

ELEMENTS

THE RACE TO SAVE THE WORLD'S DNA

*A scientific rescue mission aims to analyze every plant, animal, and fungus
before it's too late.*

By Matthew Hutson

August 9, 2023

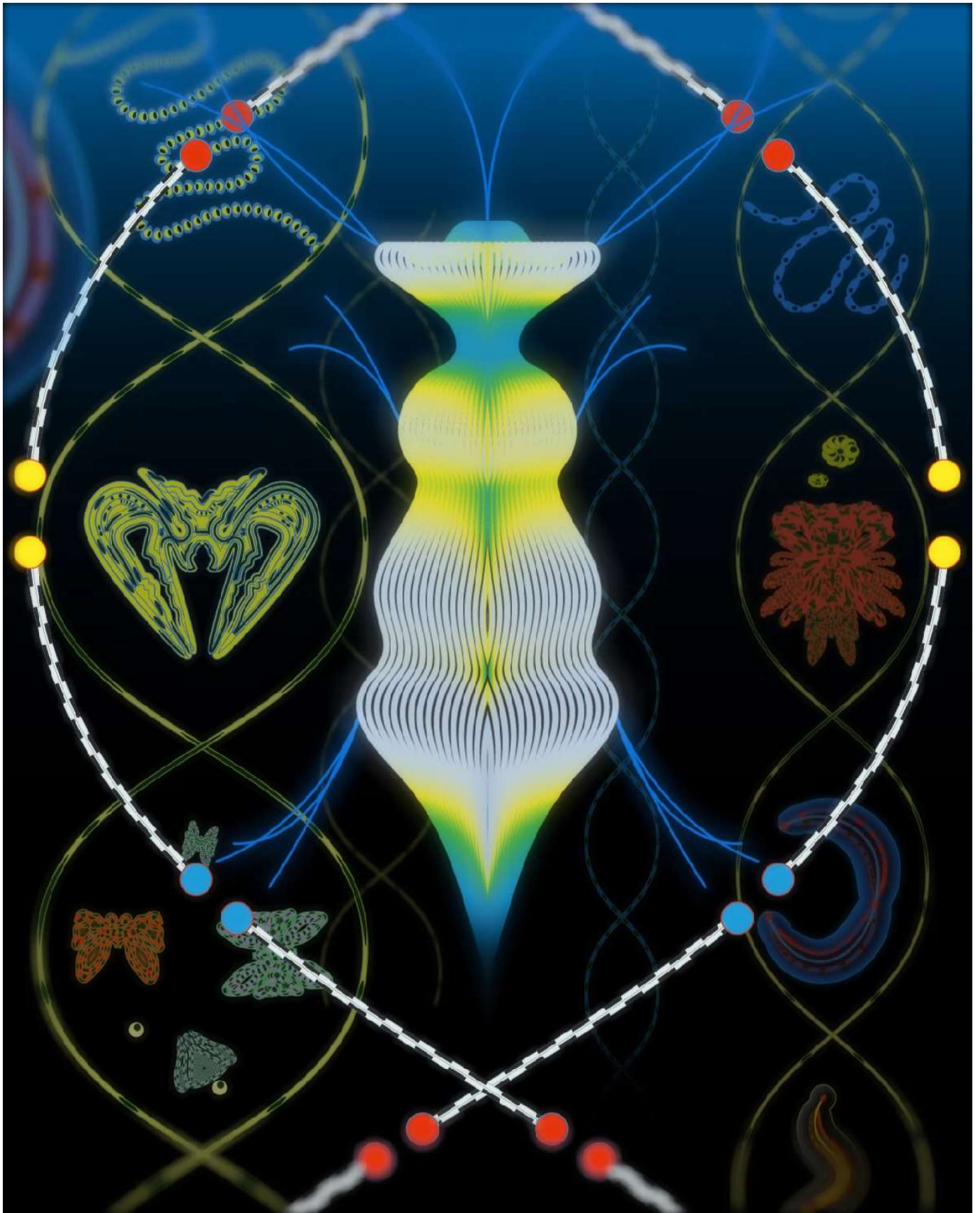
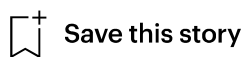


Illustration by Petra Péterffy



Four years ago, a few hundred miles off the coast of West Africa, a crane lifted a bulbous yellow submarine from the research vessel Poseidon and lowered it into the Atlantic. Inside the sub, Karen Osborn, a zoologist at the Smithsonian Institution who was swaddled in warm clothes, tried to ward off nausea. During half an hour of safety checks, Osborn watched water slosh across the submarine's round window, washing-machine style. Then the crew gave the all-clear and the vessel descended. In the waters of Cape Verde, a volcanic archipelago that is famous for its marine life, Osborn felt the seasickness dissipate. She pressed her face against the glass, peering out at sea creatures until her forehead bruised. "You're just completely mesmerized by getting to look at these animals in their natural habitat," she told me.

Osborn was on a mission to find several elusive species, including a bioluminescent worm called *Poeobius*, and to sequence their genes for a global database of DNA. "We need the genome to figure out how these things are related to each other," she explained. "Once we have that tree, we can start asking interesting questions about how those animals evolved, how they've changed through time, how they've adapted to their habitats." Eventually, such genomes could inspire profound innovations, from new crops to medical cures. Osborn was starting to worry, however: she had already made several trips in the submarine and had not seen a single *Poeobius*. Each worm measures just a few centimetres in length and feeds on marine snow, or organic detritus that falls from the surface. Because it is yellow on one end, like a cigarette, it is sometimes called the butt worm.

As the pilot steered into deeper waters, Osborn operated a suction hose at the end of a robotic arm. Whenever she spotted organisms that she wanted to sample—

crustaceans, sea butterflies, jellies—she'd suck them through a tube and into a collection box that was filled with seawater. She started to wish that the submarine had a rest room on board. Then, a few hundred metres down, she finally saw a group of *Poeobius*. “Oh, that's what we want!” she remembers exclaiming. “Go! Go get that!” The pilot slowly turned the sub and Osborn sucked up the worms.

Back on the ship, even before using the rest room, Osborn deposited her boxes in an onboard laboratory. “It's always exciting to climb out and go look at all the samplers, and take them into the lab and see what animals you've gotten,” she told me. She placed one of the *Poeobius* worms under a microscope, anesthetized it, sliced off a bit of gelatinous tissue, and placed it into a vial, which contained a liquid that would protect the DNA from deterioration. (The butt worm did not survive.) Back at the Smithsonian, a team would extract the genetic material and sequence it. It would soon become a new branch on a growing tree of life.

The evolution of life on Earth—a process that has spanned billions of years and innumerable strands of DNA—could be considered the biggest experiment in history. It has given rise to amoebas and dinosaurs; fireflies and flytraps; even mammals that look like ducks and fish that look like horses. These species have solved countless ecological problems, finding novel ways to eat, evade, defend, compete, and multiply. Their genomes contain information that humans could use to reconstruct the origins of life, develop new foods and medicines and materials, and even save species that are dying out. But we are also losing much of the data; humans are one of the main causes of an ongoing mass extinction. More than forty thousand animal, fungal, and plant species are considered threatened—and those are just the ones we know about.

Osborn is part of a group of scientists who are mounting a kind of scientific salvage mission. It is known as the Earth BioGenome Project, or E.B.P., and its goal is to sequence a genome from every plant, animal, and fungus on the planet, as well as from many single-celled organisms, such as algae, retrieving the results

of life's grand experiment before it's too late. "This is a completely wonderful and insane goal," Hank Greely, a Stanford law professor who works with the E.B.P., told me. The effort, described by its organizers as a "moonshot for biology," will likely cost billions of dollars—yet it does not currently have any direct funding, and depends instead on the volunteer work of scientists who do. Researchers will need to scour oceans, deserts, and rain forests to collect samples before species die out. And, as new species are discovered, the task of sequencing all of them will only grow. "That's a heavy aspiration that will probably never be entirely achieved," Greely, who is seventy-one, told me. "It's like, when you're my age, planting a young oak tree in your yard. You're not going to live to see that be a mature oak, but your hope is somebody will."

For hundreds of years, biologists have roamed the globe in an epic effort to collect and categorize the life on Earth. In the seventeen-hundreds, after traversing Sweden to document its flora and fauna, Carl Linnaeus helped create the system that scientists still use to classify and name species, from *Homo sapiens* to *Poeobius meseres*. In 1831, Charles Darwin set out aboard H.M.S. Beagle to collect living and fossilized specimens, which inspired his theory of natural selection. The discovery of DNA, in the nineteenth century, offered a new way to classify species: by comparing their genetic material. DNA's four building blocks—adenine (A), thymine (T), guanine (G), and cytosine (C)—encode profound differences between organisms. By studying their sequence, we might come to speak life's language.

Scientists didn't even begin to sequence a DNA molecule until 1968. In 1977, they sequenced the roughly five thousand base pairs in a virus that invades bacteria. And, in 1990, the Human Genome Project started the thirteen-year process of sequencing almost all of the three billion base pairs in our DNA. Its organizers called the endeavor "one of the most ambitious scientific undertakings of all time, even compared to splitting the atom or going to the moon." Since

then, researchers have been filling in gaps and improving the quality of their sequences, in part by using a new format known as a telomere-to-telomere, or T2T, genome. The first T2T human genome was sequenced only last year, but already scientists with the Earth BioGenome Project are talking about repeating this process for every known eukaryotic species. (Eukaryotes are organisms whose cells have nuclei.)

Because the E.B.P. does not have its own funding, it does not sample or sequence species on its own. Instead, it's a network of networks; its organizers set ethical and scientific standards for more than fifty projects, including the Darwin Tree of Life, Vertebrate Genomes Project, the African BioGenome Project, and the Butterfly Genome Project. This way, "when we get to the end of the project, it's not the Tower of Babel," Harris Lewin, an evolutionary biologist at the University of California, Davis, who chairs the E.B.P. executive council, told me. "You know—your genomes are produced this way, and mine are produced that way, and they're of different quality, so that, when you compare them, you get different results."

By 2025, the participants hope to assemble about nine thousand sequences, one from every known family of eukaryotes. By 2029, they aim to have one sequence from every genus—a hundred and eighty thousand in all. After the third and final phase, which could be completed a decade from now, they aim to have sequenced all 1.8 million species that scientists have documented so far. (Roughly eighty per cent of eukaryotic species are still undiscovered.) This database of genomes, including annotations and metadata, will require close to an exabyte of data, or as much as two hundred million DVDs. The amount of information involved is more than "astronomical," Lewin said; it's "genomical." He compared the project to the Webb Space Telescope, which received about ten billion dollars of government funding. Given how much these projects change the way that humans see the world, Lewin said, "the cost is really not that much."

Natural-history museums already have some of the samples needed to outline a genetic tree of life. The Smithsonian, for instance, has about fifty million biological samples. But, because DNA degrades quickly, it's difficult to extract a high-quality sequence from, say, a frog in formaldehyde or an old taxidermy parrot. For this reason, the E.B.P. usually restricts itself to recent samples, which are often frozen. It relies on the Global Genome Biodiversity Network to keep track of who has what; another database, called Genomes on a Tree, tracks which species have been sequenced already, and whether they meet exacting standards. Scientists such as Osborn will have to find the rest—and their jobs will only become more difficult as the low-hanging fruit is plucked.

After Osborn collected her butt worms, she had to transport them to her colleagues at the Smithsonian. This process can be more difficult than it sounds. Many researchers keep their samples intact by packing them with dry ice or liquid nitrogen in the field; airport-security workers sometimes flag these packages as suspicious, leading to delays that can spoil the DNA and waste an expedition. Osborn, for her part, checked a large insulated box on the flight from Cape Verde, and then waited a few hours in Newark for Fish and Wildlife officials to approve it for entry. As it turned out, her samples came from an entirely new species of *Poeobius*; a paper announcing the discovery is forthcoming.

The first stop in the journey from sample to sequence is a genetics laboratory such as the Vertebrate Genome Lab, at the Rockefeller University, on the eastern shore of Manhattan. On a drizzly day last May, I visited the V.G.L. to see how scientists turn a bit of animal tissue into a string of billions of letters. Olivier Fedrigo, a bespectacled geneticist who was then the lab's director, led me down a hallway decorated with photos of species that had been sequenced there: a snake, a swan, a shark. It was a kind of trophy wall on which inclusion signified not death but a kind of immortality.

Researchers extract DNA from animal tissue in a biosafety-level-two room, which requires goggles, gloves, coats, and special ventilation to protect people and samples. Nivesh Jain, a scientist who works there, told me that he minces the tissue and places it in a lysis buffer—a chemical that breaks open cells—and then uses one of two methods to get the DNA out. The first is a type of microscopic magnetic bead, which is treated with chemicals that help it stick to genetic material; magnets hold the beads and their attached DNA in place while Jain washes everything else away. The second is a glass wafer called a Nanobind disk, which similarly sticks to DNA while Jain removes the rest of the sample. When we met, Jain was standing at a lab bench, checking the concentration of DNA in a vial. The vial would then go to another room, where Jennifer Balacco, the lab-operation lead, would pipette pieces of extracted DNA into little plastic tubes. Special enzymes attach short, recognizable pieces of DNA, called adapters, to the animal DNA, which readies them for the sequencer.

Finally, the samples travel into refrigerator-size PacBio sequencing machines, which, in this case, were labelled with nicknames from “Star Trek.” Enzymes latch onto the adapters and traverse the strands, attaching a color-coded molecule to every building block of DNA. The machine detects the colors and “reads” the sequence that they represent.

It's not enough to sequence DNA in pieces: scientists must figure out how each fragment connects to make a genome. Genomes tend to be bundled up in complicated shapes. A technique called Hi-C mapping “helps you to sort out the puzzle pieces,” Fedrigo told me. The resulting map of folded DNA is crowded with colorful squiggles. At some computers down the hall from the sequencers, the maps help another team of researchers assemble sequence fragments into a full T2T genome. Nadolina Brajuka, a bioinformatician, was assembling an Asian-elephant genome. “I can physically use key and mouse controls and pick pieces of the genome up and move them around,” she said. The last step is for a “data wrangler” on the team to upload the raw-sequence data file, the final genome

assembly, and background information about the sample—including where, when, and how it was collected, and a photo of the species—to a public server called GenomeArk.

One goal of the E.B.P. is to compare and contrast large numbers of genomes, revealing how they are related. Benedict Paten, a computational biologist at the University of California, Santa Cruz, has developed software to align genomes and determine which genes correspond to one another. “It’s a really rich and difficult problem,” he told me, “because genomes evolve by a bunch of really complicated processes.” For a 2020 *Nature* paper, Paten and several collaborators used powerful computers to align more than a trillion As, Ts, Gs, and Cs and create a tree of six hundred bird and mammal species. On a typical home computer, such an undertaking could have taken more than a million hours. “If you wanted to do it for all plants and animals, it’s just a vast computational challenge,” Paten told me.

During my trip to the Rockefeller University, I visited Erich Jarvis, a well-dressed neurogeneticist who leads the Vertebrate Genomes Project, and asked him to show me the kinds of experiments that the E.B.P. will unlock. Jarvis, the son of two musicians, grew up in Harlem and originally trained as a dancer; today he studies the genes that help animals learn to imitate sounds.

We walked through Jarvis’s expansive laboratory toward a scientist who was peering through a microscope at a bird embryo. In this early stage of development, the scientist explained, it was possible to inject the embryo with cells that contain modified DNA. When the so-called transgenic bird hatched, the lab would be able to study whether the foreign genes affected its ability to learn songs.

A nearby room was filled with caged birds and mice; speakers played sounds while cameras and microphones recorded how animals responded. I bent down to look at a zebra finch, which was chirping away. A surprisingly small number of animals have been shown to imitate sounds, Jarvis told me: songbirds, hummingbirds,

parrots, dolphins, whales, seals, bats, elephants, and humans. Figuring out what these animals have in common could help us understand the genetic roots of spoken language. This kind of research, Jarvis went on, is possible only with high-quality complete DNA sequences.

“We humans would benefit so much from nature’s experiment,” Jarvis said. Some species are resistant to SARS-CoV-2. Some, including parrots and elephants, rarely get cancer. Some crops produce more food than others. “We’re going to lose that information if we don’t do something about it soon,” he said. The E.B.P. could also empower scientists to study the health of ecosystems. A researcher with access to full genomes can sample some pond water and figure out which species are living there. Such studies could help humans reverse the harms of agriculture, urbanization, and climate change—and fulfill what Jarvis called a “moral duty” to save fellow-species.

The Earth BioGenome Project “is going to blow the door wide open on conservation genomics,” Bridget Baumgartner, who works for an organization called Revive & Restore, told me. Her project, Wild Genomes, is trying to use DNA for the management of endangered species. In Bolivia, scientists are sequencing jaguars to determine which population individual jaguars came from, and also to track illegal wildlife trafficking. In the Mojave Desert, researchers are comparing the genomes of trees that survive in different temperatures, so they’ll know which individuals of that species could be planted in other places as the climate changes. And, in the archipelago of Indonesia, binturongs have been rescued from smugglers and returned to their specific island of origin, which can be determined through DNA. The other part of Revive & Restore aims for the de-extinction of lost species such as the passenger pigeon, with help from the genomes of living animals. Much of the funding for this work originally came from wealthy Bay Area tech investors—“not the typical conservation funder,” Ryan Phelan, Revive & Restore’s executive director and co-founder, said—but increasingly comes from governments.

Right now, the sequencing process is so cumbersome that scientists can't hope to repeat it a million-plus times in the coming decade. To achieve the necessary pace of hundreds of genomes a day, they will need to automate much of it, perhaps with robots that can prepare samples and improved algorithms that can assemble genomes—though the bottleneck, Lewin stressed, is still the sampling. Of course, all of this will require funding. There's little precedent for a government project that touches so many scientific fields, Lewin told me. "In the U.S., if you can eat it, U.S.D.A. will fund it. If it'll kill you, N.I.H. will fund it. If it's good for energy production, the Department of Energy will fund it. And, if you have some interesting scientific questions, the National Science Foundation will fund it. But there's no agency that owns it all." For that reason, Lewin said, the E.B.P.'s organizers are less focussed on assembling a patchwork of grants than finding what he called "a visionary philanthropist."

Sooner or later, a global database of genomes will have profound practical implications. Some creatures can regrow their limbs; others do not appear to die unless they suffer an injury. If the basis for such traits can be pinpointed in genes, humans might be able to borrow them, perhaps by using gene therapies. "Evolution has already done nearly every experiment, right?" Lewin told me. "There are organisms that'll eat oil spills, there are organisms that'll eat heavy metals. I mean, it's incredible." But, when genomes inspire new products, to whom will they belong? This question makes the E.B.P. not only a scientific project but a political one.

In the nineties, scientists from the Human Genome Project argued that DNA sequences should be in the public domain, meaning that anyone, anywhere, would be able to use them. "That has been an animating principle for genomics for the past, like, thirty years," Jacob Sherkow, a professor at the University of Illinois College of Law, told me. More recently, views have changed. " 'Public domain' is a deceptive term used to deny Indigenous peoples rights from things important to

them,” Ben Te Aika, an expert on the traditional knowledge of the Māori people, in New Zealand, told me. “It would be more honest to say ‘domain of the élites.’” In the two-thousands, many observers worried that wealthy nations would exploit biological samples without compensating the countries that they come from. This concern helped inspire the Nagoya Protocol, a piece of international legislation that encourages “benefit sharing,” and instructs countries to agree on terms before biological samples are shared. More than a hundred countries have ratified it. (The U.S. is not one of them.)

Te Aika told me that, after centuries of European colonialism, his community has been reasserting its *mana*, or traditional authority, over native species. He argues that the Māori people should have the opportunity to benefit from any scientific samples that are gathered in New Zealand. With a colleague from Ireland, Ann Mc Cartney, Te Aika has co-authored papers in support of data sovereignty, or the right of local and Indigenous people “to control data from and about their communities, land, species, and waters.” They described E.B.P. as “an opportunity to leave no one behind.” The scientific collaboration that Te Aika works for, Genomics Aotearoa, is not affiliated with the E.B.P. and has adopted an unusual structure: its data is accessible only to researchers who apply and are invited to travel to New Zealand. Outside scientists may see such restrictions as a kind of red tape, Te Aika said, but “‘red tape’ can become necessary when self-regulating systems fail.”

Several scientists told me that the Nagoya Protocol is already outdated. “Benefit sharing in the Nagoya Protocol is getting more strict and confusing,” in part because of debates about how to interpret it, Jarvis said. Currently, he argued, the protocol is discouraging scientists from developing products at all—an outcome that, in his view, helps no one. One argument for commercializing genomes is that “then you can get financial benefit going back to the people that are the caretakers of the land where the animal came from,” he said. “Something has to change.”

The most complex debate, Sherkow told me, is about whether a digital DNA sequence counts as a biological sample. If not, the Nagoya Protocol wouldn't apply to the strings of letters stored in the E.B.P., and, as Sherkow put it, "It's everyone for themselves." Any scientist, company, or country could download a sequence and use it for their own ends, without consulting or compensating the community that the sequence originated from. But, if the sequence *is* a sample, then genomes will be governed by Nagoya, and many difficult questions will follow. How should the benefits of a discovery or product be shared? Are they owed to the country that the sequence came from, or someone else, such as an Indigenous group? Communities need an opportunity to voice their own priorities: some may want to build capacity for their own research, and others may want compensation or simply credit for their contributions to a discovery. Some of the scientists I spoke to felt that new international laws would need to be written to answer these questions.

The E.B.P. has formed an Ethical, Legal, and Social Issues Committee to work through such challenges. Sherkow described its work as a balancing act: "What's best for science? What's best for the world? What's best for the particular country that we're taking samples from?" Greely, who chairs the committee, said that it also develops best practices in other areas: interactions with local communities, the humane treatment of animals being sampled, whether to sample in countries ruled by "nasty regimes," authorship on papers, and even risks of bioterrorism. He added that he was stunned to learn how many international treaties affect biological resources—treaties on food and agriculture, migratory species, whaling, the law of the sea, and more. "A lot of the hangups are not scientific or even engineering hangups," Sherkow told me. "The biggest hangup to sequencing all the world's non-human eukaryotes is humans."

The quest to document life spans scientific disciplines, continents, and generations. Darwin first drew a tree of life in his notebook around 1837; two hundred years later, the E.B.P. could finish some of what he started. Last May, Mark Blaxter, an evolutionary biologist in the U.K. who contributes to the project

and is the director of the Darwin Tree of Life, sat down in the grass in his back yard, cracked open a beer, signed on to Zoom from his laptop, and told me about the new era of biology that he foresees. Periodically, Blaxter, who has long white hair and a graying beard, interrupted himself to identify the creepies that were crawling around him: ladybug, bee, pill bug. “There’s two species of ant on this piece of grass,” he observed. “Only one of them’s biting me, though.”

Charlotte Wright, a twenty-five-year-old doctoral student who likes catching bugs, was drinking a beer with Blaxter that day. Wright studies moths, which, along with butterflies, make up a tenth of all known eukaryotic species. They, too, are mysterious. Human genomes typically have twenty-three pairs of chromosomes; *Lepidoptera* can have anywhere from five to two hundred and twenty-six. “That gives them the greatest range in chromosome number of any group of organisms on Earth,” Wright said. “They’re completely bonkers.” Because it’s difficult for animals with different numbers of chromosomes to produce offspring, studying chromosome evolution can shed light on how one species diverges into many—one of biology’s fundamental questions.

Blaxter watched a bee fly into his house. Then he reflected on the many drugs that have come from the natural world over the years. Aspirin was first derived from willow bark, which was used to relieve pain since ancient times. “We think that by sequencing, for example, fungi, there’s going to be a huge new pharmacopoeia opened up,” he told me. “Think about the transformative effect that the human genome had on our understanding of human biology and medicine and disease and health. We want that to be available for everyone.”

When Blaxter became a biologist, in the eighties, scientists had not even begun to sequence the human genome. Back then, “biodiversity” was still a new term; humans were only starting to grasp just how many species were vanishing forever, and how much our activities were transforming the planet and its climate. Blaxter, who is sixty-three, seemed conscious that he might not live long enough to see all

the impacts of the genomic revolution. “I’m on my way out,” he said. “I’m the old generation, right?” Wright’s generation would inherit unprecedented challenges, but she would also build on an unprecedented foundation of knowledge about the natural world. “Charlotte’s going to be one of the first generation of genome natives,” Blaxter told me. “What we want to do with this project is to change the way biology is done forever.” ♦

MORE SCIENCE AND TECHNOLOGY

- Can we stop runaway A.I.?
- Saving the climate will depend on blue-collar workers. Can we train enough of them before time runs out?
- There are ways of controlling A.I.—but first we need to stop mythologizing it.
- A security camera for the entire planet.
- What’s the point of reading writing by humans?
- A heat shield for the most important ice on Earth.
- The climate solutions we can’t live without.

Sign up for our daily newsletter to receive the best stories from *The New Yorker*.



Matthew Hutson is a contributing writer at *The New Yorker* covering science and technology.

More: [Biodiversity](#) [Genome](#) [DNA](#) [Science](#) [Conservation](#)

You're all set.

Thank you for signing up for this newsletter.

READ MORE

DAILY CARTOON

Daily Cartoon: Thursday, July 13th

“Well, maybe *you* shouldn't have spent a lifetime painstakingly honing a set of skills that you must have known we would eventually attempt to steal and replicate for free.”

By **Zoe Si**

BLITT'S KVETCHBOOK

Salvador Dalí's Surreal Summer in the City

What isn't melting in this heat?

By **Barry Blitt**

DAILY SHOUTS

Lies I Want to Tell My Child

But don't because I'm the grownup.

By **Trevor Alixopoulos**

ART

Cutting-Edge Image-Makers, in “New Photography 2023”

The Museum of Modern Art's program returns, after a hiatus of five years, to celebrate seven photographers connected to Lagos, Nigeria.

