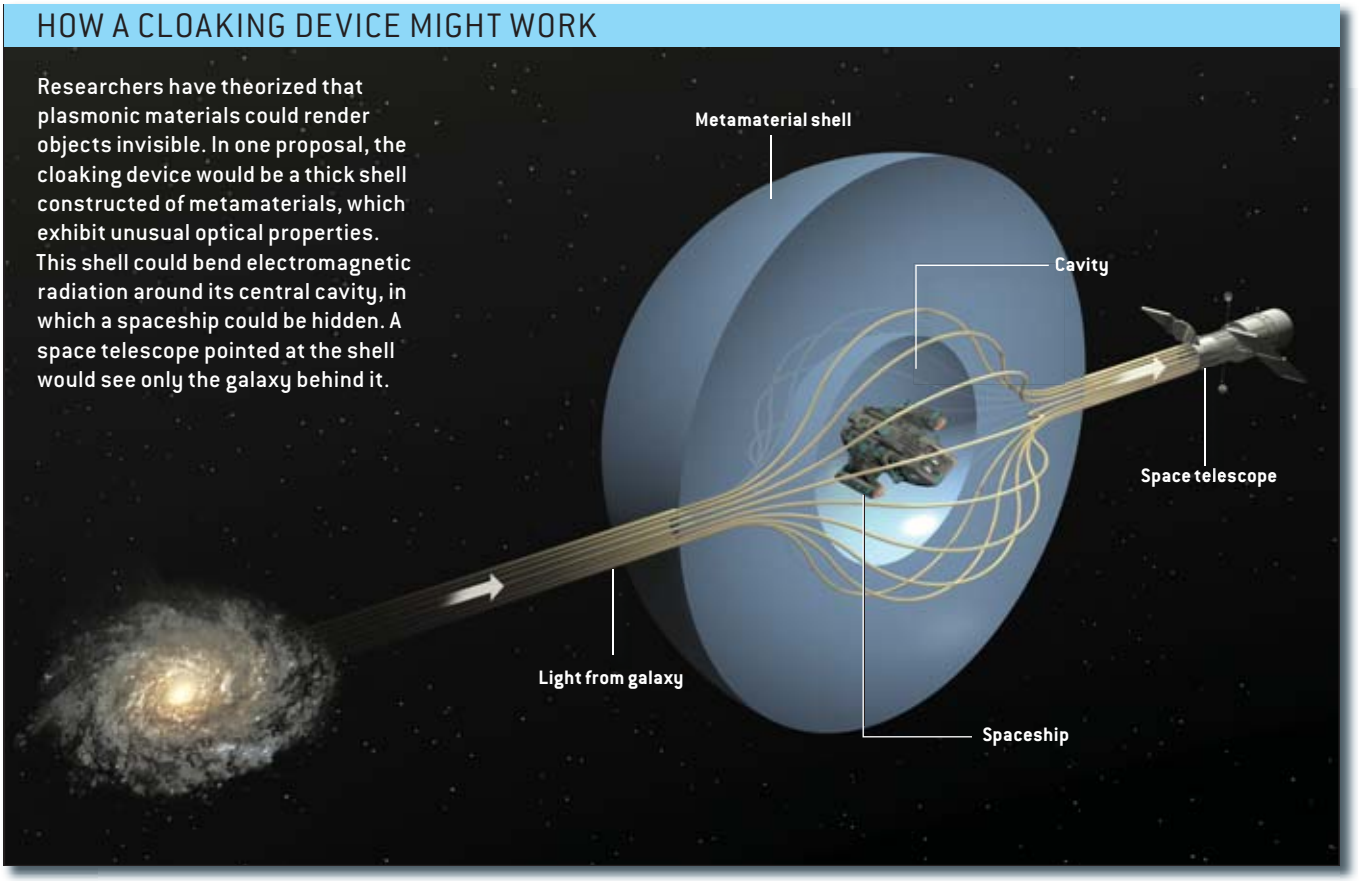
Furthermore, plasmonic nanoparticles may enable researchers to develop LEDs made of silicon. Such devices, which would be much cheaper than conventional LEDs composed of gallium nitride or gallium arsenide, are currently held back by their low rates of light emission. My group at Caltech, working with a team led by Albert Polman of the FOM Institute for Atomic and Molecular Physics in the Netherlands, has shown that coupling silver or gold plasmonic nanostructures to silicon quantum-dot arrays could boost their light emission by about 10 times. Moreover, it is possible to tune the frequency of the enhanced emissions by adjusting the dimensions of the nanoparticles. Our calculations indicate that careful tuning of the plasmonic resonance frequency and precise control of the separation between the metallic particles and the semiconductor materials may enable us to increase radiative rates more than 100-fold, allowing silicon LEDs to shine just as brightly as traditional devices.

Scientists are even working on a plasmonic analog to a laser. Mark Stockman of Georgia State University and David Bergman of Tel Aviv University have described the physics of such a device, which they called a SPASER (for *s*urface *p*lasmon *a*mplification of *s*timulated *e*mission of *r*adiation). Although the SPASER exists only in theory so far, the researchers have suggested routes to fabricating it using semiconductor quantum dots and metal particles. Radiative energy

LYCURGUS CUP, a Roman goblet dating from the fourth century A.D., changes color because of the plasmonic excitation of metallic particles within the glass matrix. When a light source is placed inside the normally greenish goblet, it looks red.

BRITISH MUSEUM, LONDON, U.K./BRIDGEMAN ART LIBRARY

HOW A CLOAKING DEVICE MIGHT WORK



PHIL SAUNDERS *Space Channel Ltd.*

from the quantum dots would be transformed into plasmons, which would then be amplified in a plasmonic resonator. Because the plasmons generated by a SPASER would be much more tightly localized than a conventional laser beam, the device could operate at very low power and selectively excite very small objects. As a result, SPASERs could make spectroscopy more sensitive and pave the way for hazardous-materials detectors that could identify minute amounts of chemicals or viruses.

Perhaps the most fascinating potential application of plasmonics would be the invention of an invisibility cloak. In 1897 H. G. Wells published *The Invisible Man,* a tale of a young scientist who discovers how to make his own body's refractive index equal to that of air, rendering him invisible. (A material's refractive index is the ratio of the speed of light in a vacuum to the speed of light in the material.) Exciting a plasmonic structure with radiation that is close to the structure's resonant frequency can make its refractive index equal to air's, meaning that it would neither bend nor reflect light. The structure would absorb light, but if it were laminated with a material that produces optical gain—amplifying the transmitted signal just as the resonator in a SPASER would—the increase in intensity would offset the absorption losses. The structure would become invisible, at least to radiation in a selected range of frequencies.

A true invisibility cloak, however, must be able to hide anything within the structure and work for all frequencies of visible light. The creation of such a device would be more difficult, but some physicists say it is possible. In 2006 John B. Pendry of Imperial College London and his colleagues showed that a shell of metamaterials could, in theory, reroute the electromagnetic waves traveling through it, diverting them around a spherical region within [*see box above*].

Although Wells's invisible man may never become a reality, such ideas illustrate the rich array of optical properties that inspire researchers in the plasmonics field. By studying the elaborate interplay between electromagnetic waves and free electrons, investigators have identified new possibilities for transmitting data in our integrated circuits, illuminating our homes and fighting cancer. Further exploration of these intriguing plasmonic phenomena may yield even more exciting discoveries and inventions. SA

MORE TO EXPLORE

Surface Plasmon Subwavelength Optics. William L. Barnes, Alain Dereux and Thomas W. Ebbesen in *Nature,* Vol. 424, pages 824–830; August 14, 2003.

Plasmonics: Localization and Guiding of Electromagnetic Energy in Metal/Dielectric Structures. Stefan A. Maier and Harry A. Atwater in *Journal of Applied Physics,* Vol. 98, No. 1, Article No. 011101, 10 pages; July 2005.

Plasmonics: Merging Photonics and Electronics at Nanoscale Dimensions. Ekmel Ozbay in *Science,* Vol. 311, pages 189–193; January 13, 2006.

Plasmonics: Fundamentals and Applications. Stefan A. Maier. Springer Verlag, 2007.

Recent experiments show that these birds use logic to solve problems and that some of their abilities approach or even surpass those of the great apes

BY BERND HEINRICH AND THOMAS BUGNYAR

A trapper in the north woods observes a common raven (*Corvus corax*) roll over on its back with its feet in the air next to a beaver carcass on the snow. A biologist laboriously climbs a cliff to band raven nestlings, and the birds' parents rain down loose rocks from above. A lone raven clamors loudly near a remote cabin, alerting a man next to it to look up and see a hidden cougar that is about to spring on him.

Each of these three people presumed to know what the ravens were up to. The trapper thought the raven was playing possum, pretending it had been poisoned to keep other ravens away so it could have the beaver carcass to itself. The biologist thought the raven pair was deliberately trying to hit him with rocks so he would go away. The man at the remote cabin thought the raven had alerted him to save his life.

These various hypotheses cannot be discounted, but most of us who have become intimate with ravens might offer other, more likely explanations. Ravens are perhaps the most playful of all birds, and they regularly roll on their back apparently for the fun of it. They often hammer the substrate in anger, wherever they may be perched, when a predator is near their nest. And they are known to lead carnivores to potential prey they cannot overpower themselves, so the bird may have been leading the cougar to the man.

Raven anecdotes are legion, and many suggest the birds are clever, but stories do not provide proof of diabolical cleverness. Even more straightforwardly sophisticated behavior by ravens—such as their habit of carving a block of suet into chunks to carry off manageable portions, their precise stacking of crackers to allow them to fly away with the whole stack, their manipulation of two doughnuts so they can carry both at the same time, and their making of false food caches that mislead raiders—does not prove that the birds are able to consciously contemplate alternative actions and choose the most appropriate ones.

Mere observations, after all, cannot rule out other possibilities, such as instinct or learning to perform specific actions by rote. Indeed, until the 1990s, probably only one careful scientific test implied logical reasoning in ravens of the type we take for granted in humans. This was a set of experiments published in 1943 by Otto Koehler of the former Zoological Institute of Konigsberg. He demonstrated that his 10-year-old pet raven, Jakob, could count up to seven by training

JEN CHRISTIANSEN

JUST HOW smart are Ravens?

TO GET A TREAT suspended on a string tied to the perch, a raven has to follow a precise sequence of steps—reach down and grasp the string, pull up on it, put the pulled-up string on the perch, step on it with enough pressure to hold it there, let go of the string, and repeat the process. Some mature ravens studied the situation for several minutes and then performed the entire procedure on their first try—a sign they used logic.

it to retrieve food from under one of several vessels with various numbers of spots on the lid. But studies undertaken in the past few years, mostly by the two of us, have finally offered some hard proof that ravens are indeed intelligent, in that they are able to use logic to solve problems. What is more, we found, to our astonishment, that they can even distinguish one individual from another. In that way, too, they are much like humans; we could not form societies (except those akin to insects') without this ability.

Proof of Problem Solving

RAVENS ARE NOT the only birds generally reputed to be smart. In the past two decades a virtual avalanche of studies have revealed that certain of their corvid relatives (which include the smaller crows, as well as jays, magpies and nutcrackers) possess surprising and sophisticated mental capacities. In some species, these abilities appear to be on par with, or exceed, those of the great apes. For example, nutcrackers have phenomenal memories that encompass thousands of food cache locations, a capacity that would challenge most humans. The New Caledonia crow (*Corvus moneduloides*) has been shown to fashion tools out of parts of pandanus leaves and to use them to extract grubs from crevices in wood. What has not been known, though, is to what extent such remarkable feats involve innate blind programming versus rote learning and memory (through past trial and error) versus reasoning (choosing from alternatives that are represented in the mind and evaluated).

The two of us have devised experiments to try to distinguish the role and the relative importance of those possibilities. In the first of these tests, we confronted individual ravens with food hanging on a string. To get the treat, they had to reach down from a perch, grasp the string in the bill, pull up on the string, place the loop of pulled-up string onto the perch, step on the string and apply the appropriate pressure to prevent slippage, then let go of the string and reach down again, repeating this sequence six or more times in a row.

We found that at least some grown-up birds would examine the situation for minutes on end and then perform this multistep procedure on their first try in as little as 30 seconds, without any preliminary efforts at trial and error. In classical "shaping" of behavior in laboratory animals, the steps in a desired behavioral sequence are typically rewarded with food, whereas false steps are punished with electric shock. The links in the sequence are established presumably without the animal needing to understand how a single one of them contributes to the overall outcome. Our animals, however, would not have encountered this task in the wild and therefore could not have learned how to do it in the past by trial and error. Hence, the simplest suggestion is that they imagined possibilities and figured out what steps to take.

Passing the test did require maturation. Young birds (a month or two past fledging) are unable to perform this complex behavior. And year-old birds require on average six minutes to solve the puzzle, during which they overtly test various possibilities (such as flying at the food, trying to rip off the string, pecking the string, or yanking and twisting it).

No one step in the pull-up sequence was rewarded with food; the raven had to accomplish the whole lengthy sequence in order to eat. One might argue, however, that each step is "mentally" rewarded and therefore reinforced simply because the food comes nearer and that the animal does not necessarily know that each step in the sequence brings it closer to its goal. But that explanation does not hold water. If each step were acquired by trial-and-error learning, it would require numerous trials, and the entire pull-up sequence would probably take months of train-

Overview/*Raven Intelligence*

- Although the clever behavior of ravens convinces most people that the birds are intelligent, it does not prove that they can consciously contemplate alternatives and choose the best one.
- The authors set out in search of such proof by devising a series of experiments that involved pulling up meat on a string and hiding food from competitors.
- They found that ravens can use logic to solve problems and that they can distinguish individuals (both humans and other ravens) and attribute knowledge to them.

JEN CHRISTIANSEN

ing. But this is not what happened. The birds acted as if they knew what they were doing.

We could only know that they knew, however, if they behaved according to certain predictions. If the ravens knew what they were doing, for example, then they should also know what they had done. They should know, for instance, that after they had pulled up the treat on the string, it was still attached to the perch. To find out if they understood, we shooed them away from the perch after they had pulled up meat. If they dropped the meat, we figured they knew it was attached to the perch; if they flew off with it (and had it ripped out of their bills), they did not know. Most dropped it, even though they always flew off with meat that had string attached to it but that was laid (and not tied) onto the perch.

Knowing requires few or no trials, whereas trial-and-error learning requires no logic. And so we sought still another test to find out if the birds might have solved the challenge of meat pull-ups by random movements that happened to be rewarding but that were unsupported by logic. This time we confronted naive birds with the same physical choices but with what we hoped would be, to them, an illogical situation, namely, a looped string that had to be pulled *down* to make the food come up.

Under this situation, the ravens were still interested in the food; they investigated the setup and pecked and yanked on the string, thus making it at times come a little closer. They soon gave up, however, and none learned to access the food even though the same pull/step/release sequences that quickly delivered the food before *could* have provided it again. We believe, therefore, that the direct pull-up was mastered quickly and sometimes almost "instantly" only because it was supported by logic. Apparently ravens have the ability to test actions in their minds and project the outcomes of those actions. That capacity is probably lacking or present only to a limited extent in most animals—and for a good adaptive reason.

The Benefits of Intelligence

BY SOME PROCESS that still remains one of the great unsolved mysteries of biology, exquisitely precise behaviors can be genetically programmed in animals with brains no larger than a pinhead. Consider, for example, a wasp that makes paper expertly from the time it is born, that fashions a nest of precise architecture with that paper, while another wasp uses mud to make a mortar nest of a very different but also very specific shape. Similarly, birds of each species are programmed to make precisely predetermined nests. All barn swallows build a shelf nest from mud that hardens when it dries. Cliff swallows construct ovenlike nests, also out of mud but with a small round entrance hole.

CONFRONTED with having to pull *down* on a string to pull food up, inexperienced ravens (those who had not mastered pulling food up on a string) seemed to decide that pulling down to make something move up was illogical and soon gave up. (The wire mesh prevented the birds from pulling up on the string.)

JEN CHRISTIANSEN

None of the most intricate of these behaviors is learned, nor do the behaviors depend on thinking (although learning and thinking can modify some genetically programmed behavior). Thinking and logic can be notoriously unreliable and can lead to much mayhem, as we all well know. The big question, then, is why, if behavior can be so precisely preprogrammed, some animals (ourselves, for example) are consigned to muddling. Why are they not endowed as most animals are to "do it right," except perhaps after experiencing the many things that can go disastrously wrong?

The usual answer is that such animals evolved in a complex and unpredictable environment in which prewired responses were inappropriate. If the animal can identify individuals, and it lives among others that can in turn identify it as a separate entity, then the environment for each of them is indeed complex. Social life among most animals that do identify individuals is thus often cited as the driving force for the evolution of intelligence: in such a context the ability to predict the responses of others, who constitute the main relevant feature of the environment, becomes extremely valuable. We were therefore led to consider the ravens' social environment to try to understand why they, more than many other animals, would have benefited from becoming intelligent.

The Natural Environment of the Raven

MUCH OF THE RAVENS' natural history suggests that they had to evolve to cope with ever changing short-term circumstances. These birds are basically opportunists that do some hunting but have specialized to live off the food other animals kill. The predators that provide them with food, however, are unpredictable and can also kill them. Lengthy conditioning through trial and error would appear to be fatally costly,

because the first mistake could cost the birds their life, and a totally programmed response to an unpredictable carnivore could be equally dangerous.

The way they compete for food with other ravens also requires coping with ever changing circumstances. Territorial pairs of ravens attempt to monopolize food bonanzas, and members of the large population of juveniles and nonbreeders mount a counterstrategy of recruiting flock mates that then overpower the territorial defenders. Significantly, however, the same behavior that gains the crowd access to the food, as well as diluting the danger by their numbers, intensifies the competition for the resource.

Food bonanzas are not only provided by carnivores, they are often quickly consumed by them. It pays the attending ravens to get an early start in the feeding cycle, preferably next to the carnivores while they are still eating. To do that, the birds need to be able to predict the predator's behavior, such as whether or when the animal might attack, how far it can jump, and how it may be distracted. Some of that knowledge needs to be in place before the raven is distracted by feeding, because in that context practice could be deadly.

Indeed, the birds acquire practice more safely early in their lives. Juvenile birds, when undistracted by feeding, routinely "test" the reactions of large animals such as wolves and other carnivores by interacting with them, usually by landing nearby and then nipping them from the rear. It is unlikely that such behavior is tactically deliberate. More likely it is a form of "play," defined in the considerable scientific literature on the subject as a behavior that has no immediately discernible function but that commonly has an ultimate function, one that is not consciously intended but that proves useful anyway.

Even youngsters recognize that nipping carnivores is dangerous (they display fear when they do it), and thus they must be wired to engage in such activity because the risky play ultimately aids survival—presumably by giving them experience in gauging how much they can get away with around their carnivore companions. By such provocation they soon learn which animals to trust and the distances required for safety. Conversely, their nearly constant presence around the carnivores accustoms the larger animals to the birds, and they gradually learn to ignore them. But getting along with dangerous carnivores is only a means to the end of getting access to a rich supply of food.

The often short time that a food bonanza lasts (deer carcasses in the Maine woods, for example, are consumed in a day or two) places a premium on hauling the food away first and eating it later. Like other corvids, ravens cache food for later use. At a contested carcass, they busily haul off one load of meat after another and hide it by burying it and camouflaging it with debris so that it is totally removed from view. Also, like many other corvids, ravens memorize the exact locations of their numerous caches and usually retrieve them within hours or days. Yet unlike most food-caching birds, ravens carefully observe the caching behavior of competitors, and they memorize the precise locations not only of the caches that they make themselves but also of those they see others make.

THE AUTHORS

BERND HEINRICH and *THOMAS BUGNYAR* share a fascination with the intellectual abilities of ravens; they investigated the birds together when Bugnyar was a research associate at the University of Vermont, where Heinrich has been a professor of biology since 1980. Heinrich received his Ph.D. from the University of California, Los Angeles, and spent 10 years in the department of entomology at U.C. Berkeley before going to Vermont. He is the author of several well-known books, including *Ravens in Winter* (Simon and Schuster, 1989) and *Mind of the Raven* (HarperCollins, 1999), which will be reissued in a new edition this summer. This is his seventh article for *Scientific American*. Bugnyar received his Ph.D. from the University of Vienna for work that he did with ravens at the Konrad Lorenz Research Station in Grünau, Austria. He is now a lecturer in the school of psychology at St. Andrews University in Scotland.

JIM BRANDENBURG *Minden Pictures*

Playing with and Hiding Food

REALIZING THAT PLAY with predators apparently helps ravens learn how to size up situations and act accordingly, we decided to test whether play really did help young birds gain the ability to adjust their behavior flexibly. Caching behavior offered a promising field for this inquiry, and a large aviary we had designed to simulate natural conditions of trees and ground cover provided a convenient setting for the experiments.

We found, as we had seen before, that ravens actively avoid one another while caching. They prefer to do their hiding in private, or they use trees or rocks to block the view from others. Cache owners also attempt to chase off potential pilferers. And we discovered that these caching skills originate from innate play responses that provoke their protagonists to react and that then permit learning of the appropriate responses. This testing and learning process starts among the siblings shortly after they leave the nest and begin to follow their parents, learning to identify the

MATURE RAVENS, which have a wingspan of 1.25 meters and weigh about 1.25 kilograms, move in on an animal recently brought down by wolves in Yellowstone National Park. The playful behavior of young ravens, the authors suggest, teaches them how to get along with much larger carnivores, on whom they depend for much of their food.

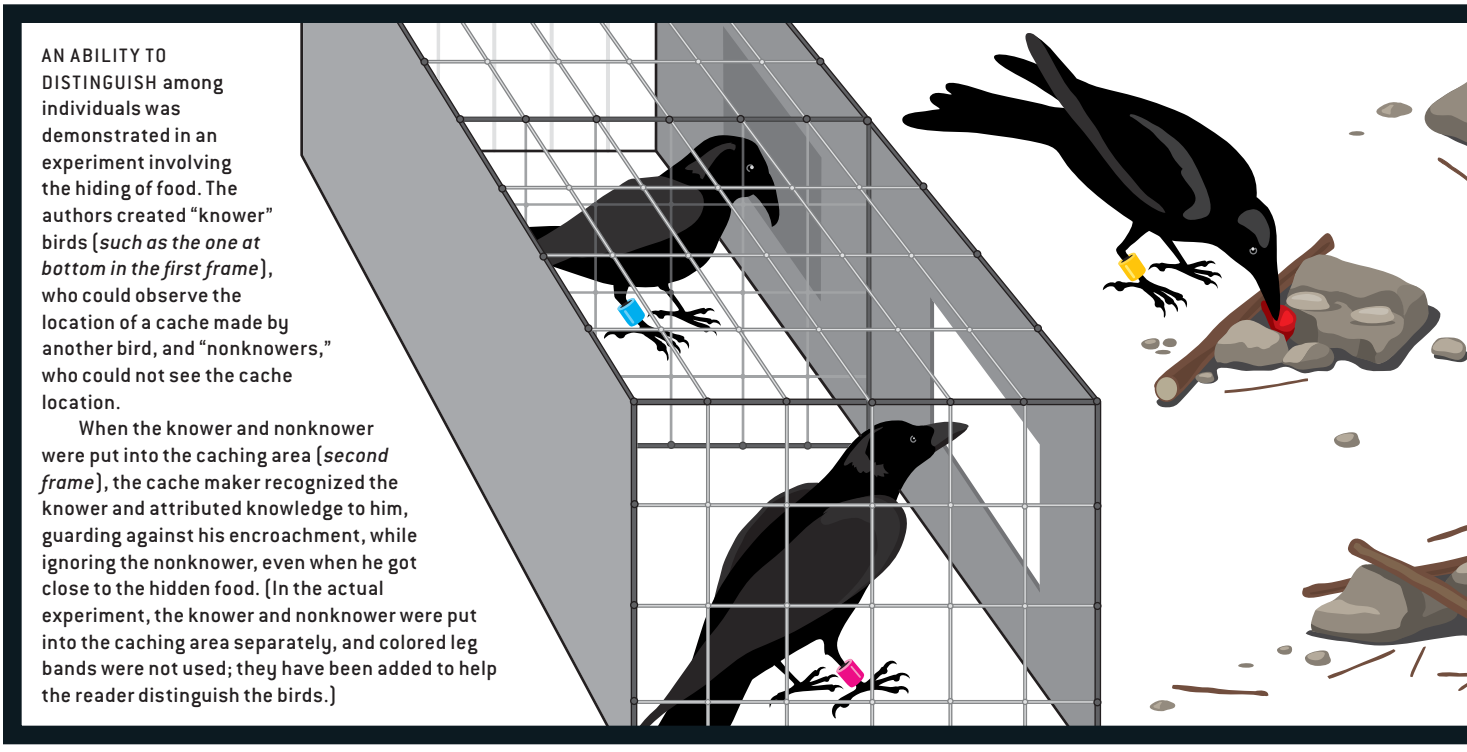
great variety of small food objects, such as insects and fruits.

Young ravens in the nest and several days out of it manipulate all sorts of objects with their bills, and like the tail tweaking of wolves, such behavior is defined as play because it brings no proximate advantages, yet it requires the expenditure of time and energy or the taking of risks. In essence, these objects are "toys." In experiments with a brood of tame ravens, one of us acted as parent and daily led the birds around. The juveniles kept themselves busy picking up twigs, leaves, flowers, pinecones, pebbles, cigarette butts, coins and other objects we had "seeded" on the ground. Within days the young ravens largely ignored the inedible items and eagerly sought out the edible. Their playful object manipulation gave them the experience of learning about their environment. Because the ravens normally would still have been fed by their parents at that time, they could afford the apparently useless behavior, whose benefit would only become manifest later.

While the young birds are learning to distinguish the edible from the inedible, they simultaneously increase and shape their caching skills. At first they indiscriminately tuck some of the items that catch their attention up against other objects. Later they shove them partially out of view into crevices, and by a month or two the still-dependent young cover cached objects with debris. Because these young ravens ordinarily cache in front of their siblings and parents, with whom they travel for several months after fledging, the siblings often recover the hidden items. We wondered whether the playful caching of the inedible items helps them gain the ability to predict others' behavior so that they can successfully hide and defend their valuable food items later on.

One problem with tests of whether early experience ultimately affects adult behavior is that it is difficult to control the experience that any one bird might have. We noticed, however, that the birds also watched *us* and retrieved "caches" of food that we, their surrogate parents and companions, hid from them. And we could control our behavior! For an experiment, we therefore designated one person as a "thief" who always stole the cached objects that the young birds hid in apparent play, whereas a second person consistently examined the birds' object caches but never retrieved any of them. In the test situation, we provided the then more mature ravens with food rather than inedible objects. This time the same two people, either the thief or the nonthief, stood by and only observed the birds' behavior without interfering.

Confronted with the potential thief, the ravens significantly delayed the time until they cached their food (as if they were waiting for a time when the thief was not looking), and they retrieved those caches that they did make when he walked near them. In contrast, the previously benign person who had not stolen cached objects did not elicit delays in stowing food, and the birds ignored him when he went near one of their caches. This experiment thus showed not only that the birds improve their food-caching skills after experiencing others' raiding their object caches but that they distinguish individuals (in this case, humans).



AN ABILITY TO DISTINGUISH among individuals was demonstrated in an experiment involving the hiding of food. The authors created "knower" birds (*such as the one at bottom in the first frame*), who could observe the location of a cache made by another bird, and "nonknowers," who could not see the cache location.

When the knower and nonknower were put into the caching area (*second frame*), the cache maker recognized the knower and attributed knowledge to him, guarding against his encroachment, while ignoring the nonknower, even when he got close to the hidden food. (In the actual experiment, the knower and nonknower were put into the caching area separately, and colored leg bands were not used; they have been added to help the reader distinguish the birds.)

Discriminating "Knowers"

WILD RAVENS in the field commonly feed in crowds, as we have described, and spend much of their time busily caching food for later use. In this situation, it would be almost impossible for any one bird to chase off every other bird that happened to wander near one of its often dozens of caches. Yet adult ravens greatly reduce the possibility of having competitors see them make their caches or having to chase others away who might be potential raiders by scattering their valuable hoards over an area of many square kilometers. In the confines of our aviary, though, it is often not possible for an individual to escape the watchful eye of competitors, and such a situation gave us the opportunity to determine experimentally whether the birds are able to discriminate among raven competitors based on what the competitors could potentially know, just as they had discriminated among different humans.

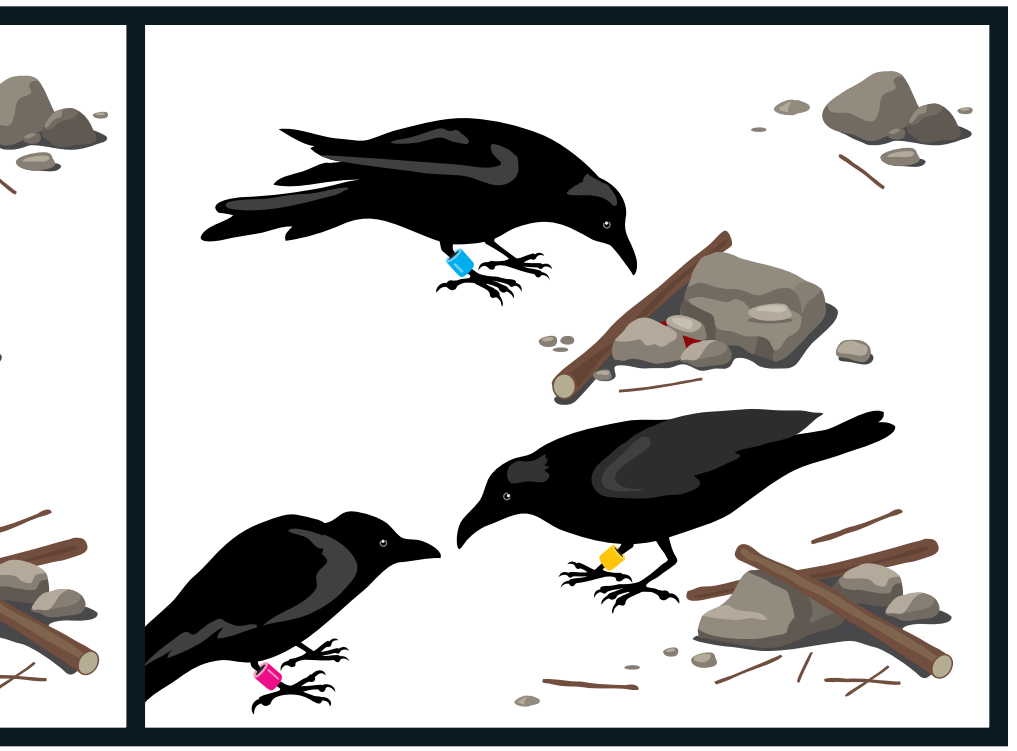
In this series of tests, we capitalized on our knowledge that ravens distinguish one another (as well as others of another species—namely, us) as individuals. We created "knower" birds—those that had observed the locations of a given bird's caches—versus "nonknowers"—those that could not have observed the cache locations. We then paired the cache maker with these different competitors, much as we had in the experiments to examine the responses of young birds to thieves and nonthieves. In this case, however, the experimental setup called for a modification of the aviary.

One large compartment of the aviary served as the caching arena. We separated a smaller compartment from this area with an opaque wall. In the wall we made a small viewing window and placed a perch in front of it, on which a bird could land and look through wire screening to observe the caching bird in the main compartment. Next to the viewing compartment a similar compartment also contained a bird, but in this case the viewing window was occluded by a curtain. Thus, two birds had the same aural access to a caching bird, but only one had visual access.

Both the birds in the small compartments would soon (within five minutes) be allowed into the caching area to search for food. They were therefore motivated to watch the cacher. Indeed, the knower bird normally perched to try to watch the cacher, and the nonknower in the curtained compartment tried to lift the curtain to watch (we found we had to secure the curtain so it could not be lifted). After the cacher had made three caches, we removed it from the large compartment and five minutes later let it back in to retrieve its hidden stores. It was allowed to retrieve either in private or in the presence of the knower or of the nonknower. (Both these potential thieves were of subordinate status to the cacher so that they would not extinguish its responses defending its hoard.)

Cachers typically retrieve their food when robbery seems imminent, and indeed the experiments showed that they retrieved significantly more of their caches when they were paired with knowers than with nonknowers or when they were alone. Furthermore, when a knower came within two meters of the camouflaged food, the cache maker chased it away, whereas he ignored nonknowers. We speculated that the cachers remembered which birds had watched them make any particular cache and later discriminated against them, as though attributing knowledge

JEN CHRISTIANSEN



to those that watched. They apparently anticipated the watcher's intentions and guarded against his expected raiding behavior. But the knowers also guarded against the defensive behavior of the cache makers: they did not go directly to the caches in their presence but waited until they were at some distance. The results of these experiments suggest the attribution of knowledge and the anticipation of a response.

In another version of the same experiment, we tried to control for the possibility that the apparent knowers had been inadvertently providing some subtle cues that the cache defenders could read, rather than that the cache defenders actually knew that the knowers had seen them. So we had a human, who would stand by passively, make the caches. As we predicted from the results of the first experiments, knowers rushed to pilfer the human-made cache if they were paired with another knower. On the other hand, when paired with an ignorant *dominant* competitor (who would attack the raider to get the cache), they delayed, by 10 times on average, the duration before they approached the cache, waiting until the dominant raven was occupied at a distance. These results do not totally exclude the possibility that the knowers provide some subtle unknown cues that the cache raiders may use, but such cueing is unlikely, and the findings strongly suggest that the birds engage in amazingly sophisticated behavior based on an ability either to interpret or to anticipate the actions of others.

What Are Ravens Thinking?

THE STUDY OF mental states of animals who cannot report their thoughts to us is beset with difficulties. Indeed, we do not know and perhaps can never know what goes on in the mind of another animal or even other individuals of our own species. Yet invoking Occam's razor and accepting the simplest explanation, as is traditional in science, we can conclude that our experiments provide a consistent affirmation that ravens use some kind of mental representation to guide their actions. The results of the string-pull experiments indicate the use of logic. And the pilfer and antipilfer tactics suggest that ravens judge their competitors on the basis of what they remember them paying attention to. They then attribute to the competitors the capacity of knowing, and they integrate that knowledge together with dominance status into strategic decisions for making and retrieving caches.

Learning occurs, but it alone cannot account for all of the observed behavior, because the behavior is exhibited very fast, almost immediately, without any trials and errors. We speculate that the birds start from an innate framework of prewired, playlike behavior, which generates the experience that is a prerequisite for learning. Learning may later translate to conscious awareness—that is, an ability to use logic—that would be useful in the highly unpredictable context of a social milieu with competitors and predators and that can then be transferred to any other novel context, such as pulling food up on a string.

We do not know how unusual the ravens' kind of ability is in nonhumans. But we suspect that although it may not be rare, it would generally be narrowed to specific kinds of tasks, because the underlying instincts and learning tendencies that are tailored to the animals' environment vary greatly. In the raven, however, it may be more general than in most. We think so because no other bird we know of is as playful as the raven and consequently exposes itself to such a variety of contingencies. These tendencies may have allowed it to become the most widely naturally distributed bird in the world, inhabiting the same continents as humans and being at home in as many diverse habitats. SA

MORE TO EXPLORE

Ravens, *Corvus corax*, Differentiate between Knowledgeable and Ignorant Competitors. Thomas Bugnyar and Bernd Heinrich in *Proceedings of the Royal Society London*, Series B, Vol. 272, No. 1573, pages 1641–1646; August 22, 2005.

Testing Problem Solving in Ravens: String-Pulling to Reach Food. Bernd Heinrich and Thomas Bugnyar in *Ethology*, Vol. 111, No. 10, pages 962–976; October 2005.

Pilfering Ravens, *Corvus corax*, Adjust Their Behaviour to Social Context and Identity of Competitors. Thomas Bugnyar and Bernd Heinrich in *Animal Cognition*, Vol. 9, No. 4, pages 369–376; October 2006.

The Movies in Our Eyes

The retina processes information much more than anyone has ever imagined, sending a dozen different movies to the brain

By Frank Werblin and Botond Roska

We take our astonishing visual capabilities so much for granted that few of us ever stop to consider how we actually see. For decades, scientists have likened our visual-processing machinery to a television camera: the eye's lens focuses incoming light onto an array of photoreceptors in the retina. These light detectors magically convert those photons into electrical signals that are sent along the optic nerve to the brain for processing. But recent experiments by the two of us and others indicate that this analogy is inadequate. The retina actually performs a significant amount of preprocessing right inside the eye and then sends a series of partial representations to the brain for interpretation.

We came to this surprising conclusion after investigating the retinas of rabbits, which are remarkably similar to those in humans. (Our work with salamanders has led to similar results.) The retina, it appears, is a tiny crescent of brain matter that has been brought out to the periphery to gain more direct access to the world. How does the retina construct the representations it sends? What do they "look" like when they reach the brain's visual centers? How do they convey the vast richness of the real world? Do they impart meaning, helping the brain to analyze a scene? These are just some of the compelling questions the work has begun to answer.

Overall, we have found that specialized nerve cells, or neurons, deep within the retina project what can be thought of as a dozen movie tracks—distinct abstractions of the visual world. Each track embodies a primitive representation of one aspect of the scene that the retina continuously updates and streams to the brain. One track, for example, transmits a line-drawing-like image that details only the edges of objects. Another responds to motion, often in a specific direction. Some tracks carry information about shadows or highlights. The representations of still other tracks are difficult to categorize.

Each track is transmitted by its own population of fibers within the optic nerve to higher visual centers in the brain, where even more sophisticated processing takes

place. (The human auditory system has a similar architecture: each auditory nerve carries information about a very limited range of pitches, and the brain combines them.) Investigators studying the visual cortex have shown that features such as motion, color, depth and form are processed in various regions and that a lesion to a given region can cause a deficit in sensing one specific feature. But the brain's ability to even sense such features in the first place originates in the retinal movies.

The diagrams on the following pages convey our best explanations for how the retina creates the surreal electrical images that inform the brain. As we continue our research, we are beginning to shed some light on how each of the movies is constructed, but by no means are we ready to offer a full model. The 12 movies carry all the information the brain will ever receive to interpret the visual world, but we cannot yet say how their patterns are integrated. It could be that the movies serve simply as elementary clues, a kind of scaffolding upon which the brain imposes constructs. This notion is not dissimilar to the well-described "mind's eye" that knits the words of a novel into a meaningful narrative.

Although the retina's representations appear to fully capture the visual facts of a scene, such as a dinner table, waterfall or talking face, essential components also seem to be missing. Nothing about the feeling, attitude, texture or focus of the scene appears to be present. Perhaps these traits are somehow inherent in the movie tracks the brain interprets. Or perhaps, by using rabbit retinas, we may have failed to find all the representations that would be captured by a human retina—"high resolution" ones that might extract qualities such as feeling in ways we have yet to uncover.

Nevertheless, it is clear that the retina's representations form a natural visual language. Understanding that language has special significance today. Groups around the world are attempting to restore vision to the blind by introducing an artificial sensor right in front of the optic nerve that would take over for the retina. The work has advanced, but the results remain relatively crude, with transmissions limited to vague versions of basic patterns. Human trials have begun at the University of Southern California's Doheny Eye Institute and are about to get under way at Wayne State University Medical School. The final goal of these trials is probably far off, but their success ultimately lies in providing the brain with patterns of activity similar to those that are normally supplied by the retina, incorporating the natural language of vision. The subsequent challenge will be to discover how to "hook up" each abstraction to the appropriate fibers in the optic nerve.

A detailed understanding of the natural language of vision formed within the retina is needed for effective prosthetic devices. At the same time, this understanding will help investigators learn much more about how the eye and brain together see clearly, are deceived by optical illusions, track fast-moving objects and fill in the missing pieces inherent in any rendering on a television, computer or drive-in movie screen. We hope our description of the retina's preprocessing power is a step toward that end.

Overview/*Surreal Vision*

- The retina does much more than pass simple signals to the brain. Surprisingly, it extracts a dozen distinct representations of a visual scene—sophisticated, ghostlike movies formed by relatively few types of neurons.
- The brain uses these abstractions to build a visual world sharp with detail and rich in meaning.
- Understanding the "visual language" that these movies carry will aid researchers who are building artificial sensors that could help the blind see. Such insights should also bolster efforts to pin down how the eye and brain see clearly, as well as how they can be deceived.

Active Anatomy

The retina's surprising behavior arises from its complex architecture. Painstaking experiments by many specialists have added physiological detail to the classic model of retinal circuitry first delineated by the great Spanish anatomist Santiago Ramón y Cajal a century ago and repeated in textbooks ever since. The transparent retina ❶ consists of a beautifully organized layering of neurons ❷. The outer layer, farthest from the lens, contains the rod and cone cells, which absorb the incoming light and convert it to neuronal activity. These photoreceptors connect to 10 different kinds of neurons known as bipolar cells, which send long signal-carrying arms, or axons, into a central "inner plexiform" layer. This band looks like a series of 10 distinct parallel strata. The axon of each bipolar cell type delivers signals to just a few of the strata.

At the innermost side of the plexiform layer ❸ are 12 different types of ganglion cells (*purple*). Most types send fingers called dendrites into one distinct stratum, where they receive excitatory input from a limited number of the bipolar neurons (*green*). The ganglion cells output the movie streams that the optic nerve carries to different brain regions for interpretation. Some ganglion dendrites branch out widely, carrying diffuse information, whereas others branch more narrowly, carrying high-resolution information. Some respond to an increasing change in the rate at which bipolar cells release neurotransmitters (messenger molecules), some to a decreasing change in that rate.

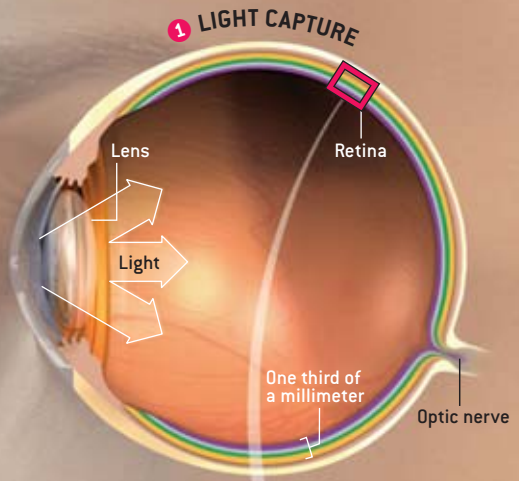
The inputs sent by the bipolar cells to the ganglion output cells within each of the strata are not enough to create the dozen movie representations, however. The signals emitted by bipolar cells are modulated by a variety of small neurons called amacrine cells (*gray*). Some of these cells operate laterally within a stratum, inhibiting communication between distant ganglion cells in that stratum. Other amacrine neurons inhibit signals vertically between strata—and therefore between different movies—as if to in-

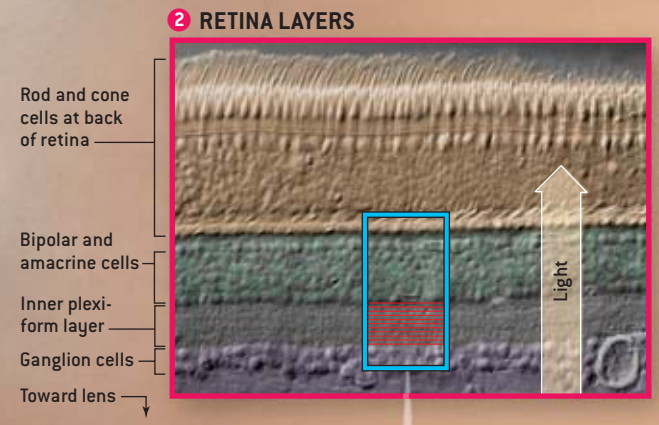
JEAN FRANCOIS PODEVIN (*preceding pages*)

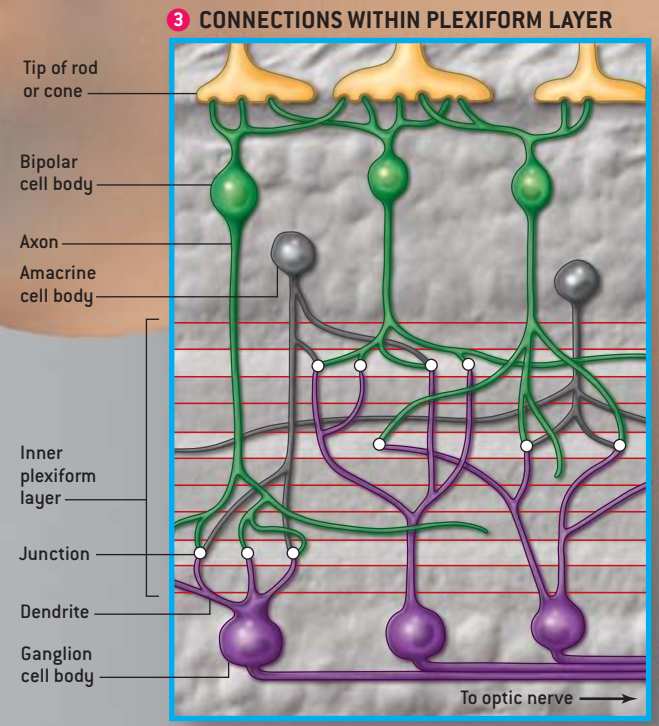
struct one stratum not to record what another stratum is recording. In this way, the amacrine cells pick up and emit signals to coordinate the movie tracks. Researchers such as Heinz Wassle of the Max Planck Institute for Brain Research in Frankfurt, Thomas Euler of the Max Planck Institute for Medical Research in Heidelberg and Richard Masland of Massachusetts General Hospital have identified at least 27 different amacrine cell types (as well as the 10 bipolar types and 12 ganglion types).

Everything we see in space is observed as time advances. Even the recording of a motionless black dot fixed in colorless three-dimensional space constitutes a movie, because the retina sees it continuously as time advances. Many cells of each ganglion type populate the retina, and the set of each type conveys a distinct movie. But unlike a box office film, which is generated frame by frame, the ganglion movies are continuous streams of signals.

The interactions among bipolar and amacrine cells that are read out simultaneously by each set of ganglion cells make up the data we receive to interpret the visual world. As we read, grasp objects, recognize faces and walk about, various combinations of these movies are the only visual clues the brain receives. They form a fundamental "visual language," with its own phrasing and grammar that embodies the neural vocabulary of vision.







DON FOLEY (*head*); STEPHEN C. MASSEY *University of Texas Medical School at Houston* (*micrograph, colorized*); JEN CHRISTIANSEN (*plexiform diagram*)

THE AUTHORS

FRANK WERBLIN and *BOTOND ROSKA* together uncovered much about the retina's functional circuitry in the early 1990s at the University of California, Berkeley. Werblin continues there as a professor of neuroscience. In 1973 he published an article in *Scientific American* after discovering unique physiological characteristics of retinal neurons with John Dowling of Johns Hopkins University. Roska is a group leader at the Friedrich Miescher Institute for Biomedical Research in Basel, Switzerland, where he is developing genetic techniques for identifying visual pathways.

Movies in a Flash

Our descriptions of the retina's complex activity are based on our own experiments. We record what is happening in individual ganglion cells with a tiny, hollow glass needle. This micropipette injects yellow dye that rapidly spreads through all the dendrites of a single ganglion cell, showing us the strata they reach. The pipette also functions as an electrode, measuring the electrical activity of the cell, which reflects the combination of excitatory signals from bipolar cells and inhibitory signals from amacrine cells.

To gain a feel for the movies that ganglion cells stream to the optic nerve, we started fairly simply: by first recording how a linear array of ganglion cells represented a square flash of light shone directly onto the retina of a rabbit ❶. The flash lasted one second and was confined to a square measuring 600 microns on each side. Thus, the flash fell on a small, well-defined region of the retina for a specific length of time.

We recorded the excitation and inhibition signals received by one type of ganglion cell over this period, repeating the procedure for each of the dozen cell types. Each type had a unique response, and the range of responses was remarkably diverse. In the plot below ❷, one box represents one second, and color indicates the magnitude of the signal current in one cell type.

Interestingly, for the ganglion cell type illustrated here, cells across the width of the flash responded, but they were not active for the whole time the light was shining. And oddly, some of them outside the 600-micron span became active after the flash had ended—behavior that appears on the plot as two lobes (*blue*) that arise after the one-second interval. A third area, within the flash region, also activates slightly, near the two-second mark.

How are we to interpret this pattern? If all the cells were sending outputs for the full second, the pattern would be "lit" across the entire span for the entire second, filling the corresponding square on our grid ❸. In reality, the output is filtered; it is as wide as the flash but is truncated in time, lasting perhaps one tenth of a second and starting about one tenth of a second after the flash began. Not only was there a slight delay before the ganglion cells responded, but they apparently responded only long enough to note how incoming light had changed—from dark to bright. Perhaps this ganglion type represents the onset of illumination but not its sustained presence. The slight activation of the cells represented in the two outlying lobes might convey some kind of "off" signals. The third blue spot at two seconds is a signal component we do not yet understand.

Each of the dozen different sets of ganglion cells creates a unique readout

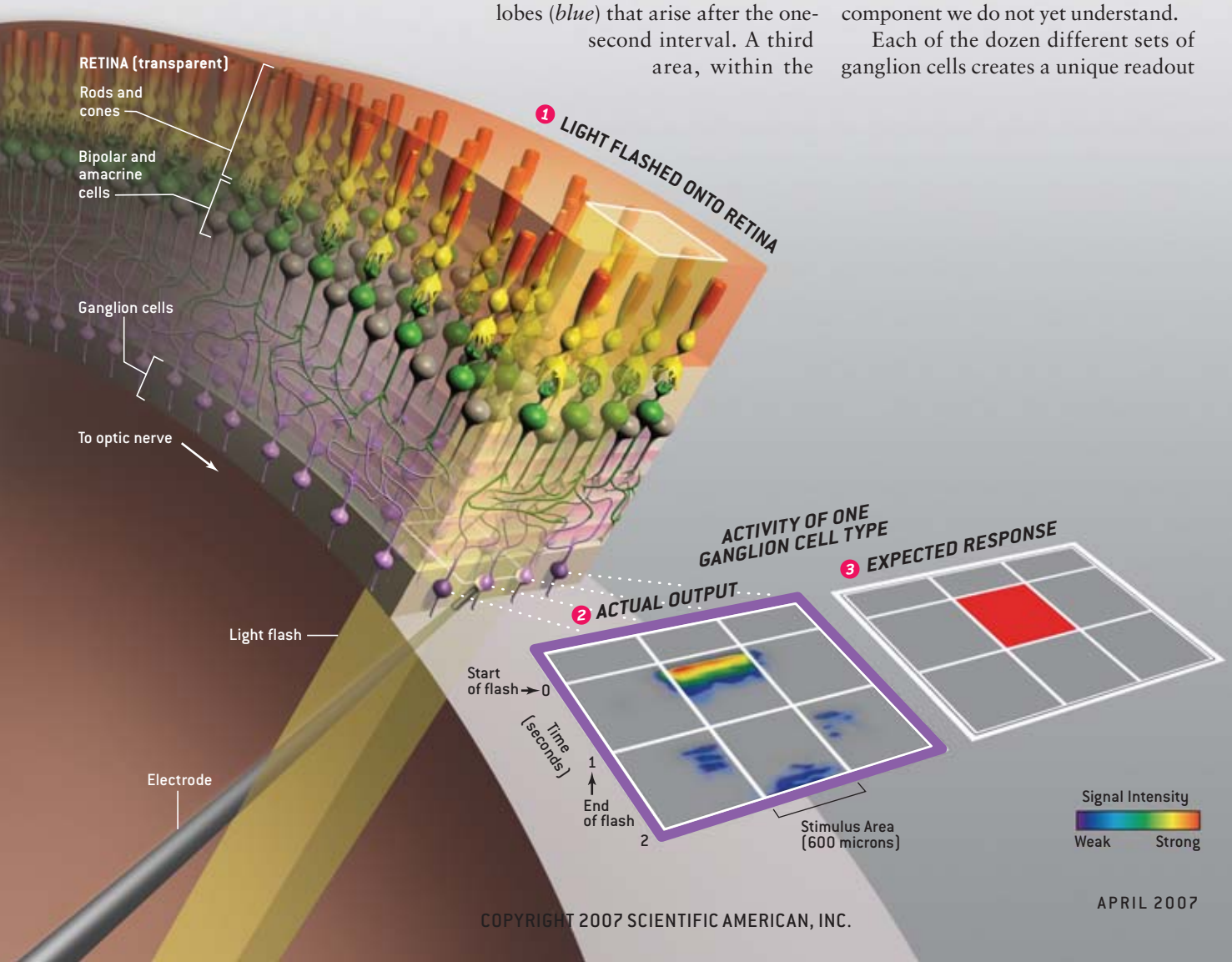
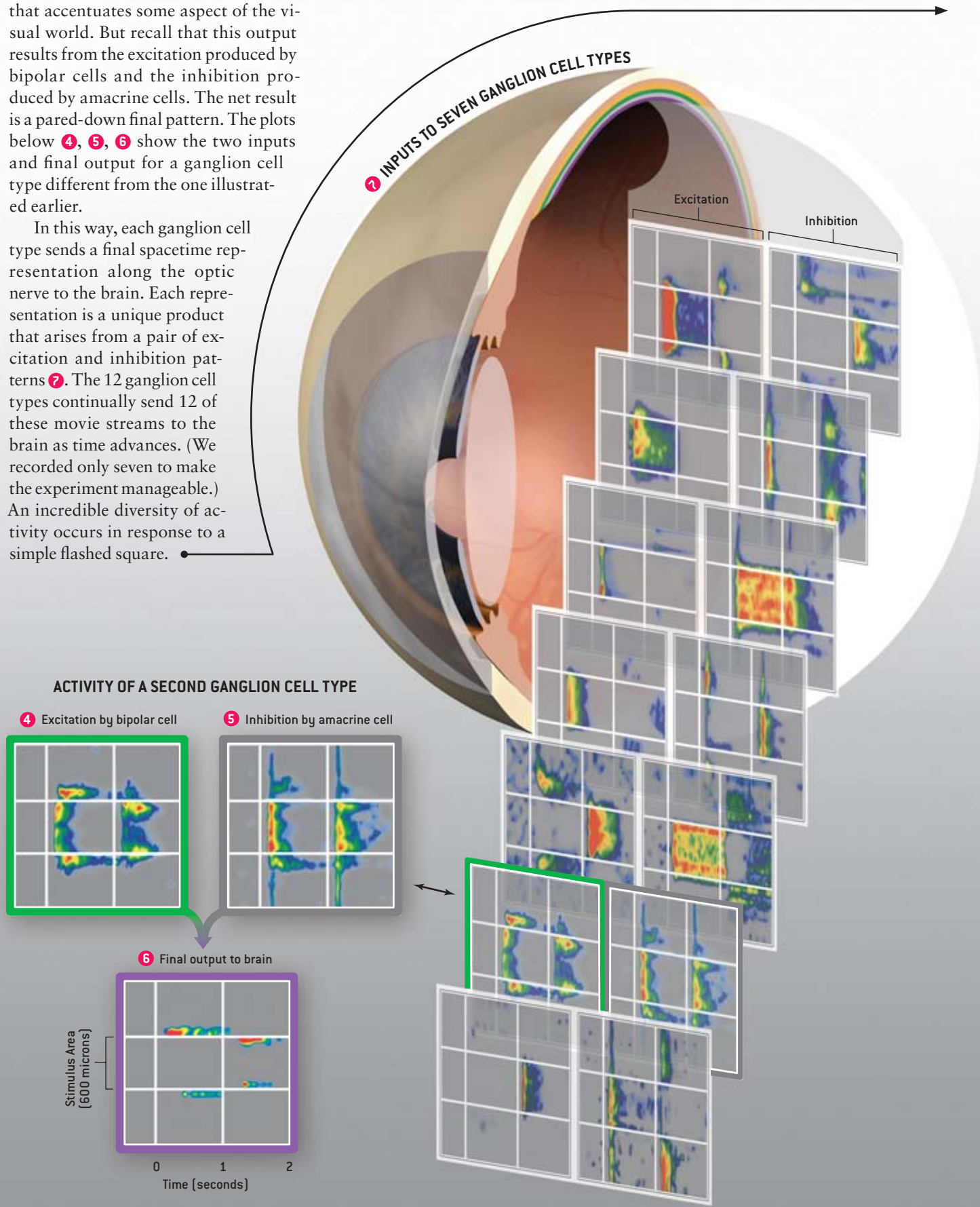


ILLUSTRATION BY DON FOLEY; IMAGE DATA FROM FRANK WERBLIN

that accentuates some aspect of the visual world. But recall that this output results from the excitation produced by bipolar cells and the inhibition produced by amacrine cells. The net result is a pared-down final pattern. The plots below 4, 5, 6 show the two inputs and final output for a ganglion cell type different from the one illustrated earlier.

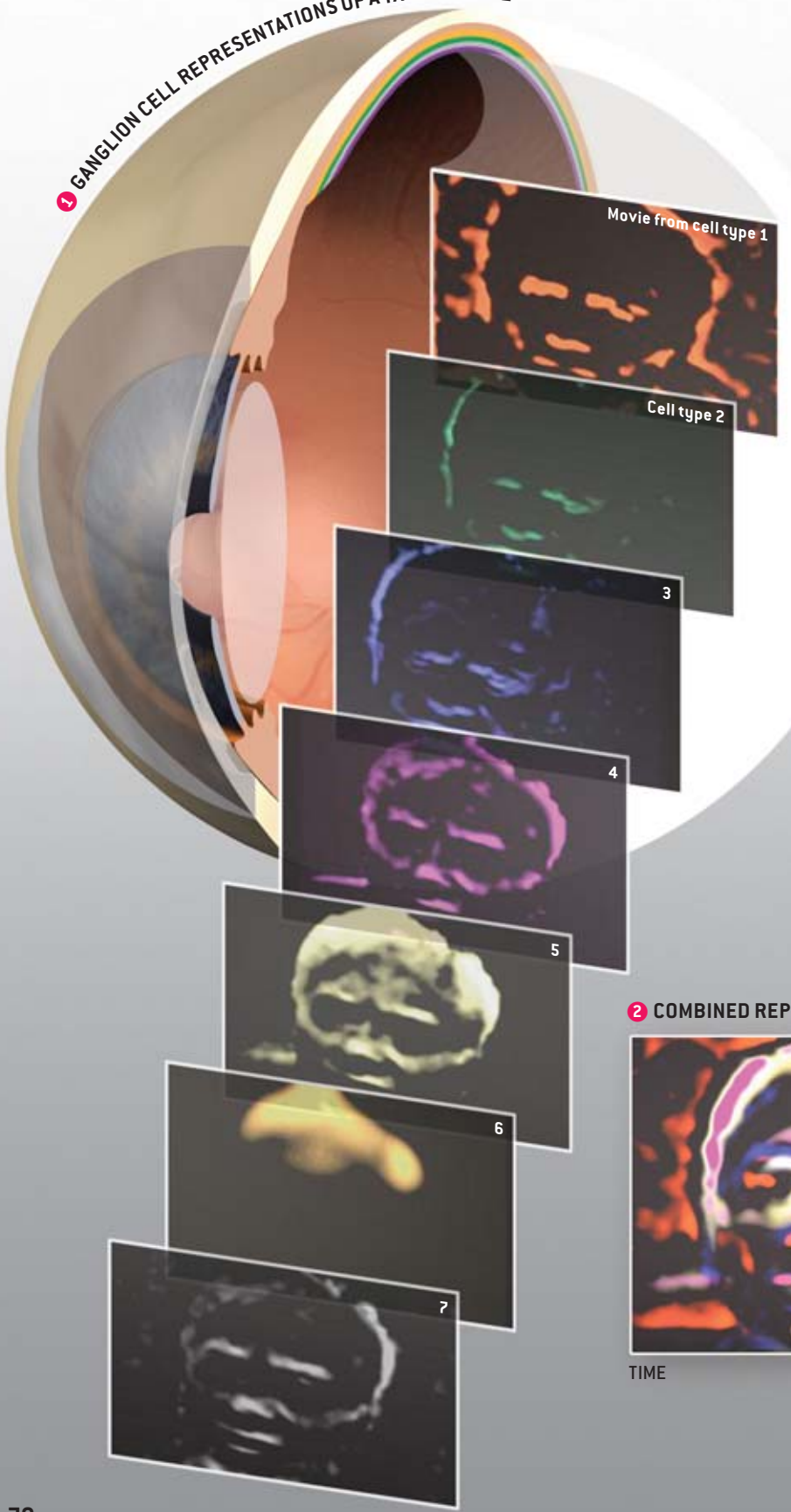
In this way, each ganglion cell type sends a final spacetime representation along the optic nerve to the brain. Each representation is a unique product that arises from a pair of excitation and inhibition patterns 7. The 12 ganglion cell types continually send 12 of these movie streams to the brain as time advances. (We recorded only seven to make the experiment manageable.) An incredible diversity of activity occurs in response to a simple flashed square.



Face Filtered

Our goal, of course, is to learn how each set of ganglion cells extracts meaning from the visual world. Because the retina is designed to handle information more interesting than a flash of light, we wondered what would happen when the retina witnessed a natural scene, such as a person talking. What would each of the 12 representations show? Would some feature be extracted by one movie but ignored by others?

Despite the seemingly straightforward explanations of how we captured the processing of a square of light, it is incredibly difficult to actually tap a living rabbit retina with enough electrodes during a simple one-second flash, much less a natural scene lasting a minute. For the latter exercise, we programmed the data from the flash experiment into a computer that simulates a famous artificial retina chip—the Cellular Neural Network—developed by Leon Chua of the University of California, Berkeley, and Tamás Roska of the Hungarian Academy of Sciences



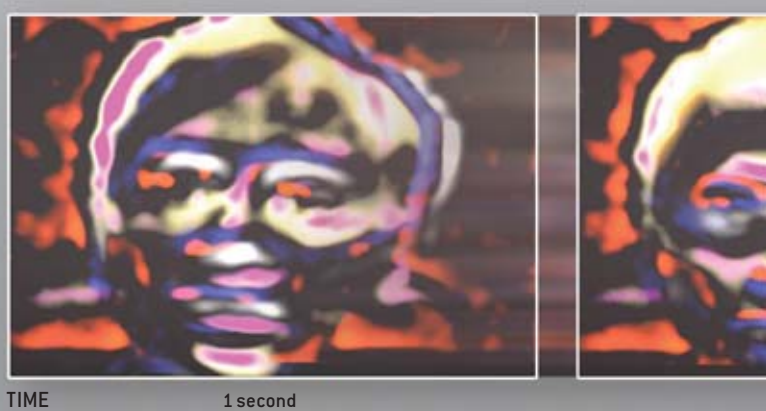


ILLUSTRATION BY DON FOLEY; IMAGE DATA FROM FRANK WERBLIN

in Budapest. [*Tamás is the father of author Botond Roska.*] The system transformed the flashed square into a dozen spacetime patterns of excitation and inhibition that very closely resembled the patterns generated by the living retina.

Encouraged, we presented the programmed retina chip with a natural scene: one of us (Werblin) sat in front of a camera and talked for just over one minute. The simulator, which was programmed for this exercise by David Balya of the Budapest University of Technology and Economics, generated movie data for seven of the different ganglion cell representations 1.

To confirm that the chip simulation was accurate, we measured the reactions of several neurons in the living rabbit retina to the talking face. It soon became evident that each population of ganglion cells acts as a filter, extracting a unique spacetime representation of the world that is sent in a unique movie to the brain. We imposed a color on each of the computer-generated representations to distinguish one from another.

For example, one filter (*opposite page, orange*) seemed to extract only the edges of features on the moving face, showing the world essentially in line-drawing form. Another filter (*purple*) accentuated the shadows underneath the eyes and nose. A third filter (*beige*) produced highlights rather than shadows or edges.

Of course, our conclusions about the information that each of the 12 filters gleaned might not be correct. Unfortunately, it is impossible to represent accurately the patterns we recorded on the printed page because they run continually as movies, but it should be noted that they contain many blank intervals. Each movie only bursts into activity for a few milliseconds at a time and is otherwise dark. Nevertheless, our method shows that each filter is sensitive to a particular quality of the face's physical appearance and movement; each type of ganglion cell has its own unique way of depicting the world.

Coloring the representations also allows us to track the contributions of each set of ganglion cells to the final, combined representation produced when the movies are superimposed. We combined the seven streams into one master movie. Four frames from different instances during Werblin's one-minute oration 2 give a sense of how his face shifts to and fro as his lips open and close, with certain representations surging and waning, making him look like a ghostly apparition. This is what the retina produces. This is what the brain receives.

Our movies are only approximations. Still, they make it clear that, remarkably, the paper-thin neural tissue at the back of the eye is already parsing the visual world into a dozen discrete components. These components travel, intact and separately, to distinct visual brain regions—some conscious, some not. The challenge to neuroscience now is to understand how the brain interprets these packets of information to generate a magnificent, seamless view of reality. SA

MORE TO EXPLORE

Directional Selectivity Is Formed at Multiple Levels by Laterally Offset Inhibition in the Rabbit Retina. Shelley I. Fried, Thomas A. Münch and Frank S. Werblin in *Neuron,* Vol. 46, No. 1, pages 117–127; 2005.

Parallel Processing in Retinal Ganglion Cells: How Integration of Space-time Patterns of Excitation and Inhibition Form the Spiking Output. Botond Roska, Alyosha Molnar and Frank S. Werblin in *Journal of Neurophysiology,* Vol. 95, pages 3810–3822; 2006.

Movie created by the retina of the talking face can be seen at ***www.sciam.com/ontheweb***

2 seconds 3 seconds 4 seconds

GASSING UP *with* HYD

ROGEN

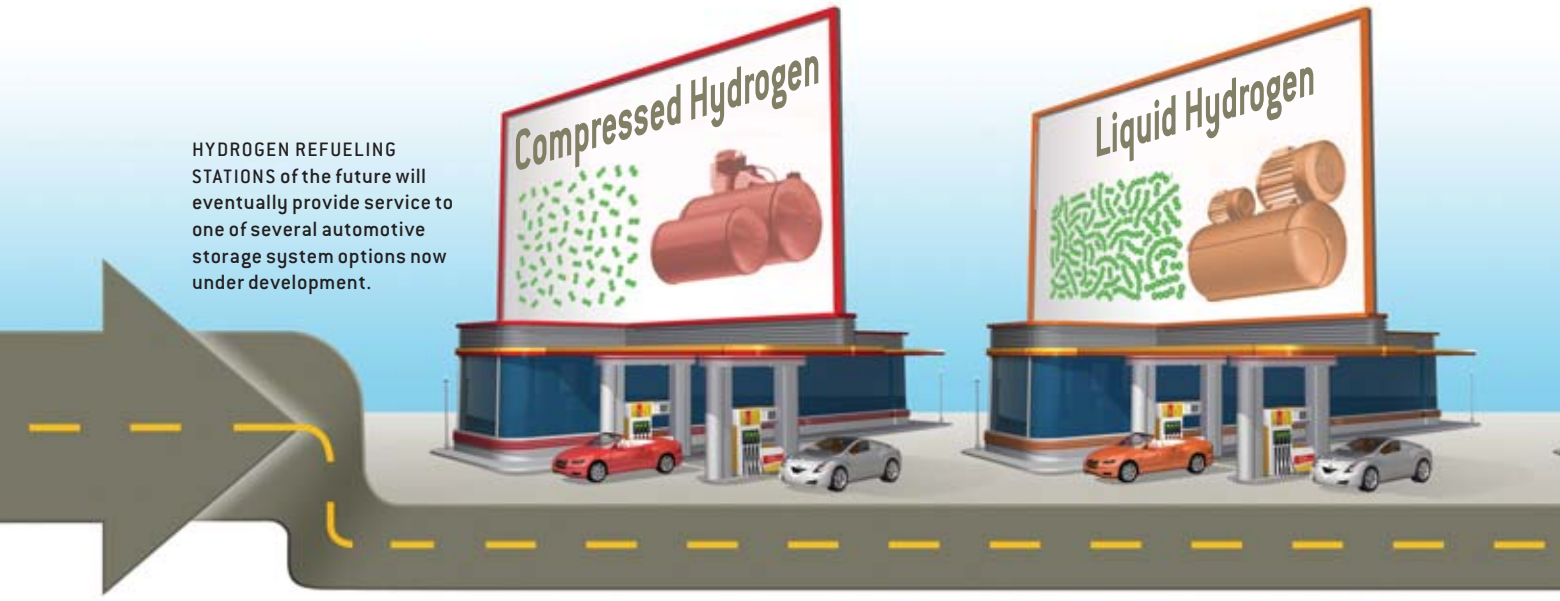
Researchers are working on ways for fuel-cell vehicles to hold the hydrogen gas they need for long-distance travel

By Sunita Satyapal, John Petrovic and George Thomas

On a late summer day in Paris in 1783, Jacques Charles did something astonishing. He soared 3,000 feet above the ground in a balloon of rubber-coated silk bags filled with lighter-than-air hydrogen gas. Terrified peasants destroyed the balloon soon after it returned to earth, but Charles had launched a quest that researchers two centuries later are still pursuing: to harness the power of hydrogen, the lightest element in the universe, for transportation.

Burned or used in fuel cells, hydrogen is an appealing option for powering future automotive vehicles for several reasons. Domestic industries can make it from a range of chemical feedstocks and energy sources (for instance, from renewable, nuclear and fossil-fuel sources), and the nontoxic gas could serve as a virtually pollution-free energy carrier for machines of many kinds. When it burns, it releases no carbon dioxide, a potent greenhouse gas. And if hydrogen is fed into a fuel-cell stack—a batterylike device that generates electricity from hydrogen and oxygen—it can propel an electric car or truck with only water and heat as by-products [see "On the Road to Fuel-Cell Cars," by Steven Ashley; Scientific American, March 2005]. Fuel-cell-powered vehicles could offer more than twice the efficiency of today's autos. Hydrogen could therefore help ease pressing environmental and societal problems, including air pollution and its health hazards, global climate change and dependence on foreign oil imports.

HYDROGEN GAS resists being crammed into the volume of a standard automotive gasoline tank.



HYDROGEN REFUELING STATIONS of the future will eventually provide service to one of several automotive storage system options now under development.

Yet barriers to gassing up cars with hydrogen are significant. Kilogram for kilogram, hydrogen contains three times the energy of gasoline, but today it is impossible to store hydrogen gas as compactly and simply as the conventional liquid fuel. One of the most challenging technical issues is how to efficiently and safely store enough hydrogen onboard to provide the driving range and performance that motorists demand. Researchers must find the "Goldilocks" storage solutions that are "just right." Storage devices should hold sufficient hydrogen to support today's minimum acceptable travel range—300 miles—on a tank of fuel in a volume of space that does not compromise passenger or luggage room. They should release it at the required flow rates for acceleration on the highway and operate at practical temperatures. They should be refilled or recharged in a few minutes and come with a competitive price tag. Current hydrogen storage technologies fall far short of these goals.

Researchers worldwide in the auto industry, government and academia are expending considerable effort to overcome these limitations. The International Energy Agency's Hydrogen Implementing Agreement, signed in 1977, is now the largest international group focusing on hydrogen storage, with more than 35 researchers from 13 countries. The International Partnership for the Hydrogen Economy, formed in 2003, now includes 17 governments committed to advancing hydrogen and fuel-cell technologies. And in 2005 the U.S. Department of Energy set up a National Hydrogen Storage Project with three Centers of Excellence and many industry, university and federal laboratory efforts in both basic and applied research. Last year alone this project provided more than $30 million to fund about 80 research projects.

Overview/*Hydrogen Storage*

- One of the biggest obstacles to future fuel-cell vehicles is how engineers will manage to stuff enough hydrogen onboard to provide the 300-mile minimum driving range that motorists demand.
- Typically hydrogen is stored in pressurized tanks as a highly compressed gas at ambient temperature, but the tanks do not hold enough gas. Liquid-hydrogen systems, which operate at cryogenic temperatures, also suffer from significant drawbacks.
- Several alternative high-density storage technologies are under development, but none is yet up to the challenge.

Infrastructural Hurdles

ONE OBSTACLE to the wide adoption of hydrogen fuel-cell cars and trucks is the sheer size of the problem. U.S. vehicles alone consume 383 million gallons of gasoline a day (about 140 billion gallons annually), which accounts for about two thirds of the total national oil consumption. More than half of that petroleum comes from overseas. Clearly, the nation would need to invest considerable capital to convert today's domestic auto industry to fuel-cell vehicle production and the nation's extensive gasoline refining and distribution network to one that handles vast quantities of hydrogen. The fuel-cell vehicles themselves would have to become cheap and durable enough to compete with current technology while offering equivalent performance. They also must address safety concerns and a lingering negative public perception—people still remember the 1937 *Hindenburg* airship tragedy and associate it with hydrogen, despite some credible evidence that the airship's flammable skin was the crucial factor in the ignition of the blaze.

Why is it so difficult to store enough hydrogen onboard a vehicle? At room temperature and atmospheric pressure (one atmosphere is about 14.5 pounds per square inch, or psi), hydrogen exists as a gas with an energy density about $^{1}/_{3,000}$ that of liquid gasoline. A 20-gallon tank containing hydrogen gas at atmospheric pressure would propel a standard car only about 500 feet. So engineers must increase the density of stored hydrogen in any useful onboard hydrogen containment system.

A 300-mile minimum driving range is one of the principal operational aims of an industry-government effort—the FreedomCAR and Fuel Partnership—to develop advanced technology for future automobiles. Engineers employ a useful rule of thumb in making such calculations: a gallon of gasoline is equal, on an energy basis, to one kilogram (2.2 pounds) of

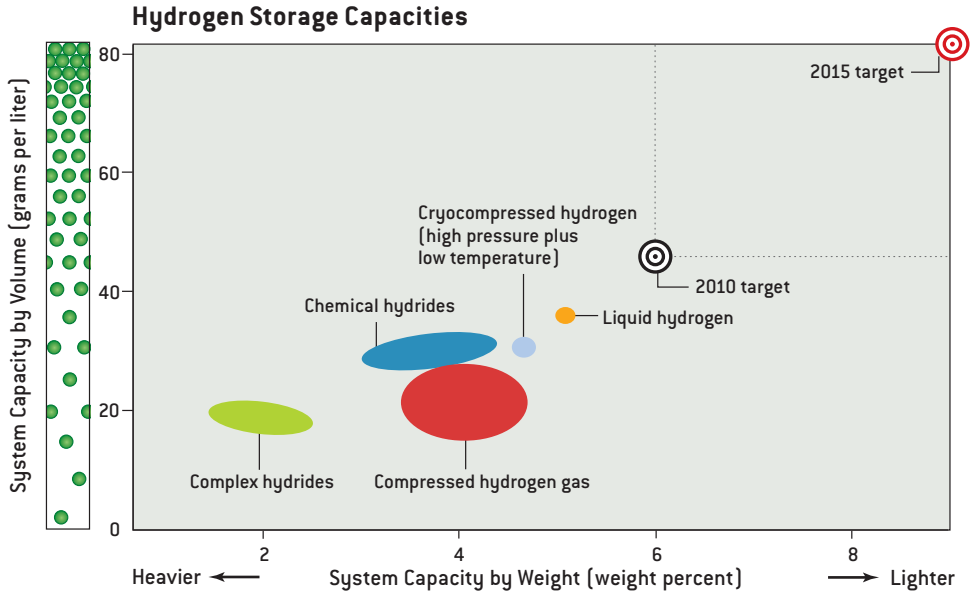
KENN BROWN (*preceding pages and these two pages*)

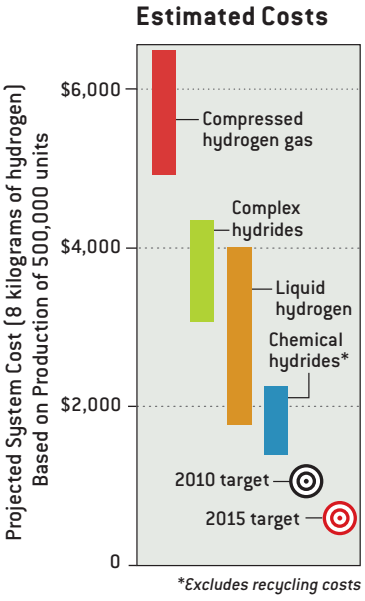
hydrogen. Whereas today's average automobile needs about 20 gallons of gasoline to travel at least 300 miles, the typical fuel-cell vehicle would need only about eight kilograms of hydrogen because of its greater operational efficiency. Depending on the vehicle type and size, some models would require less hydrogen to go that far, some more. Tests of about 60 hydrogen-fueled prototypes from several automakers have so far demonstrated driving ranges of 100 to 190 miles.

Aiming for a practical goal that could be achievable by 2010 (when some companies expect the first production fuel-cell cars to hit the road), researchers compare the performance of various storage technologies against the "6 weight percent" benchmark. That is, a fuel storage system in which 6 percent of its total weight is hydrogen. For a system weighing a total of 100 kilograms (a reasonable size for a vehicle), six kilograms would be stored hydrogen. Although 6 percent may not seem like much, achieving that level will be extremely tough; less than 2 percent is the best possible today—using storage materials that operate at relatively low pressures. Further, keeping the system's total volume to about that of a standard automotive gas-

The Storage Challenge

A hydrogen storage system must carry enough fuel for at least a 300-mile trip and also be compact and light enough to haul around in a car. A system that stored 6 percent by weight as hydrogen and 45 grams of hydrogen per liter by 2010 would likely meet the needs of first-generation fuel-cell vehicles (*black target on first graph*), but none of the current options can yet do this. Even better performance will be needed by 2015 for the wider range of vehicle types that will be available. Note that the values below include the equipment needed to run each system. Liquid hydrogen alone, for example, has a density of 71 grams per liter, but when the tank and ancillary components are included, the volume capacity falls to just under 40 grams per liter. Because hydrogen adsorbents [*see box on page 87*] are still at an early stage of development, capacity and cost data are not yet available.



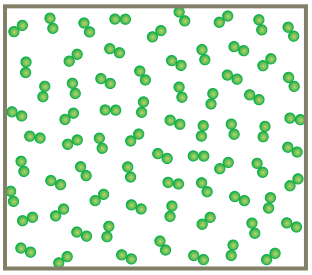


JEN CHRISTIANSEN

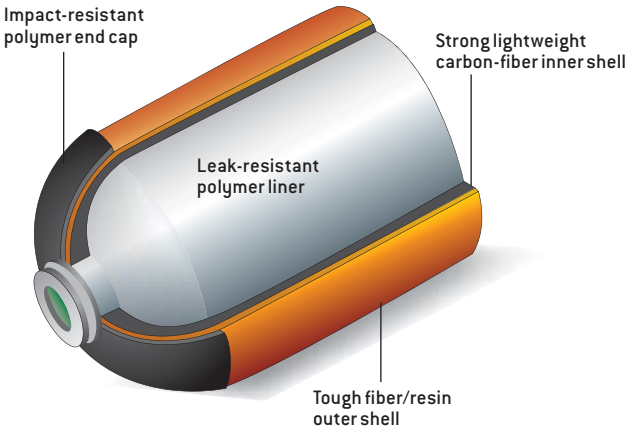
COMPRESSED HYDROGEN

Lightweight but strong, high-pressure cylinders that resemble diving tanks store the compressed gas at 5,000 and 10,000 pounds per square inch.

STORAGE DENSITY



Hydrogen



Pros	Cons	Status
Low weight	High volume; requires high-pressure compression and refueling	Available

oline tank will be even more difficult, given that much of its allotted space will be taken up by the tanks, valves, tubing, regulators, sensors, insulation and anything else that is required to hold the six kilograms of hydrogen. Finally, a useful system must release hydrogen at rates fast enough for the fuel-cell and electric motor combination to provide the power and acceleration that drivers expect.

Containing Hydrogen

AT PRESENT, most of the several hundred prototype fuel-cell vehicles store hydrogen gas in high-pressure cylinders, similar to scuba tanks. Advanced filament-wound, carbon-fiber composite technology has yielded strong, lightweight tanks that can safely contain hydrogen at pressures of 5,000 psi (350 times atmospheric pressure) to 10,000 psi (700 times atmospheric pressure) [*see box above*]. Simply raising the pressure does not proportionally increase the hydrogen density, however. Even at 10,000 psi, the best achievable energy density with current high-pressure tanks (39 grams per liter) is about 15 percent of the energy content of gasoline in the same given volume. Today's high-pressure tanks can contain only about 3.5 to 4.5 percent of hydrogen by weight. Ford recently introduced a prototype "crossover SUV" called Edge that is powered by a combination plug-in hybrid/fuel-cell system that stores 4.5 kilograms of hydrogen fuel in a 5,000-psi tank to achieve a total maximum range of 200 miles.

High-pressure tanks would be acceptable in certain transportation applications, such as transit buses and other large vehicles that have the physical size necessary to accommodate storage for sufficient hydrogen, but it would be difficult to manage in cars. Also, the current cost of such tanks is 10 or more times higher than what is competitive for autos.

Liquefying stored hydrogen can improve its energy density, packing the most hydrogen into a given volume of any existing option. Like any gas, hydrogen that is cooled sufficiently condenses into a liquid, which at atmospheric pressure occurs around –253 degrees Celsius. Liquid hydrogen exhibits a density of 71 grams per liter, or about 30 percent of the energy density of gasoline. The hydrogen weight densities achievable by these systems depend on the containment and insulation equipment they use [*see box on opposite page*].

Liquefied hydrogen has important drawbacks, though. First, its very low boiling point necessitates cryogenic equipment and special precautions for safe handling. In addition, because it operates at low temperature, the containers have to be insulated extremely well. Finally, liquefying hydrogen takes more energy than compressing the gas to high pressures. This requirement drives up the cost of the fuel and reduces the overall energy efficiency of the cryocooling process.

Nevertheless, one carmaker is pushing this technology onto the road. BMW plans to introduce a vehicle this year called Hydrogen 7, which will incorporate an internal-combustion engine capable of running on either gasoline (for 300 miles) or on liquid hydrogen for 125 miles. Hydrogen 7 will be sold on a limited basis to selected customers in the U.S. and other countries with local access to hydrogen refueling stations.

Chemical Compaction

SEARCHING FOR promising ways to raise energy density, scientists may be able to take advantage of the chemistry of hydrogen itself. In their pure gas and liquid phases, hydrogen molecules contain two bound atoms each. But when hydrogen atoms are chemically bound to certain other elements, they

THE AUTHORS

SUNITA SATYAPAL, JOHN PETROVIC and *GEORGE THOMAS* work in the U.S. Department of Energy's applied research and development program in hydrogen storage technology. Satyapal, who has held various positions in academia and industry, serves as the team leader for the DOE's applied hydrogen storage R&D activities. Petrovic, a laboratory fellow (retired) of Los Alamos National Laboratory, is a consultant for the DOE and a fellow of both the American Ceramic Society and the American Society for Materials International. Thomas, currently a consultant with the DOE, has more than 30 years of experience studying the effects of hydrogen on metals at Sandia National Laboratories. These views are those of the authors only; they do not reflect the positions of the U.S. Department of Energy.

MELISSA THOMAS; QUANTUM TECHNOLOGIES (*cylinder*)

can be packed even closer together than in liquid hydrogen. The principal aim of hydrogen storage research now is finding the materials that can pull off this trick.

Some researchers are focusing on a class of substances called reversible metal hydrides, which were discovered by accident in 1969 at the Philips Eindhoven Labs in the Netherlands. Investigators found that a samarium-cobalt alloy exposed to pressurized hydrogen gas would absorb hydrogen, somewhat like a sponge soaks up water. When the pressure was then removed, the hydrogen within the alloy reemerged; in other words, the process was reversible.

Intensive research followed this discovery. In the U.S., scientists James Reilly of Brookhaven National Laboratory and Gary Sandrock of Inco Research and Development Center in Suffern, N.Y., pioneered the development of hydride alloys with finely tuned hydrogen absorption properties. This early work formed the basis for today's widely used nickel–metal hydride batteries. The density of hydrogen in these alloys can be very high: 150 percent more than liquid hydrogen, because the hydrogen atoms are constrained between the metal atoms in their crystal lattices [*see top box on next page*].

Many properties of metal hydrides are well suited to automobiles. Densities surpassing that of liquid hydrogen can be achieved at relatively low pressures, in the range of 10 to 100 times atmospheric pressure. Metal hydrides are also inherently stable, so they require no extra energy to maintain storage, although heat is required to release the stored gas. But their Achilles' heel is mass. They weigh too much for practical onboard storage. Metal hydride researchers have so far attained a maximum hydrogen capacity of 2 percent of the total material weight (2 weight percent). This level translates into a 1,000-pound hydrogen storage system (for a 300-mile driving range), which is clearly too heavy for today's 3,000-pound car.

Metal hydride studies currently concentrate on materials with inherently high hydrogen content, which researchers then modify to meet the hydrogen storage system requirements of operating temperatures in the neighborhood of 100 degrees C, pressures from 10 to 100 atmospheres and delivery rates sufficient to support rapid vehicle acceleration. In many cases, materials that contain useful proportions of hydrogen are a bit too stable in that they require substantially higher temperatures to release the hydrogen. Magnesium, for example, forms magnesium hydride with 7.6 weight percent hydrogen but must be heated to above 300 degrees C for release to occur. If a practical system is to rely on waste heat from a fuel-cell stack (about 80 degrees C) to serve as the "switch" to liberate hydrogen from a metal hydride, then the trigger temperature must be lower.

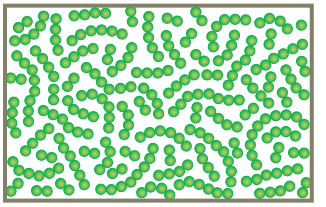
Destabilized Hydrides

CHEMISTS John J. Vajo and Gregory L. Olson of HRL Laboratories in Malibu, Calif., as well as researchers elsewhere are exploring a clever approach to overcoming the temperature problem. Their "destabilized hydrides" combine several substances to alter the reaction pathway so that the resulting compounds release the gas at lower temperatures.

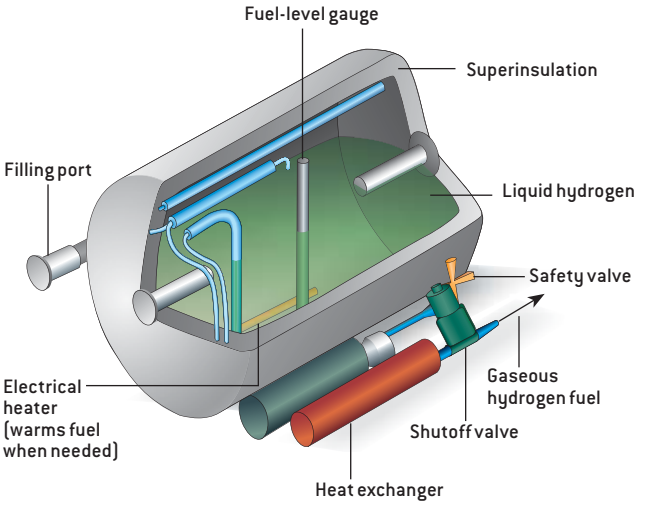
LIQUID HYDROGEN

Cooled to –253 degrees Celsius, hydrogen condenses and liquefies. Considerable amounts of insulation and ancillary equipment are required to maintain this low temperature.

STORAGE DENSITY



Hydrogen



Pros	Cons	Status
Low weight and volume	Continual loss because of boil-off; energy penalty to liquefy hydrogen	Available

Destabilized hydrides are part of a class of hydrogen-containing materials called complex hydrides. Chemists long thought that many of these compounds were not optimal for refueling a vehicle, because they were irreversible—once the hydrogen was freed by decomposition of the compounds, the materials would require reprocessing to return them to a hydrogenated state. Chemists Borislav Bogdanovic and Manfred Schwickardi of the Max Planck Institute of Coal Research in Mülheim, Germany, however, stunned the hydride research community in 1996 when they demonstrated that the complex hydride sodium alanate becomes reversible when a small amount of titanium is added. This work triggered a flurry of activity during the past decade. HRL's lithium borohydride destabilized with magnesium hydride, for example, holds around 9 percent of hydrogen by weight reversibly and features a 200 degree C operating temperature. This improvement is notable, but its operating temperature is still too high and its hydrogen release rate too slow for automotive applications. Nevertheless, the work is promising.

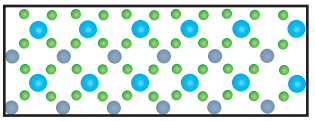
Although current metal hydrides have limitations, many automakers see them as the most viable low-pressure approach in the near- to mid-term. Toyota and Honda engineers, for

MELISSA THOMAS; LINDE (*container*)

COMPLEX HYDRIDES

Combining hydrogen chemically with metals and other substances produces metal hydrides (*below*) and complex hydrides (*right*). Heat liberates the fuel from these compounds when needed.

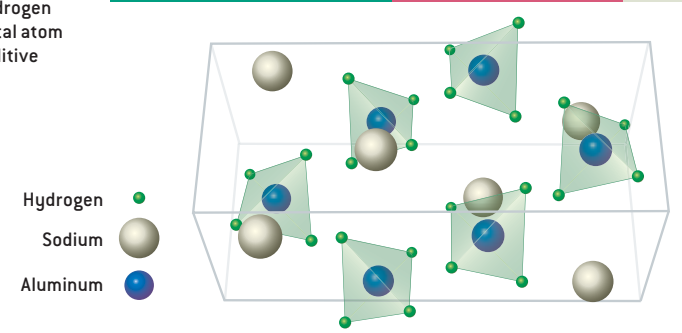
STORAGE DENSITY



COMPLEX HYDRIDE

Hydrogen
Metal atom
Additive

Pros	Cons	Status
Low volume; potential for recharging onboard; potential for moderate to low-pressure operation	High weight; high temperatures required; insufficient fuel flow	In development



SODIUM ALUMINUM HYDRIDE (example of complex hydride)

example, are planning a so-called hybrid approach in a system that combines a solid metal hydride with moderate pressure (significantly lower than 10,000 psi), which they predict could achieve a driving range of more than 300 miles. General Motors has teams of storage experts, including Scott Jorgensen, who are supporting research on a wide range of metal hydride systems worldwide (including in Russia, Canada and Singapore). GM is also collaborating with Sandia National Laboratories on a four-year, $10-million effort to produce a prototype complex metal hydride system.

Hydrogen Carriers

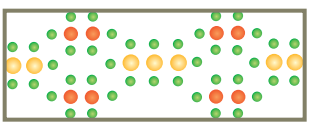
OTHER HYDROGEN STORAGE options have the potential to work well in cars, but they suffer a penalty in the refueling step. In general, these chemical hydride substances need industrial processing to reconstitute the spent material. The step requires offboard regeneration; that is, once hydrogen stored onboard a vehicle is released, a leftover by-product must be reclaimed at a service station and regenerated in a chemical plant [*see box below*].

More than 20 years ago Japanese researchers studied this approach using, for example, the decalin-naphthalene system. When decalin ($C_{10}H_{18}$) is heated, it converts chemically to naphthalene (a pungent-smelling compound with the formula $C_{10}H_8$) by changing the nature of its chemical bonds, which liberates five hydrogen molecules. Hydrogen gas thus bubbles out of the liquid decalin as it transforms into naphthalene. Exposing naphthalene to moderate hydrogen gas pressures reverses the process; it absorbs hydrogen and changes back to decalin (6.2 weight percent for the material alone). Research chemists Alan Cooper and Guido Pez of Air Products and Chemicals in Allentown, Pa., are investigating a similar technique using organic (hydrocarbon-based) liquids. Other scientists, including S. Thomas Autrey and his co-workers at the

CHEMICAL HYDRIDES

These hydrogen-containing compounds, which can be liquids or solids, release the fuel when they are heated and exposed to a catalyst (*photograph at right*). The flowchart (*far right*) shows how a representative chemical hydride must be recycled offboard to recharge it with hydrogen after use.

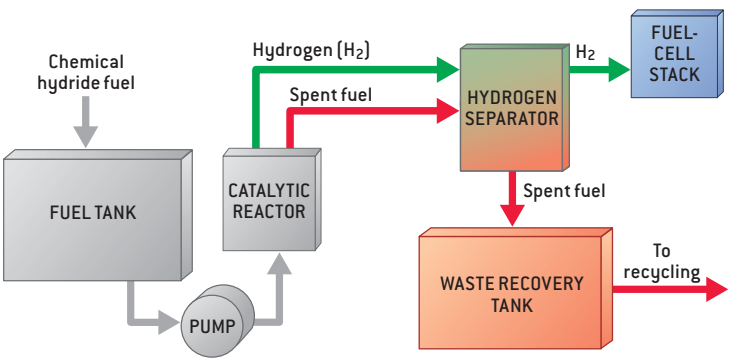
STORAGE DENSITY



Hydrogen
Other elements or chemicals

Pros	Cons	Status
Low weight and volume; potentially a liquid	Materials need to be removed from the vehicle for regeneration; cost of offboard recycling and related infrastructure	In development

N-ETHYL CARBAZOLE



MELISSA THOMAS; ENERGY CONVERSION DEVICES (*researcher with hydride*); INSTITUTE FOR ENERGY TECHNOLOGY (*complex hydride*); AIR PRODUCTS AND CHEMICALS (*test tube*)

Pacific Northwest National Laboratory and chemistry professor Larry G. Sneddon of the University of Pennsylvania, are working on new liquid carriers, such as aminoboranes, that can store large amounts of hydrogen and release it at moderate temperatures.

Designer Materials

YET ANOTHER APPROACH to the hydrogen storage problem centers on lightweight materials with very high surface areas to which hydrogen molecules stick (or adsorb) [*see box at right*]. As one might expect, the amount of hydrogen retained on any surface correlates with the material's surface area. Recent developments in nanoscale engineering have yielded a host of new high-surface-area materials, some with more than 5,000 square meters of surface area per gram of material. (This amount equates to about three acres of surface area within just a teaspoon of powder.) Carbon-based materials are particularly interesting because they are lightweight, can be low cost and can form a variety of nanosize structures: carbon nanotubes, nanohorns (hornlike tubes), fullerenes (ball-shaped molecules) and aerogels (ultraporous solids). One relatively cheap material, activated carbon, can store up to about 5 weight percent hydrogen.

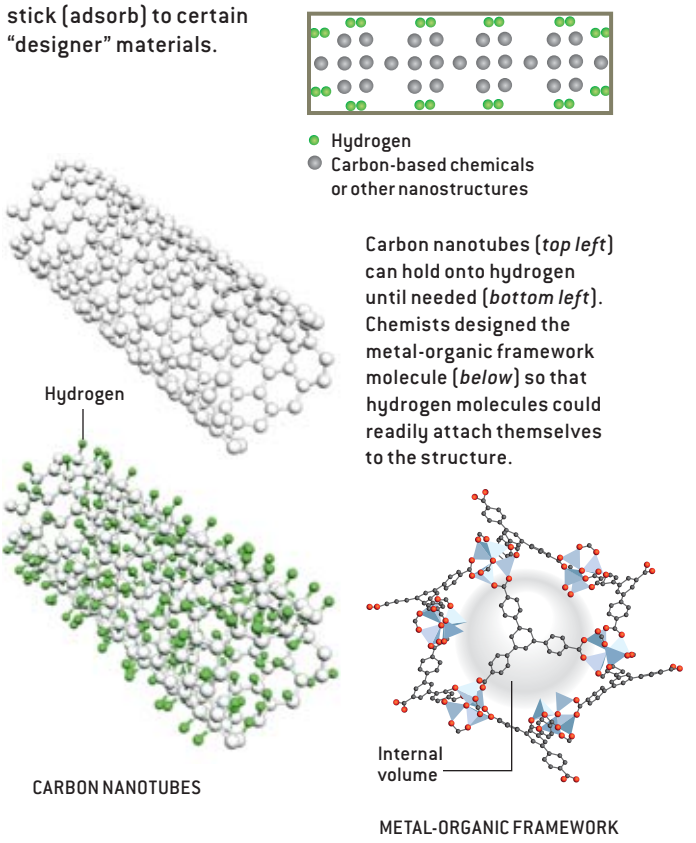
These carbon structures all share a common limitation, however. Hydrogen molecules bond very weakly with the carbon atoms, which means that the high-surface-area materials must be kept at or near the temperature of liquid nitrogen, –196 degrees C. In contrast to hydride research, in which scientists are struggling to lower the hydrogen binding energy, carbon researchers are exploring ways to raise the binding energy by modifying the surfaces of materials or by adding metal dopants that may alter their properties. These investigators employ theoretical modeling of carbon structures to discover promising systems for further study.

Beyond the carbon-based approaches, another fascinating nanoscale engineering concept is a category of substances called metal-organic materials. A few years ago Omar Yaghi, then a chemistry professor at the University of Michigan at Ann Arbor and now at the University of California, Los Angeles, invented these so-called metal-organic frameworks, or MOFs. Yaghi and his co-workers showed that this new class of highly porous, crystalline materials could be produced by linking inorganic compounds together with organic "struts" [*see box at right*]. The resulting MOFs are synthetic compounds with elegant-looking structures and physical characteristics that can be controlled to provide various desired functions. These heterogeneous structures can have very large surface areas (as high as 5,500 square meters per gram), and researchers can tailor chemical sites on them for optimal binding to hydrogen. To date, investigators have demonstrated MOFs that exhibit hydrogen capacities of 7 weight percent at –196 degrees C. They continue to work on boosting this performance.

Although current progress on hydrogen storage methods is encouraging, finding the "just right" approach may take time, requiring sustained, innovative research and development efforts. Over the centuries, the basic promise—and challenge—of using hydrogen for transportation has remained fundamentally unchanged: Holding onto hydrogen in a practical, lightweight container allowed Jacques Charles to travel across the sky in his balloon during the last decades of the 18th century. Finding a similarly suitable container to store hydrogen in automobiles will permit people to travel across the globe in the coming decades of the 21st century without fouling the sky above. SA

HYDROGEN ADSORBENTS

Hydrogen atoms readily stick (adsorb) to certain "designer" materials.



Pros	Cons	Status
Low weight; reversible onboard; potential for room-temperature operation	High volume; low temperatures for operation may be required	Early R&D stage

MORE TO EXPLORE

The Hydrogen Economy: Opportunities, Costs, Barriers, and R&D Needs. National Research Council and National Academy of Engineering. National Academies Press, 2004. Available at **www.nap.edu/catalog.php?record_id=10922**

Hydrogen Program: 2006 Annual Merit Review Proceedings. U.S. Department of Energy. Available at **www.hydrogen.energy.gov/annual_review06_proceedings.html**

United States Council for Automotive Research: **www.uscar.org**

International Energy Agency's Hydrogen Implementing Agreement: **www.ieahia.org**

International Partnership for the Hydrogen Economy: **www.iphe.net**

MELISSA THOMAS; ANTON NIKITIN *Stanford Synchrotron Radiation Laboratory (nanotubes)*; OMAR M. YAGHI *U.C.L.A. (metal-organic framework)*

JOE McDONALD *Corbis*

A CURE for RABIES?

By Rodney E. Willoughby, Jr.

The survival of a Wisconsin teenager who contracted rabies may point the way to a treatment for this horrifying disease

Rabies is one of the oldest and most feared diseases. It attacks the brain, causing agitation, terror and convulsions. Victims suffer painful throat spasms when they try to drink or eat. Paralysis follows, yet people infected with rabies are intermittently alert until near death and can communicate their fear and suffering to family and caregivers. Although vaccines against the rabies virus can prevent the illness from developing, until recently doctors could hold out no hope for patients who failed to get immunized soon after being bitten by a rabid animal. Once the symptoms of rabies appeared—typically within two months of the bite—death was inevitable, usually in less than a week.

In 2004, however, I was a member of a team of physicians at Children's Hospital of Wisconsin in Milwaukee that rescued a 15-year-old girl from a similar fate. Jeanna Giese of Fond du Lac, Wis., became the first known unimmunized survivor of rabies. (Five other people who had been immunized but developed rabies anyway have also survived.) Our novel treatment, dubbed the Milwaukee protocol, has stirred controversy among medical specialists; some claim that Jeanna's cure was a fluke. Although the few attempts to replicate the treatment have not saved the lives of any other rabies patients, I fervently hope that we are on the right track. At the very least, researchers should initiate animal studies to determine which of the elements in our protocol can help defeat rabies.

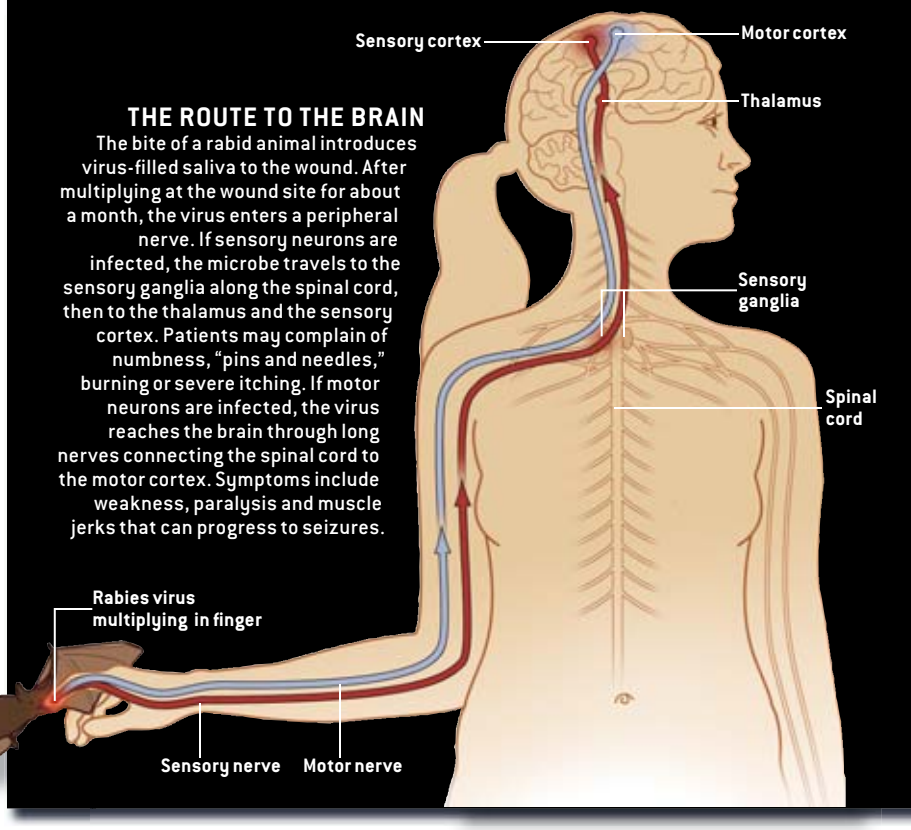
A cure for rabies would be a great boon for the developing world. The disease is quite rare in the U.S. and Europe because public health campaigns have almost eliminated the virus in domestic animals such as dogs, cats and cattle. Only two to three patients die of rabies in the U.S. every year, meaning that the chance that a person contracts the illness is only about one in 100 million. (In fact, the disease is so rare in the U.S. that it goes unrecognized in half its victims until after they die.) But the World Health Organization estimates that rabies kills 55,000 people annually in Asia, Africa and Latin America, with most of the victims infected by dog bites. If investigators can properly analyze the Milwaukee protocol and identify an inexpensive treatment that would provide similar benefits, the resulting therapy could save thousands of lives.

A Bat Bite

RABIES is an enveloped RNA virus, meaning that it has an outer membrane and uses ribonucleic acid as its genetic material rather than the deoxyribonucleic acid (DNA) used by humans and virtually all other forms of life. The bullet-shaped microbe invades human cells and forces them to churn out new viruses, inflicting all its damage by manufacturing just five proteins. Highly specialized for growth in the brain and nerves, the rabies virus is rarely found elsewhere in the body. After being transmitted by the bite of a rabid animal, which introduces infected saliva into a wound, the virus multiplies locally in muscle or skin. Because the virus is present at very low levels and does not travel through the bloodstream or lymph nodes, the body's immune system does not detect the microbe at this stage. This symptom-free incubation period typically lasts two to eight weeks but can be as long as several years. At some point the virus reaches a nerve and enters it, and then the game is over.

THE TERRIBLE COURSE OF RABIES

Vaccines against the rabies virus can prevent the illness from developing if given soon after the patient is bitten by a rabid animal. But the disease is deadly for those who fail to get immunized.



THE ROUTE TO THE BRAIN

The bite of a rabid animal introduces virus-filled saliva to the wound. After multiplying at the wound site for about a month, the virus enters a peripheral nerve. If sensory neurons are infected, the microbe travels to the sensory ganglia along the spinal cord, then to the thalamus and the sensory cortex. Patients may complain of numbness, "pins and needles," burning or severe itching. If motor neurons are infected, the virus reaches the brain through long nerves connecting the spinal cord to the motor cortex. Symptoms include weakness, paralysis and muscle jerks that can progress to seizures.

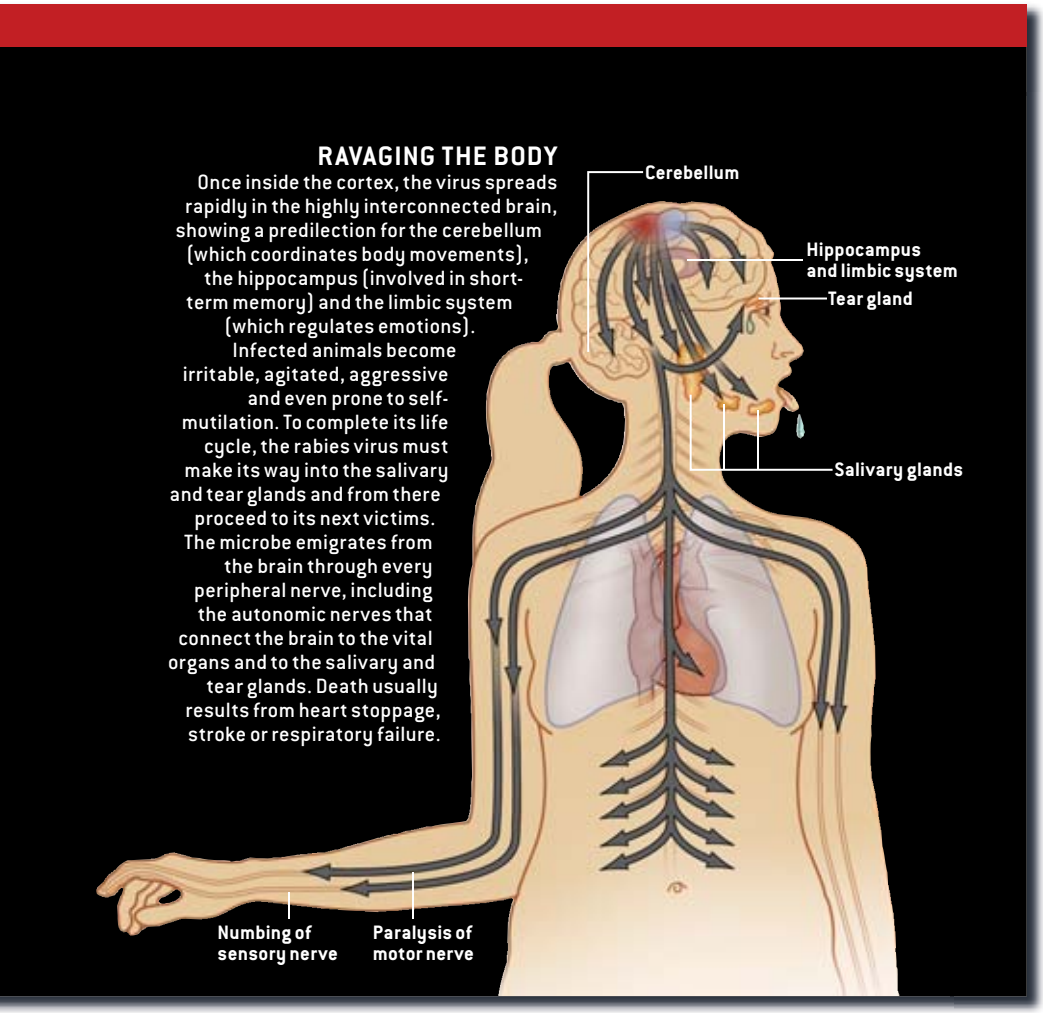
TAMI TOLPA

Overview/*A Rabies Riddle*

- In 2004 doctors saved the life of a teenager who contracted rabies from a bat bite. The treatment involved inducing coma and administering drugs that fought the rabies virus and protected the brain.
- Researchers do not know why the treatment worked, and its success has not been replicated. Scientists need to test the therapy on rabid animals, but veterinary schools are reluctant.
- An inexpensive, reliable treatment for rabies could save thousands of lives in developing countries where the disease is still common.

In the late 19th century French microbiologist Louis Pasteur discovered that injecting killed rabies virus stimulates the immune system into producing antibodies against the microbe. What is more, Pasteur realized that the time it takes the body to develop this immune response is shorter than the disease's incubation period. He injected killed virus from the spinal cords of infected rabbits into people bitten by rabid dogs, who survived by gaining immunity before they showed any symptoms. Rabies can still develop in the interval between immunization and immune response, so now doctors also inject patients with rabies-specific antibodies to bridge the gap. Properly cleaning a wound with soap and water (which kills the virus by removing its membrane) is also quite important. This postexposure prophylaxis—wound care, five shots of a very safe vaccine and one dose of antibody—has had no failures in the U.S. since its introduction in 1975.

Jeanna's battle with rabies began when a bat crashed into an interior window of her church during a service. As she picked it up by the wing tips to release it outside, the bat lunged at Jeanna's left hand, and she suffered a quarter-inch cut to her index finger. Bat teeth are small and razor-sharp, so their bites are often not felt and can be hard to find. For these reasons, health officials recommend rabies immuniza-



tion for anyone coming into contact with a bat or sleeping in a room where a bat is discovered (unless the bat is captured and tests show that it is not rabid). Jeanna cleaned her scratch with hydrogen peroxide but did not seek immunization. Had she done so, she would likely have completed her sophomore year of high school without commotion.

Instead the rabies virus multiplied in Jeanna's finger for about a month, then entered a nerve and quickly traveled toward her brain, moving at a rate of about one centimeter an hour. Because of the exclusive targeting of the nervous system, where many types of immune cells do not operate, the body does not detect the virus until it has massively infiltrated the spinal cord and brain. The rabid patient eventually develops full paralysis from infected motor nerves and loses all sensation from infected sensory nerves. The mechanism behind this loss of neural activity is not known. Nor do researchers understand exactly how rabies kills a patient. Death can come in a variety of ways: stroke, heart stoppage, respiratory failure. The rabies virus seems to induce the brain to sabotage the vital organs, and it was this observation that inspired our treatment for Jeanna.

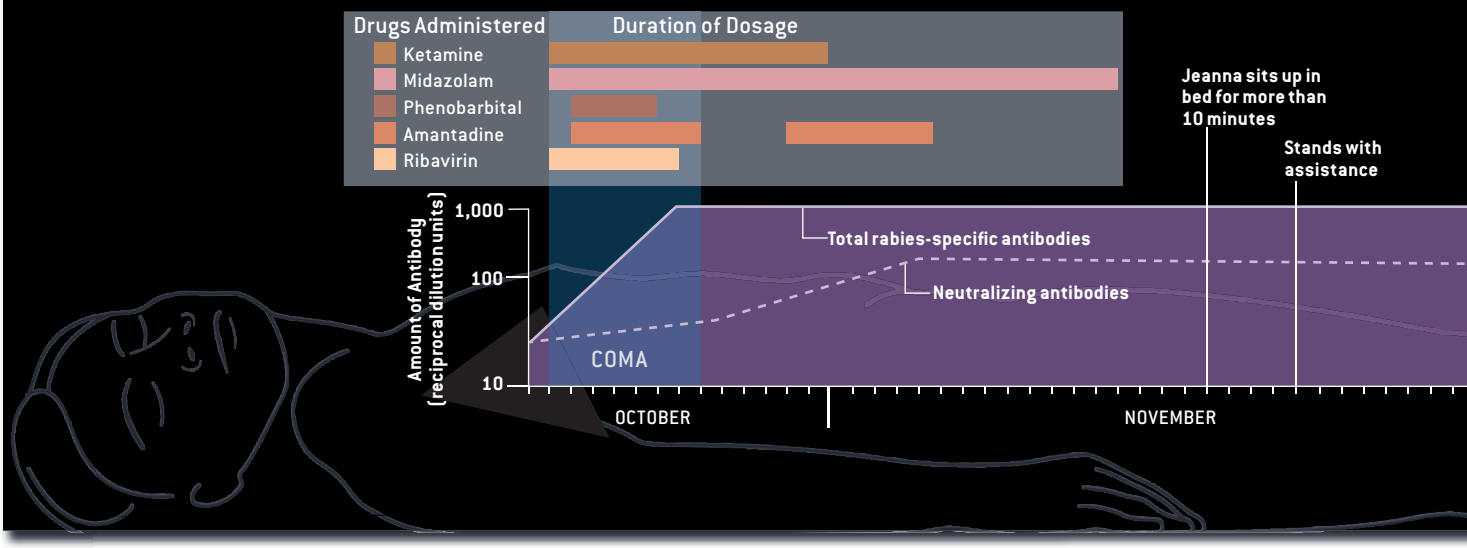
A gifted student and star of her high school volleyball team, Jeanna developed a flulike illness in October 2004, one month after being bitten. She then experienced numbness in her left hand, weakness in her left leg and double vision. She was admitted to her local hospital on a weekend and became lethargic and uncoordinated. These symptoms are typical of encephalitis, or inflammation of the brain, which is fairly common in medical practice, occurring in several thousand U.S. patients every year. Various types of viral and bacterial infections can cause encephalitis, but it can also be triggered by an immune response that goes awry and inflames the brain. Because the imaging of Jeanna's brain was normal, showing no sign of infection or stroke, her doctors assumed that she had this kind of postinfectious, autoimmune encephalitis. It appeared that she might progress to coma and need mechanical ventilation, so she was transferred to our hospital.

My colleagues and I would have almost certainly missed the diagnosis of rabies, but we got some help. Jeanna's local physician, Howard Dhonau, returned from his weekend off and did what beginning students are taught is one of the fundamentals of medicine: he took a repeat history of what had happened and heard about the bat. This history was critical to our success. I suggested that the team about to transport Jeanna to our hospital don protective equipment as a precaution. Although researchers have no proof that rabies can be transmitted from human to human, the tears and saliva of rabid animals are full of the virus, and contamination of wounds or mucous membranes (in the eyes, nose or mouth) is how rabies is perpetuated in nature. The medical staff caring for Jeanna wore visors, face masks, gowns and gloves during her first month of illness. The diagnosis of rabies required analysis of samples of her saliva, skin, blood and spinal fluid, which were shipped by air courier to the rabies section at the Centers for Disease Control and Prevention in Atlanta. This laboratory can provide first results in less than 24 hours.

Meanwhile I examined Jeanna. She was lethargic but performed simple commands. She could not maintain her balance and was weak in her left leg. Her reflexes were normal, allowing me to exclude the possibility of polio or West Nile virus. Also, her left arm was jerking intermittently. The pattern of numbness of her left hand and jerking of her left arm, correlating with the location of her bat bite, suggested rabies rather than the more common infections that cause encephalitis. But because doctors more often see an atypical presenta-

THE TREATMENT THAT SAVED JEANNA

On October 19, 2004, the day after Jeanna Giese arrived at Children's Hospital of Wisconsin in Milwaukee, doctors confirmed that she had rabies and used ketamine and midazolam to put her in a coma. Over the next week they also administered phenobarbital (a sedative), amantadine (an antiviral agent that helps to protect the brain) and ribavirin (a general antiviral drug). When Jeanna came out of the coma, her immune system was producing large amounts of antibodies



tion of a common disease than a truly rare disease, I reassured Jeanna's family, nurses and myself that it was almost impossible for her to have rabies. I was betting on another kind of encephalitis, most likely autoimmune, which is about 1,000 times more prevalent.

We had 24 hours to prepare a plan in case I was wrong. Between seeing other patients, I focused on another fundamental task of medicine: looking things up. I knew the conventional wisdom, which was that nothing works to treat rabies once symptoms appear; Jeanna was already dead if that was her diagnosis, and all we could do was limit her suffering. But I also knew that medicine is always advancing, so something new might be "out there." A search of the medical literature available online did not show any breakthroughs, but the lag between medical discovery and its publication can be as long as five years. I called Cathleen Hanlon, a rabies expert at the CDC, and received two depressing pieces of information. First, Jeanna's history and examination sounded like rabies to Hanlon, and second, nothing promising had emerged from recent scientific meetings or ongoing clinical trials.

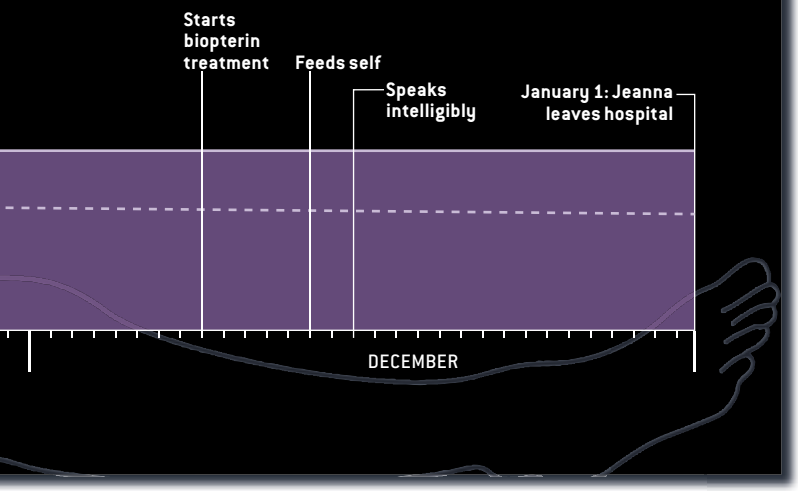
Given the limited time, I decided on a different search strategy. Virtually nobody had survived rabies, so I avoided publications about how to treat the disease in humans. Research on treatments usually starts with the effects of drugs on viruses cultivated in test tubes. Although this is a necessary first step, drugs that show promise in early studies are often toxic or cannot be delivered in adequate amounts to the site of infection, so I avoided these publications as well. What did catch my eye, when looking at the remaining articles, was the mystery that rabies experts have puzzled over for nearly 30 years: rabies patients die with virtually nothing visibly wrong with their brains. Equally important, when patients die of rabies after several weeks of intensive care, the virus can no longer be found in their bodies. The human immune system can clear the virus over time, but the eradication happens too slowly to save the patient's life.

From these two facts, we improvised a strategy. The rabies virus was apparently capable of hijacking the brain into killing the body, but without directly damaging the brain tissue itself. If we could inactivate the malfunctioning brain through careful use of drugs, putting Jeanna into a state of prolonged unconsciousness, it would restrain the terrible havoc wrought on her body and perhaps keep her alive long enough to allow her immune system to catch up.

To choose which drugs made the most sense, I searched the medical literature for studies connecting rabies with neurotransmitters (the chemicals the brain uses to signal between cells) or with neuroprotection (the science of using drugs or other interventions to shield the brain from injury). This search uncovered an astonishing pair of papers by Henri Tsiang of the Pasteur Institute in Paris. In the early 1990s Tsiang and his colleagues reported that ketamine, an anesthetic, was capable of inhibiting the rabies virus in the cortical neurons of rats. The work was reassuring for three reasons. First, the research indicated that ketamine affected a key step in the life cycle of the virus, when it is transcribing its genetic information inside the neuron [*see box on page 94*]. Second, the drug hindered only the rabies virus and not other viruses, suggesting that its effect was not a result of general toxicity to the animal. Third, a similar but more toxic drug called MK-801 also inhibited rabies in rat neurons, so the benefit most likely applied to an entire class of compounds.

DANIELA NAOMI MOLNAR (*timeline*); TAMI TOLPA (*illustration*)

against rabies, particularly the neutralizing antibodies that stop the virus from invading new cells. But her recovery was slow until physicians gave her biopterin, a compound similar to folic acid.



For more than 25 years, surgeons have used ketamine to induce and maintain unconsciousness in their patients, although it has largely been replaced because of its hallucinogenic side effects. (Illicit recreational users of the drug call it Special K.) Interestingly, ketamine's side effects appeared to offer another potential advantage for rabies patients. Ketamine acts as a neuroprotectant by blocking membrane proteins called NMDA glutamate receptors, which can kill neurons after becoming overactive following a stroke or other type of brain injury. Imagine my excitement to read about a drug that could contribute to suppressing the malfunctioning brain, while simultaneously clearing rabies virus and protecting the brain against further insult!

A Desperate Decision

AS AN INFECTIOUS DISEASE consultant, I did not have the skills to put Jeanna into a coma safely. I therefore applied another fundamental of medicine: seek help. I was new to Children's Hospital, so I asked Michael "Joe" Chusid, a senior infectious disease consultant, to help me find specialists in protecting the brain. Luckily, all the experts happened to be available that day. Kelly Tieves and Nancy Ghanayem, on call in the intensive care unit when Jeanna arrived, were experienced at minimizing injury to the brain following trauma and open-heart surgery, respectively. Our team also included Catherine Amlie-Lefond, a neurologist specializing in virus infections, and Michael Schwabe, an epilepsy expert who could provide us with continuous brain-wave monitoring to control the coma. George Hoffman, the team's anesthesiologist, recognized that our plan for inducing the coma was a routine practice for other conditions.

The team members recommended other medications to moderate the side effects of ketamine, confer additional neuroprotection and achieve the coma that was our therapeutic goal. Amantadine, an antiviral agent, would also help shut down the neuronal NMDA receptors, binding to them at a different site than the one blocked by ketamine. Midazolam, a sedative in the benzodiazepine group, and phenobarbital would help suppress Jeanna's brain activity. Charles Rupprecht, a rabies expert at the CDC, later recommended that we administer a general antiviral drug, ribavirin. Although ribavirin had been tried on rabies patients before without success, we always listen to smart people.

Having redundant consultants in each specialty permitted us to critique the hypothesis and decide whether it was safe to proceed. Whenever new things are tried in medicine or biology, they usually fail and often cause harm. That is why treatments are supposed to progress from test tubes to animal studies to clinical trials. My hypothesis appeared too simple. It might result in a medical outcome even worse than death; four of the five immunized survivors of rabies ended up with severe disabilities. We convened for an hour after Rupprecht confirmed that Jeanna had rabies. We informed her parents of the diagnosis, outlined conventional options and then proposed our treatment. With Jeanna's death all but guaranteed, Ann and John Giese asked us to try something new so that more would be known for the next child with rabies.

We estimated that Jeanna's immune system would need about five to seven days to make antibodies targeting the rabies virus. From previous human cases of rabies, we knew that the natural immune response is vigorous once it is triggered. Given that Jeanna's brain was already full of rabies virus, adding more of it with a killed-virus vaccine was unlikely to help. It might actually cause harm by skewing the natural immune response away from the virus already in Jeanna's brain and toward variants overrepresented in the vaccine. For similar reasons, we chose to avoid boosting Jeanna's immune system by injecting her with rabies-specific antibodies or interferons (proteins that enhance immune activity). We elected to induce coma in Jeanna for a week and to analyze samples of her blood and spinal fluid over that period to confirm that she was producing her own antibodies.

Instead of the extreme commotion that characterizes the terminal care of rabies—including wild swings in heart rate and blood pressure—we encountered no major crises during

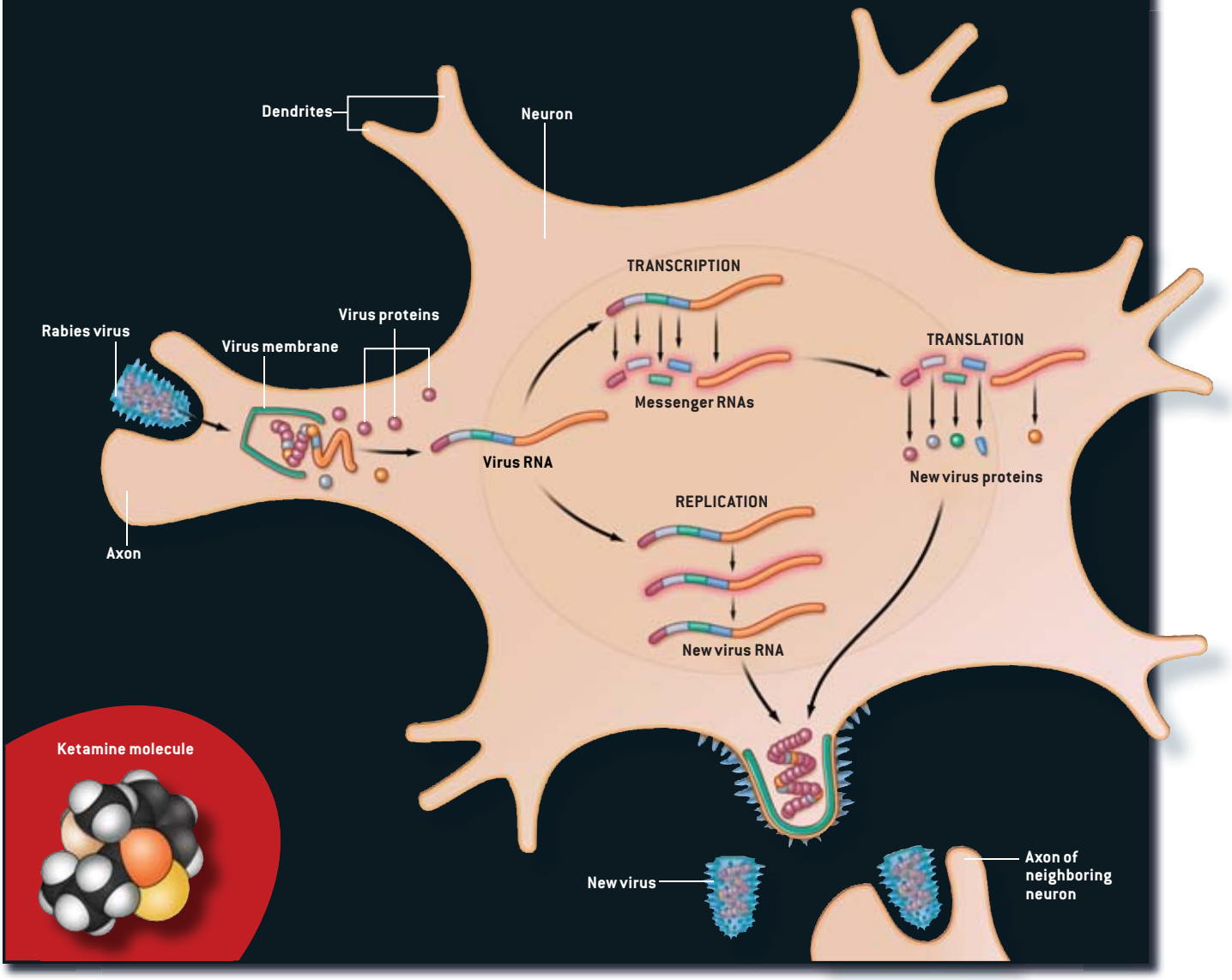
THE AUTHOR

RODNEY E. WILLOUGHBY, JR., is associate professor of pediatrics at the Medical College of Wisconsin and a consultant in pediatric infectious diseases at Children's Hospital of Wisconsin. He is a graduate of Princeton University and the Johns Hopkins School of Medicine, with postdoctoral training in pediatrics, pediatric infectious diseases, carbohydrate biochemistry and clinical investigation. His research interests include rabies, cerebral palsy, and antibiotic cycling and probiotics to minimize resistant infections in hospitals. He is a 2006 recipient of the Children's Miracle Network Achievement Award.

CURBING A DEADLY VIRUS

After the rabies virus penetrates the axon of a neuron, the microbe sheds its membrane and releases its proteins and RNA, which travel to the cell body. The viral RNA generates messenger RNAs (transcription), which in turn use the cell's machinery to produce the virus's five proteins (translation).

Then the viral RNA creates copies of itself, which are assembled with the proteins into new microbes that emerge from the neuron's dendrites to attack the next nerve cell. Studies indicate that ketamine (*inset*), a compound long used as an anesthetic, inhibits the transcription stage of the virus's life cycle.



Jeanna's coma. By the end of the week, her body had begun to produce large quantities of neutralizing antibodies that prevent the virus from invading new cells and perhaps clear the microbe by other unknown means. But the real test would come when we brought Jeanna back to consciousness. The day we removed her from coma was the worst day of my life: Jeanna was completely paralyzed and unresponsive. We had no idea whether she was still in there or what would follow. But we knew that rabies patients could appear falsely brain-dead, so we did not lose hope. The next day Jeanna attempted to open her eyes and later developed reflexes in her legs. After six days she was fixing her gaze on her mother's face (in preference to mine) and opening her mouth to help her nurses clean it. By 12 days, she was sitting up in bed.

Full paralysis results in severe deconditioning of the body. Strength, stamina and coordination are all lost, as well as the ability to swallow and speak. Jeanna's recovery took a tremendous amount of hard work. In the first two months she ran into perplexing delays: for example, she made rapid progress in walking and exercise but still could not speak or swallow. She had other residual problems (such as a buildup of lactic acid in her body) that made me think of metabolism disorders. In consultation with William Rhead, a geneticist at our hospital, we diagnosed that Jeanna had acquired a biopterin deficiency. Bi-

TAMI TOLPA

opterin is chemically similar to folic acid, the B vitamin essential to cell growth. The compound is present in limited amounts in the brain, where it is critical for making neurotransmitters such as dopamine, epinephrine, norepinephrine, serotonin and melatonin. Biopterin also controls how a brain enzyme, neuronal nitric oxide synthase, maintains the tone of blood vessels feeding the brain. In fact, we realized that low levels of biopterin could explain most symptoms of rabies other than the late effects on peripheral nerves.

This was a stunning break, because biopterin is available as an oral supplement. After receiving biopterin, Jeanna quickly relearned how to speak and swallow. The speedy improvement allowed her to leave the hospital on January 1, 2005, three months earlier than expected. We have since confirmed biopterin deficiency in the only other rabies patient from whom samples were suitably preserved. We are testing whether low biopterin levels occur in other animal species infected with rabies; if so, this deficiency may help explain how the virus ravages the body. Why rabies would reduce biopterin—most brain infections increase it—is not clear. We have prepared to test for and treat biopterin deficiency in future rabies patients who receive the Milwaukee protocol.

A Curable Controversy

BY THE FIRST ANNIVERSARY of her diagnosis with rabies, Jeanna attended an international meeting of rabies researchers in Canada as the guest of honor at their gala dinner, where she gave a speech. She rejoined her original classmates for her junior year of high school, received excellent grades and obtained a provisional driver's permit. The only lingering reminders of her bout with rabies are a small area of numbness on the bitten finger, an alteration in the tone of her left arm and a wider gait when she runs. She is graduating from high school this year and hopes to become a veterinarian.

But can the Milwaukee protocol save any other lives? Over the past two years the treatment has been tried six times, without success, in Germany, India, Thailand and the U.S. Unfortunately, several of the attempts violated key assumptions in our hypothesis or did not use most of the drugs in Jeanna's regimen. The medical community has been reluctant to repeat our therapy, with some experts publicly opposed. The resistance is understandable because Jeanna's survival appears to refute lab studies showing that the rabies virus kills brain cells. These studies may be misleading, however, because the strains of rabies used in the lab may be more prone to cause cell death than the viruses circulating in nature.

Other experts have argued that Jeanna survived because she was infected with an unusually weak variant of rabies. This argument is difficult to answer because we did not recover samples of the virus from Jeanna's body. (The CDC obtained rabies-specific antibodies from Jeanna but not the virus itself, which is difficult to isolate.) Analysis of the rabies virus carried by bats indicates that it is different from the strain in dogs, with a greater propensity for multiplying in skin rather than in muscle. But bat virus appears to be no less deadly than the canine version.

REMARKABLE RECOVERY: Jeanna Giese (*shown with the author*) is graduating from high school this year and hopes to become a veterinarian. The only lingering reminders of her battle with rabies are a small area of numbness on the bitten finger, an alteration in the tone of her left arm and a wider gait when she runs.

Perhaps the best way to address these concerns is to apply our treatment to rabid animals to determine which parts of the protocol—the induced coma, the antiviral activity, the shutdown of NMDA receptors—are crucial in battling rabies. We have asked six veterinary schools to permit these studies, but school officials are wary about treating rabid animals in their intensive care units. The lack of follow-up research is a global loss, because we will not know whether our strategy actually works until others try it. If our success is replicated, investigators can systematically address which drugs are effective and at what doses and determine whether biopterin can significantly shorten convalescence. Furthermore, doctors must find ways to reduce the cost of treatment and rehabilitation—which totaled at least $800,000 in Jeanna's case—to make it a practical option in developing countries where rabies is most common. It may be too much to expect rabies to go from 100 percent fatal to 100 percent curable, but at least now we have a chance to improve the odds. SA

MORE TO EXPLORE

Inhibition of Rabies Virus Transcription in Rat Cortical Neurons with the Dissociative Anesthetic Ketamine. B. P. Lockhart, N. Tordo and H. Tsiang in *Antimicrobial Agents and Chemotherapy,* Vol. 36, No. 8, pages 1750–1755; August 1992.

Prophylaxis against Rabies. C. E. Rupprecht and R. V. Gibbons in *New England Journal of Medicine,* Vol. 351, No. 25, pages 2626–2635; December 16, 2004.

Survival after Treatment of Rabies with Induction of Coma. R. E. Willoughby, Jr., K. S. Tieves, G. M. Hoffman, N. S. Ghanayem, C. M. Amlie-Lefond, M. J. Schwabe et al. in *New England Journal of Medicine,* Vol. 352, No. 24, pages 2508–2514; June 16, 2005.

More information about rabies and the Milwaukee protocol is available at **www.mcw.edu/rabies**

RICK WOOD *MCT/KRT*

ELECTRONIC STABILITY CONTROL

Steer Clear

Automakers are offering electronic stability control on more and more passenger vehicles to help prevent them from sliding, veering off the road, or even rolling over. The technology is the product of an ongoing evolution stemming from antilock brakes.

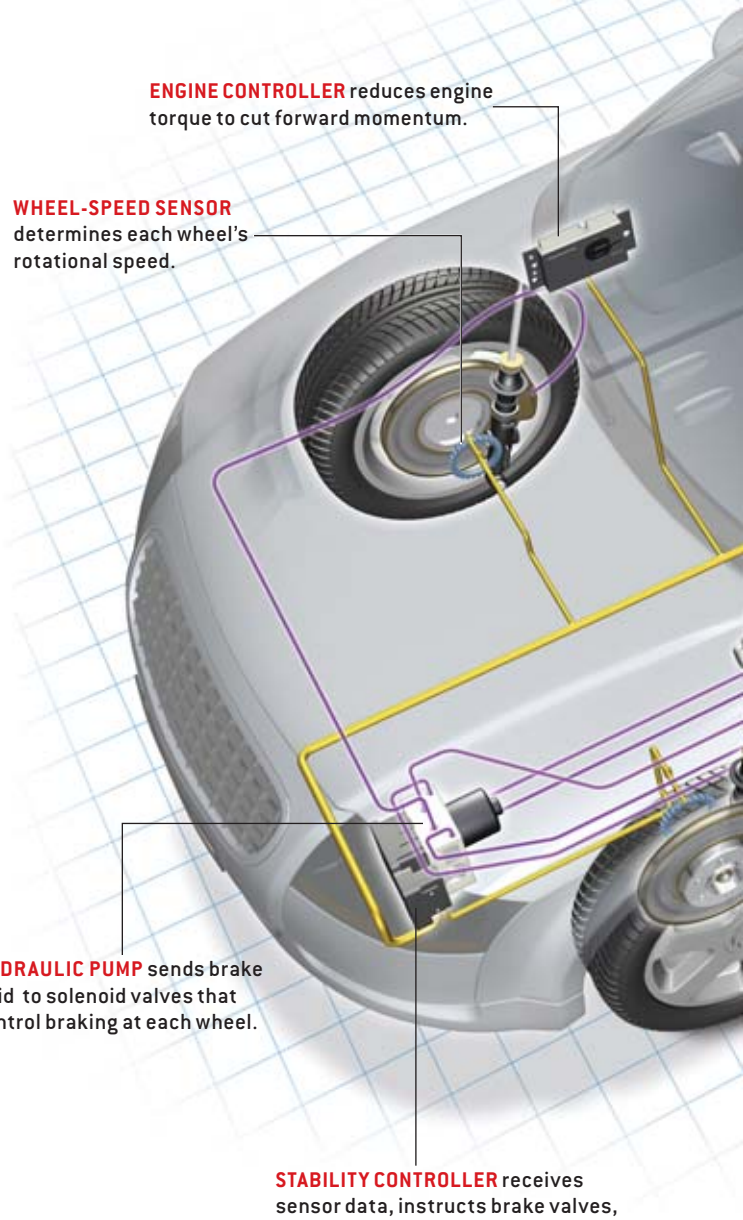
When a driver jams the brake pedal too hard, antilock hydraulic valves subtract brake pressure at a given wheel so the wheel does not lock up. As these systems proliferated in the 1990s, manufacturers tacked on traction-control valves that help a spinning drive wheel grip the road.

For stability control, engineers mounted more hydraulics that can apply pressure to any wheel, even if the driver is not braking [*see main illustration*]. When sensors indicate the car is sliding forward instead of turning or is turning too sharply, the actuators momentarily brake certain wheels to correct the trajectory [*see far right illustration*]. "Going to electronic stability control was a big step," says Scott Dahl, director of chassis-control strategy at supplier Robert Bosch in Farmington Hills, Mich. "We had to add sensors that can determine what the driver intends to do and compare that with what the car is actually doing." Most systems also petition the engine-control computer to reduce engine torque to dampen wayward momentum.

The latest systems go even further. "Basically, electronic stability control is steering with the brakes," says Philip Headley, chief engineer at Continental Automotive Systems in Auburn Hills, Mich. "The new generation adds help from the steering wheel." Actuators on the steering shaft turn the wheels (the steering wheel itself does not move). And although all the systems lessen the chance for rollover by minimizing a sideways slide, if a car does begin to lift, the newer generation can apply braking patterns to attempt to put the rising wheels back on the pavement.

The National Highway Traffic Safety Administration estimates that stability control adds $111 to a car's cost. It also says the technology would save up to 10,300 lives each year if installed in all U.S. passenger vehicles, ranking it second behind seat belts and ahead of air bags. Given that potential, the NHTSA wants the equipment to be mandatory by the 2012 model year; a final ruling is expected within months. —*Mark Fischetti*

ELECTRONIC STABILITY-CONTROL sensors compare how a driver is attempting to steer with the direction the car is actually heading; actuators control brakes and the engine to counter an errant trajectory.



SAMUEL VELASCO; SOURCES: CONTINENTAL AUTOMOTIVE SYSTEMS; ROBERT BOSCH

DID YOU KNOW...

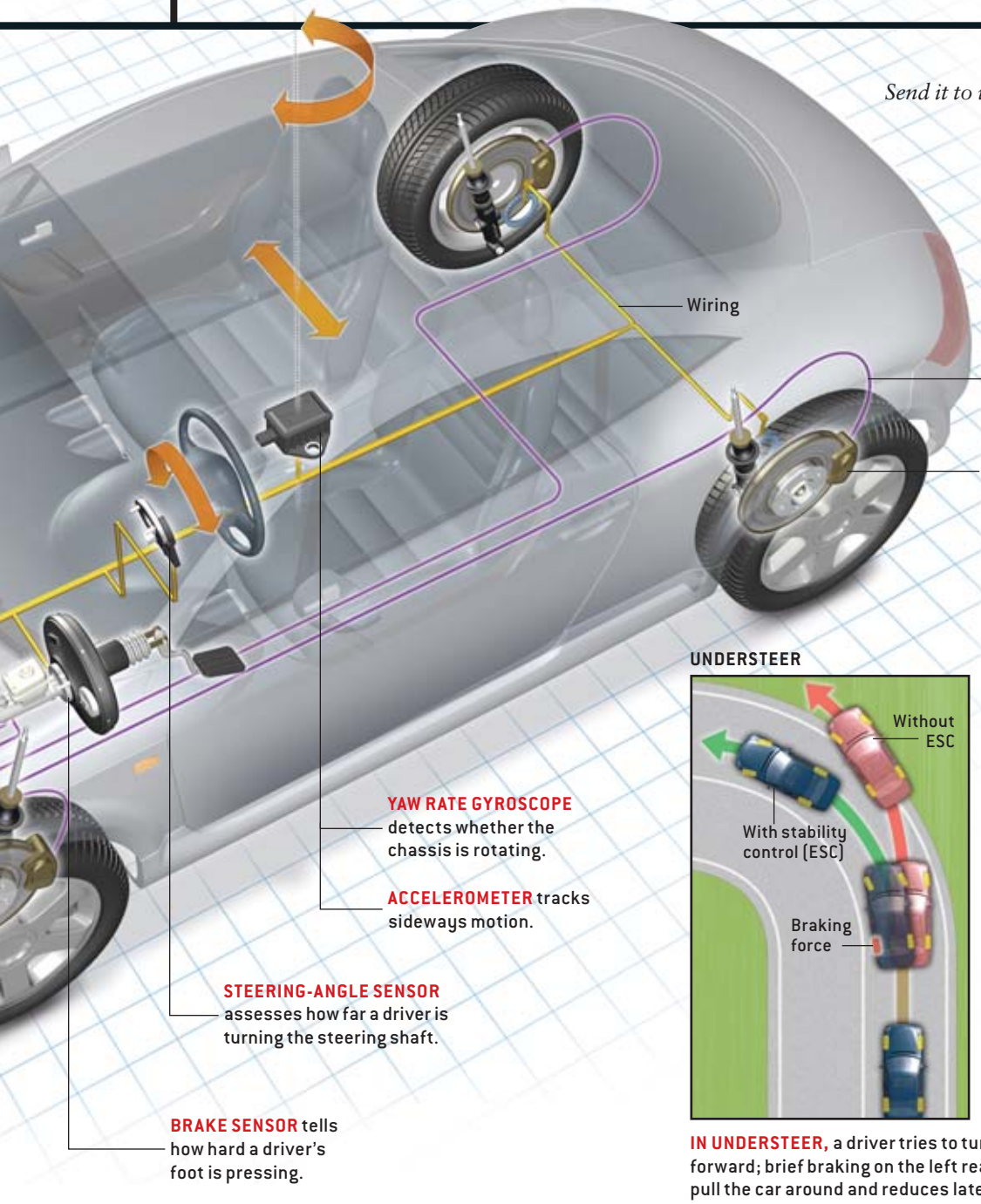
KILL SWITCH: Most automakers let a driver temporarily turn off electronic stability control so he or she can spin the wheels to get unstuck from a snowdrift—or do doughnuts in the parking lot. Tricksters might want to check their collision insurance first.

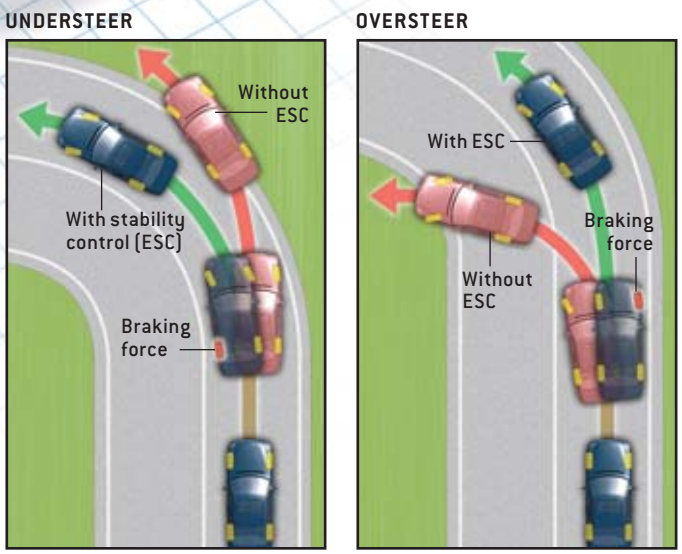
4-WHEEL DIFFERENCE: Four-wheel drive works by locking the front and rear axles together, making it impossible to brake just one wheel. A stability-control computer must instruct the differential gear that binds the axles to unlock before corrective braking can begin; some systems turn off in four-wheel drive.

NAME CONTROL: Seeking novelty, car manufacturers label their electronic stability systems with a variety of trade names. Among them: Dynamic Stability Control (BMW), StabiliTrak (General Motors) and Interactive Vehicle Dynamics (Ford).

NEXT IN LINE: Because stability systems can brake each wheel, they are being used in long-anticipated collision avoidance systems, now being offered on a few cars. The vehicle emits a radar beam to determine if its driver is encroaching too closely on a car in front of it and, if so, applies the brakes. In the future, steering control may aid so-called lane-keeping systems that could prevent a driver from wandering outside a lane.

Have a topic idea?
Send it to workingknowledge@sciam.com





IN UNDERSTEER, a driver tries to turn, but the front wheels slide forward; brief braking on the left rear wheel adds a pivoting torque to pull the car around and reduces lateral force on the tire. In oversteer, the car turns too hard, causing the back end to fishtail and possibly spin out or even roll over; brief braking on the right front wheel adds a pivot and reduces the lateral force.

The Car Doctor Is In

A WAY TO DIAGNOSE ENGINE PROBLEMS WITHOUT EVER HAVING TO LOOK UNDER THE HOOD

BY PHILIP YAM

"Uh-oh"—or another, less family-friendly phrase—is probably the first thing drivers utter when the "check engine" light flashes on. It may not mean that the engine is about to explode, but it still demands attention. Odds are, the vehicle will not pass the emissions test required for registration.

Having recently bought a new car, I was eager to try a device similar to those used by mechanics to diagnose the problems that trigger the check engine light. Offered by CarMD, a year-old firm based in Fountain Valley, Calif., the $90 tool retrieves data from the onboard computer of any vehicle sold in the U.S. after January 1, 1996. The device, which won an innovations award at the 2007 Consumer Electronics Show, will not outdo a mechanic's diagnostic equipment, but it will tell you what might be wrong in more understandable terms and how much it might cost to fix the problem. Besides, you could confidently tell the mechanic to check for a leaking exhaust manifold—even if you have no clue what a manifold is.

The driving force (ahem) behind this handy instrument is air pollution—or more specifically, the need to control it. Stiff emissions guidelines in the 1970s eventually led the Society of Automotive Engineers to devise a means of monitoring engines and their exhaust in a standard way. The culmination of that effort is the second generation of onboard diagnostics (OBD-II).

Today's automobile contains myriad sensors, explains Keith Andreasen, CarMD's director of technical services. In the mid-1990s, he says, a vehicle contained "maybe 13 to 15 sensors"; now there are "thousands in a modern-day car." Those sensors keep tabs on the emissions and relay the readings to the car's computer. For instance, if a sensor detected fumes escaping from a balky fuel gasket, it would inform the computer, which would store the data and turn on the check engine light. The computer assigns the problem a "diagnostic trouble code," such as "P0430." (Most codes start with a P, for power train; other prefix codes are B for body, C for chassis and U for the vehicle's communications network.) OBD-II employs several hundred codes, and the readers that extract this information from the computer have become indispensable to both the professional and weekend grease monkey.

CAR CARE: A device from CarMD interfaces with a vehicle's computer to diagnose "check engine" light problems. The firm also provides an estimated repair cost.

JOHN FRASER

Because CarMD targets the consumer, it has tried to make the diagnostic process simple. To communicate with the onboard computer, the handheld tester plugs into the vehicle's data-link connector—a rhomboid-shaped, 16-pin socket reminiscent of the parallel ports of yester decade. The hardest part about using CarMD, in fact, is finding the connector: I had to search online to learn the port's location in my 2006 Mini Cooper S. It was underneath the dashboard, to the left of the steering column, an area that forces you to twist your neck to the breaking point to get a direct view. Fortunately, my wife did not mind scrunch-

ing into the foot well with a flashlight to find the plastic flip-down cover, which had "OBD" stamped on it. CarMD has since posted the locations and images of the sockets on its Web site (www.carmd.com), although the description wrongly stated that my socket was uncovered.

After attaching the CarMD reader and turning the ignition to its second position, the unit began communicating with Darwin, the name my wife gave to our Mini. After four beeps from the reader, I yanked it out and saw that the green LED was on—everything was A-OK. I would have been extremely perturbed otherwise, considering that Darwin was only six months old. Connecting the reader to my Windows XP computer via a USB cable (no Mac OS X support), I uploaded the data to my account on the CarMD Web site. The unit correctly extracted the VIN number and description of my car.

But I needed to see how CarMD would do with the check engine light shining. According to the Aftermarket Industry Association, about 6 percent of all vehicles are cruising around with these indicators on. Given the statistics, at least one of those cars should belong to a staff member of *Scientific American,* which employs about 80 people in New York City. Sure enough, I shook out coworkers with car problems.

For our online project manager David Yu and his 1999 Acura CL, the tester produced trouble code P0401: "EGR System Insufficient Flow Detected." Among the possible causes, as the CarMD Web site describes them: the EGR valve source vacuum supply line was open or restricted; the EGR exhaust manifold passages were clogged or restricted; or the EGR valve assembly or solenoid valve was damaged or had failed. The most likely fix, according to CarMD, was replacing the EGR valve and cleaning the ports. Making sense of the arcane lingo, some of which is gibberish to the casual driver and grammar-

ian alike, might need some Web research. But for most people, knowing that EGR stands for "exhaust gas recirculation" is much less important than knowing the impact on their wallet. Factoring in geographically weighted labor costs, the program estimated that David should expect to pay $487.70 for repairs.

Mike Florek, our general manager, had a more elusive problem with his 2003 Nissan Pathfinder, which had been flashing its check engine light for some months. CarMD implicated the catalytic converter or the heated-oxygen sensors. When we ran the test in December, CarMD did not offer any suggested fixes—it reported that the problem appeared too infrequently. A few weeks later, though, the details of the repair and its cost appeared (technicians keep updating CarMD's database). The fix: a replacement catalytic converter. The bad news: $828.96. A skilled mechanic, though, would still need to inspect the car thoroughly, because whatever damaged the converter in the first place could still be around to wreck the replacement. CarMD bills the device as a means to "empower" drivers so that they will not, for instance, feel intimidated by sketchy repair people.

But CarMD can also pay for itself, as Silvia De Santis, our prepress and quality manager, can attest. No one likes to see the check engine light come on, but it's really inconvenient when it happens while you are in the throes of labor. (To her husband's credit, he continued driving to the hospital.) Within days after the birth of their son, they brought their Toyota to a mechanic. "After $200 of diagnostic tests," Silvia says, "they told us it was probably an improperly closed gas tank cap"—the most common reason for the light to turn on. I experimented with removing my Mini's gas cap; sure enough, after the check engine light came on, CarMD advised me to "check for loose gas cap and replace if necessary."

The idea for the company began in 1997, when Andreasen and Ieon C. Chen, both working for a developer of automotive test equipment, discussed ways to market a consumer-friendly method of interpreting OBD-II information. Manufacturing the handheld tester was not hard; the challenge was creating a database that users could tap into. Back then, the appropriate Internet technology and usage "was a long way off," Andreasen says. Chen, now the firm's CEO, tried getting venture capital, but "neither the company nor the market was ready," recounts CarMD's marketing director Kristin Brocco. Instead Chen and Andreasen spent years compiling a list of the repairs associated with the trouble codes. The database is in fact the main strength of CarMD.

The firm now has several thousand customers, each of whom can register up to three vehicles. The company also maintains a phone system of live technicians for motorists who break down on the road and cannot get online. The device knows only what the car's computer tells it, so it may not help with pesky issues such as check engine lights that repeatedly turn on and off on their own. A trusted repair shop is still essential, but communicating with the mechanic no longer has to be a one-way street. SA

CarMD READER plugs into the data-link connector, which in the 2006 Mini Cooper S hides under the driver's side dashboard. The unit displays the trouble code and has color LEDs to indicate engine health.

PHILIP YAM (*car interior*); CARMD (*handheld device*)

Is Beauty Truth and Truth Beauty?

HOW KEATS'S FAMOUS LINE APPLIES TO MATH AND SCIENCE BY MARTIN GARDNER

WHY BEAUTY IS TRUTH: A HISTORY OF SYMMETRY
by Ian Stewart
Basic Books, 2007 ($26.95)

The title of Ian Stewart's book (he has written more than 60 others) is, of course, taken from the enigmatic last two lines of John Keats's "Ode on a Grecian Urn":

> "Beauty is truth, truth beauty,"—that is all
> Ye know on earth, and all ye need to know.

But what on earth did Keats mean? T. S. Eliot called the lines "meaningless" and "a serious blemish on a beautiful poem." John Simon opened a movie review with "one of the greatest problems of art—perhaps the greatest—is that truth is not beauty, beauty not truth. Nor is it all we need to know." Stewart, a distinguished mathematician at the University of Warwick in England and a former author of this magazine's Mathematical Recreations column, is concerned with how Keats's lines apply to mathematics. "Euclid alone has looked on Beauty bare," Edna St. Vincent Millay wrote. To mathematicians, great theorems and great proofs, such as Euclid's elegant proof of the infinity of primes, have about them what Bertrand Russell described as "a beauty cold and austere," akin to the beauty of great works of sculpture.

Stewart's first 10 chapters, written in his usual easygoing style, constitute a veritable history of mathematics, with an emphasis on the concept of symmetry. When you perform an operation on a mathematical object, such that after the operation it looks the same, you have uncovered a symmetry. A simple operation is rotation. No matter how you turn a tennis ball, it does not alter the ball's appearance. It is said to have rotational symmetry. Capital "H" has 180-degree rotational symmetry because it is unchanged when turned upside down. It also has mirror reflection symmetry because it looks the same in a mirror. A swastika has 90-degree rotational symmetry but lacks mirror reflection symmetry because its mirror image whirls the other way.

Associated with every kind of symmetry is a "group." Stewart explains the group concept in a simple way by considering operations on an equilateral triangle. Rotate it 60 degrees in either direction, and it looks the same. Every operation has an "inverse," that cancels the operation. Imagine the corners of the triangle labeled *A*, *B* and *C*. A 60-degree clockwise rotation alters the corners' positions. If this is followed by a similar rotation the other way, the original positions are restored. If you do nothing to the triangle, this is called the "identity" operation. The set of all symmetry transformations of the triangle constitutes its group.

Stewart's history begins with Babylonian and Greek mathematics, introducing their basic concepts in ways a junior high school student can understand. As his history proceeds, the math slowly becomes more technical, especially when he gets to complex numbers and their offspring, the quaternions and octonions. The history ends with the discoveries of Sophus Lie, for whom Lie groups are named, and the work of a little-known German mathematician, Joseph Killing, who classified Lie groups. Through this historical section, Stewart skillfully interweaves the math with colorful sketches of the lives of the mathematicians involved.

SYMMETRY was once a synonym for beauty—recall William Blake's "Tyger! Tyger! burning bright/In the forests of the night,/What immortal hand or eye/Could frame thy fearful symmetry?"

ANDY ROUSE *Corbis*

Not until the book's second half does Stewart turn to physics and explain how symmetry and group theory became necessary tools. A chapter on Albert Einstein is a wonderful blend of elementary relativity and details of Einstein's life. Next comes quantum mechanics and particle theory, with several pages on superstrings, the hottest topic in today's theoretical physics. Stewart is a bit skeptical of string theory, which

sees all fundamental particles as inconceivably tiny filaments of vibrating energy that can be open-ended or closed like a rubber band. He does not mention two recent books (reviewed in the September 2006 issue of *Scientific American*) that give string theory a severe bashing. Lee Smolin's *The Trouble with Physics* denounces string theory as "not a theory at all," only a mishmash of bizarre speculations in search of a viable theory. Peter Woit's book is entitled *Not Even Wrong*, a quote from the great Austrian physicist Wolfgang Pauli. He once described a theory as so bad it was "not even wrong."

Is string theory beautiful? Its promoters think so. Smolin and Woit believe that its recent absorption into a richer conjecture called M-theory has turned the former beauty of strings into mathematical structures as ugly as the epicycles Ptolemy invented to explain the orbits of planets as they circle the earth.

We are back to the mystery of Keats's notorious lines. In my opinion, John Simon is right. Even beautiful mathematical proofs can be wrong. In 1879 Sir Alfred Kempe published a proof of the four-color map theorem. It was so elegant that for 10 years it was accepted as sound. Alas, it was not. Henry Dudeney, England's great puzzle maker, published a much shorter and even prettier false proof.

In *The New Ambidextrous Universe* I write about the vortex theory of atoms. This popular 19th-century conjecture had an uncanny resemblance to superstrings. It maintained that atoms are not pointlike but are incredibly tiny loops of energy that vibrate at different frequencies. They are minute whirlpools in the ether, a rigid, frictionless substance then believed to permeate all space. The atoms have the structure of knots and links, their shapes and vibrations generating the properties of all the elements. Once created by the Almighty, they last forever.

In researching vortex theory, I came across many statements by eminent physicists, including Lord Kelvin and James Clerk Maxwell, suggesting that vortex theory was far too beautiful not to be true. Papers on the topic proliferated, books about it were published. Scottish mathematician Peter Tait's work on vortex atoms led to advances in knot theory. Tait predicted it would take several generations to develop the theory's mathematical foundations. Beautiful though it seemed, the vortex theory proved to be a glorious road that led nowhere.

Stewart concludes his book with two maxims. The first: "In physics, beauty does not automatically ensure truth, but it helps." The second: "In mathematics beauty *must* be true—because anything false is ugly." I agree with the first statement, but not the second. We have seen how lovely proofs by Kempe and Dudeney were flawed. Moreover, there are simply stated theorems for which ugly proofs may be the only ones possible.

Let me cite two recent examples. Proof of the four-color map theorem required a computer printout so vast and dense that it could be checked only by other computer programs. Although there may be a beautiful proof recorded in what Paul Erdös called "God's book"—a book that, he suggested, included all the theorems of mathematics and their most beautiful proofs—it is possible that God's book contains no such proof. The same goes for Andrew Wiles's proof of Fermat's last theorem. It is not computer-based, but it is much too long and complicated to be called beautiful. There may *be* no beautiful proof for this theorem. Of course, mathematicians can always hope and believe otherwise.

Because symmetry is the glue and tape that binds the pages of Stewart's admirable history, a stanza from Lewis Carroll's immortal nonsense ballad *The Hunting of the Snark* could serve as an epigraph for the book:

> You boil it in sawdust: you salt it in glue:
> You condense it with locusts and tape:
> Still keeping one principal object in view—
> To preserve its symmetrical shape.

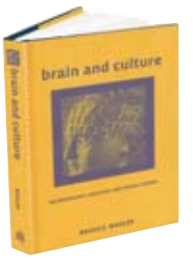
SA

Martin Gardner wrote Scientific American*'s Mathematical Games column for 25 years. His latest book,* The Annotated Hunting of the Snark, *was published last year by W. W. Norton.*

THE EDITORS **RECOMMEND**

BRAIN AND CULTURE: NEUROBIOLOGY, IDEOLOGY, AND SOCIAL CHANGE

by Bruce E. Wexler. MIT Press, 2006 ($34)



Wexler, a professor of psychiatry at the Yale University School of Medicine, offers an original, neurologically based hypothesis to explain the sectarian strife that plagues the globe today—in his words, a "neurobiology of ideology." According to his argument, input from the world around us shapes our brains, and we shape the world that shapes our brains. Through these processes the human brain molds itself to its individual environment. This early wiring then makes it difficult to accept the unfamiliar later in life. In fact, Wexler points out, disparities between the environment and the developed brain cause distress and dysfunction. His thesis raises the question in the reader's mind of why some novel experiences delight us, and one wishes this thoughtful author had addressed that puzzle. Are some experiences only apparently novel? Or do they actually conform in some subtle way to the wiring in our brains? But the book is a fascinating step forward in deconstructing the seemingly universal us/them mentality. And, who knows, once it is decoded, perhaps the next step will be to defuse it.

Doughnut Try This at Home

TINKERING WITH A CLASSIC CAN SOMETIMES GO AWRY, OR POSSIBLY EVEN PUMPERNICKEL **BY STEVE MIRSKY**

People love to tinker. And so they often combine already good things to come up with newer, seemingly even better things. Iodized salt. Vitamin D fortified milk. Fluoridated water. Now add a newcomer to the list of such hybrids: the caffeinated doughnut.

Currently being marketed under the trademarked name Buzz Donuts, the caffeinated doughnut is the brainchild of one Robert Bohannon, whom a press release about his invention describes as a "molecular scientist." (Is there any other kind?) By the way, I had been under the impression that I invented the caffeinated doughnut long ago when I cleverly dipped a doughnut into a cup of hot coffee. But apparently not. (Not to mention the time I was victimized by the false dichotomy of assuming that I could develop a self-cleaning oven by simply never cleaning the oven myself.)

Anyway, Bohannon's early attempts to add the life-affirming zip of caffeine to baked goods led to literally bitter defeat: "They were terrible, absolutely horrid," he was quoted as saying. "It would just make you puke." Eventually Bohannon hit on a process by which to mask the tart taste when the stimulating alkaloid was added to a recipe. (Bagels with a boost could also be on the drawing and cutting boards.)

The advent of the caffeinated doughnut necessarily makes one wonder what else might be improved by the addition of the right beneficial agent. Here are suggested amalgamations for other molecular scientists to develop.

The Saltpeter Oyster. Enjoy the taste, texture and possible cholera infection to be found in oysters without the pesky alleged aphrodisiacal effects. (And for an equal and opposite effect, try the Viagrafied lutefisk.)

The Premustarded Hot Dog. Ever balance a hot dog and drink in your lap at a baseball game while fighting the guys on either side of you for a precious piece of armrest? Trying to manipulate mustard packets under those circumstances would challenge Kali. A thin river of mustard within the frankfurter would be a blessing and could probably be achieved cost-effectively by using syringes confiscated from the players.

The Depilatory Deodorant. An infusion of the powerful hair remover calcium thioglycolate into deodorants or antiperspirants could be a real morning time saver if you're hitting the gym sleeveless. Should be available in small, medium, large and Robin Williams hirsute economy size.

The Nitro Cheesesteak. Love the shaved beef cheezwizardry to be found at Pat's or Geno's in funky South Philly, but hate the feeling of chest-cramping death? Cheesesteaks laced with nitroglycerin should allow customers to eat with hearty abandon. (While researching this item, I discovered that nitroglycerin, which achieves its medical effect through blood vessel dilation, is in fact being added to a brand of condom slated to soon hit the market, thus giving unfortunate potential new meaning to the expression "explosive orgasm.")

Doughnutty Coffee. Doughnuts leave powdered sugar on your shirt and present a small but serious choking hazard. But who doesn't want something sweet with that morning cup o' joe? How about adding the classic American on-the-run breakfast item to coffee—in liquid form! Doughnutty coffee would give commuters the easy-to-handle morning meal they need with their caffeine, while leaving a hand free with which to drive, apply deodorant or perform a Heimlich maneuver on anyone gagging on a chunk of dry doughnut. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

Look for Anti Gravity: Allegedly Humorous Writing from Scientific American, *an anthology of these columns, published April 1. No fooling.*

ASK THE EXPERTS

How can Bayes' theorem assign a probability to the existence of God?

Chris Wiggins, an associate professor of applied mathematics at Columbia University, offers this explanation:

In the 18th century Reverend Thomas Bayes first expressed the probability of any event—given that a related event has occurred—as a function of the probabilities of the two events independently and the probability of both events together.

Take the following example. A patient goes to see a doctor for a checkup. The doctor knows a test he performs has 99 percent reliability—that is, 99 percent of sick people test positive, and 99 percent of healthy people test negative. The doctor also knows that only 1 percent of the population is sick. Now the question is: If the patient tests positive, what are the chances the patient is sick? The intuitive answer is 99 percent, but the correct answer is actually 50 percent. Bayes' theorem relates the probability of being sick given a positive test result, p(s|+), to the probability of receiving a positive test result given that one is sick, p(+|s), and to the general probability of being sick, p(s), and the general probability of getting a positive result, p(+). (Calculating the last probability is left to the reader.) Bayes' theorem in this case would read $p(s|+) = p(+|s) \times p(s) / p(+)$.

The importance of accurate data in quantitative modeling is central to the subject raised here: using Bayes' theorem to calculate the probability of the existence of God. (Bayes, for his part, never related his theorem to the subject.) Scientific discussion of religion is a popular topic at present, with several recent books arguing against theism. In one, *The God Delusion* (Houghton Mifflin, 2006), University of Oxford professor Richard Dawkins argues specifically against the use of Bayes' theorem for assigning a probability to God's existence.

Dawkins takes exception to this usage not because he doubts the veracity of Bayes' theorem but because turning human experience into numbers is, he argues, an inherently subjective process. A Bayesian approach to evaluating the likelihood of God's existence involves enumerating possible outcomes (the presence of good, evil, religious revelations, and so on) and determining their probabilities assuming the existence or nonexistence of God. One must also express the prior belief of God's existence—the probability we would assign to the existence of God if we had no data from our experiences. Dawkins notes that these figures cannot be determined quantitatively, rendering Bayes' theorem useless in this enterprise. In applications for which data are available, however, Bayes' theorem lies at the heart of almost all statistical modeling and is a critical tool for thinking concretely about uncertainty.

How do certain hairs know to grow back when you trim them?

—C. Spain, New York City

George Cotsarelis, a University of Pennsylvania dermatologist and specialist in hair and scalp disorders, replies:

Clipping hairs on the surface of the skin does not affect their growth, because the hair above the surface is dead.

The hair visible on our bodies grows out of living hair follicles within the skin. All these follicles go through the so-called hair follicle cycle, which has three stages: growth, degeneration and rest—called anagen, catagen and telogen, respectively. During anagen, the rapid proliferation of cells at a follicle's base, or bulb, results in the constant production of a hair fiber that emerges from the skin surface. (Follicles produce hairs of different lengths by staying in anagen for varying periods.) At the end of anagen, these hair-producing cells begin to die, entering into the catagen stage. After dying back for a few weeks, the follicle rests for several weeks or even months in telogen. A new lower hair follicle then regenerates from stem cells in the telogen follicle, and anagen begins anew. The old hair fiber then falls out—often while you are brushing your hair—as a new strand replaces it. The hairs that appear to "know" to grow back after being trimmed just happen to be in the anagen stage when you cut them. SA

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

MATT COLLINS