

*Aves, Tab. 17.*



**MBL**

# Catalyst

Biological Discovery in Woods Hole

Founded in 1888 as the  
Marine Biological Laboratory

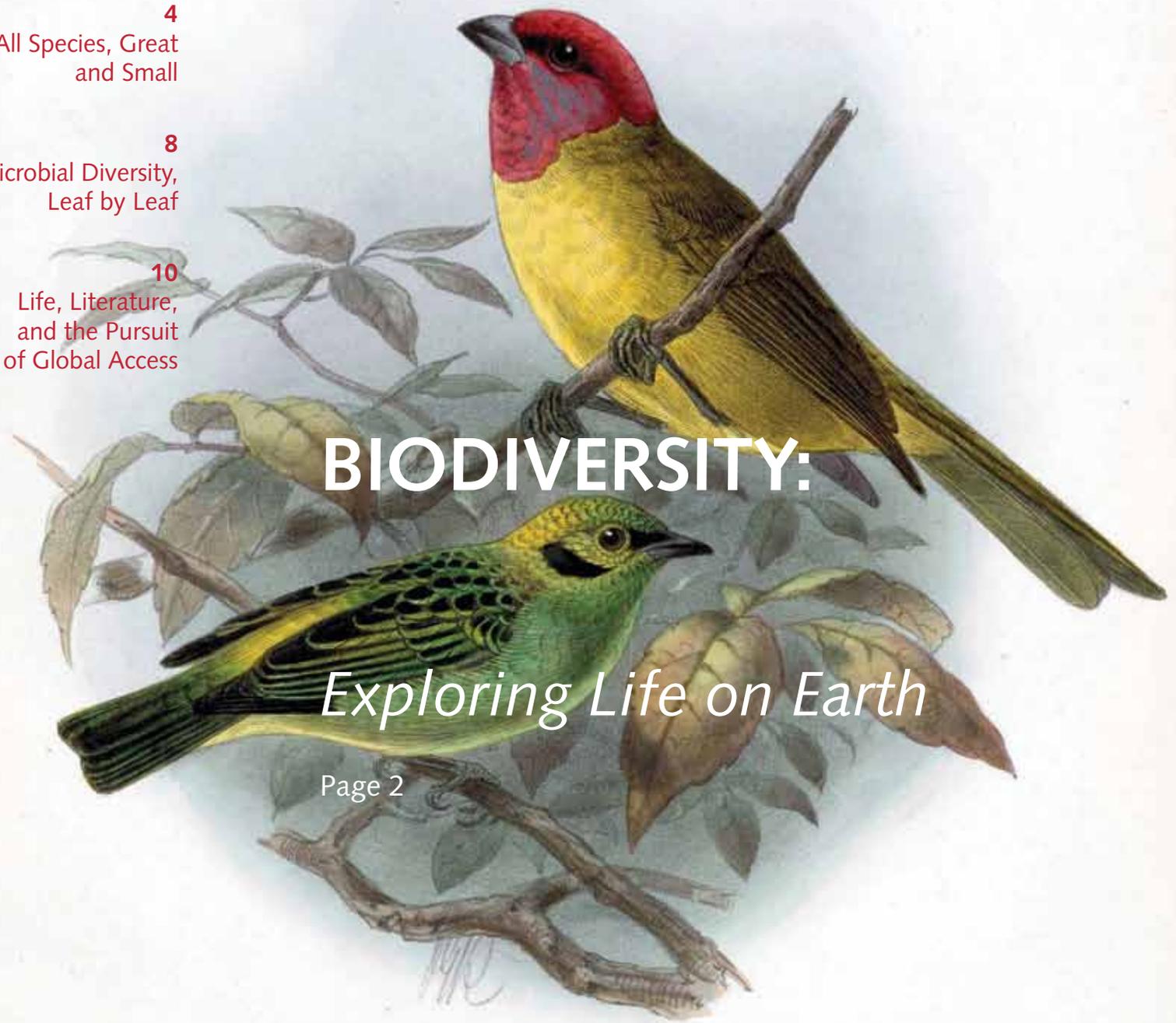
**SPRING 2012**  
VOLUME 7, NUMBER 1

**IN THIS ISSUE**

**4**  
All Species, Great  
and Small

**8**  
Microbial Diversity,  
Leaf by Leaf

**10**  
Life, Literature,  
and the Pursuit  
of Global Access



## BIODIVERSITY:

*Exploring Life on Earth*

Page 2

1 CALLISTE FLORIDA  
2 PYRANGA ERYTHROCEPHALA



FROM THE DIRECTOR

# MBL Catalyst

SPRING 2012 VOLUME 7, NUMBER 1

*MBL Catalyst* is published twice yearly by the Office of Communications at the Marine Biological Laboratory (MBL) in Woods Hole, Massachusetts. The MBL is dedicated to scientific discovery and improving the human condition through research and education in biology, biomedicine, and environmental science. Founded in 1888, the MBL is an independent, nonprofit corporation.

### Senior Advisors

President and Director:	Gary Boris
Chief Academic and Scientific Officer:	Joshua Hamilton
Director of External Relations:	Pamela Clapp Hinkle

### MBL Catalyst Staff

Editor-in-Chief:	Andrea Early
Managing Editor/Senior Writer:	Diana Kenney
Guest Science Editor:	Joseph S. Levine
Designer:	Beth Ready Liles
Contributors:	David Gallagher, Gina Hebert, Matthew Person, Sarah E. Stanley, Pamela Wilmot

**Images:** *Inside cover:* Birds (courtesy of Biodiversity Heritage Library. <http://www.biodiversitylibrary.org>); Gary Boris (E. Armstrong). *P. 1:* Birds (courtesy of Biodiversity Heritage Library. <http://www.biodiversitylibrary.org>); Butterflies (courtesy of Biodiversity Heritage Library. <http://www.biodiversitylibrary.org>); Chameleon (McBeul, Flickr, via eol.org); *Arabidopsis thaliana* (S. Simmons). *Pp. 2-3, clockwise from top:* Moon jelly (L. Viatour, Wikimedia Commons, via eol.org); *Psilocybe subaeruginosa* (M. Whitehead, Flickr, via eol.org); Florida gopher tortoise (C. O'Neal, Wikimedia Commons, via eol.org); Eastern cottontail (M. Reinbold, Flickr, via eol.org); Crocus (anemoneprojectors, Flickr, via eol.org); Jumping spider (T. Shahan, Wikimedia Commons, via eol.org); Emperor penguins (G. Grant, NSF, Wikimedia Commons, via eol.org); Blue discus (P. Farrelly, Wikimedia Commons, via eol.org); General Sherman Tree, Sequoia National Park (J. Bahn, Wikimedia Commons, via eol.org); Blue emperor (C. Fischer, Wikimedia Commons, via eol.org); Koala (Diliff, Wikimedia Commons, via eol.org). *Pp. 4-5, clockwise from left:* African elephant (Gorgol, Wikimedia Commons, via eol.org); Rhinoceros beetle (G. Gallice, Flickr, via eol.org); European green toad (M. Reinbold, Flickr, via eol.org); Biodiversity literature scanning, Library of Alexandria in Egypt (S. Casper). *Pp. 6-7, clockwise from top:* Flowers (courtesy of Biodiversity Heritage Library. <http://www.biodiversitylibrary.org>); Physiology course co-director Dyche Mullins (T. Kleindinst); Adult acorn worm (A. Pani); Tim Savas, a research assistant, measures CO<sub>2</sub> fluxes from a salt marsh (J. Tang); Marine microbe *Umbellosphaera* (ICoMM). *Pp. 8-9, clockwise L to R:* *Arabidopsis thaliana* (S. Simmons); Zoe Cardon, MBL; Stefan Sievert, Woods Hole Oceanographic Institution, and Anne Giblin, MBL (T. Kleindinst); Sheri Simmons (T. Kleindinst). *Pp. 10-11, L to R:* Flowers (courtesy of Biodiversity Heritage Library. <http://www.biodiversitylibrary.org>); Bird (courtesy of Biodiversity Heritage Library. <http://www.biodiversitylibrary.org>); Digitizing rare biodiversity books at Internet Archive Scanning Pod in Boston. *Pp. 12-13, L to R:* Nathan Wilson (N. Wilson); Violet-crowned wood nymph ("Cathy & Sam," Flickr, via eol.org); White shark (T. Goss, Wikimedia, via eol.org); Violet copper butterfly (F. Vassen, Flickr, via eol.org); Lion cub with mother in the Serengeti (D. Dennis, Wikimedia, via eol.org); Clownfish ("Jenny from Taipei," Wikimedia, via eol.org); Flowering rape (J. Hempel, Flickr, via eol.org). *P. 14, L to R:* John Valois (R. Howard); MBL SPINES students (T. Kleindinst); Cuttlefish (J. Vallino); Marine microbe (E. Haeckel). *P. 15:* A global map of ICoMM sample locations (ICoMM.mbl.edu and vamps.mbl.edu). *P. 16:* Joseph Levine (H. Chamberlain Nelson). *P. 17:* From *Voyage de la Corvette l'Astrolabe: Exécuté par Ordre du Roi, Pendant les Années 1826-1827-1828-1829, sous le Commandement de J. Dumont d'Urville* (Paris: J. Tastu and Ministère de la Marine, 1830-1834), Plate 19. *P. 18:* Cross-section of an uninjured lamprey spinal cord showing a diverse array of large and small axons, which mediate locomotor and sensory behaviors. Both populations of axons are capable of robust regeneration (P. Oliphant).

About the cover: Birds (Image courtesy of Biodiversity Heritage Library. <http://www.biodiversitylibrary.org>).

Online extras: For full image descriptions, supplemental materials, and other information related to this issue, visit:

[www.mbl.edu/catalyst](http://www.mbl.edu/catalyst)

Send correspondence to:  
MBL Communications Office  
7 MBL Street, Woods Hole, MA 02543  
508-289-7423, [comm@mbl.edu](mailto:comm@mbl.edu)

Dear Friends,

In February, a group of MBL trustees, overseers, and friends took a memorable "eco-expedition" by safari through East Africa. For most of us, the most exciting aspect was seeing the megafauna – giraffes, zebras, lions, warthogs, wildebeests, rhinoceroses—roaming wild in a pristine landscape. Except in tropical Asia, these large, charismatic animals aren't found anywhere else on the planet—not in the Americas, Europe, Australia, or New Zealand.

Why did they disappear? The reasons are debated, but there is good evidence that overkill by prehistoric humans caused major losses. Unfortunately, the near-extinction of these species 10,000 to 50,000 years ago is not the end of the story. It is generally agreed that the Earth is facing another biodiversity crisis in this century, with extinctions largely driven by destruction of habitat.

Species loss has come to the forefront of global concerns, as evidenced by such efforts as the United Nations' Decade on Biodiversity (2011-2020). But to effectively preserve biodiversity, we need to understand what we have to protect. The good news is that our ability to describe and understand existing biodiversity, and to discover new species, is rapidly advancing. At the MBL, we are committed to these efforts, and we are pioneering the technologies they require.

As a partner institution in the Encyclopedia of Life (EOL), the MBL is at the cutting edge of creating the cyberinfrastructure for vast amounts of biodiversity data (see p. 4 and p. 12). As a member of the EOL's executive committee, I can assure you that this is no small or simple effort. But gaining a comprehensive understanding of the millions of species with which we share this planet, and how they do or do not adapt to change, will provide invaluable insight into the structure of the living world and the principles of its sustainability.

Focusing the lens closer, the MBL is a leader in assessing microbial diversity, which is turning out to be immense. The International Census of Marine Microbes, co-directed by the MBL's Mitchell Sogin, ended its initial data collection phase in 2010, but the discoveries the data hold have only begun to roll out (see p. 6 and p. 15). This ambitious effort was successful due to the gene sequencing technologies that Sogin and his MBL colleagues introduced—technologies that revealed millions of new kinds of marine microbes, and the realization that rare species are surprisingly numerous.

I would like to thank Joe Levine for serving as guest science editor for this issue of *MBL Catalyst*. A scientist who is devoted to education, Joe annually brings secondary-school teachers to Costa Rican rainforests where they learn research methods to investigate tropical ecosystems and biodiversity. Through his involvement with the MBL's Board of Overseers, Joe is an enthusiastic member of the MBL's diverse, extended community.

Gary Boris  
President and Director

# Catalyst



FEATURES

**2 Biodiversity: Exploring Life on Earth**



Our planet's heritage of millions of species is constantly evolving and enormously complex. At the MBL, research to assess Earth's biodiversity is at the forefront of innovation in genomic and computational sciences.

**4 All Species, Great and Small**



There is something for everyone in the Encyclopedia of Life (EOL), which recently rolled out enhanced features for visitor interaction.



**8 Microbial Diversity, Leaf by Leaf**

How do functional communities of microbes form? MBL scientist Sheri Simmons is growing them on plants, from scratch, to find out.

**10 Life, Literature, and the Pursuit of Global Access**



The discovery and description of species over time has produced a vitally important, beautifully illustrated trove of literature. The Biodiversity Heritage Library is digitizing it for the world to see.

DEPARTMENTS

**6 NEWS & NOTES**

The latest findings from our laboratories and field sites.

**12 MBL MOMENT**

**Discovering Nature's Codes**

Nathan Wilson, director of the EOL's Biodiversity Informatics Group at the MBL, is a computer scientist with a lifelong passion for nature and biology.

**14 GIFTS & GRANTS**

**14 ACCOLADES**

**15 COOL TOOL**

**The Rise of VAMPS**

How do you visualize the data from an international census of marine microbes? MBL scientists found not one way, but many.

**16 SCIENTIST'S EYE VIEW**

**It's a Small World After All**

Joseph Levine on "rediscovering" the diversity of bacteria, and why these microbes are so important.

**17 MEMORABILIA**

**A Diversity of Fishes**

A page from the colorful history of scientific voyages to explore and document biodiversity.



# BIODIVERSITY: *Exploring Life*



“**B**iodiversity is a never-ending adventure; there is always something new to discover,” says Bob Corrigan of the Encyclopedia of Life, a website dedicated to amassing all knowledge about the Earth’s species and providing it free to the world (see p. 4).

He has a point. About two million species have been described since 1758, when Carolus Linnaeus published the tenth edition of *Systema Naturae*, which popularized his binomial system for naming plants and animals. Every year, about 18,000 new plants and animals are discovered in all kinds of environments on the planet. The Census of Marine Life, which involved several MBL scientists, encountered an “unanticipated riot of species” in the world’s oceans over a decade of exploration, and identified 6,000 species that are potentially new.

And yet we have barely scratched the surface of the Earth’s biodiversity. An estimated 7 to 10 million species are still to be discovered, and that’s not counting the tens or hundreds of millions of microbes as yet unknown. The work of exploring life on Earth has just begun.

At the MBL, scientists are deeply engaged in the global effort to assess our planet’s biodiversity. From the microbial communities that help marsh grass thrive, to a list of all the species you might find in your own backyard, the MBL is helping to build a baseline description of all life on Earth.

Why do we need this? The most obvious reason is so we can detect biodiversity changes caused by habitat conversion or climate change. For the first time in history, the rate of species extinction may now exceed that of species discovery, according

# Life on Earth



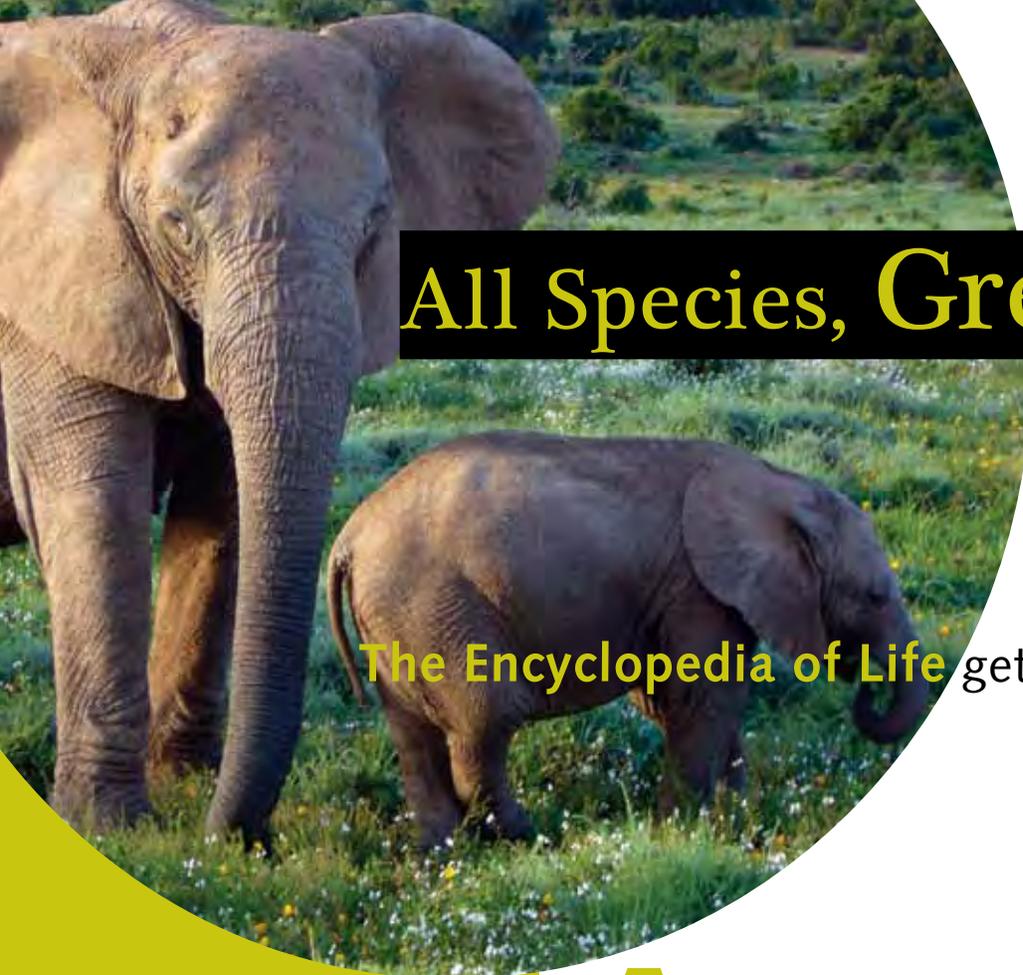
to Harvard biodiversity expert E.O. Wilson. We don't exactly know what we are losing, but it is surely thousands of species a year. Besides the negative impact that biodiversity loss has on ecosystem stability, with each extinction we lose the opportunity to discover a new material, a new molecule, or a



new design that could transform medicine, environmental engineering, or industry.

More importantly, all species are linked through our common evolutionary roots. The better we understand our co-habitants on Earth, the better we understand ourselves and the biosphere we share.





# All Species, Great and Small

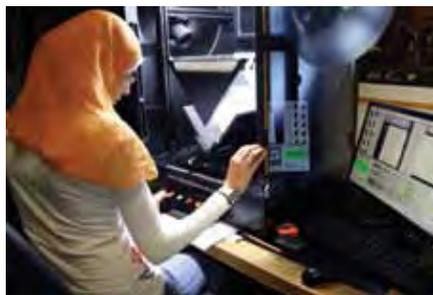
The Encyclopedia of Life gets up close and personal



## OPENING THE EOL

**THEN:** In his 2007 TED Prize speech, Harvard biologist E.O. Wilson wishes for a “key tool to inspire preservation of Earth’s biodiversity: the Encyclopedia of Life.” With funding from the MacArthur and Sloan Foundations, the EOL launches in early 2008 with the goal of creating “a webpage for every species.” At the MBL, the Biodiversity Informatics Group begins aggregating content on species from sources worldwide.

**NOW:** The EOL has vetted information on nearly 1 million species and is growing fast. Headquartered at the Smithsonian Institution in Washington, D.C., the EOL is supported by 13 global partner institutions and more than 200 content providers and 800 curators.



*Biodiversity literature being digitized at the Library of Alexandria in Egypt, one of the EOL’s global partners. Credit: Sally Casper*

**A** piglet squid that looks as funny as its name; a round-eyed lemur eating a carrot; a hedgehog ... just being a hedgehog. What do these animals have in common?

They are all photos in “Cute!”—a collection created by a visitor to the Encyclopedia of Life (EOL) (<http://eol.org>). Nathan Wilson, director of the EOL’s Biodiversity Informatics Group at the MBL, likes it because “different people have very different opinions of what is cute.”

Wilson himself has started an EOL collection that—while perhaps less cute—is of more interest to naturalists: It contains all the taxa (groups of organisms) he has personally observed on Cape Cod. Wilson is currently building a community of people on EOL who also like to observe nature, and who are interested in adding species to his “Cape Cod” collection.

“Collections” and “communities” are exciting new features in EOL version 2, which rolled out in September 2011. “Prior to version 2, even though EOL had lots of content on species, it offered no way for users to describe the relationships between them. Collections let people organize life the way they want to,” says Bob Corrigan, product manager for the EOL.

These features contribute to the goal of making the EOL more “engaging, accessible, and personal,” Corrigan says. The first version of the EOL site, which was built by an MBL team led by David Patterson, “was the equivalent of a house’s foundation. It gave us a powerful, solid infrastructure for aggregating information on species. The next challenge was to make it easy to use and extend that information,” Corrigan says.

The new features took off immediately: EOL users created more than 1,000 collections within the first month. "Someone made a definitive list of state flowers. Someone else made a list of endangered and threatened species of Costa Rica, which never before existed on the Web," Corrigan says. Now people want to share their collections via Twitter and Facebook, another new EOL capability.

Through a collaboration with developers at the Library of Alexandria in Egypt, version 2 also rendered the EOL translatable into any language. To achieve this, Wilson's group audited the EOL's open-source code, located all the text strings sprinkled around in the code, and centralized them into one file. The Library of Alexandria team could then translate the interface into Arabic. Next, Wilson's team had to make sure the EOL actually looked good in different languages. This meant redesigning the species pages from the ground up to accommodate languages that are read right to left, such as Arabic, or translations that are twice as long as the original English. Wilson's team has since integrated Spanish content into the site, which is proving to be very popular, and several other languages will be coming online soon.

While refining version 2, Wilson and Corrigan are presently focused on strategies to bring more users to the site, including improving its visibility in Google and other search engines. And they are continually thinking about "what's next" in order to realize the great potential of the EOL.

"The question most commonly asked by EOL users is, 'What is in my backyard?'" Wilson says. "Unfortunately, that's a pretty hard question to answer. But it would be very useful if EOL could give them a list of the 100 butterflies or 100 birds or 100 insects they are most likely to see. People are hungry for that functionality."

To get there, the EOL needs to incorporate maps of species distributions, an ongoing and challenging goal. The EOL is partnering with other groups working in this area, including the Global Biodiversity Information Facility and the Atlas of Living Australia. Distribution maps will also bring the EOL much closer to one of its overarching goals: to assist in the preservation of biodiversity.

And then there is the desire, voiced since the EOL was first envisioned, for the site to aid in the discovery of new species. That's not an easy problem to address, either.

"Let's say you take a picture of something and you want to know if it's a new species," Wilson says. "There is a precursor to that: You take a picture of something and you want to know what it is." For the EOL to help you, its content needs to be "computable," so if you search on a descriptive term such as "blue bird," the EOL can pull up all the blue birds referenced in its gigantic database.

But right now, the EOL has thousands of photos that someone may have briefly described, but the computer can't tell what is pictured. Similarly, the computer can locate organisms by name, but it "has no insight into what the text says," Wilson explains. However, "One of the interesting things about the (species) descriptions that come into the EOL is they are more parsable, or computable, than most data, if you apply the right technologies. The MBL Center for Library and Informatics has begun developing those technologies, and that is an important direction we want to go in," he says. "Being able to understand the historical literature, from a computer standpoint, is a very significant biodiversity informatics problem right now, and it will require a combination of computer automation and effective human review."

According to Corrigan, the Biodiversity Informatics Group is more than up to the heady challenges of reaching the EOL's goal of "Global Access to Knowledge About Life on Earth." "It really is a wonderful combination of deep thinkers and focused doers," he says. "They present cutting-edge bioinformatics best practices and terrific solutions. With the help of the team at MBL, we are building the EOL for the millions, and we are building it to last." • —DK



## Census Yields World's First Comparison of Microbial Diversity in Global Ocean Ecosystems

In 2003, microbiologists from the U.S. and Europe began an ambitious project to take inventory of all the different species of marine microbes, including those yet to be discovered. Dubbed the International Census of Marine Microbes (ICoMM), the project's leaders included co-principal investigators Mitchell Sogin of the MBL and Jan de Leeuw of the Royal Netherlands Academy of Arts and Sciences, along with MBL scientist Linda Amaral Zettler, who served as project manager. This international effort led to many unanticipated discoveries. Using technologies developed by Sogin to identify different types of microbes, in 2006 the scientists uncovered evidence of more than 500,000 kinds of marine microbes, many of which correspond to very low-abundance taxa that comprise the rare biosphere. Recently, ICoMM yielded a new comparison of microbial

diversity in oceanic ecosystems around the globe. This study analyzed 9.6 million microbial genetic sequences from 509 different marine samples from globally distributed locations. Collection sites ranged from 2,700 meter-deep hydrothermal vents in the Pacific, to waters devoid of oxygen in the Caribbean, to the open waters of the Arctic and Antarctic Oceans. The study found that microbial communities on the seafloor differ significantly from those in the water above, despite a constant "snow" of sinking microbes. On the seafloor, there is evidence for community variation based on geographical differences. (In the water column, unpublished data analyzed by Sogin and Amaral Zettler suggest that microbial communities may vary according to latitude.) (*PLoS ONE* 6: e24570, 2011) •



## Comprehensive Mission to Document Earth's Species is Planned

A global mission to describe 10 million species in less than 50 years is achievable and necessary to sustain Earth's biodiversity, concludes an interdisciplinary group that includes Gary Borisy, president and director of the MBL, and Holly Miller, director of the MBL/WHOI Library and associate director of the MBL Center for Library and Informatics. With about 30 percent of Earth's species expected to become extinct this century, "the world is in an extinction crisis," and yet "we have all the intellectual and technological capability for rapid biodiversity discovery," says co-author Johannes Vogel of the Museum of Natural History in Berlin, Germany. Following a 2010 workshop organized by Arizona State University's Institute for Species Exploration, the team published a detailed plan to begin a mission to map the Earth's biosphere. Strategic investments in cyberinfrastructure will be needed to link species identification data, the biodiversity literature, and type specimens in natural history collections worldwide. The project will also require a sizeable workforce of taxon experts to name and describe new species. This ambitious project offers many benefits, the team concluded, including detection of early responses of flora and fauna to climate change; access to evolutionary designs for solutions to practical problems; and a more comprehensive understanding of the origin and evolution of life. (*Systematics and Biodiversity* 10: 1-20, 2012). •



## Walk This Way: Scientists and MBL Physiology Course Students Describe How a Motor Protein “Steps Out”



MBL scientists and students have discovered the unique “drunken sailor” gait of dynein, a protein that is critical for the function of every cell in the body and whose malfunction has been associated with neurodegenerative disorders such as Lou Gehrig’s disease and Parkinson’s disease. Samara Reck-Peterson of Harvard Medical School led the research, which was partially conducted in the 2007 MBL Physiology

Course by students Elizabeth Villa of the Max Planck Institute of Biochemistry and David Wu of UCLA’s Geffen School of Medicine. Dynein is one of three types of “motor proteins”: tiny molecular machines that constantly shuttle materials needed to keep cells alive, allow cells to move and divide, and enable cells to talk to their neighbors. All three types of motor protein (dynein, myosin, and kinesin) are “two-footed” and use the energy from breaking chemical bonds to generate movement. “The myosin and kinesin motors work by walking more or less like we do: one foot in front of the other in a straight line,” says Reck-Peterson. “We have discovered that the third motor model, dynein, appears to be different. Its two feet are at times uncoordinated and often veer from side-to-side (think drunken sailor). This mode of walking makes the dynein motor unique and may allow it to navigate obstacles while performing its transport functions in cells. Interestingly, our data also suggest that the dynein motor becomes more coordinated when it is hauling something large, implying that the motor can become more efficient when necessary.” (*Nat. Struct. Mol. Biol.* 19: 193-200, 2012). •



## MBL and Partners to Study Climate Change in Falmouth Salt Marshes

Ecosystems Center assistant scientist Jim Tang and his partners have been awarded a \$1.3 million grant from the National Estuarine Research Reserve System to examine carbon and nitrogen cycles in salt marshes in Waquoit Bay, Falmouth, Mass., and to determine how those cycles are impacted by, and feed back to, climate change. Tang and colleagues from the Waquoit Bay National Estuarine Research Reserve, U.S. Geological Survey, and University of Rhode Island

will quantify how much greenhouse gas is stored in and emitted from coastal wetlands, and how the presence of nitrogen changes this balance. Greenhouse gases contribute to global warming by trapping heat in the atmosphere. While it is well known that forests store large amounts of carbon, thus reducing global warming, there is new focus on equivalent stores of so-called “Blue Carbon” in coastal ecosystems. Preliminary work has shown that when nitrogen is present (as it is in most coastal areas) salt marshes actually become sources of greenhouse gases, rather than sinks. Tang will use novel laser-based technology to develop a new system for measuring greenhouse gas emissions from salt marshes directly in the field to dramatically improve the accuracy and frequency of greenhouse gas measurement. In addition to producing “pure” science, the project will link researchers with end-users, who will apply the science to coastal management. •

## In a Brainless Marine Worm, Whitman Investigators Find Parts of the Developmental “Scaffold” for the Vertebrate Brain

The evolutionary origin of the vertebrate brain is a mystery, because nothing anatomically like it has been found in invertebrates. Yet a clue to its origins was recently discovered by Whitman Investigators Ariel Pani of University of Chicago, Chris Lowe of Stanford University, and their colleagues. The team discovered some of the genetic processes that regulate vertebrate brain development in, surprisingly, the acorn worm (*Saccoglossus kowalevskii*), a brainless, burrowing marine invertebrate whose lineage diverged from vertebrates more than 500 million years ago. The scientists found three “signaling centers” in the acorn worm embryo that are also present in the vertebrate embryo. In



vertebrates, Pani says, these signaling centers are major components of a gene expression program “that sets up the

foundation of how the brain develops.” In acorn worms, they direct the formation of the embryonic body plan. These signaling centers are not present in sea squirts and lancelets, vertebrates’ closest living evolutionary relatives, so it was unexpected to find them in the acorn worm, a more distantly related hemichordate. “The lancelet and sea squirts will still be the first animals we will look at if we want to understand vertebrate evolution,” says Pani. “But if we find differences, we now know it is important to look at anatomically divergent animals, where you wouldn’t have previously expected to find compelling similarities. I think this principle applies broadly to understanding animal evolution.” (*Nature* 483: 289-294, 2012). •

# MICROBIAL DIVERSITY

## Leaf by Leaf



**T**hough invisible to the naked eye, microbes are everywhere. They thrive in familiar environments like farm soils and inside the human body, but also in the harsh conditions of geysers and glaciers. Microbes are essential drivers of all ecosystems, affecting processes from decomposition to vegetation growth to climate change. Yet little is known about how microbial communities form, or the ecological rules that govern their maintenance.



Sheri Simmons with flowering *Arabidopsis*

“When we go out in nature and take a microbe sample, we’re essentially taking a snapshot of that community at a particular point in space and time,” says Sheri Simmons, a scientist at the Josephine Bay Paul Center. “We don’t necessarily know the processes that shaped that community.”

Taking many “snapshots” over time gives scientists a good working model of a particular microbial system. But Simmons is taking a novel approach. She and her colleagues are recreating natural microbial communities in the lab and watching them develop from scratch. They hope this will shed light not only on how communities form, but how they function.

Simmons is cultivating hundreds of tiny, flowering *Arabidopsis thaliana* plants at the MBL’s Research Greenhouse in West Falmouth. Seemingly bare, each leaf of each plant is actually home to its own microbial community.

“When an *Arabidopsis* leaf first emerges, it is essentially sterile,” Simmons says. “It doesn’t have any microbes on the surface, so it becomes colonized by microbes settling from the air.”

Like living petri dishes, the leaves soon become home to bacteria, archaea (another microbial domain of life), and fungi. Researchers in Simmons’s lab developed methods to sample resident leaf microbes and sequence their DNA. The sequences are matched to genes, revealing who is living on each leaf, and what they’re up to.

“*Arabidopsis* has many leaves that are quite similar to each other,” Simmons says. “When we sample individual leaves, we are looking at microbial communities that started from the same initial point, but they might have developed along different trajectories.”

Simmons aims to figure out what determines the trajectory of a microbial community. The project builds on recent research in which microbiologists have started to apply the ecological rules for ecosystems to microbial communities.

## Do the first colonizers on a leaf end up determining the future of the community? Or does each leaf eventually end up with a similar community, no matter who got there first?

Of particular interest to Simmons are two macro-ecology models that describe how a community is shaped.

The “neutral” model holds that the species composition of an ecosystem is largely determined by which species happen, just by chance, to colonize the area first. The “niche” model proposes that the habitat itself determines the set of species that can grow there, so that no matter who gets there first, the community will eventually conform to those that fit best.

Simmons hopes to determine which model—niche or neutral—plays a more dominant role in *Arabidopsis* leaf microbial community formation.

Do the first colonizers on a leaf end up determining the future of the community? Or does each leaf eventually end up with a similar community, no matter who got there first?

“What we’ll probably end up finding is that there’s always a mix of these niche and neutral forces,” Simmons says. “It’s not one versus the other.”

And, importantly, saying that different leaves host similar microbial communities doesn’t necessarily mean they accommodate exactly the same species. Different leaves could have completely different microbes that carry out identical functions, such as eating sugars and methanol produced by the plant.

“Our hypothesis is that it’s relatively random which species settle and grow on a leaf surface,” Simmons says. “But the suite of functions that is optimal for the leaf surface is not random.”

In other words, she hypothesizes that the neutral model holds true for species diversity (who is there), while the niche model guides functional diversity (what they do).

The project’s DNA sequencing methods reflect this distinction between species diversity and functional diversity. Using the Josephine Bay Paul Center’s state-of-the-art sequencing facilities, the scientists will

sequence the 16S gene, a gene often used to identify different species of bacteria and archaea. They will also sequence the entire set of microbial DNA to see which cellular functions are taking place on the leaves.

Members of Simmons’ lab, including former and current postdoctoral scientists Emelia DeForce and Loïs Maignien and research assistant Meghan Chafee, have already completed sampling for their first experiment. They planted *Arabidopsis* and took microbe samples for three consecutive months. They also sampled the greenhouse air to see how the airborne community influenced the development of leaf communities. Data analysis is ongoing.

Next, they’ll seed leaves with different kinds of microbes to compare the development of communities that have different first colonizers. They will also test how a plant’s environment affects the microbes, how *Arabidopsis* itself influences community development, and whether microbes confer pathogen resistance to *Arabidopsis*, which could have useful agricultural implications.

“Looking forward, we want to go beyond treating *Arabidopsis* as a passive substrate,” Simmons says. “Ultimately, I think there will turn out to be a lot of important direct interactions between microbes and plant leaves.” • —SES



## The Root of the Matter

While the Simmons lab explores microbial diversity on leaves, Ecosystems Center scientist Zoe Cardon will literally dig a little deeper.

*Spartina alterniflora* is a dominant grass in New England salt marshes. The sediment around its roots houses microbial communities that can help “clean up” excess nitrogen from fertilizer and sewage that often flows into marshes and threatens coastal life. But, as with *Arabidopsis* leaf communities, little is known about which sediment microbes carry out which functions.

Working with scientists Anne Giblin of MBL and Stefan Sievert of Woods Hole Oceanographic Institution, Cardon is bringing the salt marsh into the lab. They will grow *Spartina* in sediment at the MBL’s Research Greenhouse, controlling temperature, light levels, and even tidal seawater flow. They are particularly interested in bacteria that use hydrogen sulfide (the gas that smells like rotten eggs!) as fuel while processing pollutant nitrogen. DNA sequencing and biogeochemical tests will reveal both microbial species diversity and functional diversity in the sediments.

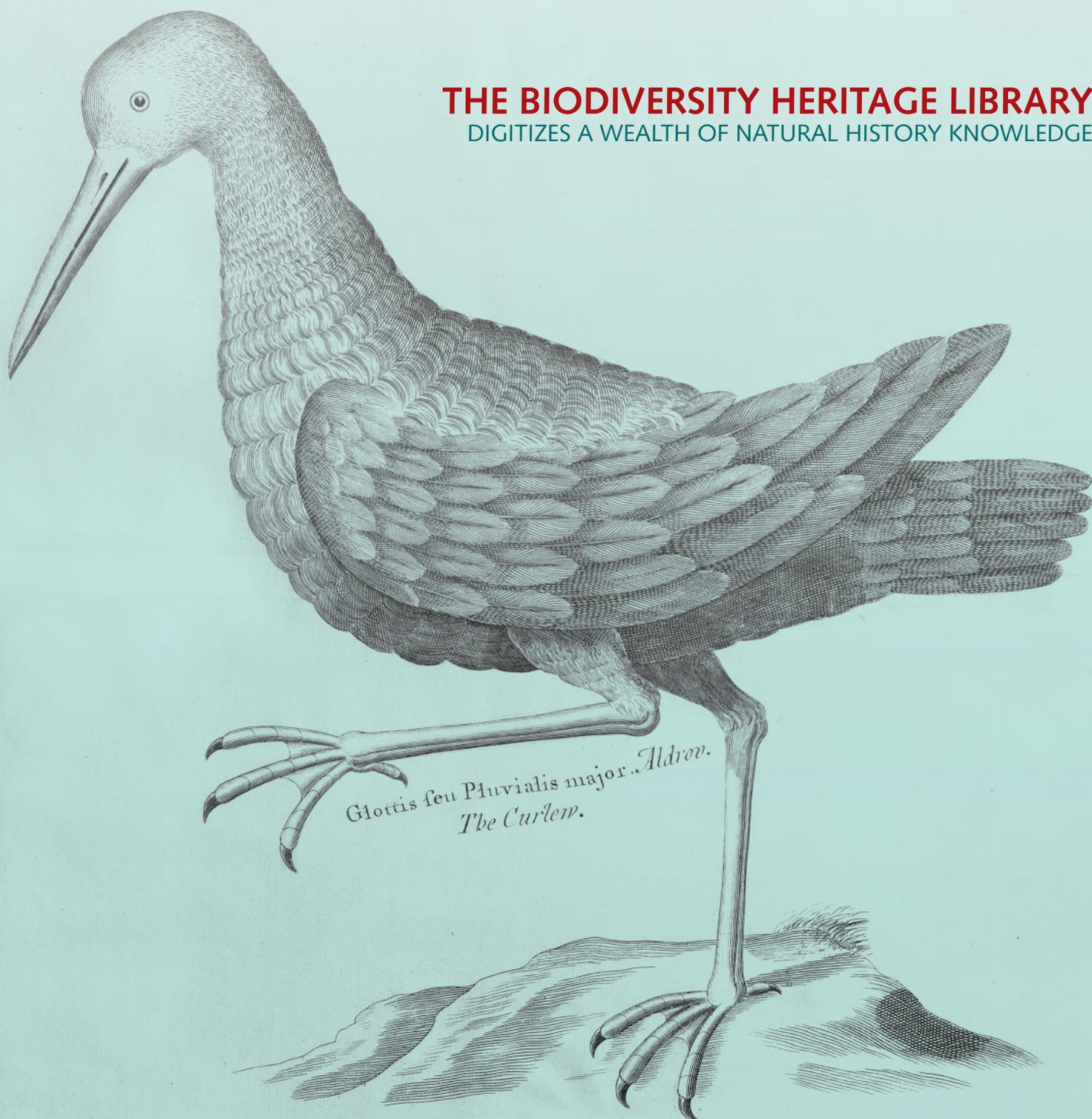
A better understanding of the microbes’ activities could aid in restoration of salt marshes, which are threatened by coastal development and rising sea level.

“Microbes are critical reservoirs of biological know-how, capable of transforming pollutants and providing new sources of fuel,” Cardon says. “Their activities are at the heart of the ecosystem services that sustain humanity.” • —SES

# Life, Literature, and the Pursuit of Global Access

Tab. 268.

**THE BIODIVERSITY HERITAGE LIBRARY**  
DIGITIZES A WEALTH OF NATURAL HISTORY KNOWLEDGE



There are millions of unique species on Earth, 1.9 million of which have been scientifically identified. What scientists have discovered about these species over centuries fills many millions of pages of literature. Now, one by one, those pages are being scanned and uploaded onto the Biodiversity Heritage Library (BHL) website.

“For many years, there has been a demand from researchers to digitize biodiversity literature,” says Tom Garnett, a Smithsonian Institution librarian and program director of the BHL. “You can’t do new research without having old literature available.”

Much of that old literature is now on the BHL website. The searchable, digital collection ranges from 600-year-old, handwritten Latin texts, to lavishly illustrated books on botany from the 1800s, to papers published by today’s scientists. Almost 40 million pages have been uploaded into the BHL since it was founded in 2005, and, despite legal and practical obstacles, new material is continuously added.

The MBLWHOI Library, one of the original 10 libraries involved in the BHL, has contributed scans of thousands of books.

“We were one of the first libraries that got this up and running,” says Cathy Norton, MBLWHOI Library Scholar and chairman of the BHL. “We started scanning immediately. I think at one point we had scanned half of the things that were in the BHL.”



The BHL is now global, with more than 40 libraries in the United States, Europe, South America, Egypt, Australia and China. Each library contributes a unique collection of biodiversity texts.

To facilitate the process, the BHL cooperates with the Internet Archive, a project to build a comprehensive online library. As part of this effort, the Internet Archive has established several scanning centers in the United States.

“We would load 500 books on trucks and send them to the Boston Public Library,” Norton says. “I like to say they went to the book spa, where they got cleaned and scanned by Internet Archive employees. Then they’d come back and we’d send up another 500.”

Norton says that some archivists were horrified that rare volumes were removed from storage and handled by others. But, she says, getting these unique books in the public eye is “the best thing in the world.”

In addition to its book collection, the MBL contributed the search engine used by the BHL. Called uBio, it allows a user to access all records of a particular species, regardless of what that species is called in the literature. For example, searching “*Monodon monoceros*” will pull up information on *M. monoceros* as well as any literature mentioning its common names—narwhal, narwhale, etc.

The BHL is also intertwined with another major project at the MBL—the Encyclopedia of Life (EOL), which strives to create a webpage for each unique species. Many species have only been described once, and those descriptions can often be found in the BHL’s historical texts.

“The BHL scans the underpinning literature of EOL,” Norton says. “We link to them, and they link to us.”

Thanks to the MBL and other BHL members, scientists worldwide can now access an ever-growing suite of biodiversity information. And the nature of biodiversity research makes the BHL particularly valuable.

“Identification of organisms requires access to prior descriptions,” Garnett says. “Imagine an entomologist in a tropical rainforest who collects an insect she can’t identify. The only information about it may be in one book in a library thousands of miles away.”

If that chapter is scanned into the BHL, however, time and research funding can be spent on science, instead of a cross-country flight.

And it’s not just scientists who use the BHL. Artists mine texts for species to incorporate into their works. Science historians use the texts to better understand 200-year-old scientific discussions. Educators gather material for classrooms. One user even explored the biodiversity of a particular time period for a historical re-enactment project.

“We get loads of unsolicited testimonials from scientists, but many other people are using this literature,” Garnett says. “It’s enormously gratifying.”

“I like to say they went to the book spa.”

— Cathy Norton

Despite its ongoing success, the BHL still faces major hurdles. Copyright law prohibits the BHL from uploading many books and articles published since 1923. Permission is secured for some texts on a case-by-case basis, but many publishers forbid the BHL to provide material for free.

“There are big chunks of important biodiversity literature that are off limits because they were published in commercial journals that don’t want to be open access,” Garnett says. “Over the years, I think this will get resolved through legislation and court cases.”

The BHL also shares a major problem with all digitization projects. The human brain can easily recognize a table of contents, an author’s name, and images in a scan. But computer programs are not yet sophisticated enough to reliably perform such tasks for older and technical literature, which impedes search efficiency. The BHL works with software developers tackling the problem.

Finally, funding is an enduring problem for the BHL. “All this biodiversity material is free,” Garnett says. “But keeping some level of funding going so we can expand is something we’re always dealing with.”

Garnett is retiring this year, handing the torch to Martin Kalfatovic, also from the Smithsonian. He leaves the BHL headed in promising directions. The BHL already collaborates with similar projects in Asia and Australia, and has been in talks with institutions in Africa.

“Many countries, especially in South America and Africa, don’t have access to information about biodiversity in their regions,” Norton says. “Now that it’s online, it’s available to everybody.”

—SES

# Discovering Nature's Codes

## MBL MOMENT

with ...

### Nathan Wilson

Director, Biodiversity Informatics Group, Encyclopedia of Life  
Director, MBL Center for Library and Informatics



Nathan Wilson is director of the MBL Center for Library and Informatics, where he oversees 12 informatics initiatives, including the EOL Biodiversity Informatics Group. While diverse, the Center's projects have the common goals of improved data management, data-driven discovery, and development of innovative visualization and analysis tools. Prior to coming to the MBL in 2010, Wilson worked for DreamWorks Animation for 12 years, where he focused on collaboration tools and open-source software. He was also active in the amateur mycology arena, serving as president of the Los Angeles Mycological Society. He has a master's degree in computer science from the University of California, Santa Cruz, and a master's degree in experimental psychology from the University of Pennsylvania.

### MBL: Which came first: your interest in nature and biology, or in computers?

**NW:** I've been a naturalist since I could run out into the fields and collect frogs and snakes and butterflies. I got interested in mushrooms around age 10 and started reading books on mycology. In the seventh grade I became interested in computers, but I was sure I was going to be a biologist until my second year of college. Then I discovered that biology was primarily about chemistry at that time (early 1980s) and chemistry was my least favorite science. I ended up getting a master's degree in computer science, and my thesis was on identification of biological species using computers, with fungi as an example case.

*"I can combine my passion for nature and biology with my passion for computers and technology. It's the perfect job," says Nathan Wilson of his role as director of the Biodiversity Informatics Group for the Encyclopedia of Life (EOL). Wilson leads the MBL technical team that develops the software and hardware infrastructure for the EOL. He is also an avid field naturalist specializing in fungi (mushrooms). Wilson enjoys innovating with international collaborators in the growing field of biodiversity informatics, while at the same time delivering a high-quality product—the EOL website—to scientists and the general public.*



Professionally, I focused on computers, working at SRI International, a couple of Silicon Valley startups, Apple Computer, Digital Domain, and most recently at DreamWorks Animation. While at DreamWorks, in my spare time I put together Mushroom Observer, which is a web site for people who study fungi. My goal was to organize

the information on fungi in a collaborative way so people could work together to create a common picture of mycology. It's really grown; we now have more than 1,500 professional and amateur mycologists contributing from all over the world.

**MBL: What do scientists most value about the EOL, and what does the public most value?**

**NW:** Interestingly, what scientists say is most compelling is the EOL provides a vehicle for outreach. Many federal grants require that the scientist demonstrate the broader impacts of their research, and the EOL provides a public face and support for that. Scientists can just contribute their (species) information to the EOL and then it is part of a public resource that thousands of people visit on a regular basis.

The general public values the breadth of information in the EOL. As far as we know, we have data on more species than any other online system. Also, in contrast to a site like Wikipedia, they like that we clearly indicate the status and origin of the information. They know whether it has been vetted by scientists, either before coming to the EOL or through the EOL curatorial processes. Wikipedia has a similar review process, but it's less clear what

**MBL: What is the main challenge in the field of biodiversity informatics right now?**

**NW:** The main focus for biodiversity informatics is still the process of understanding how to bring all this species information together and have it make sense. There is a vast amount of data out there, only some of which we have tapped into with EOL. Several other groups besides the EOL are working on organizing the planet's species, and we are all working together to present that information. Different groups present different organizations of the tree of life, and EOL tries to remain agnostic on that by supporting multiple trees, multiple ways of organizing organisms. But how we manage those different organizations is a challenge.

**MBL: What does this mean for the EOL?**

**NW:** The challenge we have is our content can be confusing to users, because it's an assemblage of content that comes from many places. How do we present the information in a way that makes sense to most people, but also allows someone to drill down in a particular area? That is a key problem, and it is related to the more general problem of generating a more authoritative taxonomic classification for species.

But we are working on giving the curators better tools for reviewing and improving content. So a curator might decide to present a good, brief summary of the organism first or a certain image first, maybe because it really characterizes that branch of the tree of life, even if it isn't the most popular image.

**MBL: As a side project, you are developing an alternative framework for naming organisms that groups them by observable traits, rather than by evolutionary relationships (taxonomy). Why is there a need for this?**

**NW:** Taxonomists are pursuing the question, "What are the self-propagating populations of organisms that we call species?" That is a really important question for a lot of reasons. For one, if I want to understand the preservation of biodiversity, and I don't know what is in a given population, I may kill off some group without even knowing that it was distinct. But there are also many important questions and socioeconomic issues related to organisms for which species classifications aren't relevant. For instance, if zebra mussels are imported to my region, what will be the impact on boat channels? I may not care that there are three species of zebra mussels, because there isn't any functional difference between them, from my perspective. I may only care what happens when zebra mussels are in my boat channels.



the status of the information is. Also, in the EOL, we present the information in the scientists' own words in a way that it can't be changed. Unlike Wikipedia, we preserve the ability for a user to know, "What did this scientist actually say about this species?"

Ultimately, what we need to do is enable the EOL curators to make more sense of the species pages. Right now, the EOL users rate content from 1 to 5, and the higher-rated content is presented on the page first. The content can also propagate up the tree and compete to be shown first, so the top image in the kingdom Animalia, for instance, will almost certainly have all ratings of 5.

The system I'm working on is called the Semantic Vernacular Naming System. It is complementary to the traditional, scientific naming systems and shares many of the same features. It also uses some of latest computer science approaches to describing entities to create a practical, reliable way to describe and name meaningful groups of organisms.

• —DK

## GIFTS & GRANTS



An **anonymous donor** pledged \$1.2 million to support the John J. Valois Marine Resources Center Endowment and the Stanley W. Watson Education Fund.

**The National Institutes of Health** awarded \$1,039,523 for the Summer Program in Neuroscience, Ethics & Survival (SPINES) course.

**Dart Neuroscience LLC** awarded \$1 million to provide lab space for visiting investigators conducting research at the MBL in the area of learning and memory.

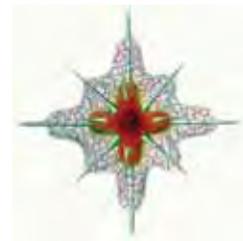
**The Ellison Medical Foundation** pledged \$798,712 to support the Molecular Biology of Aging course.



**The National Science Foundation** awarded \$763,954 for a project titled "Collaborative Research: Development of Compact, Waterproof Hyper-Spectral Imaging and Multi-Channel Spectro Polarimeter Instrumentation

for Marine Biological Studies of Coloration." Roger Hanlon is the principal investigator.

**The National Science Foundation** awarded \$506,194 for a project titled "Collaborative Research: Microbial Interactions with Marine Plastic Debris: Diversity, Function and Fate." Linda Amaral Zettler is the principal investigator. •



## ACCOLADES

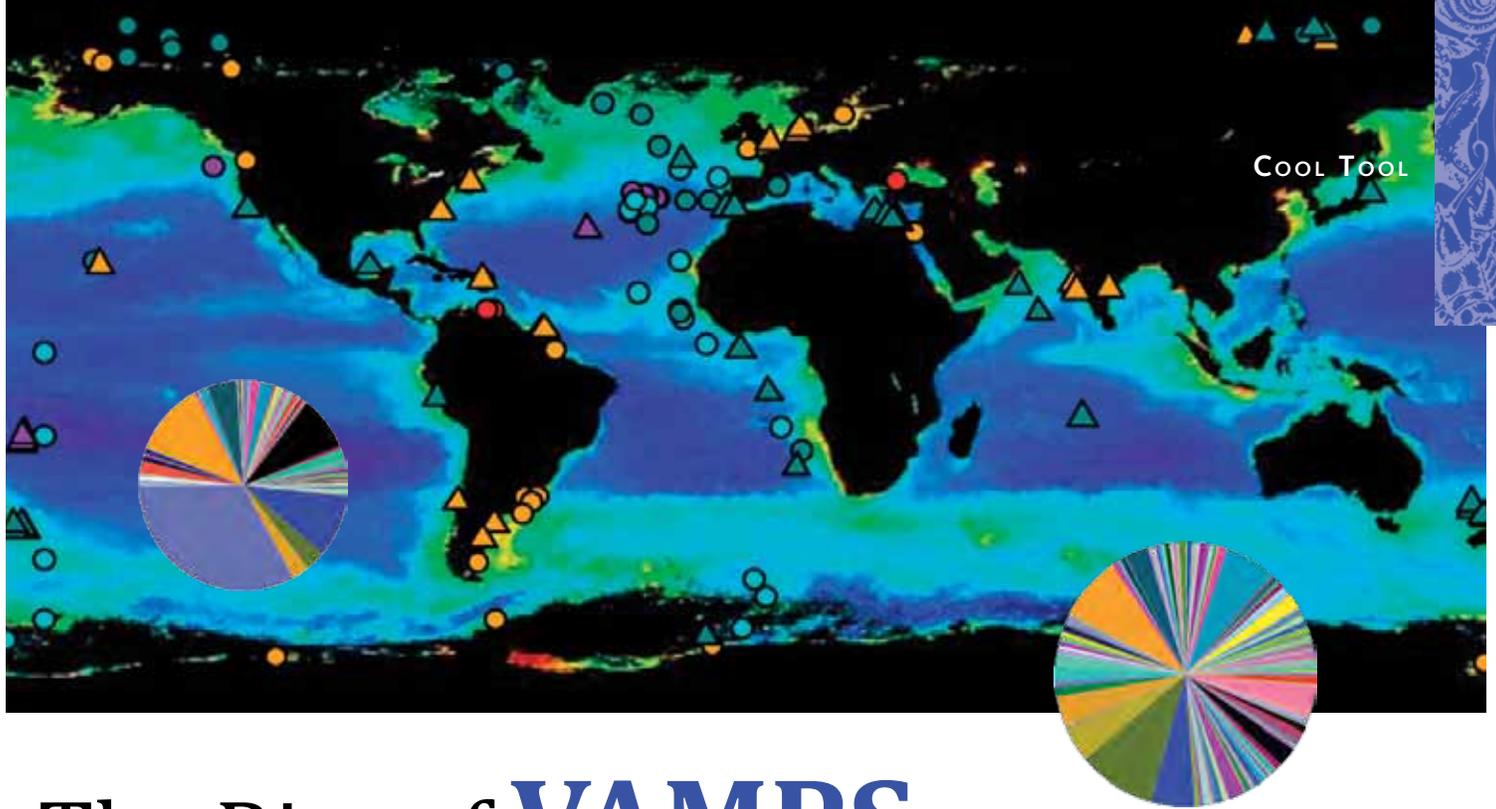
MBL Ecosystems Center director **Hugh Ducklow** was appointed to the U.S. Antarctic Program Blue Ribbon Panel. The 12-member panel, formed at the request of the White House Office of Science and Technology Policy and the National Science Foundation, is tasked with examining the status and capabilities of the U.S. Antarctic Program. Ducklow directs the Long Term Ecological Research Project at Palmer Station, on the west Antarctic Peninsula, a collaborative endeavor to study and understand the Antarctic marine ecosystem.

**Ralph Brinster**, an alumnus of the MBL FERGAP course (now Frontiers in Reproduction) was awarded the 2011 Presidential Medal of Science. Brinster, a professor at the University of Pennsylvania School of Veterinary Medicine, was honored for his research on the manipulation of the mammalian germline.

The **Census of Marine Life Scientific Steering Committee** was awarded Japan's International Cosmos Prize. **Researchers from the MBL's Bay Paul Center** were among the scientists contributing to the Census through their leadership of the International Census of Marine Microbes, a research project of the larger Census that focused on the biodiversity of microscopic life forms in the world's oceans.

MBL trustees recently elected six new Board members. **Dr. Ethan A. Lerner** of Chestnut Hill, MA was elected to the Class of 2013. **Dr. Rita R. Colwell** of Bethesda, MD; **Dr. Robert Haselkorn** of Chicago, IL; **Mr. William I. Huyett** of Concord, MA; **Dr. Walter G. Massey** of Chicago, IL; and **Dr. Dyann R. Wirth** of Boston, MA were each elected to the Class of 2015.

**Dr. E. Peter Greenberg**, former MBL Microbial Diversity course director, faculty, and lecturer, received the 2012 D.C. White Research and Mentoring Award from the American Society for Microbiology. Greenberg, a professor in the Department of Microbiology at the University of Washington School of Medicine, was honored for his interdisciplinary scientific approach and for being a dedicated and inspiring mentor. •



# The Rise of VAMPS

*Above: Each slice of the pie represents the abundance of a particular type of bacteria. VAMPS users can explore the data in each pie slice, such as individual DNA sequences.*

*Background: A global map of ICoMM sample locations. The origin of the samples is indicated by triangles (benthic or seafloor realm) and circles (pelagic or ocean water realm) while ecosystem types are indicated by the color orange (coastal), light blue (open-ocean surface waters), dark blue (deep sea), red (anoxic), and purple (hydrothermal vents). Further visualization of sample distributions are available at <http://vamps.mbl.edu/mapper/index.php>.*

**W**hen scientists began collecting tens of thousands of genetic sequences for the International Census of Marine Microbes (ICoMM), the project's co-director, MBL microbiologist Mitchell Sogin, knew they would need better software to handle all that data. So he oversaw the development of a free online tool called Visualization and Analysis of Microbial Population Structures (VAMPS).

VAMPS allows scientists to upload microbial genetic sequences, interpret what kinds of microbes harbor those sequences, and use sophisticated paradigms to compare the microbial communities represented by the sequences as comprehensive charts and color-coded graphs. The website is distinctive in many ways, but is particularly unique in allowing scientists to easily convert between taxonomic levels—phylum, family, genus, etc.

"You can really control at what level the analyses are being done," Sogin says. "You don't see that in other software."

VAMPS also links maps of ICoMM's ocean sampling locations to metadata (temperature, salinity, pH, geospatial and temporal information). In addition, it links to external websites such as GenBank and the Encyclopedia of Life, which provide details about particular species. And Sogin and colleagues, including the MBL's David Mark Welch and Susan Huse, continue to develop increasingly sophisticated tools for VAMPS users.

These days, VAMPS users still include ICoMM scientists. But it is now also a vital tool for other projects, including the Microbiology of the Built Environment Network and the Deep Carbon Observatory, which examines microbes deep inside the earth. As of early this year, VAMPS had more than 750 registered users from around the world.

"VAMPS can be used for any molecular microbial ecology project," Sogin says. "It's becoming very popular." • —SES

## It's a Small World After All: "Rediscovering" the Diversity and Importance of Bacteria

By Joseph S. Levine



Joseph S. Levine is an educator and author who works to improve science education and public understanding of science. A member of the MBL's Board of Overseers, he is an alumnus of the Boston University Marine Program (BUMP), formerly based at MBL. He has published in journals ranging from *Science* to *Scientific American*; produced science features for NPR's *Morning Edition* and *All Things Considered*; and served as scientific advisor to PBS's *NOVA* and as science editor for two PBS series: *The Secret of Life*, and *The Evolution Project*. He co-teaches professional development courses for teachers through the Organization for Tropical Studies in Costa Rica, where he serves on the Board of Visitors. He co-authors *Biology* (Pearson Education), the most widely used high school biology textbook in the United States. Levine has a B.S. from Tufts University, an M.S. from Boston University, and a Ph.D. from Harvard University.

Bacteriologists today are in a frenzy of discovery, identifying new species and documenting mind-boggling diversity we never knew existed. But why the buzz now? It's hardly breaking news that bacteria are everywhere.

It turns out that 99 percent of bacteria cannot be grown in the lab, so traditional "swab and culture" methods reveal only a tiny fraction of bacterial species in a sample. These techniques are useless for studying bacterial diversity, abundance, and ecology.

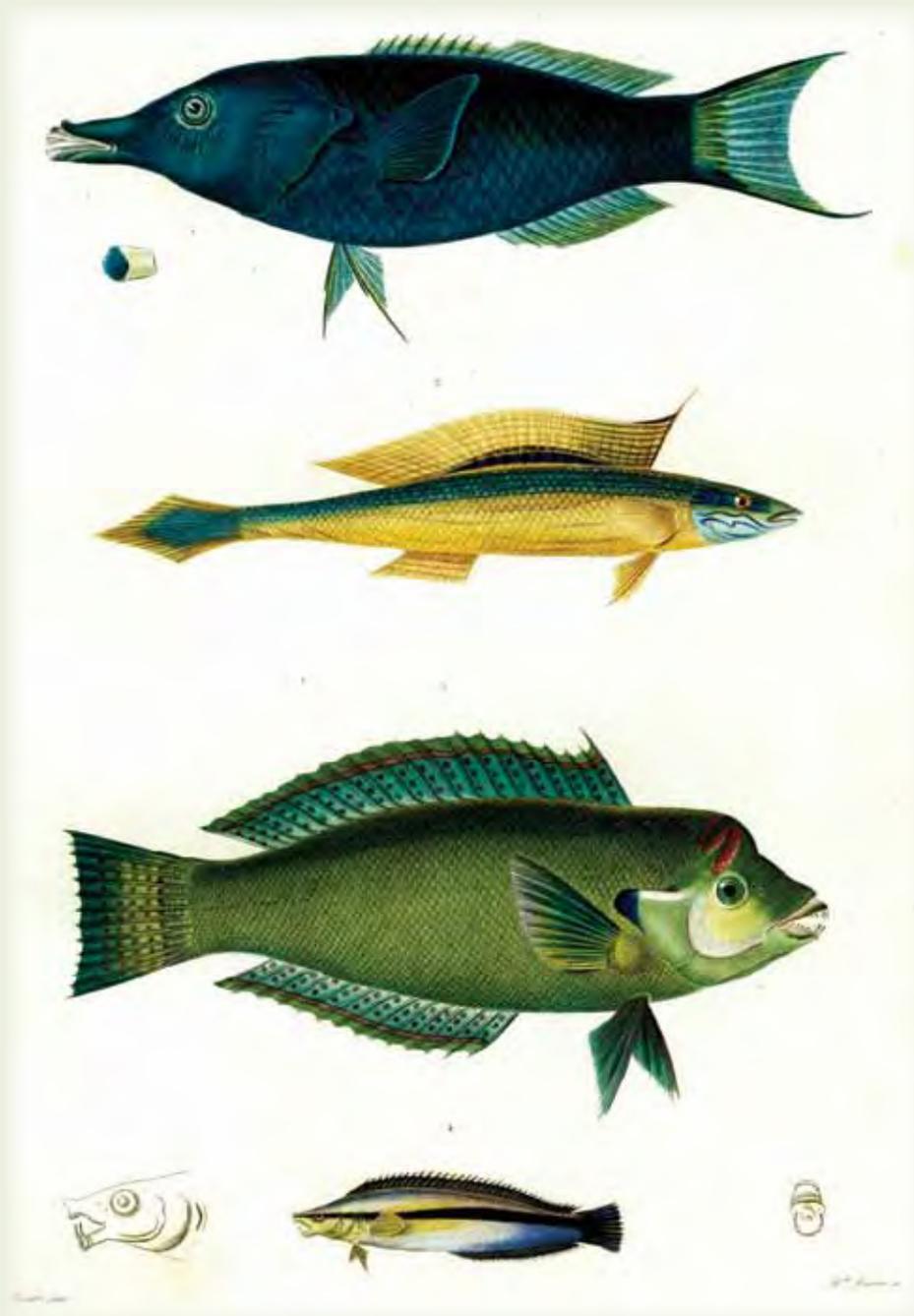
DNA analysis offered alternative ways to study bacterial communities, but time and price initially limited its use. In the last few years, however, the cost of DNA sequencing has plunged. As recently as 2006, it cost \$10,000 to identify 100 bacteria in each of 25 samples. Today, \$10,000 funds identification of more than 10,000 bacteria in 100 samples. A bacterium's genome—all of its DNA—can now be sequenced in less than two hours. And bioinformatic tools (including the MBL's VAMPS, see page 15) are being developed to analyze, organize, and distribute the staggering amounts of new data generated by these techniques. Welcome to metagenomics—the study of microbial communities, their genes, and their functions.

The result has been an explosion in our understanding of bacterial diversity. We now know that as many as 300 bacterial species inhabit a typical human face; that bacteria living in and on our bodies outnumber our own cells 10:1; and that the genes of those bacteria outnumber our genes 100:1. Plant leaves in nature host 100 to 500 different bacterial species each, with differences found from leaf to leaf and even greater among leaves of different plant species. A cubic meter of soil can house between 1,500 and 10,000 species of bacteria.

A devil's advocate might ask, "So what?" So there are lots of bacteria around. What are they doing? Lots of important things. Bacteria in our bodies perform metabolic tricks that we never evolved on our own, such as synthesizing essential vitamins. Gut bacteria affect our ability to extract calories from foods, influencing our propensity to produce body fat and our susceptibility to obesity and diabetes. Gut bacteria also affect the activity and toxicity of medications ranging from chemotherapeutic agents to acetaminophen—contributing to inter-patient differences in drug efficacy and side effects. Small wonder that many researchers suggest approaching nutrition (and medicine in general) from a metagenomic, rather than a strictly human genomic, perspective.

Now consider these kinds of interactions scaled up to the ecosystem level. Some members of microbial and fungal communities help to cycle nutrients through an ecosystem; others assist plant roots in absorbing water and nutrients, and create networks through which the products of photosynthesis are shared among members of the network. What other ecological functions might microbial communities perform? How do they vary among ecosystems? How do they affect ecosystem responses to nitrogen pollution from sewage and fertilizers? To pollution of different kinds? Or climate change? Or other human-induced environmental changes?

Metagenomic surveys and analyses at the ecosystem level should help answer these questions. Some preliminary results are tantalizing; some are counterintuitive. (Certain deserts appear to host higher microbial diversity than tropical rain forests.) Researchers have barely scratched the surface, yet it is already clear that the ecosystems we can see depend on unseen microbial communities—in more ways than anyone expected just a few years ago. •



From *Voyage de la Corvette l'Astrolabe: Exécuté par Ordre du Roi, Pendant les Années 1826-1827-1828-1829, sous le Commandement de J. Dumont d'Urville* (Paris: J. Tastu and Ministère de la Marine, 1830-1834), Plate 19.

## A DIVERSITY OF FISHES

The voyage of *L'Astrolabe* (1826-29) was a scientific expedition to the South Pacific under the command of French explorer and naval officer Jules-Sébastien-César Dumont d'Urville. The explorers brought back hundreds of plant specimens, rock samples, and a wealth of observations, which were published in 31 volumes by J. Tastu and Ministère de la Marine of the French government. The books contain finely detailed illustrations of terrestrial and aquatic organisms that showcase the biodiversity the explorers encountered, as well as maps, charts, and other scenes of discovery.

These hand-colored fishes were rendered by scientific illustrators who accompanied the voyages to record their discoveries. The fishes are described as (top to bottom): "*Comphose Bleu, Malacanthè Rayè, Girelle à front Bombè, and Aspidonte à Ruban.*" • —MP

**MBL**

Biological Discovery in Woods Hole

MBL  
7 MBL Street  
Woods Hole, MA 02543 USA

[www.mbl.edu](http://www.mbl.edu)



Smartphone users:  
Scan QR code for  
online content.

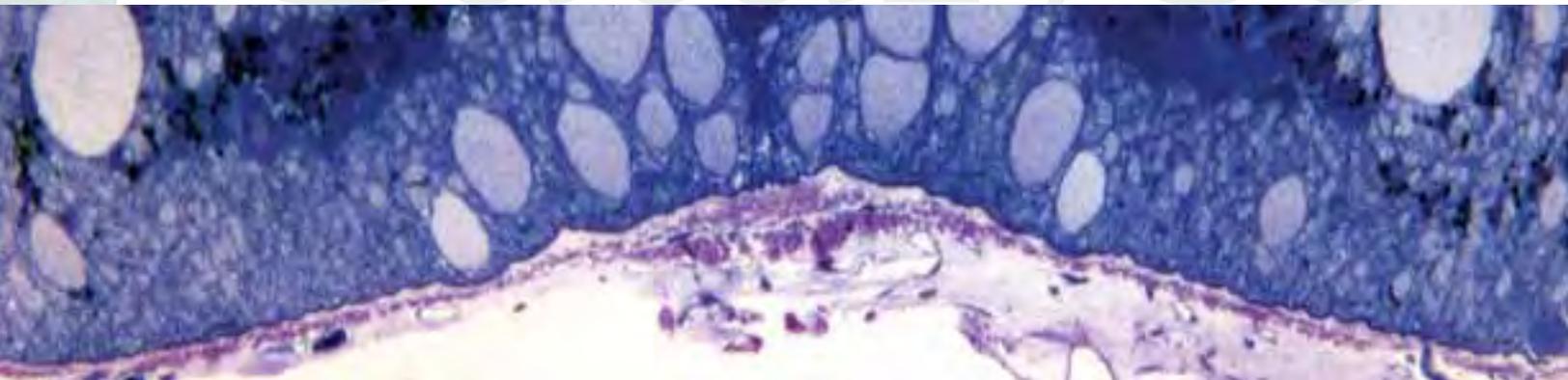
NON-PROFIT ORG.  
U.S. POSTAGE  
PAID  
Marshfield, MA  
PERMIT NO. 51



Printed with recycled materials.

IN THE NEXT *MBL CATALYST*

# Catalyst



## Neuroscience U

What are the biggest discoveries and mysteries in neuroscience today? The MBL is the place to find out. Our campus becomes "Neuroscience Central" during the summer, thanks to a convergence of hundreds of the world's top researchers, faculty, and students in this exciting and rapidly advancing field.