DNA barcodes for ecology, evolution, and conservation

W. John Kress¹, Carlos García-Robledo¹,²,³, Maria Uriarte⁴, and David L. Erickson¹

The use of DNA barcodes, which are short gene sequences taken from a standardized portion of the genome and used to identify species, is entering a new phase of application as more and more investigations employ these genetic markers to address questions relating to the evolution and ecology of natural systems. The suite of DNA barcode markers now applied to specific taxonomic groups of organisms are proving invaluable for understanding species boundaries, community ecology, functional trait evolution, trophic interactions, and the conservation of biodiversity. The application of next-generation sequencing (NGS) technology will greatly expand the versatility of DNA barcodes across the Tree of Life, habitats, and geographies as new methodologies are explored and developed.

DNA barcodes: what, why, and how

In reference to the universal coded labels found on commercial products, the term ‘DNA barcode’ is now commonly applied by biologists to a standardized short sequence of DNA that can be recovered and characterized as a unique identification marker for all species on the planet [1] (see Glossary). For many users of DNA barcodes, identification of an unknown sample by correctly matching a specific genetic marker to a reference sequence library is the primary goal (Box 1). However, DNA barcodes can also be applied as tools for addressing fundamental questions in ecology, evolution, and conservation biology, such as: how are species assembled in communities? What multispecies interactions occur in previously poorly known environments (e.g., soils)? And, where are the most evolutionarily rich habitats to be targeted for protection? These uses of DNA barcodes, which have only recently been considered and offer some of the most exciting prospects for using this new taxonomic tool, are the primary focus of this paper.

A DNA barcode, in its simplest definition, is one or few relatively short gene sequences taken from a standardized portion of the genome and used to identify species. The use of DNA sequences for biological identifications is not new [2], but the concept of a ‘DNA barcode’ for quick and reliable species-level identifications across all forms of life, including animals, plants, fungi, and microorganisms, was first proposed just over a decade ago [3]. A short DNA sequence of 600 base pairs (bp) in the mitochondrial gene encoding cytochrome c oxidase subunit 1 (CO1; [3]) has been accepted as a practical, standardized, species-level DNA barcode for many groups of animals (Figure 1). Given that CO1 does not work as a DNA barcode in plants [4] and fungi [5], a concerted search was required for more effective gene regions for these major groups of organism (Figure 1). The DNA barcode loci now most commonly used for plants [a combination of plastid rbcl, matK, and trnH-psbA with nuclear internal transcribed spacer (ITS)] and fungi (nuclear ITS) may never be as efficient and successful in identification as CO1 in animals. However, even for some groups of animals (e.g., some invertebrates [6], amphibians, and reptiles [7]), CO1 works poorly and other gene regions are being used as DNA barcodes (Figure 1).

The introduction of any new method of analysis in science often brings some controversy and concern, which has been the case with DNA barcoding, especially in the field of taxonomy (e.g., [8]). In some taxa, DNA barcode markers were not as effective as first proposed (e.g., [9]). Plants, which have inherently lower rates of nucleotide substitution in mtDNA compared with animals [10], were especially problematic during the early stages of developing DNA barcodes. In addition, the concern that DNA barcodes will give poor results or faulty identifications because of the complications of ancestral polymorphisms, hybridization, and/or introgression certainly applies to both plants and animals [10,11]. These complications can be particularly acute in some groups of plant in which hybridization is widespread and pseudogenes in the nuclear genome are common [12]. More concerted work is needed on taxa with extensive hybridization to verify whether DNA barcodes can successfully provide accurate identifications across all species.

However, it may also be true that poor species resolution with DNA markers is due in part to our imperfect and variable definition of species across the major lineages of...
Glossary

Alpha and beta species diversity: alpha diversity is calculated as the total species diversity in any single site or unit. Beta diversity quantifies site-to-site variability in community composition.

Community assembly: ecological communities are created through the arrival, reproduction, and local extinction of individual species. Community assembly processes drive colonization and extinction dynamics, and range from entirely stochastic to entirely deterministic (e.g., competition or environmental filtering).

Community phylogeny: an evolutionary framework for co-occurring species in a community that is based on historical relations among the member taxa.

Cryptic species: a complex of species in which members are reproductively isolated from one another yet morphologically indistinguishable.

DNA barcode: in the strict sense, a DNA barcode is one or more short gene sequences taken from a standardized portion of the genome and used to identify species. In a broad sense, a DNA barcode is any DNA sequence used for identification at any taxonomic level.

DNA barcode library: a set of DNA barcode sequences compiled from species of known taxonomic origin used to compare and identify DNA barcode sequences recovered from unknown samples.

DNA metabarcoding: the use of NGS to identify multiple species in a sample using DNA barcodes.

Ecological forensics: a specialized version of DNA metabarcoding in which trophic or network interactions among species are resolved by genotyping complex mixtures of individuals. These mixtures may take the form of an individual organism in which parasites, mutualists, diet items, and symbionts are recovered and sequenced, or complex mixtures of propagules, such as pollen grains and seeds.

Functional trait: a measurable property, phenotype, or characteristics of an organism that may influence its fitness (i.e., survival, growth, reproduction, or dispersal).

Habitat filtering: the process where only species that carry a specific trait are able to successfully colonize a habitat.

Mini-barcode: a portion of a standard DNA barcode that is targeted for use in ancient or degraded samples in which the complete DNA barcode sequence may be unable to be recovered.

Morphospecies: species recognized by taxonomists based on their morphology, but not yet classified or named as novel species or varieties.

Next-generation sequencing (NGS; or massive parallel sequencing): sequencing methods that differ from Sanger sequencing in the ability to sequence thousands to millions of small DNA fragments simultaneously. Several related but distinct technologies are referred to as NGS, including Illumina, Roche454, and Ion Torrent, among others whose unifying similarity is that they are not based on the Sanger Di-Oxyx method.

Niche differentiation: the process by which different resource use or different niches allows the co-occurrence of phenotypically different species.

PCR: a method to target and copy a segment of DNA from among many sequences in a genome or sample. The result is billions of copies of the target sequence, which may then be analyzed to infer the exact nucleotide sequence.

Phylogenetic diversity: the sum of genetic distances connecting all taxa in a dated or molecular-clock scaled phylogeny. It focuses on the overall evolutionary divergence among taxa rather than species number (Simpson’s Diversity) alone in a community or habitat.

Tree of Life: a phylogenetic reconstruction of living organisms that demonstrates the shared evolutionary origins and divergences of the included lineages.

Trophic interactions: a network of interactions among autotrophs, predators, and their prey.

Box 1. Building the DNA barcode library using Sanger sequencing

The workflow for generating DNA barcodes for individual species entails two basic steps: (i) building the DNA barcode library of known species; and (ii) matching the DNA barcode sequence of an unknown sample against the barcode library for identification (Figure 1).

The DNA barcode library is a collection of DNA sequences associated with verified taxonomic identification and ideally with voucher specimens. A comprehensive DNA barcode library, which will be the most useful for multiple applications, is a recognized limiting factor because of the overwhelming number of species of plants, animals, and fungi already described by taxonomists. In this respect, museum specimens are a critical source of tissue for generating DNA barcodes with known vouchers. Nearest neighbor algorithms are usually used to assign an unknown sample to a known species by finding the closest database sequence to the sample sequence [76]. Basic Local Alignment Search Tool (BLAST) is a common matching tool, provided through NCBI, that searches for correspondence between a query sequence and a sequence library [77].

Life. Even those who profess little confidence in DNA barcoding accept that our species concepts are not universal and often inconsistent among taxonomists [8]. In some taxa, the numbers of species recognized by different taxonomists may vary by up to 50% [8]. Although we have no comparative evidence across phyla that indicates that all
species should be genetically parallel, the variation in success of DNA barcodes across lineages suggests that rates of evolution and the processes of speciation are not uniform.

The primary application of DNA barcodes will continue to be the identification of unknown samples. With the adoption of NGS technologies in many DNA barcode investigations (e.g., metagenomic applications [13]), uses are already being expanded to answer both applied and basic biological questions. Even if DNA barcodes are not uniformly successful for unambiguous identifications across the entire Tree of Life, ecologists, evolutionary biologists, and conservationists are already adopting DNA barcodes as a tool in their respective fields (e.g., [14,15]).

**Current evolutionary, ecological, and conservation research with DNA barcodes**

**Taxonomy, systematics, and species discovery**

A primary goal of evolutionary biologists and ecologists is to understand the origin of species and the factors causing the disparity in species richness in different biomes across the globe. In many cases, the full diversity of species in a given region is still unknown, especially in the most biodiverse habitats [16]. DNA barcodes have been particularly useful in the discovery of cryptic and previously unrecognized species of animals [17]. For insects, it has been demonstrated that new species can be revealed through a combination of ecological field observations and DNA barcode markers [18,19]. For example, a single species of common skipper butterfly found throughout Central America defined by morphological features of the adults in fact comprises numerous species that are clearly delineated by DNA barcode sequences in congruence with diets and features of the larvae [20] (Box 2). Similarly, cryptic species of hispine beetles and larval stages that were identified through a DNA barcode survey of plant–herbivore interactions in Costa Rica linked the adults with the larvae found on the host plants [21] (Box 2). An extensive DNA barcode study on Microgastrinae wasps demonstrated significant insights into their taxonomy and species discovery [22]. Such cryptic diversity has been uncovered using DNA barcodes in other animal taxa, including crustaceans, diatoms, and fish [23–25]. The use of DNA barcodes for the discovery of new species is emerging as a powerful tool to clarify species boundaries and to quantify species diversity [26]. In many cases, these genetic markers serve as the starting point for the discovery of new taxa (Figure 1).
Box 2. Discovering new species, cryptic life stages, and ecological interactions in natural populations

DNA barcodes are a valuable genetic tool to reveal cryptic species previously unrecognized through the analysis of standard morphological variation. When combined with ecological data, the DNA barcodes provide additional evidence for determining species boundaries (Figure I). The *Astraptes fulgerator* species complex is a classic example of the use of DNA barcodes as a tool to discover cryptic biodiversity. In the Guanacaste Conservation Area (Costa Rica), DNA barcodes, combined with host plant records and larval morphology, demonstrated that one species of skipper (*A. fulgerator*), was in fact a complex of ten different species [20].

Likewise, the potential exists for new plant species to be discovered and described as a result of genetic inventories based on both plastid and nuclear DNA barcode markers. For example, in the complex tropical plant family Lauraceae, the community phylogeny generated for the tree species on Barro Colorado Island with DNA barcode sequence data supported the recognition of a previously undescribed, but suspected, new species of *Nectandra* [27]. Furthermore, an ongoing DNA barcode survey of the trees in a forest dynamics plot located in the heart of the Amazon near Manaus, Brazil, suggested that many of the 'morphospecies' recognized by local taxonomists that do not currently have scientific names may have congruent support from the DNA barcode sequence data (Kress et al., unpublished, 2014). This hyperdiverse research plot in the Amazon with over 1400 species of trees will be a test case for the utility of DNA barcodes in identifying potentially new species in a poorly known flora.

Community ecology and phylogeny
Over a decade ago, Webb [28] introduced a new set of methods and metrics to evaluate processes affecting community assembly in a phylogenetic context. This work demonstrated how phylogenetic relations among species within a community can be analyzed to infer processes such as competition, environmental filtering, and trait evolution. Subsequent studies have been fruitful in mining phylogenetically structured community data to investigate alpha and beta species diversity [29,30], the linkage between phylogenetic diversity and dispersion [31,32], and the role of functional traits and evolutionary history in structuring communities [33,34] (Figures 2 and 3). One constraint on the use of phylogenetic data in community ecology has been the availability of phylogenies that accurately reflected the evolutionary relations among community members, particularly at lower taxonomic scales. Given that standard phylogenetic analyses are taxon driven, inclusion of all members of a specific community is rare. Thus, the availability of highly resolved phylogenetic trees that contain all members of a community has been critically lacking. The ability of DNA barcode data to help reconstruct evolutionary relations within targeted communities is helping to resolve this problem [27].

The first application of DNA barcodes in plant community phylogenetics was the reconstruction of the relationships of 281 tree species found in the Barro Colorado Island Forest Dynamics Plot in Panama [27]. That study demonstrated how the multilocus plant DNA barcode could robustly
reconstruct evolutionary relations among phylogenetically disparate community members (Figure 2). The Panama study also showed that a DNA barcode-based community phylogeny has greatly improved topological resolution relative to super-tree methods that had been adopted previously. The improvement of topological resolution through incorporation of DNA barcode data for all or most species in the community improves the power of evolutionary-based hypothesis testing in an ecological framework. Since that first DNA barcode-based community phylogeny, similar investigations have focused on other forests in the tropics [27,35].

A major challenge of reconstructing a phylogenetic tree using DNA barcode sequence data is to capture the proper evolutionary relations among highly divergent as well as closely related species. Given the DNA barcode requirement for relative short genetic markers (less than 1500 bp in total for plants), resolving phylogenetic relations among the basal branches and the tips of the branches is difficult. In this sense, one can regard these community phylogenies as a small part of a Tree of Life project, which seeks to assemble and reconstruct species relations. It has been shown that, despite limited nucleotide content, taxon relations can be well estimated when species density is high [36]. As more DNA barcode data are assembled, an increased number of taxa will be woven into increasingly larger phylogenies. This approach results in more well-supported phylogenetic reconstructions of the constituent species from which individual community phylogenies may be pruned out for targeted analysis [35]. Furthermore, the use of a constraint tree derived from existing Tree of Life studies will enforce topological relations at deeper phylogenetic levels, where the limited signal from the small amount of nucleotide data provided by the DNA barcode markers is most problematic (Figure 3) [27]. The DNA barcode data are then relied upon to primarily resolve generic- and species-level relations towards the tips of the evolutionary branches.

**Species assembly and functional trait evolution**

The rapid increase in phylogenetic information based on DNA barcode sequence data [27] coupled with advances in functional trait-based community ecology [37,38] have been used to address questions of community assembly and diversity [39,40]. This merging of evolutionary and ecological perspectives has led to new insights into the role of ecological, biogeographical, and evolutionary processes in the distribution of biodiversity [41] (Figure 3).

Despite these advances, several conceptual and methodological challenges remain. One of the most salient issues is the difficulty of dynamically linking biophysical processes and species interactions that unfold over vastly different spatial and temporal scales (R. Muscarella et al., unpublished, 2014). Most empirical studies that have
simultaneously investigated phylogenetic and functional community structure as a means to provide insights into community assembly processes have been focused at either the local (e.g., [29]) or global scales (e.g., [42]) rather than at intermediate (i.e., regional) scales, where local and global factors interact. Global-scale studies have examined the role of historical biogeography on community assembly without much attention to ecological processes. By contrast, local-scale studies have generally focused on disentangling the influence of habitat filtering and niche differentiation on community composition (e.g., [29]) or demographic rates (e.g., [34]). In most of these local studies, the use of phylogenies has been ‘corrective’, that is, largely limited to quantifying how the assumption of phylogenetic dependence between species [43] might influence results.

A greater emphasis on intermediate scales (i.e., beta diversity) that reflect how historical and local processes unfold over environmental gradients in space is likely to yield deeper insights into the role of evolutionary processes on community assembly. Recent efforts are moving in this direction (e.g., [44]). Several approaches have been introduced to characterize the relations between particular traits, evolutionary lineages, and variation in environmental conditions across sites (e.g., [44–46]). For instance, Pavoine [47] developed a novel method to identify the association of trait states and phylogenetic trees with spatially variable environmental factors. At coarse spatial scales, the association of phylogenies with space, but not with environmental factors indicates that historical processes (e.g., colonization) predominate over environmental factors in explaining community composition. Advancing our understanding of the processes that structure beta diversity will require data on species performance and community composition collected across environmental gradients (e.g., soil or precipitation) and at multiple spatial and temporal scales. The relative ease of generating universal DNA barcode data for plants in a community will facilitate these broad spatial comparisons and is currently being carried out across forest dynamics plots around the globe [35].

**Diet analyses, trophic interactions, and ecological forensics**

DNA barcodes represent a unique opportunity to understand trophic interactions among organisms, especially in habitats that are difficult to access, such as the forest canopy or underground zones. On Barro Colorado Island, a tropical rain forest in Panama, DNA barcodes were used to identify the species of roots collected from soil cores [48]. By combining DNA barcode-based root identifications with spatial records of individual trees from a mapped vegetation plot, the relation between belowground and aboveground diversity was assessed in a diverse plant
DNA barcodes can also assist in determining the diets of invertebrates, frugivorous birds, and elusive large mammals in the wild [50,51]. The first step in this DNA barcode approach to diet analyses is the collection of DNA tissue samples of the prey or host taxa under investigation (Box 1). For invertebrate herbivores, once the library is complete, DNA of host species is isolated directly from the gut contents of the animals [52]. By contrast, analyses of vertebrate diets are not invasive and are based on DNA extracted from scats [53,54]. The barcode markers utilized for amplification will depend on the diet of the organism [55]. The standard DNA barcode loci used to investigate herbivore diets differ markedly in their ability to identify plants at different taxonomic levels because of varying base-pair substitution rates across taxa. The plant DNA barcode markers rbcL and trnL-trnF-intron, although easy to amplify from gut contents and scats, are usually only useful to identify plant tissues to the level of taxonomic family or genus (Figure 1) [52]. The DNA barcode loci ITS2, matK, and the noncoding intergenic spacer trnH-psbA are able to identify diet items to genus and species [52]. Unfortunately, the amplification of these DNA regions can be challenging or even impossible for some plant groups [52]. For animal prey, the most broadly used DNA barcode marker to identify diets is mitochondrial COI (Figure 1) and when employed with host-specific blocking primers, can enable identification of extensive prey diversity [56].

DNA sequences recovered from gut contents and scats are usually short (~100–400 bp) and of low quality as a consequence of DNA degradation during digestion [57,58]. The implementation of extraction techniques originally developed for ancient and antique DNA may improve the quality of the resultant DNA sequences [52]. However, DNA degradation during digestion is still a major limiting factor in diet analyses. When the DNA recovered is degraded, a mini-barcode version of the full DNA barcode marker [52] or a DNA barcode with more forensic applications [59] is required, both of which can not only improve amplification success, but also reduce rates of species-level identification.

The DNA sequencing approach to be used depends on the diet breadth of the consumer. For species that feed on one or only a few species during each feeding bout, it is possible to use traditional Sanger sequencing techniques (Box 1). For polyphagous species in which all diet items are more difficult to identify, it is now possible to accurately determine all consumed species using NGS methodology [60] (Box 3).

The most efficient method to identify the taxa consumed by a particular animal or group of animals is to compare DNA sequences extracted from gut contents and scats directly to a reference DNA barcode library of the food items (Box 4). The quality of identifications will depend on the completeness and the accuracy of the sequences included in the reference library. Researchers should be cautious when using public DNA sequence repositories, such as GenBank, because of their incompleteness. When DNA sequences for potential food items are missing in the database, the contents of guts and scats can be misidentified. The error rate may also increase because of faulty taxonomic identifications and lack of verifiable taxonomic voucher specimens [57]. However, the scope of these taxonomic errors in open-source databases has not been quantified.

A solution proposed by recent studies is to generate a priori a comprehensive and well-curated DNA barcode library of potential diet items [21,50,52]. By generating complete DNA reference libraries for food plants, diet identifications can be accurate, even to the species level, and the error rate will be reduced or even eliminated [52] (Box 4). As NGS technologies become more accessible and cost-effective, more researchers will rely on these techniques to investigate and understand the complexity of trophic interactions in nature.

Conservation biology: quantifying species richness and evolutionary diversity within and among communities

The level of biological diversity present in an environment can be quantified by either enumerating numbers of species (e.g., Simpson’s diversity) or estimated evolutionary divergences among species in which genetic distances have been calculated [28]. Although most measures of alpha and beta diversity across plant communities are based on numbers of species, DNA sequence data, if available, will provide an evolutionary dimension to diversity estimates that incorporate genetic distance among species (i.e., phylogenetic branch lengths). Evolutionary diversity, also called phylogenetic diversity [61,62], is correlated with, but not equivalent to, species richness. For example, geographic areas that harbor relatively few species may have high phylogenetic diversity if the species present are broadly dispersed across the Tree of Life. It has also been shown that a discontinuity between species richness and evolutionary diversity occurs when there are concentrations of closely related species in an environment [63].

DNA barcodes can provide a universal marker across species in a community or a region by which genetic distance, hence phylogenetic diversity, can be quantified within and across ecological communities at varying geographic scales [64]. When compared with species richness in the same communities, these genetic measures can also be used to evaluate species boundaries, can serve as clues to assist in documenting new species, and can identify targeted habitats for conservation [61,65]. In geographic regions known for their especially unique lineages of plants and animals, such as northeast Queensland in Australia and South Africa, phylogenetic diversity defined with DNA barcode sequence data may be the most important measure for comparing diversity and establishing protected areas across the landscape (A. Shapcott et al., unpublished, 2014). As the DNA barcode library becomes populated with species across the globe, comparative measures of phylogenetic diversity will become standard metrics for conservation assessment.

In addition to the assessment of biodiversity hotspots for conservation, DNA barcodes are now also being used for the reliable identification and detection of illegally traded and often endangered species [66]. Similarly, genetic identifications using DNA barcodes for wild-collected medicinal
plants [67], market adulterants in certified natural and commercial products [68], and genetically modified crops [69] are becoming more common. These applied uses of DNA barcodes for conservation and commercial purposes will undoubtedly increase in the future, especially as sequencing technology becomes simpler and less expensive.

Concluding remarks and future contributions of DNA barcodes

Here, we have focused mainly on the basic scientific applications of DNA barcodes to increase our understanding of species relations and boundaries, community ecological processes and networks, and the assessment of biodiversity for effective conservation. In addition, the forensic use of DNA barcodes for identification of endangered species and commercially useful plants and animals is being expanded by local, state, and national governments. DNA barcodes are proving useful as evidence in criminal cases and investigations of natural and manmade disasters. For example, a library of CO1 markers for birds is now routinely used to identify avian species involved in airplane strikes [70]. These applications are only in their infancy, but may eventually be a major source of data for the growing DNA barcode library.

The most significant advance in DNA barcoding in the near future will be the application of new technologies for generating and analyzing DNA barcode sequences. Several studies [71,72] and reviews [72,73] have addressed the marriage of DNA barcoding and NGS, especially with regards to environmental sampling. In general, NGS platforms are not used for collecting and constructing DNA barcode reference libraries for a specific set of taxa. Rather, NGS will enable the capture of all representative sequences present in a complex mixture of species and then the mapping of those sequences to a reference DNA barcode database [72]. These complex mixtures can be environmental samples of water or soil used for biodiversity assessment [74], or the mixtures can be targeted samples, such as animal scats or pollen loads on pollinators, for examining diet choice or pollinator foraging behavior (Box 3) [73,75]. This application of DNA barcoding to identify component species of an environmental mixture is termed ‘DNA metabarcoding’ [75]. As is true in other fields, the application of NGS to DNA barcoding will lead to tremendous growth in available sequence data, which themselves will require new tools for analysis as well as new systems for information storage.

Over the past 10 years, DNA barcoding has become an invaluable addition to our suite of tools to better understand nature and the environment. As the DNA barcode library expands across the Tree of Life, habitats, and
Box 4. Understanding ecological interactions in complex networks of species

Novel molecular techniques based on DNA barcodes can reconstruct interactions in complex networks of species in communities, including antagonistic and mutualistic plant-animal interactions. At La Selva Biological Station (Costa Rica), interactions between plants in the order Zingiberales and associated herbivores (chrysomelid rolled-leaf beetles) were determined by combining field records and DNA barcode analyses (Figure I). Here, each black rectangle represents a plant or insect herbivore species. Lines represent single and multiple plant–herbivore associations. These molecular techniques have the potential to become a standard methodology for a detailed understanding of plant–herbivore interactions.

At the Garrapilos Mediterranean lowland forest in Spain, seed dispersal by birds was determined using a novel DNA barcode technique (Figure II). After assembling a comprehensive DNA barcode library (using COI), mini-barcodes were developed to amplify degraded DNA for this particular bird assemblage. The bird species responsible for dispersing seeds were then identified by amplifying bird DNA obtained from the defecated propagules. The identification of interacting species represents a unique opportunity to understand the process of fruit selection by frugivores and the contribution of each frugivorous species to seed dispersal to different microsites [50]. These techniques have the potential to connect frugivory with seed dispersal in ways previously unattainable through traditional field techniques.

Acknowledgments

We are grateful to Susanne Renner and two anonymous reviewers of this manuscript for their constructive suggestions. Authors were supported by grants from the NSF-RCN program for ‘Tropical Forests in a Changing World’ (Award ID: 0741956) to D. McLaughlin, the Smithsonian Institution Postdoctoral Fellowship Program, Global Earth Observatories Program, and the Office of the Under Secretary for Science, and the National Geographic/Watt Institute Grant (W149-11) to C.G.-R. M.U. acknowledges support from NSF DEB-0862091 to the Luquillo Long-Term Ecological Research Program. The assistance of I. Lopez in constructing the illustrations was invaluable. We would like to thank D. Janzen, N. Swenson, P. Hollingsworth, L. Weigt, A. Driscoll, P. Taberlet, J. Chave, C. Sick, C. Meyer, D. Schindel, and many others for helpful discussion on methods and uses of DNA barcodes.

References

2. Pecknair, Z.F. and Buzan, E.V. (2014) 20 years since the introduction of DNA barcoding: from theory to application. J. Appl. Genet. 55, 4–52
19 Bertrand, C. et al. (2014) Mitochondrial and nuclear phylogenetic analysis with Sanger and next-generation sequencing shows that, in Area de Conservación Guanacaste, northwestern Costa Rica, the skipper butterfly named Urbanus belli (family Hesperiidae) comprises three morphologically cryptic species. BMC Ecol. 14, 15
26 Costion, C. et al. (2011) Plant DNA barcodes can accurately estimate species richness in poorly known floras. PLoS ONE 6, e26841.
49 Kesavanur, P.R. et al. (2011) Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. Mol. Ecol. 20, 1289–1302
54 Clare, E.L. et al. (2009) Species on the menu of a generalist predator, the eastern red bat (Lasiurus borealis): using a molecular approach to detect arthropod prey. Mol. Ecol. 18, 2532–2542
68 Swetha, V.P. et al. (2014) DNA barcoding for discriminating the economically important Cinnamomum verum from its adulterants. Food Biotechnol. 28, 183–194
82 Powers, T.O. et al. (2014) COI haplotype groups in Mesocricinema (Nematoda: Criconematidae) and their morphospecies associations. Zootaxa 3827, 101–146
83 Pillai, P.M. et al. (2014) Description, DNA barcode and phylogeny of a new species, Macrobrachium abrahami (Decapoda: Palaemonidae) from Kerala, India. Zootaxa 3768, 546–556
99 Deagle, B.E. et al. (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. Biol. Lett. 10, 20140562