

Mammoth Achievement: Researchers at the forefront of molecular biology

January 26 2010, by David Pacchioli



Woolly Mammoth. (ExhibitEase LLC - Steven W. Marcus)

Forget Jurassic Park. By successfully sequencing the DNA of a long-extinct species, Stephan Schuster and Webb Miller have helped push back the boundaries of molecular biology.

Stephan Schuster was never all that interested in [ancient DNA](#). As a young genomicist at the Max Planck Institute for [Developmental Biology](#) in his native Germany, his forte had always been bacteria. By deciphering and comparing the genomes—the genetic blueprints—of various [microbial species](#), he sought to unlock the secrets of these ubiquitous creatures: how they evolve and interact with the organisms that play them host.

Schuster's early work had attracted considerable attention. In particular, a study done with colleagues in Germany and England in 2004 laid bare the fascinating life cycle of *Bdellovibrio*, a predatory microbe whose efficient dispatching of its rivals suggests the promise of a "living antibiotic." But when Schuster accepted an offer to join Penn State's Center for Comparative Genomics and Bioinformatics in 2005, he knew he had a decision to make. "I had to rebuild my lab," he said, "and I had already learned that there was a big change in technology about to happen." This change was the emergence of a next-generation DNA-sequencing machine, brainchild of a biotech start-up in Connecticut named 454 Life Sciences.

The automated "reading" of [DNA sequences](#)—the paired strands of nucleotides, or bases, that make up our genetic alphabet—had long depended on a chemical process developed by the British biochemist Frederick Sanger back in 1977. The so-called Sanger method had transformed biology, birthing the field of genomics and culminating in the successful decoding of the entire human genome, completed in 2003. But the sheer costliness of Sanger sequencing had placed strict limits on its use.

The emergent 454 machine, employing a new technology called sequencing by synthesis, allowed for "massively parallel" sequencing of [DNA](#) fragments, which meant a vast increase in speed and a corresponding drop-off in cost. Its developers envisioned that this approach would open up whole new frontiers of basic and biomedical research.

For all its promise, however, initial reaction to the new machine was "surprisingly reserved," Schuster remembers. In truth, the first 454 did seem to have some drawbacks. It was capable of reading only about 100 bases at a time, compared to 800 or so for the latest Sanger sequencers, and this shorter read length would make it harder to reassemble

fragments of DNA into a complete genome. There were also questions about its accuracy. Established researchers, and funding agencies, moreover, were heavily invested in the existing technology.

For Schuster personally, the moment was an important one. DNA sequencers cost in the range of \$500,000 a piece, and a successful enterprise would certainly require multiple machines. The path he chose would be defining. But as he crossed the Atlantic to embark on the next stage of an already flourishing career, that choice became clear. His goal, after all, was to "explore the limits" of DNA sequencing, Schuster told himself. "I decided I would rather risk my money on the new technology than continue to work with the established one," he recalls. "I made what I called my half-million dollar bet."

The 454 GS20 sequencer he requested for his brand-new lab at University Park was only the fourth one off the production line—the first purchased by a university— and it was the 454 that led Schuster to the woolly mammoth.

Faded genes

The study of ancient DNA, which began in the mid-1980s, has always been deviled by two realities. First, the genetic trail erodes over time. Over hundreds and thousands of years, the DNA molecules of a defunct organism inevitably disintegrate, leaving only a welter of fragments. These faded traces, in turn, are mashed up with sundry other bits and pieces, the equally degraded DNA of the plants, animals and microbes that, over millennia, happened to die on top of or near—or inside—the body in question. This many-layered presence of competing information is excruciatingly hard to interpret: Lifting the original fingerprints from a recently unearthed Roman coin might actually be easier.

To minimize their handicap, researchers in ancient DNA must seek out

the most pristine of specimens, remains either mummified or frozen (preferably both) of the most enduring of biomaterials: bones and teeth. "Hydroxyapatite—that's a mineral contained in bone—binds the DNA and stabilizes it," Schuster said, "but bone is also a highly porous material, and in the process of putrefaction, bacteria grow deeper and deeper into it, using the last remaining organic materials, amino acids, as a body decomposes. In the end, these bacteria also die, and they deposit their own DNA on top of the animal's or person's DNA."

Even the best specimens, therefore, have yielded little in the way of useful information. Using traditional Sanger sequencing, Schuster says, only a tiny fraction—at most 2 percent—of the DNA picked up from a sample of ancient bone would be likely to be the DNA of the creature to whom the bone actually belonged. "And that was not enough to sustain a large-scale project."

To Schuster, however, in the blue-sky excitement of trying out a new technology, the notoriously poor quality of ancient DNA smelled like opportunity. To the extent that it can be recovered at all, he knew, the stuff turns up as alphabet soup, snippets only dozens of base pairs long. You don't get long strands of intact code. "I thought that might be a good match for the 454's short read lengths," Schuster says simply. What others had seen as a flaw might turn out to be an asset. Acting on this hunch, Schuster set about lining up as many samples of ancient DNA as he could. "We systematically explored all kinds of animals that went extinct within the last 100,000 years," he recalls. "One of these samples happened to be from a mammoth, and it worked for us immediately."

Joining forces

"I first saw sequence data from a woolly mammoth on Nov. 18 of 2005, at about 3:30 in the afternoon," Webb Miller recalled. "Stephan walked into my office and said, 'Hey, I've got something here that I think you're

going to find really interesting. And that was it."

The two had been looking for a way to collaborate since Schuster's arrival at Penn State some months earlier. Miller, 18 years Schuster's senior, had been a pioneer in the now-exploding field of bioinformatics (his trailblazing efforts were recently recognized with a career award from the International Society for Computational Biology). After starting as a computer scientist in the late 1970s, he became intrigued by early reports of the Human Genome Project, and, looking for a new challenge, decided to take the plunge into biology.

One of Miller's early successes, the Basic Local Alignment Search Tool, or BLAST, for which he and two colleagues developed the computer algorithms, is still one of the most widely used programs for searching databases of genetic sequences. In the years since, Miller, now professor of biology and of computer science and engineering at Penn State, has made a specialty of developing and applying methods to compare longer and longer sequences of DNA, most recently complete vertebrate genomes.

"Webb has played an essential role in nearly every vertebrate genome sequence project," said colleague David Haussler of the University of California at Santa Cruz.

He was, in short, the perfect match for Schuster and the woolly mammoth.

"Coming from the microbial world, I found mammalian genomes very intimidating," Schuster said. "Mammal genomes contain a lot of repeat elements—less than two percent of the genome is coding, where the actual information is stored. This compares to 90 to 95 percent in a bacterial genome. You could say that almost nothing is coding in a mammalian genome, and almost everything is coding in a bacterial

genome. You need very different computational tools to be able to assess them."

Their first paper together was published in the journal *Science* in December 2005. Working with Hendrik Poinar of McMaster University, a leading expert in ancient DNA, Schuster and Miller presented sequence data retrieved from a 28,000-year-old mammoth jawbone that had been frozen in the permafrost of northern Siberia. Using the present-day African elephant for comparison, they were able to identify 13 million DNA base pairs—a tiny fraction of the beast's genome, but by far the largest piece that had ever been sequenced. More importantly, they were able to show that fully 50 percent of what they had gleaned was actual mammoth DNA, and not that of an environmental contaminant. No prior study involving an extinct mammal, Schuster said, had ever yielded more than a few percent.

A whole new field

Some months later, Schuster was in Europe visiting with another leader in ancient DNA research, Tom Gilbert of the University of Copenhagen, when the lunchtime talk got around to specimens. Gilbert, having tired of the contamination issues he encountered working with fossil bones, had begun experimenting with hair as a source material. Although it is routinely analyzed for evidence in present-day crime labs, hair had been pretty much ignored by the ancient DNA crowd. "When people thought of sequencing DNA from hair, the usual assumption was that the material must come from the hair root, or follicle, because the hair shaft appears to be dead," Miller said. Skin cells attached to the follicle make juicy tidbits for crime scene investigators, but they degrade rapidly.

Gilbert's trials, however, had revealed that the hair shaft itself contains DNA. Even better, this DNA is encased in keratin, the tough fibrous protein that Miller calls "a kind of biological plastic." Thus protected, it

should remain viable much longer than DNA from even bone. And unlike bone, Schuster said, it could be easily decontaminated—"by shampooing and then soaking in ordinary household bleach."

In September 2007, the three researchers, working with a large international consortium, published in *Science* the complete mitochondrial DNA for 10 woolly mammoths taken entirely from tufts of hair, some of them 50,000 years old. Importantly, these samples had been stored away in institutions, not frozen in ice. One of them, in fact, came from the famous Adams mammoth, which had been kept at room temperature in a Russian museum for over 200 years. That such material could yield such rich genetic information suggested that their sequencing method might be applied to specimens of other extinct and non-extinct species held in collections around the world. This broad new application for DNA analysis even inspired them, only half in jest, to coin a term for this new field of study: museomics.

Going nuclear

Mitochondrial DNA, or mtDNA, the strange scrap of genetic information found outside the cell nucleus, is valued by researchers for a number of reasons. Hundreds of copies of this information are present in every cell, which makes it that much easier to recover. And mtDNA evolves much faster than its nuclear counterpart, which makes it useful for spotting differences within a population. But mtDNA makes up only a tiny fragment of an individual's [genetic blueprint](#) (in the mammoth, only 13 of some 20,000 genes). To get the bigger picture requires unraveling the entire genome.

No one had ever attempted this feat for an extinct animal. Sample quality aside, with traditional Sanger sequencing the task was simply too expensive. With the next-generation machine, however, it was suddenly feasible to sequence the same stretches of DNA over and over (and over,

up to 20 times), which is critical for spotting mistakes and getting a true read. In the case of the mammoth, there was the added advantage of having a close relative available as a reference. "We had a pretty good sequence of the African elephant to map onto," Miller said. "That greatly simplifies the job of analyzing these little fragments."

In November of last year, after months of effort, Miller and Schuster published in the journal *Nature* a paper that riveted the scientific world: Using hair taken from two mummified specimens, they had successfully sequenced over 4 billion bases of DNA, roughly 140 bases at a time. By comparing against their elephant guide, they could confirm that 3.3 billion of these bases were mammoth DNA. In all, they estimated they had accounted for 50 to 70 percent of the entire mammoth genome, with the rest waiting only for additional funding. Whatever the exact percentage, this was a dataset "100 times more extensive" than any yet seen for an extinct species, Schuster said. "This really is the first time that we have been able to study an extinct animal in the same detail as the ones living in our own time."

These results, combined with those of the earlier mtDNA study, yielded several new insights into mammoth—and elephant—evolution. Woolly mammoths apparently separated into two groups around 2 million years ago, and these groups eventually became genetically distinct sub-populations, Schuster said. One of these groups died out approximately 45,000 years ago, while the other lived on until the last Ice Age, about 10,000 years ago. The data also show a closer relationship between mammoths and modern-day elephants than was previously suspected: "Their genomes are over 99 percent overlapping." In that remaining fraction of a percent, he and Miller have begun to look for the genetic causes of some of the mammoth's unique traits, including its adaptation to extreme cold.

Ice Age 2?

These revelations met with keen interest throughout the ancient-DNA community, but it was something else Schuster said, quoted at the tail end of a Penn State press release, that caught the attention of the wider world. "By deciphering this genome," he allowed, "we could, in theory, generate data that one day may help other researchers to bring the woolly mammoth back to life."

"Ice Age Mammal May Walk Again," boomed one of the resulting headlines. "Jurassic Park-style breakthrough," blared another. The story was picked up and trumpeted by dozens of news outlets around the world, and Schuster was interviewed on Fox News and Good Morning America, following video clips from Mammoths to Manhattan and Ice Age 2. Although raising a mammoth was not the object of the study, he said, blinking a little under the studio lights, "most experts would agree" that, in the wake of the new data, "for the first time it is not entirely impossible to think about" doing so. Not surprisingly, his careful qualifications seemed lost on his TV hosts.

Miller, for his part, was even more dismissive. "At first I thought it was a stupid idea," he admits. "But I'm starting to get more interested. I'd like to see more research being done in reproductive technology, for the possibility of human benefit down the road, and this might be a relatively safe way to do that." He muses. "It would be sort of like a moon shot."

Schuster's current argument is that, given the theoretical possibility, rapid advances in the practice of genetic engineering over the last five years make it inevitable that scientists will one day have at least the capability of cloning a mammoth. "Just look at the amount of manipulation that is already being done in farm animals," he says.

The easiest way to proceed would be to alter the genome of a modern-day elephant by introducing mutations—inserting mammoth DNA at the

approximately 400,000 sites (out of 4.5 billion) where elephants and mammoths differ. This hybrid genome would then be injected into an elephant embryo and carried to term in an elephant mother. "You would get what we call a mammothified elephant," Schuster said. "We have no idea what it would look like." A more radical approach would be to use a completely re-assembled mammoth genome to synthesize a set of actual mammoth chromosomes. As far-fetched as that may sound, Schuster points out, genomics pioneer Craig Venter has already succeeded in synthesizing the chromosome of a bacterium.

"This field of synthetic biology is unfolding as we speak," Schuster said. "We will be able to design entire organisms, and as a side product we will one day be able to synthesize the chromosomes of extinct animals. However—and here is my word of caution—at the moment when we are actually capable of doing this, the technology will have such profound impacts on human society that I don't think we will have much interest in a folly like resurrecting a mammoth."

Extinction biology

In April of this year, Schuster and Miller were named to Time magazine's list of "Top 100 Most Influential People" along with Michelle Obama, Energy Secretary Steven Chu, and the Twitter guys. Craig Venter, who wrote their citation, discounted the possibility of bringing a mammoth back to life. The real accomplishment, he wrote, was in "pushing the limits of DNA analysis, both to explore our past and perhaps predict our future."

Boutique science aside, the real benefits of the mammoth genome project, Schuster and Miller agree, will likely come in the here-and-now realm of extinction biology. One of their immediate goals is getting a better handle on just what forces killed off this mighty creature. "There are many hypotheses," Schuster said, "but all of them are hard to

substantiate when you look closely." Their sequencing data already rule out humans as culprits, he says, at least for that first big wave of extinction 45,000 years ago. "There were no human hunters in Siberia at that time."

The mtDNA data also have revealed a surprisingly low level of genetic diversity across mammoth populations, which may have made the species especially susceptible to environmental threats. "We're actually thinking about three separate extinction events," Miller said. "The one at 45,000 years ago, the famous one at 10-to-12,000 years ago, and then there were actually some woolly mammoths that survived on isolated islands up until about 3,700 years ago. It could well be that they're not due to the same causes."

Their techniques, they believe, can yield important answers for other long-lost mammals too, and even, said Schuster, for reptiles and amphibians, "particularly if we can get parts similar to hair that contain keratin, such as scales, horns and claws. This is a very robust and widely usable approach."

Already, he and Miller have turned their attention to more recent cases of extinction, like the Tasmanian tiger, a wolf-like marsupial also known as the thylacine. "One of the things we want to see is what does a population look like 10 years before it goes extinct, or 20, or 30 years," explained Miller. "We can't do that with the woolly [mammoth](#), not at that resolution. But with the Tasmanian tiger, we know exactly when it went extinct: Sept. 7, 1936. There are something like 700 known specimens of this animal. We can sequence all of them, and know when they were collected. We can really watch the endgame of a species."

Clues to the Future

Such data provide valuable points of comparison for present-day

endangered species, the researchers say, such as the Tasmanian tiger's legendary relative, the Tasmanian Devil. Currently teetering on the edge of extinction, the Devil is being wiped out by an infectious facial cancer whose spread is facilitated by inbreeding—a lack of genetic diversity so acute that it knocks out immune response. By sequencing animals that have the cancer and comparing those sequences against those of animals that have resisted the disease, and careful outbreeding based on the results, they suggest, wildlife biologists might create a new starter population that could be held in captivity with the hope that someday the cancer will have run its course. "We hope the Tasmanian Devil becomes the first instance where [genome](#) technology has been put to work in order to try to save an endangered species," said Schuster.

Understanding the genetic underpinnings of past extinction events, he and Miller argue, may be crucial for protecting other potentially threatened species, including, perhaps, even our own. "What makes us so sure that we cannot go extinct?" Schuster asks. "We are so happily messing around, even actively contributing to a change in our environment, believing that we are untouchable. By reconstructing the biological history of the last 10,000 years—the big change that has happened since the last Ice Age—we may find a message stored in the fossil record that is very important for our future."

"This is my fascination with genomics, these final answers," he said. "You can sequence genomes down to the very last base pair. And by then making comparisons you have an excellent way of really understanding the biology that is going on—in evolution, in function, in disease. This is why I'm convinced that next-generation sequencing is the biggest thing that has happened in biology in a long time."

Provided by Pennsylvania State University

Citation: Mammoth Achievement: Researchers at the forefront of molecular biology (2010, January 26) retrieved 26 April 2024 from <https://phys.org/news/2010-01-mammoth-forefront-molecular-biology.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.