300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants

Robyn S. Cowan[1], Mark W. Chase[1], W. John Kress[2] & Vincent Savolainen[1]

[1] *Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3DS, U.K. r.cowan@kew.org (author for correspondence).*
[2] *Department of Botany, United States National Herbarium, National Museum of Natural History, Smithsonian Institution, P. O. Box 37012, Washington, D.C. 20013-7012, U.S.A.*

DNA barcodes have been successfully applied to a limited number of animal groups with the application of the mitochondrial gene, cytochrome *c* oxidase subunit 1. Recently two DNA regions, the plastid *trnH-psbA* spacer and nuclear ribosomal ITS region, have been shown to have potential as an identification barcode for land plants, although with some significant drawbacks. The ideal barcode should be relatively short in length (~700 bp), more variable between than within species, and easily amplifiable with universal primers. Building on current success, ongoing investigations are searching for the best barcode to apply to all land plants. Once established, a plant barcode may be effectively used in biodiversity inventories, conservation assessments, and applied forensic investigations. Advances in sequencing technology and the completion of the DNA barcode library have the potential to provide the public with increased access to information about the natural world.

KEYWORDS: angiosperms, biodiversity, cytochrome *c* oxidase, herbarium, ITS, land plants, plastid genome, *trnH-psbA*.

INTRODUCTION

The term "DNA barcoding" has been coined to describe the use of a short gene sequence from a standardized region of the genome as a tool for species identification (Hebert & al., 2003a, b). For many years biologists have been using a wide range of DNA fingerprinting techniques such as plastid and nuclear microsatellites, random amplified polymorphic DNA, amplified fragment length polymorphisms and DNA sequencing as tools in the study of population dynamics (Fay & Kraus, 2003), species delimitation (e.g., Richardson & al., 2003), hybridization (e.g., Clarkson & al., 2004), and evolutionary relationships (Savolainen & Chase, 2003). They have also been applied, albeit on an *ad hoc* basis, as a tool to aid identification in situations in which it is difficult or unachievable using morphological characters, e.g., the identification of an unknown plant without reproductive characters (Bradford & Barnes, 2001). Despite the great utility of many of these techniques, a major drawback, especially when used for identification purposes, is that they are not standardized and in many cases not standardizable across a wide range of different organisms or for identifying the same organisms in different laboratories.

As a result of discussions and meetings on the potentially enormous benefits of a standardized molecular identification tool for plants and animals and with funding from the Alfred P. Sloan Foundation, the Consortium for the barcode of life (CBOL) was formed in 2004. The aims of the CBOL initiative are to develop "an accurate and reliable tool for scientific research on the taxonomy of plant and animal species, a practical, cost-effective tool for assigning unidentified specimens to their correct species, and a system for expanding interest and activity in taxonomy" (see http://www.barcoding.si.edu/). CBOL envisages this tool being a short gene sequence from a universally amplifiable region. The main aspiration of this initiative is not to create a molecular (DNA) taxonomy as controversially suggested by advocates of the phylocode (http://www.ohiou.edu/phylocode/) or even a DNA taxonomy in which DNA sequence data would be part of species typification (Tautz & al., 2002, 2003), but rather a tool to allow quick identification of organisms for a wider group of users than is possible at present. Several authors have expressed concerns that DNA barcoding may be seen as a replacement for taxonomy (e.g., Ebach & Holdrege, 2005; Moritz & Cicero, 2004), whereas others have pointed out that it is an identification tool which will at the same time provide additional data for use by taxonomists, by highlighting taxa for which further taxonomic effort would be profitable (e.g., Gregory, 2005; Schindel & Miller, 2005).

The papers of Hebert and colleagues' (Hebert & al., 2003a, b) proposed the use of the mitochondrial gene, cytochrome *c* oxidase subunit 1 (*COI* or *cox1*), as a suitable sequence for barcoding of animals, and there have been a series of publications since then illustrating its usefulness in a range of taxa (e.g., DeSalle & al., 2005; Hebert

& al., 2004; Janzen & al., 2005; Lorenz & al., 2005; Monogahan & al., 2005; Smith & al., 2005; Ward & al., 2005), although in some cases, for example amphibians (Vences & al., 2005), an alternative or additional region may be necessary. For several reasons as outlined below, this gene sequence is not suitable for use as a plant DNA barcode. Recently Kress and co-workers (2005) proposed two DNA regions (one in the plastid genome and one in the nuclear genome) that have the potential to serve singly or in tandem as a barcode for flowering plants. Here we discuss the progress that has been made in identifying a land plant DNA barcode, possible alternatives to those that have been proposed, as well as ongoing activities to use plant barcodes in biodiversity studies.

FINDING A DNA BARCODE FOR LAND PLANTS

Unlike animals for which mitochondrial DNA is highly conserved in terms of gene content and order, but with a high rate of sequence substitution (e.g., Brown & al., 1979; Moritz & al., 1987; Boore, 1999), higher plant mitochondrial genomes exhibit frequent rearrangements, transfer of genes to the nuclear genome, and incorporation of foreign genes (e.g., Palmer & al., 2000; Mower & al., 2004). Substitution rates are both much slower in plants than is usual in animals and with some notable exceptions in specific taxa (Bakker & al., 2000; Cho & al., 2004; Parkinson & al., 2005) the slowest of the three plant genomes (mitochondrial, plastid and nuclear). The success of a DNA sequence as a species identification tool—the barcode—depends on the existence of unique substitutions that distinguish among closely related species. The nuclear and plastid plant genomes therefore offer the best hope of yielding a suitable sequence(s) for DNA barcoding, i.e., ones that will be variable enough to differentiate among species and at the same time have less intra- than infra-specific variability (Chase & al., 2005; Kress & al., 2005).

Additionally the ideal barcode would be short enough to sequence in one piece with current technology (~700 base pairs [bp] or less), technically simple to sequence (e.g., without any long repeat regions that challenges the accuracy of DNA polymerases), length-conserved to mitigate alignment problems (although indels would undoubtedly provide useful diagnostic characters), and recoverable from herbarium samples and other degraded DNA samples such as alcohol-preserved (pickled) tissue, wood, fruits, etc. (Kress & al., 2005). Also, an effective barcode would use universal primers that can be applied across all land plant groups. Identifying hybrids would obviously be desirable, and in the case of long established natural hybrid species this should not be problematic. In cases of recent hybridization or ongoing introgression it is not possible to make a reliable identification using a single or even two plastid DNA regions and would require the ability to "disentangle" different allelic sequences from a single sample (Chase & al., 2005), which increases the technical demands and cost of barcoding. In plants plastid genes are uniparentally inherited (most often maternally); therefore identification of hybrids would necessitate inclusion of multiple single copy nuclear genes in the barcode. These single copy genes are at present technically demanding to sequence and generally not retrievable from herbarium or other degraded samples because their amplification is highly subject to DNA quality.

The lack of resolution typically encountered when only single DNA regions are used in phylogenetic reconstruction (especially in angiosperms) has led to concerns that it may not be possible to pinpoint a single, short plastid DNA region that could be used as the DNA barcode in plants. However, unlike phylogenetic reconstruction in which phylogenetically informative characters (those shared among two or more taxa) are necessary, the main requirement for a DNA barcode is sufficient DNA identifiers that are unique and universal within a species.

To search for an effective plant barcode, Kress & al. (2005) compared the total plastid genomes of tobacco (*Nicotiana tabacum*) and deadly nightshade (*Atropa belladonna*), close relatives in the family Solanaceae, to identify the most variable sequence regions. Trials of the resultant nine most variable plastid regions across a set of widely divergent angiosperm taxa, including closely related species, identified the *trnH-psbA* spacer as a good barcode candidate. This spacer, although short (~450 bp), has been found to be exceptionally variable in many plant groups (e.g., Shaw & al., 2005) and is easily amplified across a broad range of land plants. The second region identified by Kress and colleagues (2005) is the nuclear internal transcribed spacer (ITS) of ribosomal DNA, which is the most commonly sequenced locus in plant phylogenetic investigations at the species level and shows high levels of interspecific divergence (although problems with paralogues do occur in certain taxa). Together, these two DNA regions have the potential to discriminate among the largest number of plant species for barcoding purposes. Furthermore, extraction of DNA and amplification of the *trnH-psbA* spacer from herbarium specimens (some over 100 years old) was highly successful (Kress & al., 2005).

The trials by Kress & al. (2005) on the *trnH-psbA* plastid spacer and ITS included a total of 99 species in 80 genera encompassing 53 families of flowering plants, but did not adequately address within species variability nor the application of the plastid spacer to groups of land

plants other than angiosperms. It has already been demonstrated that intraspecific variation is low in ITS (Baldwin & al., 1995). With regards to the *trnH-psbA*, spacer recent trials in 33 species for which two to five individuals per species were sampled (W. J. Kress & I. Lopez, unpubl.) have shown that within-species variability is non-existent or exceptionally low (0.00–0.80% sequence divergence due to nucleic acid substitutions; Table 1), thereby enhancing the potential of this spacer as a barcode. In 24 of the 33 species sampled the intergenic *trnH-psbA* spacer showed no variation among individuals; in the remaining nine species most of the intraspecific variation was due to insertion-deletion events (see below). Equally important for a barcode, a search of GenBank revealed that the *trnH-psbA* spacer has been successfully amplified in gymnosperms, ferns, mosses, and liverworts, but not always using the same primers that work on angiosperms. The success of amplification of this spacer across a broad set of land plants has been verified in a more extensive survey of lycophytes and monilophytes (Small & al., 2005).

Results of trials in 33 species in 26 families of flowering plants and one family of conifers on intraspecific sequence variation of the *trnH-psbA* spacer. These species include many of the same taxa studied by Kress & al. (2005) and cover the same study site. Two to five individuals were sampled for each species. Variation only includes nucleic acid substitutions and not insertion-deletion events. Extraction, amplification, and sequencing procedures follow that of Kress & al. (2005).

Plant Family	Number of species sampled	Percent intraspecific sequence variation
Alliaceae	1	0.00
Apiaceae	1	0.00
Apocynaceae	1	0.00
Asteraceae	2	0.00
Balsamiaceae	1	0.00
Betulaceae	2	0.00–0.66
Boraginaceae	1	0.00
Brassicaceae	2	0.00–0.28
Caprifoliaceae	2	0.00
Celastraceae	1	0.80
Crassulaceae	1	0.00
Cupressaceae	1	0.00
Ericaceae	1	0.00
Fabaceae	1	0.00
Juglandaceae	1	0.00
Lamiaceae	1	0.00
Lauraceae	1	0.25
Malvaceae	1	0.00
Oleaceae	1	0.00
Platanaceae	1	0.00
Polemoniaceae	1	0.25
Rubiaceae	1	0.00
Sapindaceae	3	0.00–0.26
Simaroubaceae	1	0.00
Solanaceae	1	0.00
Ulmaceae	1	0.00
Vitaceae	1	0.00

In another case study we used the dataset of Perret & al. (2003) with complete species sampling of closely related taxa within the Gesneriaceae family (*Sinningia* s.l.) and found that the commonly used phylogenetic markers in this datatset do exhibit high discrimination among species and could also be suitable as barcodes in many cases (Table 2; see also Chase & al., 2005 for similar results in other taxa).

Although these studies have provided evidence that a plant DNA barcode with a high degree of utility can be implemented, they also illustrated problems that need to be overcome to achieve the ideal barcode. Probabilities of identifying the correct species in the *Sinningia* study were high but not 100% (Table 2), and the regions with the highest probabilities are too long to sequence in one piece. The *trnH-psbA* spacer used in Kress & al. (2005) ranged from 119 to >1,000 bp in length among the angiosperm taxa studied, and it is likely that this range would increase as other groups of land plants are included. This length variation, due to the high number of insertion-deletion events known to occur in *trnH-psbA* (Shaw & al., 2005), can lead to difficulties in alignment, which could be problematic for its use as a barcode. Although the spacer exhibits significant interspecific variation in most groups sampled so far, it still may fail to discriminate among species in those taxa of land plants with notoriously low genetic differentiation between species (e.g., palms; Carl Lewis, Fairchild Tropical Garden, pers. comm.; orchids, Mark Whitten, University of Florida).

In some groups, ITS will also be difficult to use because multiple, uncorrected copies are maintained, making it impossible to sequence without cloning and less likely to be an effective barcoding tool. New algorithms for combining barcoding sequences to yield species-level unique identifiers will have to be developed to deal with these problems of spacer length variation and the necessity of combining more than one DNA region for an effective plant barcode.

WORK IN PROGRESS

A Plant Working Group has been initiated under the auspices of CBOL, and an international consortium of 11 institutes is undertaking a project to expand on previous results with the aim of identifying the most suitable plastid DNA regions (one or perhaps two) to serve as a universal DNA barcode for all land plants. This initiative followed the first international conference on "Barcoding of Life", hosted by the Natural History Museum, London, on behalf of CBOL (for a selection of papers presented at the conference see *Philosophical Transactions of the Royal Society B: Biological Sciences* 360, 29 Oct. 2005). Expanding on the methodology and taxon sam-

Table 2. Probability (p) of identifying the correct species based on DNA sequences in the example of *Sinningia* s.l. (n = 96; Perret & al., 2003). Pair-wise distance matrices of absolute numbers of differences were computed using PAUP* 4.0b10 (Swofford, 2001). The probability of identifying the correct species was calculated as the proportion of comparisons in which at least one nucleotide difference was found between species pairs. Ideally, we would want genes that have more than just one nucleotide difference between species.

DNA regions (number of base pairs in the alignment)	p
trnS-G region (611 bp)	p = 0.95
trnT-L region (846 bp)	p = 0.96
*rpl*16 intron (1161 bp)	p = 0.98
trnL-F region (1042 bp)	p = 0.98
atpB-rbcL spacer (1099 bp)	p = 0.98
ncpGS (670 bp)	p > 0.99

pling of Kress & al. (2005) the project initially involved screening additional regions of the plastid genome and developing primers for the most promising regions that are universal across land plants. Screening plastid regions only, overcomes the problems of multiple variable copies as found in ITS, and the technical difficulties of sequencing single copy nuclear regions, especially when using poor quality or degraded DNA. An assessment of length variability is also being made in the candidate regions to try and overcome the difficulties associated with alignment and analysis of highly length variable regions such as the *trnH-psbA* spacer. These regions are being evaluated in a range of taxa with complete or near complete sampling at the species level to evaluate the probability of any one region providing correct identifications among many closely related species. The taxonomic groups being used in this evaluation, include all major lineages of land plants, and have been chosen for a variety of reasons, e.g., recent radiations with a difficult taxonomy expected to be more challenging versus older radiations with greater genetic divergence and a generally agreed upon taxonomy. More information on this project with ongoing updates until its completion can be found at www.rbgkew.org.uk/barcoding.

At the same time, several projects have been launched that include plant barcoding in a biosurvey or inventory-based setting. These include a "Darwin Initiative for the Survival of Species" funded project to RBG Kew on DNA barcoding of the orchids of Costa Rica in collaboration with the Lankester Botanical Garden, Costa Rica, and a project in collaboration with the University of Johannesburg and Royal Botanic Gardens, Kew, that aims to DNA barcode the flora of the Kruger National Park in South Africa. Projects are also being considered or already underway at the Smithsonian Institute to generate DNA barcodes for all economic plants, especially medicinals and poisonous plants that will have direct impact on the commercial sector and society.

PROSPECTS FOR USING A LAND PLANT DNA BARCODE

DNA barcoding *per se*, however, is not the only end point activity to be achieved in these planned projects. These barcode efforts can be clearly linked with interrelated objectives of basic biodiversity endeavors, such as: to (*i*) provide measures of biological diversity, (*ii*) provide plant DNA barcoding for conservation and trade surveillance, (*iii*) achieve high standards for research and training in basic taxonomy and link with global efforts to build the tree-of-life, and (*iv*) assist in implementation of the Global Strategy for Plant Conservation (Global Strategy for Plant Conservation, http://www.bgci.org.uk/policies/globalstrategystatement.html) developed by the Convention on Biodiversity and the various action plans of the World Conservation Union (IUCN).

Given the increasing ease and speed of DNA sequencing alongside decreasing costs, DNA barcoding will facilitate rapid and large-scale biodiversity surveys, both for inventory purposes and ecological studies. These could be performed without presorting of samples or the necessity for taxonomists to devote their time to highly repetitive identification rather than additional scientific research. It would allow the identification of different life stages, e.g., seeds or seedlings, and fragments of plant material that do not bear the requisite morphological characters for identification. There is an increasing demand for plant identification in the fields of international trade (CITES regulations), forensics, herbal medicines, and commercial foodstuffs. At the Royal Botanic Gardens, Kew, and the United States National Herbarium at the Smithsonian, we are regularly asked to identify potentially illegally imported plant materials seized by customs, the contents of herbal medicines, and plant material associated with scenes of crime.

Of course to be able to implement these potential uses and make them widely available a reference species database of plant DNA barcodes must be built using sequences from fully verified and vouchered samples, with multiple sampling per species to assess within-species variability. Hence it is desirable that the DNA barcode is retrievable from already existing herbarium specimens; this will bring an immense task into the realms of achievability. This level of sampling is also expected to highlight potential cryptic species and provide data for studies of species delimitation (see Hebert & al., 2004).

As discussed above, a DNA barcode based on one or two short plastid DNA regions will not identify hybrid taxa, but only one of the two parents (usually the maternal parent). In taxonomically complex groups (TCG), those with blurred species delimitations because of factors such as introgression, apomixis and backcrossing (Ennos & al., 2005), identification would likely be limit-

ed to a complex or group of taxa within the TCG. Provided the identification system includes information on the possibility of hybrids, complex groups, etc. for any individual barcode (see Chase & al., 2005) this should not be a problem. However, in situations or for particular taxa in which this would be insufficient, a local multi-region "secondary" barcode might need to be implemented.

DNA BARCODES OF THE FUTURE

So far, we have discussed the implementation of a DNA barcoding system for land plants that is based on one, or if necessary two, short plastid and/or nuclear DNA sequences from universally amplifiable region(s) across land plants. Sequencing technology already allows production of some 400,000 barcodes from just one DNA sequencing machine per year, assuming bidirectional reads on a 96 capillary machine running 24 hours per day, 50 weeks per year. The major bottleneck in the system is therefore collection, verification, and DNA extraction of samples, with the use of herbaria specimens being critical. However the technology involved is advancing rapidly; not only the rate of barcode production, but also the portability and accessibility of a sequencing device will undoubtedly increase. Pyrosequencing and chip based sequencing technologies already allow the rapid sequencing of many short (~ 60 bp) barcode regions. This opens the way for an enhancement of the "basic" barcode proposed to incorporate short multiple regions of different genomes to be incorporated in a single barcode that would not only provide species-level identification, but also provide data valuable to improving our understanding of hybridization, speciation, and population dynamics.

The three billion biological specimens now housed in the world's natural history museums and herbaria offer an immense resource for building the DNA barcode library for both plants and animals. The application of barcodes to these specimens will therefore greatly increase their value. However, there is a considerable time and cost investment required to extract DNA from the many vouchered and verified samples that will be required to build a plant DNA barcode reference database. We therefore urge, that wherever feasible, these samples are banked as a valuable resource for future biodiversity research (Savolainen & Reeves, 2004; Chase & al., 2005; Savolainen & al., 2006).

"Imagine what it would do to any and all aspects of human interactions with wild plants if you could walk up to any plant anywhere—seedling, sapling, 40 m tree, grass, root, pressed leaf, or fallen log—and know in a few seconds its scientific name. …. That capacity would transform far more than the science of plant biology, the conservation of plants, and the superficial ways we currently make use of the incredible diversity of form, physiology, genetics and chemistry of plants. It would be to plants what the printing press was to stories, education, science, law, medicine and communication" (Janzen, 2005). Once fully developed, DNA barcoding has the potential to completely change not only how biologists understand and monitor biodiversity, but also, as emphasized by Janzen in the above quote, the relationship of the general public to nature. When a hand-held DNA "barcorder" (portable DNA sequencing device) becomes available in a few years, the new technology will help many non-scientists, whether they are in the field, garden, or market, to quickly and inexpensively identify known species and retrieve information about them. If implemented successfully, barcoding will provide a vital new tool for appreciating and managing the Earth's immense and changing biodiversity.

ACKNOWLEDGEMENTS

We thank the Consortium for the Barcode of Life, the Alfred P. Sloan and Gordon and Betty Moore Foundations, the Darwin Initiative, the South African NRF, the University of Johannesburg, and the Smithsonian Institution for funding. Laboratory and field assistance to WJK was provided by I. Lopez, K. Wurdack, and N. Bourg.

LITERATURE CITED

Bakker, F. T., Culham, A., Pankhurst, C. E. & Gibby, M. 2000. Mitochondrial and chloroplast DNA-based phylogeny of *Pelargonium* (Geraniaceae). *Amer. J. Bot.* 87: 727–734.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S. & Donoghue, M. J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Ann. Missouri Bot. Gard.* 82: 247–277.

Bradford, J. C. & Barnes, R. W. 2001. Phylogenetics and classification of Cunoniaceae (Oxalidales) using chloroplast DNA sequences and morphology. *Syst. Bot.* 26: 354–385.

Boore, J. L. 1999. Animal mitochondrial genomes. *Nucl. Acids Res.* 27: 1767–1780.

Brown, W. M., George, M. & Wilson, A. C. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* 76: 1967–1971.

Chase, M. W., Salamin, N., Wilkinson, M., Dunwell, J. M., Kesanakurthi, R. P., Haidar, N. & Savolainen, V. 2005. Land plants and DNA barcodes: short-term and long-term goals. *Philos. Trans. Royal Soc., Ser. B.* 360: 1889–1895.

Cho, Y., Mower, J. P., Qiu, Y.-L. & Palmer, J. D. 2004. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc. Natl. Acad. Sci. U.S.A* 101: 17741–17746.

Clarkson, J. J., Knapp, S., Garcia, V. F., Olmstead, R. G., Leitch, A. R. & Chase, M. W. 2004 Phylogenetic relationships in *Nicotiana* based on multiple plastid loci. *Molec. Phylog. Evol.* 33: 75–90.

DeSalle, R., Egan, M. G. & Siddall, M. 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philos. Trans. Royal Soc., Ser. B.* 360: 1905–1916.

Ebach, M. C. & Holdrege, C. 2005. DNA barcoding is no substitute for taxonomy. *Nature* 434: 697–697.

Ennos, R. A., French, G. C. & Hollingsworth, P. M. 2005. Conserving taxonomic complexity. *Trends Ecol. Evol.* 20: 164–168.

Fay, M. F. & Krauss, S. L. 2003. Orchid conservation genetics in the molecular age. Pp. 91–112 in: Dixon, K. W., Kell, S. P., Barrett, R. L. & Cribb, P. J. (eds.), *Orchid Conservation*. Natural History Publications, Kota Kinabalu, Sabah.

Gregory, T. R. 2005. DNA barcoding does not compete with taxonomy. *Nature* 434: 1067–1067.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & De Waard, J. R. 2003a. Biological identifications through DNA barcodes. *Phil. Trans. Roy. Soc., Ser. B.* 270: 313–321.

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

Hebert, P. D. N., Ratnasingham, S. & De Waard, J. R. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Phil. Trans. Roy. Soc., Ser. B.* 270: S96–S99.

Janzen, D. H. 2005. Foreword. Pp. x–xii in: Krupnick, G. A. & Kress, W. J. (eds.), *Plant Conservation: A Natural History Approach.* Univ. Chicago Press, Chicago.

Janzen, D. H., Hajibabaei, M., Burns, J. M., Hallwachs, W., Remigio, E. & Hebert, P. D. N. 2005. Wedding biodiversity inventory of a large complex *Lepidoptera* fauna with DNA barcoding. *Phil. Trans. Roy. Soc., Ser. B.* 360: 1835–1845.

Kress, J. W., Wurdack, K. J., Zimmer, E. A., Weigt, L. A. & Janzen, D. H. 2005. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. U.S.A.* 102: 8369–8374.

Lorenz, J. G., Jackson, W. E., Beck, J. C. & Hanner, R. 2005. The problem and promise of DNA barcodes for species diagnosis of primate biomaterial. *Phil. Trans. Roy. Soc., Ser. B.* 360: 1869–1877.

Monogahan, M. T., Balke, M., Gregory, T. R. & Vogler, A. V. 2005. DNA-based species delineation in tropical beetles using nuclear and mitochondrial markers. *Phil. Trans. Roy. Soc., Ser. B.* 360: 1925–1933.

Moritz, C. & Cicero, C. 2004. DNA barcoding: promise and pitfalls. *Plos Biol.* 2: 1529–1531.

Moritz, C., Dowling, T. E. & Brown, W. M. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annu. Rev. Ecol. Syst.* 18: 269–292.

Mower, J. P., Stefanovic, S., Young G. J. & Palmer, J. D. 2004. Plant genetics: gene transfer from parasitic to host plants. *Nature* 432: 165–166.

Palmer, J. D., Adams, K. L., Cho, Y., Parkinson, C. L., Qiu, Y.-L. & Song, K. 2000. Dynamic evolution of plant mitochondrial genomes: Mobile genes and introns and highly variable mutation rates. *Proc. Natl. Acad. Sci. U.S.A* 97: 6960–6966.

Parkinson, P. L., Mower, J. P., Qiu, Y.-L., Shirk, A. J., Song, K., Young, N. D., dePamphilis, C. W. & Palmer, J. D. 2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evolutionary Biology* 5: 73.

Perret, M., Chautems, A., Spichiger, R., Kite, G. & Savolainen, V. 2003. Systematics and evolution of tribe Sinningieae (Gesneriaceae): evidence from phylogenetic analyses of six plastid DNA regions and nuclear *ncpGS*. *Amer. J. Bot.* 90: 445–460.

Richardson, J. E., Fay, M. F., Cronk, Q. C. B. & Chase, M. W. 2003. Species delimitation and the origin of populations in island representatives of *Phylica* (Rhamnaceae). *Evolution* 57: 816–827.

Savolainen, V. & Chase, M. W. 2003. A decade of progress in plant molecular phylogenetics. *Trends Genet.* 19: 717–724.

Savolainen, V., Powell, M. P., Davies, K., Cothals, A. & Reeves, G. (eds.) 2006. *DNA Banking for Biodiversity and Conservation*. Kew Publishing and IUCN, Kew.

Savolainen, V. & Reeves, G. 2004. A plea for DNA banking. *Science* 304: 1445.

Schindel, D. E. & Miller, S. E. 2005. DNA barcoding a useful tool for taxonomists. *Nature* 435: 17–17.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J. Siripun, K. C., Winder, C. T., Schilling, E. E. & Small, R. L. 2005. The tortoise and the hare II: Comparison of the relative utility of 21 non-coding chloroplast DNA sequences for phylogenetic analysis. *Amer. J. Bot.* 92: 142–166.

Small, R. L., Lickey, E. B., Shaw J. & Hauk, W. D. 2005. Amplification of noncoding chloroplast DNA for phylogenetic studies in lycophytes and monilophytes with a comparative example of relative phylogenetic utility from Ophioglossaceae. *Molec. Phylog. Evol.* 36: 509–522.

Smith, M. A., Fisher, B. L. & Hebert, P. D. N. 2005. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Phil. Trans. Roy. Soc., Ser. B.* 360: 1825–1834.

Swofford, D. L. 2001. PAUP* 4.0: *Phylogenetic Analysis using Parsimony (* and other Methods)*. Sinauer Associates, Sunderland, Massachusetts.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2002. DNA points the way ahead in taxonomy—in assessing new approaches, it's time for DNA's unique contribution to take a central role. *Nature* 418: 479.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2003. A plea for DNA taxonomy. *Trends Ecol. Evol.* 18: 70–74.

Vences, M., Thomas, M., Bonett, R. M. & Vieites, D. R. 2005. Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Phil. Trans. Roy. Soc., Ser. B.* 360: 1859–1868.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. N. 2005. A start to DNA barcoding Australia's fish species. *Phil. Trans. Roy. Soc., Ser. B.* 360: 1847–1857.