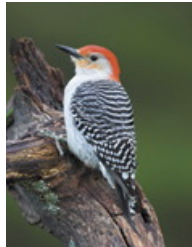


Barcoding life



Evolutionary biologist Paul Hebert, overshadowed by the projection of a DNA barcode (OPPOSITE) at the Biodiversity Institute of Ontario in Guelph, is leading the Canadian effort to use a standardized gene region to classify all species and chart the diversity of life.

Identifying each of the 15 million or so species on the planet may seem an impossible task. Can a genetic shortcut developed by a Canadian scientist do the job in mere decades?

BY SIOBHAN ROBERTS

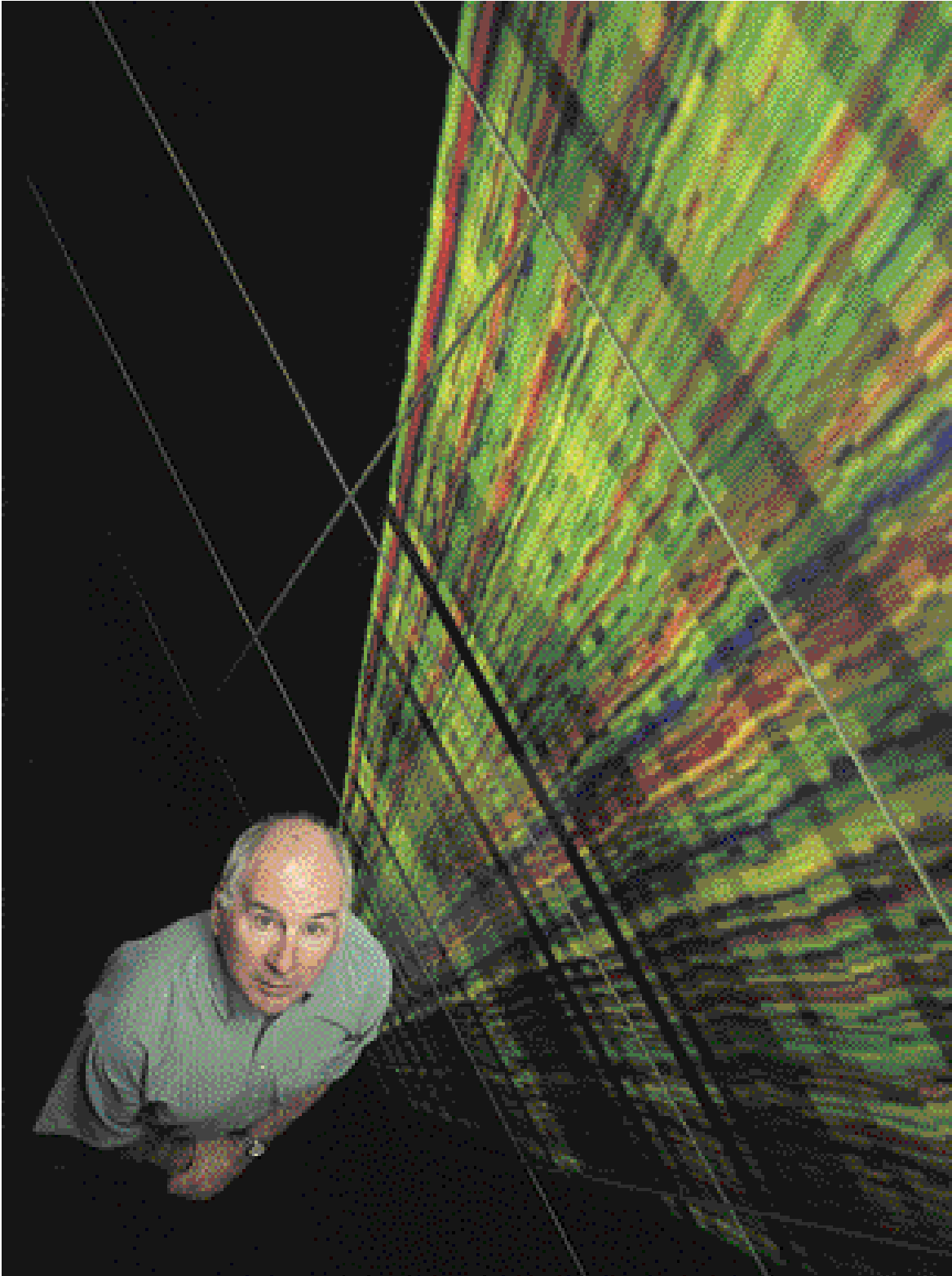
Life is short in Churchill, Man., where ice lingers on Hudson Bay until July and by September, it's snowing again. Even with these limitations, the tundra teems with activity and beckons biodiversity hunter Paul Hebert like a pet store to a wide-eyed child. Over three weeks last summer, he conducted a "biodiversity blitz" in Churchill — a census of all the organisms he could get his hands on.

Hebert, an evolutionary biologist at the University of Guelph, along with about 30 staff and students grabbed samples from the flotsam parts they found lying about: feathers from the ptarmigan and hawk-eyed snowy owl, tufts of hair from the woodland caribou, skin swabs from the beluga whale, along with specimens of caplin, fairy shrimp and tiny jet-black water fleas.

"Our prize catch occurred just two days ago," Hebert reported in an e-mail from the field. "We found a specimen of the black witch moth on a rock bluff right beside Hudson Bay. This is a migratory species that breeds in Mexico. It now stands as the most northern record ever for this species; it beats a 1957 record from Juneau, Alaska!"

Back in the lab, the team is reading a snippet of DNA from the tens of thousands of invertebrate specimens and the hundred or so vertebrate samples it collected. "We are not out slaughtering organisms," explains Hebert. The Churchill expedition deployed what some consider a revolutionary new taxonomical tool: a standardized method for identifying species using a short DNA sequence from a common locality on the genome. Hebert debuted this technique in a 2003 article in the *Proceedings of the Royal Society of London*, where he proclaimed that "these sequences can be viewed as genetic 'barcodes' that are embedded in every cell."

With that passing metaphor, conceived in a flash during a trip to the supermarket, Hebert's invention has



DAVID TRATTLES, OPPOSITE, TOP TO BOTTOM: ROBERTMCCAW.COM, WWW.MIKEGRANDMAISON.MB.CA, GLEN AND REBECCA GRAMBO, ROBERTMCCAW.COM, RANDY MORSE/MAXX IMAGES

Some 650 lines — each colour representing one of four nucleotides, the building blocks of DNA — make up a DNA barcode (BELOW). Whether looking at the barcode of a scarab beetle (BOTTOM) or of a merganser (OPPOSITE), the sequence is unique for each species.

HUMAN



He describes his laboratory as an international clearing house, receiving ‘body bits’ FedExed from all latitudes and longitudes.



become widely known as “barcoding life.” In four years, it has gained considerable currency. The Consortium for the Barcode of Life (CBOL), with members from more than 130 organizations representing 40 nations, is based at the Smithsonian Institution in Washington, D.C. Hebert’s Barcode of Life Data Systems, an online resource listing all species barcoded so far — almost 26,000 and counting — is housed at the Canadian Centre for DNA Barcoding, a CBOL member and key occupant of the spiffy new Biodiversity Institute of Ontario, which opened its doors at Guelph last October.

The facility is equipped with a four-by-four-metre plasma wall and comfy theatre seats for reviewing results, conducting workshops and teleconferencing. “I’m tired of getting on planes,” says Hebert, who is swamped daily with e-mails from colleagues on all continents. “It’s barcoding pyromania,” he says, noting that science at its best is pursued by coalition. He describes his laboratory as an international clearing house, receiving “body bits” FedExed from all latitudes and longitudes. Other sites doing similar work based on Hebert’s protocols have processed a few thousand specimens; his laboratory has rendered more than 150,000.

Hebert’s initiative combines the science of “CSI” with Wal-Mart’s superstore efficiency. The superficial connotations of that analogy, however, don’t respect the complexity of an organism, suggesting that any form of life could be taxonomically classified in such a one-dimensional, assembly-line manner. This is what makes DNA barcoding not only a potentially revolutionary taxonomic tool but also a controversial one.

FOR HEBERT, who carries himself with the energy of someone constantly high on adrenaline, the equation is simple: DNA barcoding offers a solution to a taxonomic problem.

“I’m fascinated with biodiversity, but I’m intellectually inadequate for the challenge,” says Hebert, his cheeks full of colour and his legs elegantly crossed but antsy, jostling his lanky frame to and fro in his office chair. “The human mind is not capable of dealing with biodiversity on a global context.”

The issue is simple but vast. In the nearly 250 years since Swedish botanist and physician Carolus Linnaeus laid the foundations of a system for describing, enumerating and classifying organisms, fewer than two million of the estimated 15 million species on the planet — some estimate 100 million — have been identified. Traditional, or “integrative” taxonomy, comprehensively analyzes all aspects of an organism — its appearance, behaviour, feeding and migratory patterns and molecular makeup — before drawing conclusions about classification. It is a labour-intensive and time-consuming endeavour. “Species are dying off before we know they exist,” says Hebert.

The kaleidoscopic diversity of life underpins all biological studies and makes nature the awe-inspiring wonderland that it is, but diversity is also an onerous burden. “Physicists deal with a cosmos assembled from 12 fundamental particles,” says Hebert, “but biologists confront a planet populated by millions of species. Their discrimination is no easy task. A community of 15,000 taxonomists will be required in perpetuity to identify life if our reliance on morphological diagnosis is to be sustained.”

Recognizing and worrying about that crisis led the biologist to develop the DNA barcoding tool.



DNA barcoding revealed that a single species of butterfly, known for 250 years as *Astrartes fulgerator*, is, in fact, 10 different species.

HEBERT CAN QUICKLY summon from memory the genesis of his current obsession. He grew up in Kingston, Ont., and in 1951, at age four, he was already a lover of critters of all kinds. Catching a bumblebee perched on a peony, he ran to show it to his mother but tripped and fell, jamming a piece of glass from the bug jar into his hand. He still bears the scar, thrusting his palm out for inspection. “I wasn’t smart enough to stop,” he says of the path he’s been on ever since.

By the time he was six, Hebert was breeding moths, and later, at Queen’s University, he studied biology. His doctorate in population genetics at Cambridge in England looked at what happens to animals when they give up sex (essentially, they begin to clone). His studies of animal asexuality would prove to be a fortuitous introduction to molecular science.

Through it all, his interest in insects abided. He set up ultraviolet lights in his Cambridge dining room, curious to see what unusual English fare he could procure. The lamps lured bugs with a come-hither glow, then trapped them on glue board. To his delight (and his wife’s dismay), moths and their ilk took up residence in the curtains and in crevices and once flew out en masse during a dinner party, besieging his guests. To this day, Hebert finds it difficult to sleep at night without knowing he has bug lights to check when he awakens. From March to December, visitors to his rural property in Guelph are greeted with the blue glow of UV lights in his front yard, while a farm of mercury vapour lamps illuminate the backyard. “There is nothing more enjoyable than turning on a light — well, maybe a few things are more enjoyable — and seeing what magical things congregate,” he says.

Hebert is particularly into moths, part of the second largest insect order, Lepidoptera. “I’m interested in moths

that do unusual things. That’s all magical for me. All moths are interesting. Even the little one of them,” he says, adding a punchline, “What is considered dementia in normal activity of daily life is acceptable in academia.”

During his career, Hebert has done stints at the University of Sydney in Australia and the Natural History Museum in London, England, and he spent more than a decade at the University of Windsor before migrating to Guelph, where, 16 years ago, he began prying into the molecular constitution of organisms. “Using molecules to tell species apart was happening 30 years ago,” he says, “but it wasn’t scalable. The science could tell five things apart, but five million was not doable. When you’re like me, you’re greedy.”

By the early 1990s, polymerase chain reaction (PCR) technology began to replace the ultra centrifuge-based technology used in isolating DNA for sequencing. The latter could process 12 samples every three days. PCR — essentially a photocopy of DNA — allowed 100 samples every three hours. Genomic technology advanced around the same time, providing more efficient DNA sequencing.

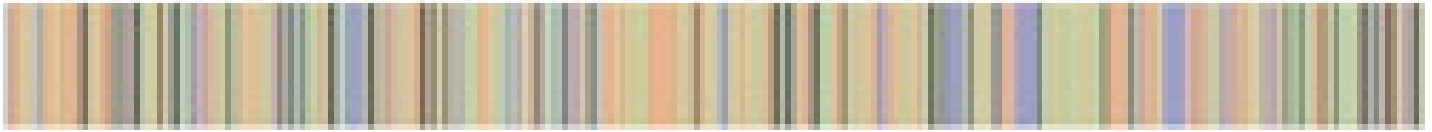
Hebert initially tested the technology on 200 Lepidoptera species from his bug lights. He was hardly the first to suspect there was a species-specific DNA tag. But it was a haphazard science, with molecular miners choosing to study the gene regions they pleased. Hebert was the first to suggest that focusing on one standard gene region was an effective marker for differentiating species. He chose the mitochondrial genome of animals, specifically the cytochrome *c* oxidase 1 (CO1) gene. He decided that mitochondria, the so-called power plants of cells, were the best target for analysis, since they have an isolated DNA loop and the gene sequences change quickly.

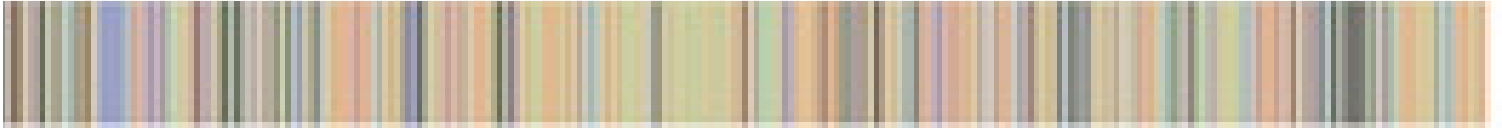
A single hair from a Rocky Mountain bighorn (OPPOSITE) or a tiny piece of wing from a giant brown grasshopper (BELOW) is all that is needed to produce a species-specific barcode. This allows scientists to easily gather a range of plant and animal specimens for inclusion in the Barcode of Life Data Systems, based in Guelph.



Researchers and amateurs alike could soon be identifying species, such as the *Protesilaus* swallowtail (BELOW) and the blue starfish, using a hand-held DNA barcoder (rendering, OPPOSITE). Equipped with a GPS system, such a device would enable field scientists to log the geographical location of each specimen they barcode.

IMPERIAL MOTH





Within a decade, Hebert envisions a hand-held device that could identify species in minutes. Making amateur taxonomists of us all can only help the cause of classifying biodiversity.

THE BARCODING BLITZ in Churchill marked the latest expedition in what Hebert calls the “Canada Campaign,” a five-year plan that began in 2005 to classify one-tenth of the country’s species. “Canada is the first country on the planet to do this,” says Hebert. DNA barcode research in Canada, headquartered at the centre in Guelph and involving 50 researchers across the country, has garnered about \$27.5 million in three years from various federal and Ontario government organizations, along with Genome Canada in Ottawa and the Gordon and Betty Moore Foundation in San Francisco, one of the biggest players in biodiversity science. The budget shows impressive momentum, though in the early stages, Hebert says, the project was “starving in the wilderness.”

“The funding is a good poster case for the state of science in Canada,” he says. “It is a bit of a shock when Canada can lead the world in science. Most things end up like the Avro Arrow.” Still, the venture needs broader international support before it is on target to meet Hebert’s goal of classifying all life on the planet within 20 years.

The funds fuel what Hebert calls his “barcode hopper,” a conglomerate of human labour, chemistry and infrastructure, such as liquid-handling robots and two automated DNA sequencers capable of amplifying and sequencing a million individual samples a year. After passing through the assembly line, specimens receive a DNA tag, which is entered into the Barcode of Life Data Systems archive and slotted into the genealogical taxon identification tree, along with the specimen’s photograph and the GPS location where it was found. Hebert has demonstrated that CO1 DNA sequences are 99.75 percent identical within members of a species, while a match of 97.5 percent or less likely indicates another species. “Is that not an amazing system?” says Hebert. “I think it is totally amazing. Any idiot could do it. It’s data entry.”

Indeed, in Hebert’s opinion, one of the most exciting outcomes of DNA barcoding will be the democratization



Traditional taxonomists study the life cycles, habitats, physical features and other attributes of a species, such as the raccoon (BELOW), to determine where it fits in the genealogical tree of life. Barcoding is poised to offer genetic detail that will fine-tune or even redraw the tree's terminal branches.

RUDDY TURNSTONE SANDPIPER

of taxonomy. "We're repositioning humanity's relationship with life," he effuses.

Within a decade, Hebert envisions a hand-held device smaller than a BlackBerry that will do the same work as his laboratory hopper — a device for both professionals and hobbyists in which a sample from an ant, for instance, could be inserted and, within minutes, classified as not just any anonymous ant belonging to one of the 20,000 ant species, but, specifically, as *Formica rufa*. Making amateur taxonomists of us all can only help the cause of classifying biodiversity and conserving endangered species.

"I want to motivate the citizenry to help us," says Hebert. "I see millions of Canadians on Canada Day with UV lights in their backyard sending in specimens," he proclaims, perhaps with tongue only slightly in cheek. "Soon, there will be other people playing in our sandbox."

But, for some, this democratization is cause for concern. Barcoding has been "demonized," says Hebert, by those traditional taxonomists who fear that DNA barcoding will replace the integrative approach. Critics, such as insect systematist Kipling Will at the University of California, Berkeley, say DNA barcoding will create "a telephone book of life rather than an encyclopedia of life."

Will believes DNA barcoding is a regression to the worst sort of taxonomy. Following the Linnaean tradition, a trend developed in taxonomy where, in the 1950s and 1960s, species classification had devolved almost to the level of storytelling and lacked scientific rigour. Taxonomists often named species by the artful whims of their authority, rather than using a critical study of data. "The rest of the biologists looked at [taxonomy] as if someone was collecting stamps. The whole field was definitely having difficulty justifying itself as a science," says Will, adding that taxonomy has since become more scientific and testable in its methodology and thus was just regaining its credibility as barcoding made its debut.

Part of that historical authoritative mentality, says Will, was a reliance on a kind of phenotypic essentialism, or the notion that a species can be recognized by a single characteristic. "It's not too different from saying we can go out and sample just CO1 from all kinds of plants and animals on the planet and that will tell us what the species are," he says. "If we know anything, we know that the history of life on this planet is very complicated, and it just isn't that simple. It's going to take more than one data type to tell us what species are out there."



What barcoding the CO1 gene does provide, notes Will, is a molecular operational taxonomic unit, which some have called "gene species." Essentially, the molecular unit measures the phenetic distance between organisms, or their similarity through observable attributes alone. "The question is," says Will, "is that a meaningful thing? And the only way to really answer that is to have a nice integrative study to know that there is something worth recognizing [as a species]."

Yet there is some consensus, even among the doubters, on the merits of DNA barcoding as a taxonomic tool. At the very least, it successfully flags specimens that have not previously been classified. "What we are doing is shining a spotlight on places where taxonomists should devote some time," says Hebert. At a conservation area in Costa Rica, for example, DNA barcoding revealed that a single species of butterfly, known for 250 years as *Astraptes fulgerator*, is, in fact, 10 different species. As adults, the butterflies are rarely identical, and the fledgling caterpillars vary their stripes and food sources considerably. "That's the place the taxonomists need to rush in and think about the meaning of it," says Hebert. "We think we are setting the table for traditional taxonomy."

Jesse Ausubel, director of the Program for the Human Environment at The Rockefeller University in New York City and a program director at the Alfred P. Sloan Foundation, attended one of Hebert's early presentations on DNA barcoding and immediately saw great prospects for the Sloan's Census of Marine Life, a 10-year assessment of the diversity and distribution of ocean life. A census expedition, says Ausubel, might yield 60,000 specimens, but even a great naturalist would know only a few hundred species by name.

All the hyperbole touting the promise of DNA barcoding is already siphoning funding from traditional taxonomists.

OCTOPUS

“A shrimp expert wouldn’t know fish,” he says, “and a fish expert wouldn’t know copepods, and a copepod expert wouldn’t know nematodes.” Faced with the daunting task of identifying such numerous specimens, Ausubel points out that the criticism of DNA barcoding’s 96.4 percent accuracy rate does not worry him as a user. “If I can identify 964 specimens out of 1,000 [with barcoding], I’m in heaven,” he says. “That means I have to identify only 36 specimens [by traditional methods].”

Ausubel also points out that taxonomists often complain they are inundated with routine identification requests from amateur naturalists, police and customs officers — requests that they say are boring and a waste of their skills. DNA barcoding would reduce the burden, freeing taxonomists to focus on those 36 unidentifiable specimens that need comprehensive study for classification. But Will has a counter-argument. He thinks all the hyperbole touting the great promise of DNA barcoding is already siphoning funding and resources from traditional taxonomy, which will result in fewer taxonomists to do that follow-up work.

THE BROADER APPLICATIONS of DNA barcoding, however, are undeniable. The Feather Identification Lab at the Smithsonian Institution, working with the Federal Aviation Administration (FAA) and the U.S. Air Force, is using the tool to identify bird species that collide with aircraft, a phenomenon the FAA estimates costs about US\$345 million every year. But “snarge,” a Feather Lab term for the goop wiped from an aircraft following a bird strike, is often anatomically unrecognizable. Being able to identify a species allows airfields to implement appropriate habitat management programs, as well as warn crews of particular bird dangers and help engineers design better airplanes.

Barcoding could also be used to identify BSE (bovine spongiform encephalopathy), along with agricultural and forestry pest invasions. “If we stop a single invasive species, our network will pay for itself,” says Hebert. “We will get back many times over the \$30 million [invested].”

But perhaps most compelling is how the biodiversity crisis intersects with the emerging infectious-disease crisis. Dan Brooks, a parasitologist at the University of Toronto, says that knowing what is out there is crucial to tracking disease transmission. “It’s a question of finding them before they find us,” he explains. “The idea is that there would be no emerging infectious-disease crisis if we had a complete inventory of everything on this planet. There would still be disease

problems, but they wouldn’t be these enormous surprises that we have to scramble to catch up to. That’s an immediate breakthrough. Everyone I talk to who works with parasites — even if they are not all that sold on the idea that barcoding is good for basic taxonomy — agrees that tracking life cycles of disease is a really great use of the technique.”

Beyond these practical applications of the technique, DNA barcoding also cuts to the heart of the “big science” debate about the origin of life.

Barcoding has revealed one big surprise: there is very little genetic variation among members of the same species. Because species are millions of years old, it is natural to expect that the years would show in their DNA. Like an old face marked by wrinkles and age spots, gene pools should be scarred with variations and mutations.

In *Homo sapiens*, the lack of genetic diversity has been explained by the so-called Eve hypothesis, which suggests that humanity coalesced into a single female lineage in Africa approximately 150,000 years ago. Many scientists believe that this was the result of population bottlenecks, when lean times reduced human populations to 1,000 or so members and stripped the genetic pool of variation. But barcoding has shown that every animal species seems to have a similarly shallow ancestry. Hebert says it is implausible to think that every species — “every rotifer, every fish, every tiny little soil invertebrate” — went through a similar population crisis. “We are then left with a slightly uncertain mystery,” he says. “What stripped the variation?”

Answering that question could change theories on how life evolves. Hebert suggests natural selection might involve a scouring mechanism, cleansing mitochondrial genomes of their diversity on a regular basis for adaptive purposes. Then the question becomes exactly how the scouring of diversity occurs.

“At this point, that is a scientific mystery,” he exclaims. In getting on with the ontological pursuit for an answer, for the master key of life, Hebert makes a simple demand: “More data collection, please!”

Siobhan Roberts is a Toronto writer whose first book, King of Infinite Space: Donald Coxeter, The Man Who Saved Geometry, was published in 2006 by House of Anansi Press.



To comment, e-mail editor@canadiangeographic.ca.
Visit www.canadiangeographic.ca/web/ma07.