Beyond Human: New Faces, Fields Exploit Genomics

Fast new genomics technology is not just for human geneticists and biomedical researchers anymore

Find a podcast by the author of this feature, as well as other genome stories, at http://scim.ag/genome10.

Twenty years ago, the proposal to sequence the entire human genome was met with skepticism, even derision. Ten years later, the completion of the first human genome sequence was a source of awe worthy of presidential recognition. Today, it's a paragraph in a 3-year grant proposal.

The Human Genome Project drove a technological revolution in DNA sequencing that has continued since the full draft sequences were published in *Science* and *Nature* in 2001. And as sequencing DNA has become faster and cheaper, opportunities have grown for fields outside human biology.

Analyzing the genome of an organism once required extensive mapping of its chromosomes and the development of genetic tools specific for that species—a time-consuming and, usually, too expensive process. And decoding just one genome per species was the norm. Now researchers can skip those steps and even think about generating genomes or partial genomes for many members of a species. The new technology also lets researchers track gene activity on an unprecedented scale. "There's been a quantum change in what can be done and the number of organisms that can be studied," says evolutionary biologist William Cresko of the University of Oregon, Eugene.

That's why, in the last of our news features commemorating the 10th anniversary of the human genome, we look beyond human biology (although we peer inside the human gut in one case) to profile five research teams that have embraced genomic-scale science to tackle questions they could not have easily addressed before.

–ELIZABETH PENNISI

CREDIT: GONZALO GIRIBET

On a quest. Casey Dunn went to Greenland in search of rare animals for his phylogenetic studies.

Tracing the Tree of Life

In 1994, researchers came across a completely new animal in a spring on a large island off the west coast of Greenland. Looking a bit like a worm, this microscopic creature had complex jaws and was so unusual that the biologists could not assign it to a known phylum. Instead, they put it in a group of its own, called Micrognathozoa. The oddball, officially named *Limnognathia maerski*, has been little studied since. Indeed, researchers in Greenland have documented only a few hundred specimens.

Yet with the help of next-generation sequencing, evolutionary biologist Casey Dunn and his colleagues plan to shine a scientific spotlight on micrognathozoans. During a recent field trip to Greenland, they collected specimens of the tiny invertebrate and have stored them in a freezer in Dunn's lab at Brown University. In a project that would have been unthinkably expensive for a single lab just a few years ago, they will now decipher much of the creature's genome and identify its genes. And that's just the beginning. Dunn and his colleague Gonzalo Giribet, a Harvard University invertebrate biologist, have freezers full of other unusual organisms whose genomes they plan to sequence so that they can refine the much-debated animal tree of life.

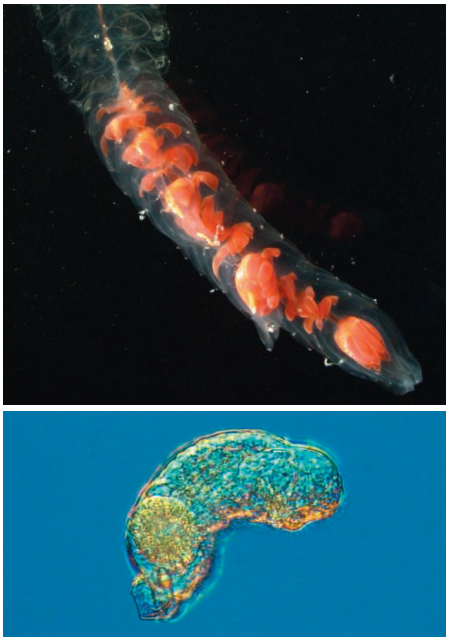
Molecular systematists have long used individual genes to assess relationships

Published by AAAS

among organisms, identifying differences within equivalent pieces of DNA in various species. Controversy often ensued when those early results didn't agree with more traditional classification schemes, such as those based on fossils or morphology. To add clarity, researchers have, over time, increased the amount of genetic material they compare, creating a field called phylogenomics in which many hundreds of genes are evaluated in each analysis (*Science*, 27 June 2008, p. 1716).

A few years ago, it cost about $12,000 per animal to sequence 1000 or so genes, says Dunn. Now, a few thousand dollars delivers many more genes. "We can do now what we couldn't do before," Dunn says. That includes sequencing little-studied organisms, such as micrognathozoans, so accurately that scientists may resolve the relationships of invertebrates whose lineages split off from a common ancestor 500 million years ago.

There are often challenges to sequencing unusual organisms. Sometimes researchers don't have enough DNA to work with; other times the organism has odd ratios of DNA's four bases that make decoding samples difficult. But Giribet has already sequenced and analyzed 20 of these animals, including a whip scorpion, a ribbon worm, and several mollusks. Dunn is excited about the prospect of resolving the animal tree as never before: "It's clear we are going to be able to base our tree on lots of data from lots of species."



Odd creatures. Siphonophores (*top*) and micrognathozoans may clarify animal evolution.

Next-generation sequencing technologies are also allowing Dunn to explore the evolution of animals by documenting differences in gene expression patterns across closely related species. The goal is to find out how these changes influence shifts in traits and behaviors across the tree of life. To do this, Dunn and his colleagues have turned to a new technique, known as RNA-Seq, that can gauge genetic activity in a sample by sequencing the complementary DNAs (cDNAs) that represent specific genes. The busier a gene is in a sample, the more times its cDNA will be sequenced.

Dunn and Stefan Siebert, one of his postdocs, have already compared the gene activity of the swimming and feeding forms of a siphonophore, a marine colonial organism. That analysis yielded thousands of genes potentially responsible for the differences in the animal's two structures. By repeating this experiment with multiple related siphonophore species, Dunn hopes to home in on those key to, say, the swimmer's development. "This will allow us to identify which genes have changes in expression that are associated with evolutionary changes," he says. **–E.P.**

Using DNA to Reveal a Mosquito's History

Ten years ago, the mosquito *Wyeomyia smithii* lived a largely anonymous life inside the "pitchers" of the purple pitcher plant common in bogs along the eastern United States, the Great Lakes, and southeastern Canada. Unlike some of its nastier relatives, the insect isn't known to transmit diseases to people or livestock. Larvae feast on microbes and detritus inside the pitcher plant, and adults sip on nectar, not blood, for the most part. Then in 2001, husband-and-wife evolutionary geneticists Christina Holzapfel and William Bradshaw of the University of Oregon (UO), Eugene, made the mosquito a poster child for climate change when they demonstrated for the first time that an animal had evolved in response to global warming.

Now the same researchers are applying next-generation DNA sequencing tools to probe further details of this species' evolutionary history—tools that have become so cheap and widely available that they can be applied to other poorly studied organisms as well. It's a "transformative technology," says Mark Blaxter of the University of Edinburgh in the United Kingdom.

Mosquito hunters. Christina Holzapfel and William Bradshaw embraced next-generation sequencing last year.

Holzapfel and Bradshaw began studying *W. smithii* 30 years ago, curious about how the mosquito had made its way so far north, because its relatives tend to reside in the tropics. In the course of their studies, they found that from 1972 to 1996, the mosquito's larvae in Maine had gradually delayed the start of hibernation by a week. Mosquitoes from farther north had postponed hibernation even later, whereas those in Florida had stuck to the same schedule as 25 years earlier. The pair concluded that the change in this genetically controlled trait was triggered by the longer growing season that resulted from gradual warming in the northern United States (*Science*, 23 November 2001, p. 1649).

Although the finding drew headlines, it still didn't explain how the mosquitoes had ended up in the north. To address that,

CREDITS (TOP TO BOTTOM): CASEY DUNN; REINHARDT MØBJERG KRISTENSEN; CHRISTINA HOLZAPFEL AND WILLIAM BRADSHAW

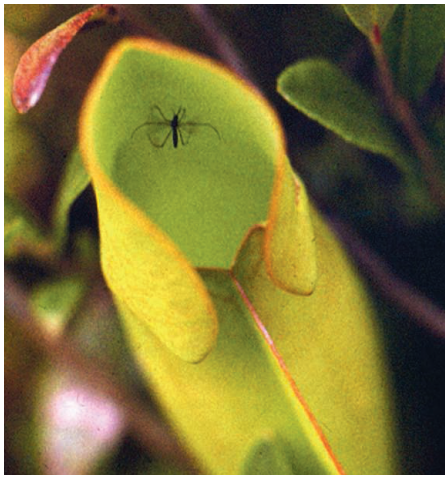
Published by AAAS

Holzapfel and Bradshaw wanted to know where the mosquitoes were in the past, particularly following a glacial period 20,000 years ago, when a warming trend had allowed them to migrate to new habitats. And to trace the migratory history of the species, the couple needed to establish the relatedness of populations from across the mosquito's range.

For years, they had tried to do this, but existing techniques were not able to resolve the differences between populations clearly enough. The mosquitoes from the various populations look too much alike to be distinguished morphologically, for example. In the 1990s, they tried in vain to reconstruct the biogeographical record by comparing proteins called allozymes among populations. Later, they fruitlessly looked at population differences in the insect's mitochondrial DNA. Even microsatellites, short stretches of DNA used in constructing genetic fingerprints, weren't up to the task. "We needed a better tagging or sorting system," Holzapfel recalls.

In 2009, they found one down the hall. UO colleague William Cresko had just teamed up with UO molecular biologist Eric Johnson to study the evolution of sticklebacks. They had genetically characterized populations of this fish by developing a catalog of single-nucleotide polymorphisms (SNPs), individual bases that vary frequently within a species. That work was made possible because a year earlier, Johnson's and Cresko's labs had developed a shortcut SNP-discovery method known as restriction-site-associated DNA sequencing (RADSeq).

This approach takes advantage of the speed and low cost of next-generation sequencing to quickly generate thousands of SNPs that distinguish populations and individuals. Researchers start by taking animals from multiple populations of a species and using so-called restriction enzymes to, at specific DNA sequences, chop up the genomes of each one into short fragments. Each animal's DNA fragments are then joined to a unique "bar code," a synthetic five-base strand of DNA whose sequence reveals which animal the non-bar-code DNA came from. All the fragments are then pooled together for mass processing by a next-generation sequencing machine. Because the bar codes allow the resulting sequences to be associated with specific animals, researchers aided by bioinformatics software can quickly identify genetic differences among individuals or populations.



Test case. Researchers didn't need a sequenced genome to make a dense genetic map of the pitcher plant mosquito.

For the mosquitoes, the researchers found 13,000 SNPs, 3700 of which helped to finally determine the relatedness of various populations of *W. smithii*. "This gave us the resolution to discriminate between postglacial populations," says Bradshaw. Based on that information, the researchers deduced that after glaciation, a remnant population of the pitcher plant mosquitoes gradually expanded out of the mountains of North Carolina—not out of the Gulf Coast, as some had presumed. The expansion proceeded gradually northward, then westward, they reported online 26 August 2010 in the *Proceedings of the National Academy of Sciences*.

When Cresko and Johnson's team tested RADSeq on the stickleback, they were able to match the fish's already sequenced genome to the newly generated sequence to help look for differences. No one had the resources to sequence the genome of *W. smithii*, and yet RADseq still worked effectively on the mosquito, demonstrating that the technique could be useful for a variety of organisms, even those for which little is known about their genetics. "This tagging system is definitely the wave of the future," says Holzapfel.

Furthermore, the cost for the entire mosquito study—examining all 23 populations of *W. smithii*—was just $3000. "The RADSeq method is cheaper, faster, and delivers thousands of markers," says Blaxter. He and his collaborators now have 18 RADSeq projects under way in snails, moths, nematodes, butterflies, salmon, ryegrass, sturgeon, beavers, beetles, oaks, elms, and spruce. Already for the diamondback moth, a crop pest, they have used newfound DNA markers to help pinpoint a gene that makes this moth resistant to a certain insecticide. Says Bradshaw, "This is an awesome technique." **–E.P.**

CREDITS (TOP TO BOTTOM): CHRISTINA HOLZAPFEL AND WILLIAM BRADSHAW; ROLAND KNAPP

Tackling the Mystery of The Disappearing Frogs

For more than a decade, Roland Knapp has watched and agonized as the mountain yellow-legged frog, which normally thrives in high-altitude lakes and ponds too cold for other amphibians, disappears from the Sierra Nevada. In 1997, Knapp counted 10,000 tadpoles in a single mountain lake—the frogs seemed to "occupy every possible bit of water," he recently recalled on his blog. This past summer there were almost none. Surveys of 15,000 sites by Knapp, a field ecologist at the Sierra Nevada Aquatic Research Laboratory in Mammoth Lakes, California, and others have shown that this frog—which is actually two species—

Going, going. The mountain yellow-legged frog has disappeared from 90% of its Sierra Nevada habitat.

Published by AAAS

is now missing from more than 90% of its former habitat.

There are multiple explanations for the frog's disappearing act, but a key one is the chytrid fungus, *Batrachochytrium dendrobatidis*, which has wiped out amphibians around the globe, including many populations of the mountain yellow-legged frogs. Yet every so often, some of these frogs survive the fungus, and Knapp has been unable to discern whether the amphibian's immune response or some environmental factor made the difference. "It's been pretty clear that our field experiments and observations only take us so far," he explains. "We needed to go to an entire new level of investigation."

So he was thrilled when Erica Bree Rosenblum, an evolutionary biologist now at the University of Idaho, Moscow, approached his team about collaborating on the endangered amphibian. In the past, Rosenblum, who studies the genetic basis of animal traits such as color or limb length, had been limited to what she calls "spearfishing": sequencing specific genes already suspected of influencing the trait. But about 5 years ago, she realized that new sequencing technologies would make it affordable to directly decipher all the active genes of a species without doing the extensive, and expensive, presequencing legwork required in the past. Thus, she could try "net-fishing," casting a net that could ensnare more than just suspected genes.

Gone. Roland Knapp's genomic studies may help explain the mountain yellow-legged frog's die-off.

Rosenblum, Knapp, Cherie Briggs of the University of California, Santa Barbara, and ecologist Vance Vredenburg of San Francisco State University are now using this approach on wild populations of the frogs, comparing ones that persist despite exposure to the fungus to nonexposed ones that ultimately prove susceptible to it. The key step, which next-generation sequencing greatly facilitated, was elaborating the frog's transcriptome, its full repertoire of expressed genes, by sequencing the so-called complementary DNAs (cDNAs) that represent each gene. With these cDNAs in hand, the researchers could construct a device known as a microarray to assess which genes were active in various organs of exposed and unexposed frogs. Results so far suggest that in susceptible frogs, the immune system doesn't go on the defensive, says Rosenblum; the fungi somehow evades the body's defenses.

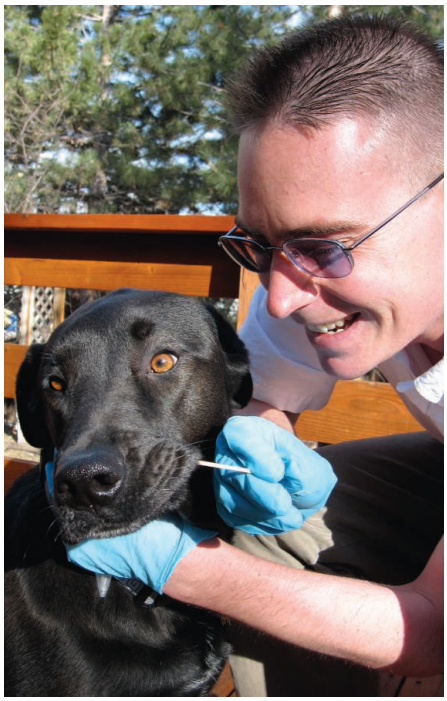
The researchers are also using the same microarray to evaluate gene activity in the amphibian's skin to understand why it degrades during infection. And by sequencing microbial DNA swabbed from frog skin, they are examining whether resistant frogs have an unusual repertoire of surface bacteria, as some microbes have been found to make an effective antifungal compound. Such genomic insights are much welcomed, says Vredenburg; the sequencing projects have "affected my work immensely."

–E.P.

Digging Deep Into The Microbiome

It isn't only animal studies that have benefited from the explosion in genomics tools. Next-generation DNA sequencing has transformed microbial ecology studies as well. The past decade has seen the growth of metagenomics, in which researchers sequence DNA from a soil sample, the gut, even a computer keyboard, to learn what bacteria live there. With the new technologies, "you can sequence at a level deep enough that you can understand what's going on in the community," says Rob Knight, a microbiologist at the University of Colorado, Boulder.

The microbial makeup of our gut is a case in point. In the past decade, scientists have come to realize that animal intestines naturally harbor diverse microbial communities that help provide nutrients and sustain good health. A landmark 2005 study by Stanford University's David Relman and colleagues (*Science*, 10 June 2005, p. 1635) concluded that the bacterial communities in the human gut vary tremendously from one individual to the next. But that work looked at the guts of just three people, using traditional sequencing technology to probe for different variants of ribosomal RNA genes, each of which represented a different microbe.



Bug hunt. Rob Knight studies the microbiomes of humans, dogs, and other animals.

A new analysis of 146 people, made possible by the lower cost and higher efficiency of DNA sequencing, is now telling a much more detailed story. Junjie Qin of BGI Shenzhen in China and colleagues recently collected fecal samples from 124 Europeans, some healthy, some obese, and a few with inflammatory bowel disease. They not only identified and sequenced all available ribosomal RNA genes in the samples but also deciphered more than 3 million other genes from the bacteria in the people's guts. (The 576.7 gigabases of DNA sequence data was

CREDITS (TOP TO BOTTOM): NEIL KAUFFMAN; ROLAND KNAPP; AMANDA BIRMINGHAM

Published by AAAS

almost 200 times the amount generated in any previous study.)

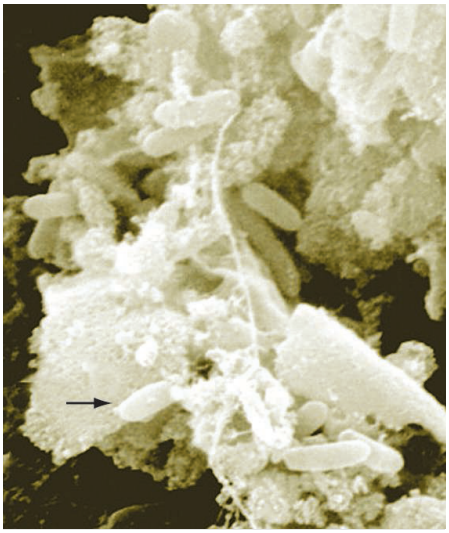
This more comprehensive analysis revealed limits to how much the common gut microbiome varies among people. There is a core set of gut bacteria, indicating that the "prevalent human microbiome is of a finite and not overly large size," Qin's team wrote in a 4 March 2010 *Nature* report. Certain bacterial gene sets and species in the gut also correlated with obesity, Knight adds. "When you look at a lot more people, you see systematic patterns of variation" in the gut microbiome, he says.

A twist on next-generation sequencing is also providing a new way to evaluate which genes are "must-haves" for the microbes. To find these genes, Knight and his postdoc Cathy Lozupone are working with Andrew Goodman and Jeffrey Gordon of Washington University in St. Louis, Missouri, who have developed a technique called insertion sequencing. This involves using mobile DNA elements called transposons to introduce mutations into tens of thousands of bacteria. Before adding the transposons to the bacteria, the researchers tag each transposon with an identifiable DNA "bar code" that allows each mutant bacterium to be tracked—and for the gene disrupted by the transposon to be characterized. With the new sequencing technology, researchers can follow mixed populations of these mutant strains on various growth mediums or in different environments. The relative number of copies of a bar-coded DNA sequence will reflect the success of a particular mutant; bacteria carrying mutations in genes required for a specific environment or medium wind up poorly represented.

The researchers first tried the insertion-sequencing technique on a human gut bacterium, *Bacteroides thetaiotaomicron*, introducing transposon-mutated strains into the guts of various kinds of mice. Some mice were normal and had their own gut microbes; others had various immune defects, were germ-free, carried human gut microbes, or had a combination of those characteristics. Two weeks later, the scientists took stock of how these bacteria fared in their different rodent hosts.

In the germ-free mice, they saw a decrease in the abundance of 280 distinct bacterial strains, suggesting that the gene mutated in each had been essential to staying alive in the gut. Defects in about 90 genes seemed to provide a competitive advantage, as the corresponding mutant strains colonized the germ-free mice better than other strains did. Whether the mice carried other human gut microbes made a difference to the survival of various strains of *B. thetaiotaomicron*, as a different subset of mutants disappeared in those mice, the researchers reported in the 17 September 2009 issue of *Cell Host & Microbe*. "We were able to find genes that determine the ability of a bacterium to thrive in the mammalian gut in specific microbial community contexts," says Goodman, now at Yale University.

Goodman calls the insertion-sequencing approach "exciting." Others agree. Several have begun to use it to characterize key genes for various microbes in different organisms and tissues. **–E.P.**



Gut life. Researchers have pinpointed the must-have genes in these capsule-shaped intestinal microbes.

CREDITS (TOP TO BOTTOM): J. L. SONNENBURG *ET AL.*, *SCIENCE* **307**, 5717 (25 MARCH 2005); JOHN A. BYERS

Probing Pronghorn Mating Preferences

Pronghorns, the American antelope, are the fastest animals on the North American continent, yet coyotes still kill many of the fawns, catching them before they develop the quickness to run away. Animal behaviorist John Byers of the University of Idaho, Moscow, has shown, however, that if a female pronghorn picks the right male, her fawns will grow faster than normal and have a much better chance of surviving. Since 1981, he and colleagues have tracked six pronghorn generations at the National Bison Range in Montana.

Byers suspects that female pronghorns, which he found favor males best able to fight off other males, are actually choosing mates with the lowest burden of so-called deleterious mutations. Byers hasn't had a good way to prove his theory, but thanks to the growing availability of next-generation DNA sequencing, he may finally have a chance. He and his colleague have over the years collected tissue samples from 835 pronghorns across the generations, and they now plan to genetically profile each animal to determine whether female pronghorns do indeed pick genetic studs. "I think it's going to be ultracool," Byers says. **–E.P.**

Published by AAAS