In 2004, University of Pennsylvania biologist Dan Janzen used snippets of DNA to prove that a common American butterfly actually represents at least ten different species. The study demonstrated that "DNA barcoding" can radically speed up species identification.

The Consortium for the Barcode of Life (CBOL) now aims to use barcoding to tag every living organism on Earth. For starters, CBOL plans to barcode 10,000 species of birds and 12,000 marine fishes by 2010. Scientists are testing barcoding's applicability to other animals, plants, and plankton as well.

Biologists are split on whether barcoding represents a great advance in taxonomy or a diversion of scarce research dollars. Here, Academy scientists Peter Roopnarine and Brian Fisher take up the debate.

Today Is Too Soon

by PETER ROOPNARINE

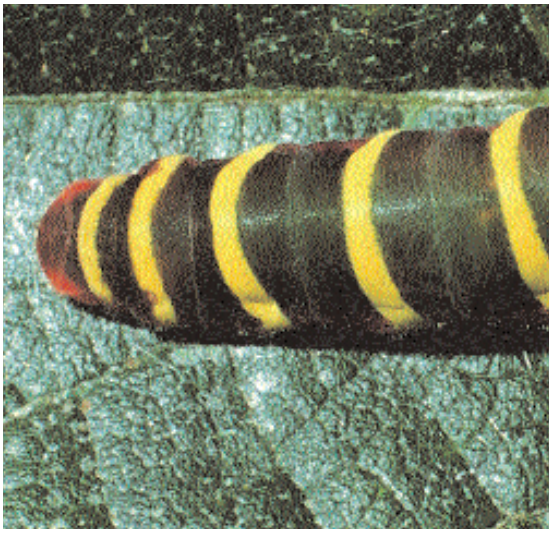
A SET OF DNA AND DATA analytic techniques, known collectively as DNA barcoding, has been proposed recently as a solution to the problem of enumerating the Earth's biodiversity. Estimates of the number of species on Earth range broadly from fewer than 10 million to upward of 50 million. With fewer than 2 million species actually identified by scientists, these estimates are the results of very different interpretations of the uncertainty and incompleteness of our biodiversity surveys.

Given the logistical difficulties of sampling Earth's myriad corners, and the time and expertise normally required for the identification of species, any proposed revolution in methodology must be considered seriously. But is DNA barcoding the revolutionary solution that it purports to be? To answer this question, we need to examine the goals of DNA barcoding, how scientists go about barcoding organisms, and then decide if barcoding is really up to the job.

The basic assumption behind DNA barcoding is that every biological species has a short sequence of DNA that, like a fingerprint, is unique to that species. The sequences should come from parts of the genome that evolve quickly enough to separate species that share a recent common ancestor, but slowly enough to minimize differences among members of the same species. This is a tall order, but supporters of DNA barcoding have identified several candidates. Foremost among these is the mitochondrial gene, cytochrome *c* oxidase I (*cox1*). This gene codes for an enzyme so critical to metabolism that apparently every creature in the animal kingdom has it. Once a *cox1* sequence has been obtained from an animal, that sequence is then compared to a database of already established sequences for identification.

Proponents of barcoding claim that it will help biologists more rapidly identify species, provide a better way to classify them, and serve as the basis for phylogenies (family trees) of groups of species. The claim has also been made that barcoding will allow the efficient identification of previously undescribed species. Sadly, these claims are exaggerated.

Imagine that you are a biologist on a



PHOTOS: DAN JANZEN

These caterpillars look very different from one another but mature into physically identical adults. DNA barcoding recently revealed that the tropical skipper butterfly (Astraptes fulgerator) (over) actually consists of at least ten different species.

AT loggerheads

DNA BARCODING

Tomorrow Is Too Late

by BRIAN L. FISHER

SYSTEMATISTS ARE CHARGED with documenting and describing the history of life on Earth. They search for answers to several fundamental biological questions: What kinds of living things exist? Where do they live? How are they related?

With only an estimated 10 percent of life described on this planet so far, the thought of being able to identify all or most of the world's species might seem like an impossible, idealistic dream. In today's era of accelerating species extinctions, the quest to identify all or most of the world's species is even more daunting: scientists must discover and describe biodiversity before it disappears. Identifying what's out there is key to protecting the future of these species. Taxonomy can help ensure that the wildlands that are conserved will protect the widest array of species possible. In addition, understanding the planet's life forms will undoubtedly put humankind in a better position to understand the essential ecosystem services they provide, and foster the development of new uses for natural products.

Modern technology has presented us with a new and exciting means to identify diversity on this planet. This technique, based on DNA sequencing, will complement the more traditional and painstaking work of morphological taxonomy—describing species by their physical traits. Known as DNA barcoding, it involves reading and comparing the same small segments of genetic data between species. It provides a new source of data that can easily be used to describe species. In addition, large volumes of barcoding data can be generated for relatively low cost. For all of these reasons, barcoding represents a major step forward in the race to describe and conserve biodiversity in the face of rapid species extinctions.

I am now testing the utility of DNA barcoding for uncovering diversity in an ecologically important group: ants. I have teamed up with colleagues at the University of Guelph, Alex Smith and Paul Herbert, to test whether DNA barcoding can accelerate our inventory of

> Traditional taxonomy will not provide enough data in the short term to address Madagascar's urgent conservation needs.

the ants of Madagascar. Our results convince me that the union of DNA barcoding and traditional systematics mark a major advance in twenty-first century science.

Madagascar is one of the world's outstanding biodiversity hotspots. It is populated by a unique biota whose composition and origins are helping scientists piece together the course of evolution since the breakup of the ancient supercontinent Gondwana. Gondwana consisted of what are now Madagascar, India, Africa, Australia, Antarctica and

field expedition to a remote location. Few of the animals you will see have been previously described. You have just collected an unusual animal, one you have never seen before. Without barcoding, you might be flummoxed. But with this technique, proponents say, your problem is easily solved. By taking a small sample of tissue from this animal, you could read a section of its *cox1* gene right there in the field! Your handheld sequencer then uploads the data, via satellite, where it is compared to sequences in the master database. Minutes later, you will receive an identification of your species, or learn that the sequence is not currently in the database. You have either added another species to your list of the area, or just discovered a new species.

Such scenarios, if feasible, would revolutionize the documentation of Earth's biodiversity. The technology isn't so far off, but first let's consider how your DNA sequence will be interpreted.

First, suppose that an exact match to the sequence was not found in the database. Do you have a new species? Maybe, maybe not. To arrive at that conclusion, you had to assume that all animals belonging to the same species have the same *cox1* sequence. This means, for example, that all humans on Earth have identical *cox1* sequences. Now, one could argue that we cannot possibly draw that conclusion, since we have no practical way of sampling *cox1* from every human on the planet. But we could sample many humans, and then estimate how variable *cox1* actually is within our species. If *cox1* is variable in humans or other animals, then how do we know that the *cox1* that we sampled and sequenced from our specimen is indeed unique?

This question is made all the more difficult when you realize that there are actually very few individuals of any single species that have been sequenced and are represented in the database. Barcoding supporters address this problem by pointing out that even though members of the same species will vary in their sequences, the variation among those individuals is far smaller than the differences between species. For example, that would mean that human sequences, though not necessarily identical, will be far more similar to each other than to sequences from, say, chimpanzees. So-called "thresholds" could be established, meaning that two sequences differing by more than, say, a few percent, must come from different species. Where to draw this line is where barcoding runs into its first serious obstacle.

IN THE OLD DAYS, TAXONOMISTS relied almost exclusively on examinations of morphology (skeletal characters, soft-tissue anatomy) to establish species identities (paleontologists in fact still operate in this manner, since fossils almost always only show morphology). This approach requires a great deal of time and expertise, quite in contrast to DNA barcoding. A set of methods known as numerical taxonomy were developed in the 1950s and 1960s, to quantify the process of comparing morphologies and assist taxonomists in their decisions. Numerical taxonomy can actually work quite well for discriminating species.

However, one of the caveats is that individuals of the same species can vary, sometimes dramatically, in their morphologies. When faced with this variation, how does one decide if the collection of specimens belong to a single species? The notion of establishing thresholds of morphological differences was suggested. But as scientists developed ways to analyze the genetic blueprints of organisms, it became clear that you first have to understand how variable morphology is within a species.

As it turns out, there is no simple correspondence between morphological variation and genomic variation. Chimpanzees and humans are quite clearly different species, and we know this from morphology alone, but we share more than 99 percent of our genomes. Different breeds of domestic dog, all belonging to a single species but morphologically very different, also share more than 99 percent of their genomes. Yet other species might have significantly different segments of their genomes, but be morphologically nearly identical. So there is a classic Catch-22 problem—how can one understand variability within a species if one is not even certain what belongs to the same species? Today, taxonomists generally attack the problem

PHOTOS: DAN JANZEN

South America. Madagascar and India both split away from Africa around 120 million years ago. (India then broke away from Madagascar and slammed into Asia, a collision that formed the Himalayas.) Madagascar's long isolation has resulted in a unique set of flora and fauna. But since humans colonized Madagascar approximately 2,000 years ago, it is estimated that as much as 90 percent of Madagascar's original habitat has been destroyed.

To help stem these losses, the Malagasy government plans to more than triple the number of protected areas over the next five years. It now needs to prioritize the remaining patches of natural habitat for conservation. If we are really serious about "zero biodiversity loss" in Madagascar and elsewhere, conservation planning needs to be based more fundamentally on science. Researchers must conduct detailed inventories of what species exist and precisely where they are found, and protect the remaining habitat fragments that possess the greatest biodiversity.

Yet at present, scientists have only an incomplete knowledge of the island's patterns of diversity. What is known is based mostly on vertebrates—which represent only a small proportion of the island's species. Vertebrate data is generally on a scale too coarse to assess habitat quality or uncover diversity differences among the remaining fragments of natural habitat.

Insects, on the other hand, are generally a better gauge of a habitat's biodiversity. They often exhibit far higher rates of spatial change than larger animals. For example, while one species of lemur might range over hundreds of square miles, different ant species might populate each small valley. Insects therefore provide a measure of biodiversity on the same spatial scale at which conservation decisions are typically made.

In 1999, the Academy initiated in Madagascar one of the largest arthropod inventory programs ever undertaken in the world. From 2000-2005, a field crew of Malagasy taxonomists inventoried 85 sites across Madagascar, and processed over 3.5 million specimens. A team of 15 trained Malagasy students sorted specimens at the processing facility in the capital, Antananarivo, and sent them on to the Academy for distribution to over 100 collaborating taxonomists around the world. Processing such massive collections posed a major challenge—how to quickly recognize known species and identify new ones.

In traditional, morphology-based taxonomy, discrete "forms" are tentatively recognized and hypothesized to be species. Taxonomists search for consistent differences in physical traits that might indicate reproductive isolation.

Identifying and describing the species from the Madagascar arthropod inventory will take decades of work. It takes countless hours of careful observation through a dissecting microscope to measure and study morphological variations such as head width and length when describing ant species. Traditional morphological taxonomy will not provide enough data in the short term to address Madagascar's urgent conservation needs. If nothing is done to change the slow pace of current taxonomic efforts, it will take centuries to complete even a preliminary map of the insect diversity of Madagascar.

To determine whether DNA barcoding might eliminate these bottlenecks, Smith, Herbert and I began testing whether diversity patterns based on DNA barcode sequences are significantly different from patterns based on traditional morphological taxonomy. In our study, recently published in the *Philosophical Transactions of the Royal Society*, we tested ants collected from four critical forest patches in northeastern Madagascar. By comparing the sequence of each specimen's cytochrome oxidase I (*cox1*) gene, we have been able to rapidly group specimens with similar *cox1* sequences. These sequence groupings are termed Molecular Operational Taxonomic Units, and can be used to assess species richness and changes in species composition across landscapes.

We found that data from DNA barcoding grouped the ants in the same way

with a multitude of complementary tools, including detailed examinations of morphology, life-histories, behavior, ecology, and, of course, genomic structure.

Barcoding shortcuts this process. In doing so, it will generate tremendous, and unmeasurable, uncertainty in species identifications. The counter-argument is that the differences between individuals of the same species will always fall below established thresholds. We simply do not know that.

Back to you sitting in the wilderness with your barcoder. If the sequence of your specimen did indeed fall above the threshold, should you conclude that the species is new to science? That could be a valid conclusion, noting that further verification, probably by expert taxonomists, should be performed at some point in the future. This approach could in fact accelerate biodiversity description and documentation. Given our current biodiversity crisis, this sort of acceleration is something that we absolutely need.

But why are we so concerned with counting all the species on the planet? If barcoding worked, and we were able to tally all the species over the next ten years (an impossible feat given available funding, the number of biologists, and our mortal limitations), how would we use that information? Would humans be more conscious of our impact on ecosystems and biological communities if there were 20 million, instead of 5 million, species on Earth? Would there be greater awareness and increased conservation efforts if 10 million out of 20 million species were threatened, instead of 2.5 million out of 5 million?

It is important for all of us to understand that diversity is measured in different ways. While the total number of species is important, I argue that ecological diversity is of greater importance. We need to know what these species are, how many of them are out there, and their functions in ecosystems. Barcoding can help, by suggesting species identities and numbers, but it tells us nothing else about the nature of these species.

Barcoding's statistical approach to describing diversity is also potentially dangerous. It is unlikely that we will be able to count all the species on the planet, accurately, in time to address catastrophic habitat destruction and climate change. Scientists working on these problems are faced constantly with poorly known factors, and while great effort goes into accumulating increasing quantities of data of known accuracy, the ticking clock forces us to incorporate uncertainty into our predictions of the future.

The uncertainties inherent in the barcoding scheme are no worse than those generated by more traditional approaches. Whether or not they are significantly better remains unanswered, and we cannot risk assuming that they are better. Verifying species with complementary approaches will, and must, take time. Dedicating more scientists, staff, and materials to barcoding efforts could inadvertently detract from the overall goal of biodiversity documentation.

DNA barcoding is an innovative addition to the taxonomist's toolbox, and will speed up the discovery of new species, but it is not the panacea some of its supporters claim. We may have only one chance to truly understand the nature of Earth's biodiversity, so let's get it right.

PETER ROOPNARINE *is Associate Curator of Invertebrate Zoology and Geology at the California Academy of Sciences.*

as the study of their traditional morphological traits. Both approaches discerned the same relative patterns of diversity within and between forest patches. However, DNA barcoding achieved results much faster. Our DNA analyses took only three weeks, whereas detailed morphological analyses of each specimen would have required many years.

We concluded that barcoding can rapidly help create biodiversity maps—a boon for groups such as insects, where experts are scarce and identifying specimens is time-consuming. In this case, barcoding allowed the results of insect inventories to be applied immediately towards conservation.

Our experience showed that DNA barcoding can speed up the description of new species as well. In Madagsacar, up to 75 percent of the insects we collected may represent new, undescribed species. Barcoding allowed us to quickly highlight specimens of particular interest, such as those with unusual sequences. Those have been culled for further morphological study to assess whether they represent variants of a single species, or a novel species altogether.

Sequence data are particularly helpful for sorting through insect specimens. For example, when we set out to describe the ant species belonging to the genus *Anochetus*, we used both DNA barcoding and traditional morphology. The workers, queens, and soldiers had very different morphologies, but were easily ascribed to the correct species with DNA barcoding.

In sum, we found that DNA barcoding works in concert with more conventional morphological approaches to taxonomy. It neither competes with nor replaces the traditional study of physical characteristics.

This is a world where we cannot cherish what we do not know exists, where we cannot conserve what is of no known use. In this environment, the documenting of life will help create a bioliterate society, a society that can for the first time understand and hopefully value all of the components of life on this planet, from species to ecosystems. Armed with a new tool such as DNA barcoding, enthusiasm for the exploration of the planet will return.

Little time remains to document global biodiversity. DNA barcoding—a simple, standardized data format which will eventually expand to include multiple genes—is helping to change taxonomy. Collaborating taxonomists, equipped with modern tools, have a chance to move systematics to the forefront of conservation and the public's attention. As more taxonomic information is produced, in a more visible and accessible manner, public and political support for the conservation of life on this planet should follow.

BRIAN L. FISHER *is Associate Curator and Chair of Entomology at the California Academy of Sciences.*