

---

# Reading the Trees

---

The giant sequoia and the California redwood live thousands of years and can reach the height of a skyscraper, but climate change threatens to topple these mighty trees. Now scientists hope to transcribe their genome and unlock the story inside the trees themselves.

---

ELIZABETH EVITTS DICKINSON

Illustrations ELEANOR TAYLOR



In California's Yosemite National Park, you will find one of the oldest living organisms on earth: *Sequoiadendron giganteum*. The giant sequoia. Within the park's three groves of sequoia trees, the Grizzly Giant is the oldest. At 2,900 years, give or take a few centuries, this tree has been around since the Roman Empire. Giant sequoias can live more than 3,200 years and can grow to nearly 300 feet.

On a clear, sunny day in September of 2017, Winston Timp wasn't calculating the age or height of the sequoias in Yosemite's Tuolumne Grove. He was snapping photos of his 6-year-old daughter, Emma, standing inside a tree. Emma had placed her palm on the sequoia's fragrant, wrinkled bark, which can grow as thick as 3 feet, and had climbed inside a cave carved out of the trunk by a wildfire. These trees grow so wide and their trunks are so hardy that in 1881 people cut a tunnel through the base of a living sequoia so tourists could drive through.

Timp, a 39-year-old biomedical engineer at Johns Hopkins Whiting School

of Engineering, was hiking Yosemite that day with researchers from the Save the Redwoods League, a nonprofit working to protect giant sequoia and redwood trees. They were joined by a coterie of fellow scientists embarking on a new project to help conserve and restore these forests. Before this trip, Timp had never seen such big trees. "We weren't even seeing the tallest in the park, and yet these were still enormous," Timp says. The giant sequoia, as the naturalist John Muir wrote, "seems the very god of the woods."

While massive, the giant sequoia is not the tallest; that honor goes to the coast redwood, *Sequoia sempervirens*, which grows to the northwest of where Winston and Emma Timp were hiking that day. Edging the coast of California and into Oregon, groves of redwoods tower over the land, with the tallest measuring the height of a 35-story skyscraper. Redwood ancestors are among the oldest species of tree on our planet. Fossils of their seed cones were found in Europe and date back more than 200 million years to the Jurassic period.

Sixty-six million years ago, when a meteor slammed into the Yucatán Peninsula of Mexico and caused global tsunamis and deep freezes that killed off entire species including the dinosaurs, the redwoods remained.

---

Redwoods, then, are remarkable survivors. Sixty-six million years ago, when a meteor slammed into the Yucatán Peninsula and caused global tsunamis and deep freezes that killed off entire species including the dinosaurs, the redwoods remained. Even when charred to black by fire, a tree can renew itself as long as its roots aren't damaged.

What the mighty redwood cannot survive, it turns out, is *Homo sapiens*.

In 1848, flakes of gold were found shimmering in a river at the foothills of the Sierra Nevada mountain range, and the infamous rush was on. Prospectors flocked to California for gold; they discovered the wealth of the redwoods. Two million acres of fertile old growth forest from Big Sur to Oregon. Soon, sawmills dotted the landscape and the redwood was felled at a staggering clip for its dense and durable wood. Public forests in California passed into private hands during decades of shady land-grab deals. Meanwhile, in the Sierra Nevada mountains near Lake Tahoe, a giant sequoia named the Discovery Tree was chopped down in 1853, and its remains were turned into a dance floor. A cotillion of 32 people once waltzed across its stump.

Government action and public advocacy have protected some groves of giant sequoias and redwoods. In 1864, even as

The delicate ecosystem protecting the few remaining redwood forests is eroding at the edges from development, drought, fires, and pests.

civil war fractured the country, the majesty of the trees that Timp and his daughter saw spurred an act of Congress signed by President Lincoln. The government created the first protected scenic land in America, what would become Yosemite National Park. In 1918, a group of citizens alarmed at the destruction of old-growth forests formed the Save the Redwoods League. The group secured 200,000 acres in 66 state parks and reserves, but logging on unprotected lands continued apace. By the 1960s, humans had felled 95 percent of the redwoods.

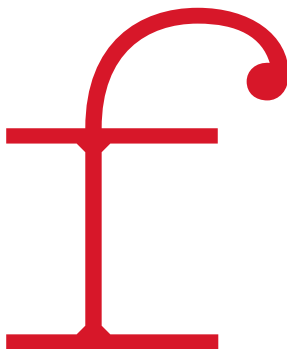
Today, people continue to threaten the coast redwoods and the giant sequoias, this time with development, pollution, and climate change. The diminishing of snow packs in the Sierra Nevadas means less precious groundwater for the thirsty giant sequoias. The delicate ecosystem protecting the few remaining redwood forests is eroding at the edges from development, drought, fires, and pests. "Climate change is going

---

to be a major stressor going forward,” says Emily Burns, an ecologist and head scientist at Save the Redwoods. In order to protect and restore the giant sequoia and redwood forests, she says, “we need to understand the trees on a genomic scale.”

To that end, the Save the Redwoods League launched the Redwood Genome Project in 2017 under the direction of Burns and David Neale, a forest geneticist at the University of California, Davis, who has studied trees for over 40 years. They have partnered with Johns Hopkins computational biologist Steven Salzberg, and with Timp, to attempt something never done before: sequencing the genome of two of the world’s largest organisms. This time, it’s not an act of Congress that promises to save some of the wildest, oldest living species on the planet; it will be the secrets held inside the trees themselves.

Sillett sat on  
a limb hundreds  
of feet off the  
ground and snacked  
on fresh berries  
plucked from a  
huckleberry bush.



From our human perspective on the forest floor, it’s impossible to fathom the full complexity of the redwood canopy. Until recently, the tops of the redwood trees went unexplored, a true terra incognita. Strong branches rarely grow low on a redwood trunk, so one might have to scale 25 stories into the air before reaching a steady purchase. People simply couldn’t imagine climbing into them.

Then, in 1987, a 19-year-old named Stephen Sillett free climbed (as in no ropes) into the redwood canopy with a friend. Redwoods branch off high above the ground into hundreds of aerial trunks that twist and arc and fuse together. Cradled within the pockets of these many trunks, the redwood creates its own canopy soil, which supports an ecosystem. Sillett found ferns and mosses and bonsai trees—what Burns now calls “a whole garden in the air.” Sillett sat on a limb hundreds of feet off the ground and snacked on fresh berries plucked from a huckleberry bush.

Today, Sillett is a botanist at California’s Humboldt State University, and along with a handful of elite scientists, he ascends these trees, climbing well over 300 feet. The canopy structures of the redwoods are so sizable that these

---

climbers have occasionally gotten lost. What the canopy scientists have discovered is a habitat comparable to the coral reefs; within its branches the redwood provides a home for untold species of animals and organisms. Some live their entire existence in the upper reaches, including salamanders and a form of plankton previously believed to live only in the ocean. Botanists continue to find new epiphytes—plants that grow on other plants—whose roots never touch the ground, including “a 40-foot-tall western hemlock growing in soil accumulated on a redwood branch hundreds of feet off the ground,” Richard Preston wrote in his 2007 best-selling book, *The Wild Trees*.

How are the trees capable of such feats? Scientists have long looked to the redwood and giant sequoia and wondered at their longevity and grandeur. How do they adapt? What allows some of them to thrive for centuries? The questions have always been there but were “frustratingly difficult to answer,” David Neale says. “Now the technology is finally catching up to our questions.”

The Human Genome Project, when it launched in 1990, became one of the largest collaborative science projects ever. It cost in excess of \$2.7 billion and took 13 years to complete. Today, we can spit into a tube, mail it to a lab, and in a couple of weeks get a report that reveals not only our ancestry but also our genetic health risks and our likelihood of carrying inheritable diseases. “People quite rapidly appreciated the value of



having done the human genome, in terms of the impact on medical research and medical diagnostics that contribute to the understanding of human disease and treatment,” Neale says. “It’s really no different when it comes to plants.”

Starting in the early 2000s, a next generation—or nextgen—of sequencing technology meant that DNA could be run faster and cheaper. Scientists began sequencing the genomes of agricultural crops. Every living organism has its own unique set of genetic operating instructions, and knowing that information allows breeders to produce varieties of plants that grow larger, taste better, and survive drought or pests.

Mapping the genome of the giant sequoia and the redwood, however, is not about breeding bigger, better plants. “It’s along the lines of how we use the information in human health,” Neale says.

---

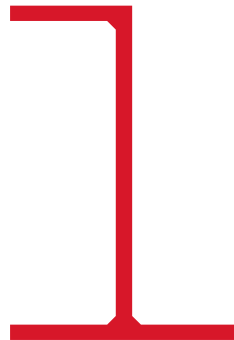
“It’s about conservation and restoration and maintaining healthy forests.”

Trees, like humans, possess different traits. Some are better at surviving drought, disease, and blight. Preserving the forests requires active management, including regular thinning to prevent wildfires. Knowing the genetic composition of a tree would ensure that forest managers don’t cut down those most capable of survival.

Also like humans, trees require genetic diversity in order to thrive. Over their many decades of active logging, timber companies did plant new trees; there are now nearly 1.5 million acres of young redwood forests. But some of the trees are clones bred by the companies. “You could have hundreds of thousands of redwoods, but if they are all the same clone, they could be susceptible to the same disease,” Burns says, bringing to mind the devastation of the American chestnut to a strain of virulent fungus.

Sequencing the genome would allow for the development of a device that could quickly assess the genotype of individual trees in the field, the same way we are now able to spit in a cup and learn about ours. If ecologists discover that a redwood forest is composed mainly of the same family, for instance, “we could work to introduce diversity back in by planting genetically different trees,” she says.

With the Redwood Genome Project, Timp says, “we’re creating a 23andMe for trees.”



Last November, two months after his visit to Yosemite, Timp was back in his lab on the Homewood campus of Johns Hopkins University in Baltimore. Over the summer, Neale had shipped Timp needles from a giant sequoia and a redwood tree, and now Rachael Workman, a senior research specialist in Timp’s lab, donned a pair of gloves and pulled the samples from a high-powered freezer. Vibrantly green and thick, the leaves resemble a robust spray of rosemary.

Sequencing a genome for the first time so that it can be scientifically useful requires three steps: sequencing, assembly, and annotation.

Before you can sequence a species, you need a single specimen, in this case, one tree from each species. The project team is extracting DNA from two sources of each tree: a seed cone (think of a pine nut) and the needles. The seed, which is how these trees reproduce, offers a small but intricate picture of its DNA. A close-up snapshot. The needles, on the other hand, yield lots of DNA, more like a panorama portrait.

---

Giant sequoias and redwoods are ever-green conifers and don't shed annually; needles on their lower sections could be decades old, and tissue from those would likely have degraded over time. You need new growth from the tops of the trees, which is also where new seeds are found.

It was Stephen Sillett who collected the samples Workman is using. After his free climb back in college, he customized a method developed by arborists of using ropes and gear to move within the giant trees. It's called "air walking." Sillett and Save the Redwoods won't disclose exactly which trees these samples were taken from. There's a policy among botanists to guard species from overly curious nature tourists, who may "trample off trail and unwittingly love a tree to death," Burns says. What Sillett will share is that the sample redwood tree is, conservatively, 1,300 years old; the giant sequoia, 1,340.

With samples in hand, it was time to sequence the DNA. The information in DNA is a code made up of four chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T). DNA bases like to pair up, creating what's known as base pairs (AT and CG). Humans have 3 billion of these pairs. A chromosome is composed of DNA that contains the genes.

Sequencing a conifer's genes is far more complex than sequencing a human's. The giant sequoia, like humans and other animals, is a diploid—it has only two pairs of chromosomes. But the sequoia is big, with a genome roughly three times the size of a human's. The redwood, meanwhile, is a hexaploid; it

contains six sets of chromosomes and is believed to be over 10 times the size of the human genome. The redwood will be one of "the largest and most difficult genomes ever sequenced," Neale says.

Several years ago, Neale and Salzberg became the first to sequence a conifer, the loblolly pine. Through that experiment, Neale determined the DNA sequence from a pine cone seed with nextgen technology. The sequences produced—those lines of AGCT—were accurate, but short. Almost like a few words instead of a sentence. With a genome as large as the redwood's, longer sentences are needed to understand its full story.

At Timp's lab last November, several computers churned away. Squares of primary colors shimmered and shifted across the screens. These, Workman explained, represented giant sequoia DNA being actively sequenced. Connected to the computers was a brand-new device that looks something like a portable charger for a

*"You could have hundreds of thousands of redwoods, but if they are all the same clone, they could be susceptible to the same disease."*



smartphone. Called a nanopore sequencer (“nanopore” refers to the microscopic hole into which a researcher can feed a DNA sample), the device is able to read longer sequences than previous technology.

This is the first time nanopore technology has been used for such a large genome. Once Timp and his team finish sequencing both trees, they will have a bulk of data so immense that it requires offsite storage on a cloud server.

The nextgen sequencers are capable of reading and spitting out pieces of raw DNA, but what they can’t do is put those pieces into any kind of comprehensible order. The next step is to build these fragments into a cohesive picture of the trees’ full genome. And this, Timp says, is where things really get complicated.

Across town from Timp’s lab, on the East Baltimore campus of the Johns Hopkins School of Medicine, is the lab

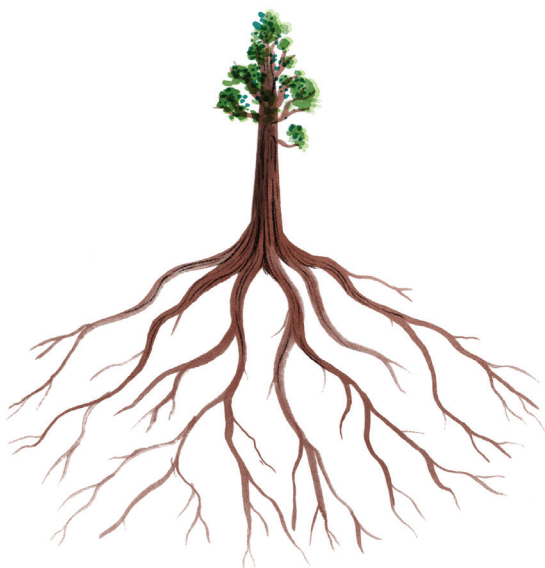
The redwood is a hexaploid; it contains six sets of chromosomes and is believed to be over 10 times the size of the human genome.

of Steven Salzberg. A computational biologist, Salzberg, age 57, worked on the Human Genome Project in the 1990s and co-authored one of its papers. In 2001, he helped sequence the anthrax strain used in the deadly bioterrorism attacks, allowing the FBI to close in on the lab where it originated. Over his career, Salzberg has assembled genomes for hundreds of species, ranging from berries to woolly mammoths. He’s also assisted epidemiologists in tracking flu outbreaks: In 2004, Salzberg helped launch a flu genome project capable of quickly sequencing the ever-mutating virus.

Salzberg is what’s known in the field as a gene finder. Buried within all those bits of DNA that were read by the sequencers are the genes that compose a map of the genome. Salzberg is the one who will take the raw data and piece them back together into an accurate picture. This process, called assembly, is “like putting together a giant jigsaw puzzle, only the pieces are made up mostly of blue sky,” Salzberg says.

While nextgen sequencers revolutionized how machines can read DNA, the advancements in the computational capacity to find the genes within that raw data “lie largely with Steven and the algorithms that his lab has produced,” Neale says.

Among his algorithms, Salzberg wrote a popular gene-finding program in



1998 known as GLIMMER. He made it open source. Today, it's been downloaded and used by scientists across the world to assemble thousands of genomes.

"Making genomics software open source lets people focus on the science without worrying about licensing fees or other restrictions," Salzberg has said in the past.

Salzberg now develops computational methods for finding the genes within the morass of DNA data coming out of these high-throughput sequencers. The software programs—what are known in the computer science world as string alignment algorithms—are able to sift through the DNA, find the genes, and then piece them together into a cohesive picture of a genome.

Salzberg will use software that he and others have created to read the short DNA data from the seed, which is very accurate, and the long data from the needles, which is bountiful, to locate the genes. "When we're done, the redwood genome will be the biggest genome ever assembled," Salzberg says, "and that's a serious technical challenge."

Once the tree genomes have been sequenced and assembled, the data will go back to Neale's lab for the third and final step, annotation, which makes the genome more readily legible for those who wish to use it in the field. The cost to sequence the genomes is around

Once Timp and his team finish sequencing both trees, they will have produced a bulk of data so immense that it requires offsite storage on a cloud server.

\$1 million—a fraction of the human genome's cost—and Neale says they hope to have both trees completed by this October, in time for the 100th anniversary of the Save the Redwoods League.

"We're hoping that in the next two years we will have the first tools that will allow us to do genotyping work in the field," Burns says. "As we wait for those tools, we're already collecting samples from the forests."

As Burns and her team walk the forests, they encounter the many tourists who flock to see these giant trees. The visitors' reactions, Burns says, are often spiritual. The forests "invoke in many the feeling of being in a cathedral because the canopy is so high above you and because of the way the light filters through it," she says. "The forests feel prehistoric and show us what it looked like on the planet before people were around."

Standing in the forests today, with shafts of crepuscular light piercing the thick canopy and the distinctly piquant citrus scent of these trees filling the air, humans get to commune with some of the oldest living organisms on earth. Every year, thousands make the pilgrimage. We seek out the redwoods and the giant sequoias because, as Muir wrote of these forests years ago, "I am learning nothing in this trivial world of men. I must break away and get out into the mountains to learn the news." These wise stanchions of time support a whole world in their branches; they contain the story of centuries. And they have much to teach us. /