Mongolian Journal of Biological Sciences 2004 Vol. 2(2): 3-11

Barcoding Biological Diversity: A New Microgenomic Identification Approach

Bazartseren Boldgiv

Department of Ecology, Faculty of Biology, National University of Mongolia, Ulaanbaatar 210646, Mongolia
Present address: Department of Biology, University of Pennsylvania, Philadelphia, PA 19104-4207, USA
e-mail: boldgiv@sas.upenn.edu

Abstract

Morphology-based taxonomy suffers from its inherent limitations, even though most of biological research depends on reliable identifications of species. A recent microgenomic identification approach, which is now being called the "DNA-barcoding," presents a promising potential of developing into a real-time, on site tool for identification of organisms, especially animals and of providing an added insight into evolutionary history. For animals, the DNA-barcode seems to have been found in the mitochondrial genome and researchers are in quest of developing similar microgenomic DNA-barcoding systems for other domains of biological diversity. This article discusses the DNA-barcoding technique and considers some of the implications of this approach.

Key words: DNA-barcode, cytochrome *c* oxidase subunit I (COI), mitochondrial DNA

Taxonomy is hard

Reliable on-site, real-time identification of species has always been a burden for biologists, as well as for conservationists, environmentalists, collectors, tourists, farmers, law enforcement, security and customs officials, nature enthusiasts and so on. All these groups of people with different interests have the same need when it comes to species identifications. To mention just a few, for example, farmers want to know what insects have been infesting their crop without losing much time, whereas homeland security or customs officials need to know whether a particular organism is bringing any threat to the well-being of country. Moreover, it can be a great tool for law enforcement when the rarest one of the two closely-related species differing in their conservation status (such as CITES) was illegally harvested, but perpetrators argue otherwise. Therefore, the ability of correctly identifying species has far-reaching implications not only in ecology and biodiversity research, but also in many aspects of environmental management and policy.

Unfortunately, the morphology-based taxonomy has always had its limitations. First of all, traditional dichotomous taxonomic identification keys always require high level of expertise. In many cases they are not easily understood by users due to their specialized jargons (Hebert *et al*., 2003b). Even with the help of glossary of the taxon-specific terms, it is hard to know what the key is describing unless one is an expert in that field. Let us for a minute imagine that you needed to identify an insect specimen using identification keys which reads as "forewings membranous, hind wings forming halteres and tarsi three-segmented." There are people who are sufficiently literate in the field to understand what this means. But majority of people, even majority of biologists, would have no idea what this sentence is describing.

Second, taxonomic keys are effective only for a certain sex or developmental stage of a life cycle. Take an example of birds. Even though birds are the most well-known class of organisms, imagine trying to classify young nestlings of leaf-warblers (genus *Phylloscopus*) into species. It is next to impossible because they all look the same. Take another example, crane flies (family Tipulidae) in this case. Taxonomic keys for this group are mostly based on adult male crane flies. Therefore, you would run into a trouble if you have a crane fly larva (leatherjacket e.g.) or an adult female specimen.

Third, cryptic species and highly variable species are not very uncommon in nature (Knowlton, 1993) and in such cases morphology-based identification is useless. And finally, it is impossible to use a single

http://dx.doi.org/10.22353/mjbs.2004.02.12

trait as an identification criterion. Taxonomy thus has to take a multivariate approach, taking many traits into consideration. For all these reasons, morphology-based identification keys can commonly lead to misidentification and therefore to misjudgment. These arguments certainly illustrate how hard to achieve a predictive, efficient, and stable science of taxonomy. Cataloging even less complicated human-made objects such as books, for example, has proven to be very challenging and there is still no clear-cut catalogue system created by library researchers. In fact, most of the books have to be cross-referenced under different subjects, for example, simultaneously in ecology and evolution or in population biology and mathematics. In taxonomy, species cannot be cross-listed, even though interspecific hybridization or the gene flow among species is not uncommon especially in some groups of organisms, a fact that complicates classification of organisms.

On top of all these, not very many students want to pursue a career in taxonomy nowadays: the very picture of spending a whole career among museum and herbarium collections is not very appealing in this age of information and technology. Therefore, the number of hardcore taxonomists has been dwindling in recent years. On the other hand, there are arguably about 10-15 million species on the world and researchers have discovered only *ca.* 10% of the estimated diversity in the 250-year history of Linnaean taxonomy (Besansky *et al*., 2003). It seems an impossible task to finish inventorying all the species by morphology-based taxonomy.

All these reasons lead scientists to develop a new approach of taxon identification system. A recent new approach of DNA-based taxonomy proposes to use a small segment of a genome to classify all life forms (Hebert *et al*., 2003a). Although the DNA-based taxonomy has been with us for some time, this new approach suggests using a tiny fragment of a genome, which is now analogously called a "barcode," to classify individuals into their corresponding taxonomic groups. This idea of DNA-barcoding comes in the light of technological advance which enabled us to quickly and cheaply determine exact sequences in DNA fragments via polymerase chain reaction (PCR). In fact, the PCR-based sequencing has become so easy a laboratory exercise as to be included in high school biology curricula around the world.

DNA-barcoding of animals

In molecular systematics, scientists compare genes and proteins of organisms because the evolutionary divergence of species parallels the accumulations of differences in their genomes. In doing so, scientists have accumulated a great deal of information compiled into large databases of DNA and amino acid sequences which are available via the Internet. For example, the NCBI database contained nearly 90,000 entries for animal mitochondrial genes by August 2002 (Hebert *et al*., 2003c). The basis of such molecular systematic approach is that cladograms can be constructed with branch points defined by mutations in DNA sequence that mark each lineage. In other words, DNA-based identification takes advantage of diversity among DNA sequences. Recently, a quick and easy diagnostic technique in which the sequence of a portion of a single gene is used for rapid species identification of animals was proposed (Hebert *et al*., 2003a). In a sense, this DNA fragment can be seen as a genetic "barcode" that are contained in every cell. It is analogous to the Universal Product Codes (UPC barcodes), which have been successfully used in the retail industry in the United States and Canada since 1973 (Fig. 1A). By using this system, it is possible to uniquely identify a product and its manufacturer as the 12-digit combinations can generate 100 billion unique identifiers. If only 12-digit barcodes could generate that many unique identifiers, then why could we not use only a few hundred nucleotide sequence combinations for uniquely identifying organisms, if we could find such an ideal portion in a genome? Once we find such a DNA fragment, we can use it as a "DNA-barcode" to uniquely identify individuals into species.

The mitochondrial genome of animals (Fig. 1B) was a better target than the nuclear genome in the search of such a unique DNA sequence because of the following reasons. First of all, individuals tend to be homoplasmic [*i.e.*, single mitochondrial DNA (mtDNA) sequence predominates in all tissues; (Avise, 1994)]. Here I should point out that a paternal inheritance resulting in heteroplasmy was found in humans recently (Bromham *et al*., 2003). Although it seems rare, such paternal inheritance can complicate analyses using mtDNA. Second, it has been shown that even protein-coding sequences in mtDNA evolves at more rapid rate due to lack of known repair mechanisms (Wilson *et al*., 1985), making it possible to reconstruct more recent

evolutionary divergences. Third, mtDNA is transmitted only matrilineally (Avise and Vrijenhoek, 1987), except for a few known cases (Avise, 1994). And the fourth, mtDNA is nonrecombining, *i.e.*, mitochondrial genes lack introns and haploid (Avise, 1979; Avise *et al*., 1987). Although nonrecombination is assumed in phylogenetic reconstruction analyses using mtDNA, there is now clear evidence that recombinant mtDNA occurs in human mtDNA, the finding that makes scientists to reconsider the robustness of conventional analyses (Kraytsberg *et al*., 2004; Slate and Gemmell, 2004).

Past phylogenetic work has often focused on mitochondrial genes encoding ribosomal (12S, 16S) DNA, but their use in broad taxonomic analyses is constrained by the prevalence of insertions and deletions that greatly complicate sequence alignments (Hebert *et al*., 2003c). Looking at mitochondrial genomes of animals, (Hebert *et al*., 2003a) discovered that cytochrome *c* oxidase subunit I (COI) profiles ordinarily assign newly analyzed taxa to the appropriate phylum and order. They also demonstrated that species-level assignments can be obtained by creating comprehensive COI profiles and using the COI portion as a DNA-barcode has a great potential of rapid microgenomic identification system (Hebert *et al*., 2003b). From the time of these findings, an international effort of DNA-barcoding animal diversity was initiated of Paul Hebert at University of Guelph (Canada) and his colleagues (Janzen, 2003a, b) and an organized research toward barcoding other domains of biodiversity is underway.

Reasons for using cytochrome *c* oxidase subunit 1 gene (COI) as a DNA-barcode were following. First, the universal primers for this gene are very robust, enabling recovery of about 658 base pairs of its 5' end from representatives of most, if not all, animal phyla (Folmer *et al*., 1994; Zhang and Hewitt, 1997). Second, COI appears to possess a greater range of phylogenetic signal than any other mitochondrial gene. It shows a high incidence of nucleotide substitutions, leading to a rate of molecular evolution that is about three times greater than that of 12S or 16S rDNA (Knowlton and Weigt, 1998). In fact, the evolution of this gene is rapid enough to allow the discrimination of not only closely related species, but also populations within a single species [significant spatial subdivisions were revealed in freshwater crustaceans (Cox and Hebert, 2001); in a mayfly species (Hughes *et al*., 2003); in freshwater shrimps (Hurwood *et al*., 2003); and in blepharicerid midges (Wishart and Hughes, 2003)]. Third, although COI may be matched by other mitochondrial genes in resolving such cases of recent divergence, this gene is more likely to provide deeper phylogenetic insights than alternatives such as cytochrome *b* (Simmons and Weller, 2001) because changes in its amino-acid sequence occur more slowly than those in this or any other mitochondrial gene (Lynch and Jarrell, 1993). And finally, insertions and deletions (indels) seem to be rarer in the 5' end of the COI gene, which is the portion that is being used as a barcode (Hebert *et al*., 2003c). It makes an analysis much easier because DNA sequence analysis is dependent on the ability to identify and compare

A

B

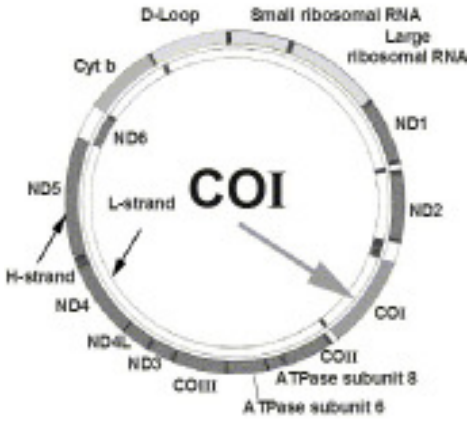

Fig. 1. A: A 12-digit, dummy UPC barcode used in the US and Canada. Barcodes of this kind are used in retail industry for correctly identifying products and their manufacturer (the barcode was generated by an online tool). B: Animal mitochondrial genome. The position of cytochrome *c* oxidase subunit 1 gene, which is being analogously used as a DNA-barcode for animals, is shown by large arrow (courtesy of Dan Janzen).

homologous nucleotide positions and this task is complicated by indel occurrences (Doyle and Gaut, 2000).

Ability of the DNA-barcode, *i.e.*, a portion of COI gene sequence, to discriminate among species has been witnessed by numerous studies on various animal taxa. The list is covers a wide range of animal taxa, including bats of South East Asia, birds of North America, saturnid moths of Costa Rica (Janzen, 2003a, b), gastropods (Remigio and Hebert, 2003), moths of New Guinea (Brown *et al*., 2003), springtails from the Canadian Arctic (Hogg and Hebert, 2004), nematodes (Blaxter, 2004) and many other groups of animals [*see* (Hebert *et al*., 2003a; Hebert *et al*., 2003b; Hebert *et al*., 2003c) for more published data]. An impressive example comes from Prof. Dan Janzen of University of Pennsylvania (Hebert *et al*. 2004). A Neotropical butterfly species *Astraptes fulgerator* was recorded to have a broad distribution range spanning from Mexico to Argentina, South America. Based on the adult morphology, individuals from this broad geographic range looked all the same and were traditionally regarded as a single species. However, analysis of COI was able to discriminate 10 species among the morphologically identical individuals within the range. And larva-rearing experiments confirmed that there were in fact 10 different species with striking differences in larval morphology and behavior (Hebert *et al*., 2004).

Barcoding other taxonomic groups

Because mitochondrial genome of animals has properties that better suited for DNA-barcoding as discussed above, the research in this area is much more advanced than in the rest of the domains of life. But one might ask "what about other taxa?" As for the plants, polyploidy and interspecific hybridization have been important factors in their evolution and it is this that complicates their phylogenetic history reconstruction (Bergthorsson *et al*., 2003). However, researchers believe that the impact of these processes on species identification will be small unless they have led to many cases of recent speciation (<100 thousand years) or to the regular genesis of F_1 hybrids. Because this does not appear to be the case, DNA-based approaches for species identification should be effective. Projects such as Deep Green have provided a comprehensive understanding of the patterning of genetic diversity in the plant kingdom (Barkman, 2000), information useful in designing a DNA-barcoding system for plants. For example, it was shown that rates of mitochondrial evolution are far slower in plants than animals, making COI less useful. Fortunately, rates of evolution in the chloroplast genome are higher (Wolfe *et al*., 1987). Much past work has focused on the analysis of sequence diversity in rbcL and this gene is able to deliver generic-level identifications. Other chloroplast genes, such as mat-K, show promise in generating reliable species identifications (Hebert *et al*., 2003c).

For prokaryotic organisms such as archaea and bacteria, the molecular identification is complicated because there have been horizontal DNA transfers in their evolutionary history that affected their genome (Koonin *et al*., 2001). For example, comprehensive genomic analysis on three *E. coli* strains revealed that they shared only 39.3% of their protein-coding genes (Welch *et al*., 2003). Despite such fluid genomes, disparate species units do exist in these prokaryotic groups. By focusing on fundamental genomic regions, it is possible to create an effective DNA-based identification system. It is now accepted that the sequence analysis of just 6-9 genes provides sufficient information to discriminate closely related microbes (Unwin and Maiden, 2003).

For protistans, DNA-based methods have long been used to diagnose their diversity. Much of this work has employed sequence diversity in 18S rDNA to delineate major protistan lineages (Cavalier-Smith, 2002). However, the 18S rRNA gene is rather conservative to discriminate closely related species (Hebert *et al*., 2003c). Ehara *et al*., (2000) have shown that COI diversity is substantial among protistan lineages, while Lynn and Strueder-Kypke (2003) have established that COI is very effective in discriminating species of *Tetrahymena*. One should recognize that a barcoding approach based on COI will not provide a universal solution to protistan classifications because there are anaerobic protistans that lack mitochondria (Henze and Martin, 2003). However, specialized diagnostic approaches can be developed to probe protistan diversity in these settings (Hebert *et al*., 2003c).

Identification of viruses is a special challenge because of their unusual biological properties, which makes the taxonomy of viruses difficult. It is sufficient to mention that the standard system of Linnaean nomenclature was only recently adopted (Mayo and Horzinek, 1998). However, the intractability of morphological approaches led viral

taxonomists to an early involvement in DNA-based approaches to species description as well as routine identification (Ward, 1993). These involvements have led (Gibbs *et al*., 2004) to suggest that all future species descriptions for viruses should include a summary of short DNA sequences that allow the unambiguous separation of the new taxon from existing species. There are also ongoing efforts to build a web-based database of the diagnostic sequence information needed to support routine identifications (Onodera and Melcher, 2002).

It is clear that major advances have been made toward creating more effective DNA-based diagnostic tools for identification of all major groups of biodiversity. Much of the work is still needs to be done and it remains to be seen whether the DNA-barcoding approach can work for all domains of biodiversity.

Scientific and social implications

Indeed biologists have been accumulating phenotypic information about living and extinct species for centuries. The development of cladistics provided more objective method for comparing morphology and incorporating the data into phylogenetic hypotheses (or cladograms). Molecular systematics has added a powerful new tool in comparative biology, extending the analysis of phylogenetic relationships down to the level of DNA. Cladistic analysis and molecular systematics, complemented by a revival of interest in paleontology and comparative biology in the past few decades, are stimulating a reassessment of phylogeny that is bringing us closer and closer to understanding the history of life on Earth. In many cases, independent approaches, such as paleontology and DNA sequencing, converge in supporting a particular phylogenetic hypothesis. For example, the fossil record, comparative anatomy, and molecular comparisons all concur that crocodiles are more closely related to birds than to lizards and snakes (Gauthier *et al*., 1988, Hedges *et al*., 1990, Hedges and Maxson, 1991), a conclusion that probably would have surprised some of the early evolutionary biologists.

In some cases, molecular data conflict with other evidences, such as fossil records. For example, the oldest fossils of mammals date back 220 million years, into the Triassic period. However, fossils documenting the origin of most modern mammalian orders are much younger, dating to the early Tertiary period, about 60 million years ago, after the extinction of dinosaurs. In contrast to the fossil evidence, molecular clocks push the origin of the major mammalian orders back to almost 100 million years ago (Benton, 1999). Many researchers place more trust in the fossil evidence and express their doubts about whether molecular clocks are reliable. Additional research is required to resolve this debate. So, evolutionary theory has evolved as new methods, new data and new ideas have continued to refine our view of life.

Although some researchers strongly argue to give the DNA-based taxonomy a central role of taxonomy (Tautz *et al*., 2003), it should be pointed out that the barcoding is not intended to supplant or otherwise invalidate existing taxonomic practice. It is not "DNA-based taxonomy," but it is an extension of the existing taxonomic system. Moreover DNA barcoding should adhere to established professional standards for specimen and data management, including routine deposition of voucher specimens in institutional collections, and freely accessible electronic databases and specimen images. By providing a simple and convenient molecular diagnostic tool, DNA barcoding is intended to enhance both the identification of existing species and the discovery of new ones, as well as provide other applied benefits to science and society.

However, some researchers point out shortcomings and have some reservations. Most importantly, using a single gene for assessing phylogenetic relationships is prone to errors because the sequence is short to give unambiguous resolution at all taxonomic levels and often misleading as genes do evolve at different rates in different groups of organisms. Also, this invokes the problem of "what constitutes a species?" Although DNA barcoders advocated a species threshold of 3% COI gene divergence (Hebert *et al*., 2003a), the amount of sequence divergence between the ten *Astraptes* species in the example above were <0.5% (Mallet *et al*., in press). This illustrates the difficulty of quantitative criteria for delimiting species (Harris and Froufe, in press) and barcoding could lead to taxonomic inflation with wide-ranging scientific and applied implications (Isaac *et al*., 2004, Mallet *et al*., in press). Furthermore, maximum parsimony approach, the widely used method in phylogenetic tree construction, is more reliable if based on a large database of DNA sequence comparisons for the set of species in a

tree (Avise, 1994). Additionally, DNA sequence data for COI is not available for very many organisms. Therefore it is difficult to identify organisms based on comparison with organisms that are not very closely related. And finally, it becomes complicated to use in case of interspecific hybrids because mitochondrial genome is transmitted only matrilineally and only maternal species is representing in the DNA-barcoding probe. Robustness of the DNA-barcoding approach is still needed to be rigorously tested. If proven adequate, this technique will greatly advance biology, ecology and biodiversity research by allowing a more accurate estimate of species numbers within a study system.

Most importantly, the DNA-barcoding approach will enable anybody without an expertise in morphological identification to identify specimens at any stage of their life cycle (Hebert *et al*., 2003a). The name of specimen is the most important link to the tremendous amount of information available on the Worldwide Web. Prof. Dan Janzen, who has been the enthusiastic spokesperson for the DNA-barcoding effort, envisions that using on-site, real-time identification of species will mark a revolution of "democracy in biodiversity." It is analogous to what literacy has been to society and it is probably the best way to conserve biological diversity because "everybody will be able to read it" (Janzen, 2003a).

The future

As mentioned above, COI sequences are not known for most animal species. Therefore, the very first thing to do is to create a large database that serves as a global bioidentification system (GBS) for animals (Hebert *et al*., 2003a). Universal protocols are already available for such an effort. This will be a substantial undertaking and will provide revolution in access to basic biological information and a newly detailed view of the origins of biological diversity (Hebert *et al*., 2003a). The Alfred R. Sloan Foundation has given a $669,000 grant to a consortium of herbaria, museums and research institutes to jump-start the Barcode Life Initiative, which aims to create an online catalog of the world's flora and fauna. Institutions against threats of bioterrorism also have a great interest in quick and reliable means of identification of biological agents and they support the initiative. Initially, barcoders will focus on herbaria and museum collections, taking advantage of new techniques for using old DNA in collections. Eventually, they hope to develop a COI database within 20 years for the 5-10 million animal species on the planet (Hebert *et al*., 2003a).

Barcoders also hope that technological advance will help produce a portable device containing the database which can be used in the field. It will have to be able to extract DNA from specimen, amplify it via PCR, align the sequence with the database contained in computer microchip and provide a user with results, all on a real-time basis in the field situation. This device can be uplinked via satellite to a center for verification (Janzen, 2003a, b). The database would have to be updated as more information is accumulated for a fee and can even be regionally-operated (*i.e.*, by biogeographic regions).

It will be interesting to witness whether or not the DNA-barcoding efforts catch on and whether or not it is proved to be an accurate means of species identification. If a portable DNA-barcoding device is ever produced, I will certainly be a user.

Acknowledgments

I thank Prof. Dan Janzen for his generosity, enthusiasm and willingness to provide much needed information and Dr. Linda Robinson for her discussions on the topic and for developing an introductory biology laboratory using the DNA-barcoding technique. I also thank the Nick Isaac whose suggestions improved the manuscript better.

References

Avise, J. C. 1994. *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

Avise, J. C., J. Arnold, R. Martin Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst* 18: 489-522.

Avise, J. C., C. Giblin-Davidson, J. Laerm, J. C. Patton and R. A. Lansman. 1979. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proc. Natl. Acad. Sci* 76: 6694-6698.

Avise, J. C., and R. C. Vrijenhoek. 1987. Mode of inheritance and variation of mitochondrial DNA

in hybridogenetic fishes of the genus *Poeciliopsis*. *Mol. Biol. Evol.* 4: 514-525.

Barkman, T. J., *et al*. 2000. Independent and combined analyses of sequences from all three genomic compartments converge on the root of flowering plant phylogeny. *Proc. Natl. Acad. Sci* USA 97: 13166-13171.

Benton, M. J. 1999. Early origins of modern birds and mammals: molecules vs. morphology. *BioEssays* 21: 1043-1051.

Bergthorsson, U., K. L. Adams, B. Thomason, and J. D. Palmer. 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* 424: 197-201.

Besansky, N. J., D. W. Severson, and M. T. Ferdig. 2003. DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you. *Trends in Parasitology* 19: 545-546.

Blaxter, M. L. 2004. The promise of a DNA taxonomy. *Phil. Trans. R. Soc. Lond. B.* 359: 669-679.

Bromham, L., A. Eyre-Walker, N. H. Smith, and J. Maynard Smith. 2003. Mitochondrial Steve: paternal inheritance of mitochondria in humans. *Trends in Ecology and Evolution* 18: 2-4.

Brown, J. W., S. E. Miller, and M. Horak. 2003. Studies on New Guinea moths. 2. Description of a new species of *Xenothictis meyrick* (Lepidoptera: Tortricidae: Archipini). *Proc Entom Soc Washington* 105: 1043-1050.

Cavalier-Smith, T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. *Int. J. Syst. Evol. Microbiol* 52: 297-354.

Cox, A. J., and P. D. Hebert. 2001. Colonization, extinction and phylogeographic patterning in a freshwater crustacean. *Mol. Ecol.* 10: 371-386.

Doyle, J. J., and B. S. Gaut. 2000. Evolution of genes and taxa: a primer. *Plant Mol. Biol.* 42: 1-6.

Ehara, M., Y. Imaga, K. Kazuo, I. Watanabe, and T. Ohama. 2000. Phylogenetic analysis of diatom coxI genes and implications of a fluctuating GC content on mitochondrial code evolution. *Curr. Genet* 37: 29-33.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3: 294-299.

Gauthier, J., A. G. Kluge, and T. Rowe. 1988. Amniote phylogeny and the importance of fossils. *Cladistics* 4: 104-209.

Gibbs, A. J., J. S. Armstrong, and M. J. Gibbs. 2004 (in review). Combinations of gene sequence features describe organisms efficiently: examples from tobamoviruses.

Harris, J. D., and E. Froufe. (in press). Taxonomic inflation: species concept or historical geopolitical bias? *Trends in Ecology and Evolution*.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003a. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* 270: 313-322.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proc. Natl. Acad. Sci* USA 101: 14812-14817.

Hebert, P. D. N., S. Ratnasingham, and J. R. deWaard. 2003b. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B* 270: S96-S99.

Hebert, P. D. N., S. Ratnasingham, and R. Dooh. 2003c. Barcodes of life. *in*.

Hedges, S. B., and L. R. Maxson. 1991. Pancreatic polypeptide and the sister group of birds. *Mol. Biol. Evol*. 8: 888-891.

Hedges, S. B., K. D. Moberg, and L. R. Maxson. 1990. Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Mol. Biol. Evol*. 7: 607-633.

Henze, K., and W. Martin. 2003. Essence of mitochondria. *Nature* 426: 127-128.

Hogg, I. D., and P. D. N. Hebert. 2004 (in press). Biological identification of springtails (Collembola: Hexapoda) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can. J. Zool.* 82.

Hughes, J. M., P. B. Mather, M. J. Hillyer, C. Cleary, and B. Peckarsky. 2003. Genetic structure in a montane mayfly *Baetis bicaudatus* (Ephemeroptera: Baetidae), from the Rocky Mountains, Colorado. *Freshwater Biology* 48: 2149-2162.

Hurwood, D. A., J. M. Hughes, S. E. Bunn, and C. Cleary. 2003. Population structure in the freshwater shrimp (*Paratya australiensis*) inferred from allozymes and mitochondrial DNA. *Heredity* 90: 64-70.

Isaac, N. J. B., J. Mallet, and G. M. Mace. 2004. Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology and Evolution* 19: 464-469.

Janzen, D. H. 2003a. How to conserve wild plants? Give the world the power to read them. *in* G. Krupnick and J. Kress, editors. *Plant Conservation: A Natural History Approach*.

Janzen, D. H. 2003b. Now is the time. *Phil. Trans. R. Soc. Lond. B.*

Knowlton, N. 1993. Sibling species in the sea. *Annu. Rev. Ecol. Syst* 24: 189-216.

Knowlton, N., and L. A. Weigt. 1998. New dates and new rates for divergence across the Isthmus of Panama. *Proc. R. Soc. Lond. B* 265: 2257-2263.

Koonin, E. V., K. S. Msakarova, and L. Arvind. 2001. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol*. 55: 709-742.

Kraytsberg, Y., M. Schwartz, T. A. Brown, K. Ebralidse, W. S. Kunz, D. A. Clayton, J. Vissing, and K. Khrapko. 2004. Recombination of human mitochondrial DNA. *Science* 304: 981.

Lynch, M., and P. E. Jarrell. 1993. A method for calibrating molecular clocks and its application to animal mitochondrial DNA. *Genetics* 135: 1197-1208.

Lynn, D. H., and M. Strueder-Kypke. 2003. Cytoplasmic genes and evolution of species of Tetrahymena: evidence from cytochrome *c* oxidase I. *in* Ciliate Molecular Biology, FASEB Conference, Vermont.

Mallet, J., N. J. B. Isaac, and G. M. Mace. (in press). Response to Harris and Froufe, and Knapp *et al.*: Taxonomic inflation. *Trends in Ecology and Evolution.*

Mayo, M. A., and M. C. Horzinek. 1998. A revised version of the International Code of Virus Classification and Nomenclature. *Arch Virol*. 143:1645-1654.

Onodera, K., and U. Melcher. 2002. VirOligo: a database of virus-specific oligonucleotides. *Nucl. Acids Res* 30: 203-204.

Remigio, E. A., and P. D. N. Hebert. 2003. Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships. *Mol. Phylogenet. Evol.* 29: 641-647.

Simmons, R. B., and S. J. Weller. 2001. Utility and evolution of cytochrome *b* in insects. *Mol. Phylogenet. Evol*. 20: 196-210.

Slate, J., and N. J. Gemmell. 2004. Eve 'n' Steve: recombination of human mitochondrial DNA. *Trends in Ecology and Evolution* 19: 561-563.

Tautz, D., P. Arctander, A. Minelli, R. H. Thomas, and A. Vogler. 2003. A plea for DNA taxonomy. *Trends in Ecology and Evolution* 18: 70-74.

Unwin, R., and M. C. J. Maiden. 2003. Multi-locus sequence typing: a tool for global epidemiology. *Trends in Microbiology* 11: 479-487.

Ward, C. W. 1993. Progress towards a higher taxonomy of viruses. *Res. Virol*. 144: 419-453.

Welch, R. A., and coauthors. 2003. Extensive mosaic structure revealed by the complete genomic sequence of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci.* 99: 17020-17024.

Wilson, A. C., R. L. Cann, S. M. Carr, J. George, M., U. B. Gyllensten, K. M. Helm-Bychowski, R. G. Higuchi, S. R. Palumbi, E. M. Prager, R. D. Sage, and M. Stoneking. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc*. 26: 375-400.

Wishart, M. J., and J. M. Hughes. 2003. Genetic population structure of the net-winged midge, *Elporia barnardi* (Diptera: Blephariceridae) in streams of the south-western Cape, South Africa: implications for dispersal. *Freshwater Biology* 48: 28-38.

Wolfe, K. H., W. S. Li, and P. M. Sharpe. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci*. 84: 9054-9058.

Zhang, D. X., and G. M. Hewitt. 1997. Assessment of the universality and utility of a set of conserved mitochondrial primers in insects. *Insect Mol. Biol*. 6: 143-150.

Хураангуй

Биологийн ихэнх судалгаа морфологийн таксономиос хамааралтай боловч морфологи-суурьтай таксономид дутагдалтай тал цөөнгүй байдаг. Сүүлийн үед шинээр гарч, "ДНХ-ийн савхан код" (DNA-barcode) гэж нэрлэгдээд буй микрогеномийн тодорхойлох систем нь организмыг, ялангуяа амьтдыг газар дээр нь түргэн хугацаанд найдвартай тодорхойлох хэрэгсэл болон хөгжих боломжтойгоо харуулсаар байна. Амьтдын хувьд тэднийг зүйлийн түвшинд хүртэл

тодорхойлоход митохондрийн геномд байрлах “ДНХ-ийн савхан код” буюу нэгэн генийн хэсгийг ашиглаж болохыг олон судалгаа харуулсан бөгөөд үүнтэй төстэйгөөр бусад организмыг (ургамал, прокариотууд г. м.) тодорхойлох микрогеномийн “савхан код” боловсруулах судалгаа ид өрнөж байна. Энэхүү богино өгүүлэлд “ДНХ-ийн савхан код” гэж юу болох, түүний арга зүйн үндэслэл болон түүнийг хэрэглэхийн ач холбогдолтой болон дутагдалтай талын тухай өгүүлсэн болно.

Received: 05 October 2004
Accepted: 28 November 2004