The Beep of the Barcode of Life

Foreword to W. John Kress and David L. Erickson (eds.),

DNA Barcodes: Methods and Protocols, Humana, 2011.

Jesse H. Ausubel[1]

The diversity of life in a hectare of reef, a county of grassland, or a shipload of imports challenges biologists called to identify the species comprising biodiversity, functioning as ecosystems, or invading ports. The sequences of black-and-white barcodes that empower a newly hired clerk to wave a wand over a cart full of goods swiftly, print an itemized receipt infallibly, and order replacements invisibly call forth a vision of an analog for identifying species. The resemblance of commercial barcodes to sequences of DNA shown as black-and-white bars on electrophoretic gels reinforced the vision back in 2003 in the founding meetings of the barcode of life movement. This book edited by early adopters John Kress and David Erickson proves the barcode of life has arrived in environmental science. In less than a decade they and the other authors have realized the vision of a short DNA sequence on a uniform locality of the genome to identify species. We start to hear the DNA barcoder beep.

Because the currency of biology is species, their identification is no academic diversion. Biologists count the rise and fall of biodiversity in species. Regulators designate endangered species by their identified populations and reserve land where they identify the endangered. Governments appraise the success of preservation in the currency of species. Inspectors define quarantines in identified species. Biologists carry the weight of these consequences as they pick a name from almost two million known species names plus millions more unknown.

Written as a sequence of four discrete nucleotides - CATG – along a uniform locality on genomes, a barcode of life provides a “digital” identifying feature, supplementing the more analog gradations of shapes, colors, and behaviors. A library of digital barcodes will provide an unambiguous reference that will facilitate identifying species invading and retreating across the globe and through centuries.

Making a difficult task harder, many species metamorphose into different forms as they cycle through stages in their lives. Eggs may become caterpillars and caterpillars become butterflies but, of course, remain the same species carrying the same genes. Different species may resemble one another or be too small to distinguish easily but carry different genes and thus barcodes that can unmask them. Furthermore, the inspector of unloaded cargo on a dock or

[1] Vice President (Programs), Alfred P. Sloan Foundation, Director, Program for the Human Environment, The Rockefeller University, and Chair, international Barcode of Life (iBOL) initiative; thanks to my colleagues Mark Stoeckle and Paul Waggoner.

analyst of the remains of diets in a stomach may be called to identify species from only a snippet, a hair, or a fin. The fragment may be unrecognizable but will faithfully carry the barcode of the source.

Since Carl Linnaeus (1707-1778) developed systematic naming, ranking, and classifying of organisms, biologists have master keys to all knowledge about a species in the form of binomial names. Biologists use distinguishing features such as shape, color, or number of legs in taxonomic keys first to assign binomials like *Homo sapiens* and then to associate organism names with biological knowledge about the species and its relatives. Of course, the bank of names suffers some problems such as several names for one species.[i] And, biologists continuously debate criteria for species. The diversity of life from bacteria to whales renders any single rule inadequate for defining all species. Nevertheless a few basic criteria, such as that distinct species do not interbreed and meld their genetic sequences, serve for many groups.

Since Charles Darwin (1809-1882) proposed a branching pattern of evolution, biologists have sought to arrange a phylogenetic system of species on an evolutionary tree of life. A tree of life illustrates every introductory biology text. Barcoding will reveal whether a newly collected specimen belongs to a species already on the tree. And if instead a specimen is a truly new species, barcoding will help place it as a new leaf among known species on the proper branch of the tree of life.

Whatever the criteria for species, their inheritance and genes must differ to maintain species distinctions, generation after generation. Isolation by no interbreeding and evolution means that genes distinguish species. Since the molecular discoveries of the mid-20th century, genes intimate a code comprised of sequences of the four nucleotides that constitute DNA. Even before the barcoding movement now embodied in the 200 member organizations from 50 countries of the Consortium for the Barcode of Life (CBOL), revisions of species boundaries included DNA analysis, and separating out new species included DNA divergences.

The barcode analogy leant impetus to the continuing matching of species and genetic differences. Commercial barcodes must be uniform across shelves and warehouses. For animals, concentration on the single segment of the mitochondrial COI gene across the far wider shelves of life imparted the necessary uniformity to avoid a Tower of Babel. Conceiving the series of nucleotides CATG as bars and their presence and order as digital bars opened the door to rapid and unambiguous connection of specimens. Instead of connecting biological specimens to shelves and suppliers, the barcode of life would connect them to curated collections in museums and herbaria, lifting their importance. It would also connect specimens to the biological literature of binomial names. Additionally, barcodes offer a globally consistent way to present provisional or candidate species that experts have not yet honored with a full description and binomial name.

Worries at first evoked by DNA barcoding have not been realized. It has heightened the nuance of the species concept, not diminished it. It has widened humanity's view of diversity, not reduced diversity to ciphers. It has excited wonder at the knowledge hard-won through earlier techniques and accessible through the master key of binomial names. It has enhanced the need for systematists to match the flood of barcodes with a sound array of binomial names. Barcoding is not a mere slogan and an inadequate analogy.

Recurring to the need for uniformity to avoid a Tower of Babel, the choice of a segment of the mitochondrial COI gene has excelled for almost all animal taxa. This barcode region meets the four basic specifications. The locality must be present in all barcoded species and be shaved as short as possible. The locality must have sequences stable within a species through many generations, but nevertheless have sequences variable enough to distinguish species. As this book reports, botanists have now also found barcode regions proving successful from carrots and chamomile to oats and pines. Fungal barcodes are not far behind.

Some observers do ask a single, searching question about the barcode of life arriving in environmental science. When will it be small, cheap, and convenient enough for non-experts, even children? In particular, when will the needed equipment shrink to the size of a laptop or a handheld barcoder? In fact, the key machines have shrunk until they fit comfortably on a desk or tabletop.

The analogy of the newly hired clerk faultlessly pricing the cart of goods suggests making expertise go further, and very far if there were a handheld barcoder or equivalent means of providing identification services conveniently and cheaply. Clues that such a goal will be achieved lie in reports of students detecting endangered marine species on sale, identifying insect traces in their homes, and analyzing tea leaves with inexpensive equipment. As well as enabling specialized scientists to do more and lift the value of specimen collections, barcoding promises to enable laymen to appreciate the diversity of life.

The cleverness of barcoding will truly empower the millions only if what a barcode device, large or small, reads can be matched against first a reference sequence and then a binomial species name. The master key of a binomial name can access knowledge, for example, through the Encyclopedia of Life (www.eol.org), which aims to offer freely on-line a website for every species. How soon can taxonomists and museums array the requisite matching names and barcodes?

The array must rest on a sound foundation of binomial names, which vouchered and identified specimens must provide--preeminently. A sound foundation of binomials based on new and existing natural history collections stands as the first priority to make a handheld barcoder beep cheerily. Fortunately, the Global Names Architecture project associated with the Encyclopedia of Life has already amassed 19 million common and scientific names and is reconciling them for the 2 million or so species estimated already to be known. Within five years, the beep should celebrate achievement of the international Barcode of Life (iBOL) project: access to the barcodes of an array of 5 million specimens sequenced from 500,000 species. Voucher specimens prepared, curated, databased, often digitally imaged, and stored in natural history collections will support the beep.

Already, in just a handful of years, the barcode of life database (www.boldsystems.org) has soared to over 1.2 million specimens from more than 100,000 species. Now, even before the celebratory beep, the chapters in this book show how the arrival of barcoding multiplies what a marine ecologist knows about a reef. And, as authors explain, early sampling of organisms like invasive species that have great environmental impact or like the sea life that enters our food supply will motivate and sustain the building of the reference library of barcodes and the removal of obstacles for its quick, frugal realization.

Our vision inspired by a wand in the hand of a supermarket clerk is comparable magic for an ichthyologist on a research vessel with featureless fish larvae, a child on a woodland trail, or an inspector at a port infallibly identifying a species. Reading this book, we hear the beep.

---

[i] D.J. Patterson, J. Cooper, P.M. Kirk, R.L. Pyle and D.P. Remsen, Names are key to the big new biology, *Trends in Ecology and Evolution 25* (2010), 686-691.