

Nós Realmente Precisamos de *DNA Barcoding*?

Do We Really Need DNA Barcoding?

Rosane Garcia Collevatti*

Laboratório de Genética e Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal de Goiás – UFG, GO, Brasil

O conhecimento atual sobre a riqueza global de espécies ainda é bastante inadequado. Estima-se que o número global de espécies esteja entre 2 a 100 milhões de espécies. Entretanto, somente cerca de 1,4 milhões de espécies foram reconhecidas até hoje, sendo cerca de 280.000 plantas e 950.000 insetos. Novas espécies vêm sendo descobertas continuamente, mas não são necessários muitos argumentos para mostrar a imensa lacuna entre o nosso conhecimento e o que ainda precisamos conhecer. Com certeza, este será o grande desafio para a humanidade neste século.

O desconhecimento se deve, principalmente, a nossa incapacidade de descrever formalmente os espécimes já coletados e estocados nas coleções biológicas, fenômeno que é conhecido como “o déficit Linneano” (“*Linnean shortfall*”) (Whittaker *et al.* 2005). Igualmente, entendemos muito superficialmente a distribuição geográfica da maioria das taxa, o que compromete uma boa delimitação da ocorrência das espécies e estudos biogeográficos, fenômeno denominado de “déficit Wallaceano” (“*Wallacean shortfall*”) (Whittaker *et al.* 2005). Uma das restrições para contornar os déficits Linneano e Wallaceano é o número de taxonomistas, profissionais considerados atualmente como “espécie em extinção”: o GTI – *Global Taxonomy Initiative* (<http://www.cbd.int/gti/>) estima que há entre 20.000-30.000 pessoas com alguma qualificação em taxonomia no mundo inteiro (incluindo autônomos e amadores) e, provavelmente, entre 5.000-7.000 taxonomistas profissionais (acadêmicos). O declínio no número de espécies novas descritas e no número de taxonomistas dentro da comunidade acadêmica tem sido amplamente discutido e muitos cientistas reconhecem atualmente que a taxonomia é uma “ciência em crise” (*e.g.* Lee 2000; Hopkins & Freckleton 2002). Além da falta de taxonomistas, projetos de taxonomia ou que visem a formação de novos taxonomistas recebem pouco financiamento, principalmente quando comparados a outras áreas do conhecimento, e a infraestrutura dos museus e coleções é inadequada e mesmo incipiente em muitos países.

Todos estes fatores restringem nosso acesso à Biodiversidade, o que vem sendo chamado de “impedimento taxonômico à conservação da biodiversidade” (Wheeler *et al.* 2004). No Brasil, apesar das últimas iniciativas do MCT-CNPq (Ministério de Ciência e Tecnologia-Conselho Nacional de Desenvolvimento Científico Tecnológico) e também do MEC-CAPES (Ministério de Educação-Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), com programas de financiamento específicos, como o PROTAX, a situação não é diferente. O provisionamento de recurso ainda é incipiente perante o enorme desafio de alcançar uma infraestrutura mínima e a formação de recursos humanos adequadas. Esta situação é crítica porque, ao mesmo tempo em que as estimativas de diversidade global estão alcançando dezenas de milhares de espécies desconhecidas, as ameaças à diversidade global em diferentes escalas e de diferentes fontes (como perda de habitat, mudanças climáticas) estão aumentando.

Certamente, o desenvolvimento de ferramentas moleculares e de outras ferramentas analíticas, como a comunicação pela internet e a disponibilidade de *hardware* e *softwares* mais avançados para o gerenciamento de coleções levou a uma aceleração do reconhecimento de novas espécies (veja para uma discussão detalhada Graham *et al.* 2004). Mas, como apontado por Wheeler *et al.* (2004, p. 285), “[...] *some naively see the information technology challenge as liberating data from cabinets*”. Ou seja, a disponibilidade de tecnologias de ponta de alto desempenho possibilita uma aceleração na obtenção de dados e de informação de modo geral, mas a aceleração na interpretação dos dados e, conseqüentemente, sua real “liberação dos gabinetes”, depende de outros fatores que vão além da obtenção destes dados, mas passam pela disponibilidade de recurso humano qualificado.

DNA Barcoding – Uma Inovação “Chique” Para a Identificação de Espécies?

Hebert *et al.* (2003) propuseram o uso de uma sequência de DNA padrão como uma “etiqueta” para proporcionar a identificação das espécies, chamada de *DNA Barcoding* (código de barras de DNA). Os autores garantiram que o

*Send correspondence to: Rosane Garcia Collevatti
Laboratório de Genética e Biodiversidade,
Instituto de Ciências Biológicas,
Universidade Federal de Goiás – UFG,
CP 131, CEP 74001-970, Goiânia, GO, Brasil
E-mail: rosanegc68@hotmail.com

uso do *DNA Barcoding* permitirá a identificação rápida, inequívoca e automatizada das espécies, aumentando a velocidade de descoberta de novas espécies.

Não há, fundamentalmente, nada novo na proposta de *DNA Barcoding* exceto o aumento na escala de geração de dados e a padronização. Os pesquisadores já vêm utilizando seqüências para delimitar espécies críticas e para estudos filogenéticos e filogeográficos baseados em amostras vinculadas a vouchers (Avice 2004). Especificamente, o *DNA Barcoding* foi definido como uma seqüência padrão de tamanho mínimo (i.e. pode ser facilmente seqüenciada usando somente um iniciador senso e um antisenso), posteriormente depositado em um banco de dados principal ou central, e vinculado a um voucher cuja origem e *status* atual são também registrados no banco de dados (Hebert &

Gregory 2005). Resumidamente, a atribuição do *DNA Barcoding* a um grupo de espécies deve incluir protocolos padronizados, como os procedimentos para a coleta de tecidos e para o laboratório, tudo vinculado ao número do voucher e a informações relevantes sobre o voucher (Figura 1). A região de DNA escolhida para ser *Barcoding* deve conter variabilidade e divergência genéticas significativas em nível de espécie, ser amplamente distribuída nos taxa, ser passível de seqüenciamento com iniciadores universais e ser pequena o suficiente para ser seqüenciada facilmente (Kress & Erickson 2008). A seqüência do gene mitocondrial *co1* (*citocromo c oxidase subunidade I*), que codifica uma das subunidades da enzima *citocromo c oxidase* da cadeia de transporte de elétrons, presumivelmente atende a estes critérios, e foi designado como *DNA Barcoding*. Neste sentido, um consórcio foi criado para fornecer uma biblioteca

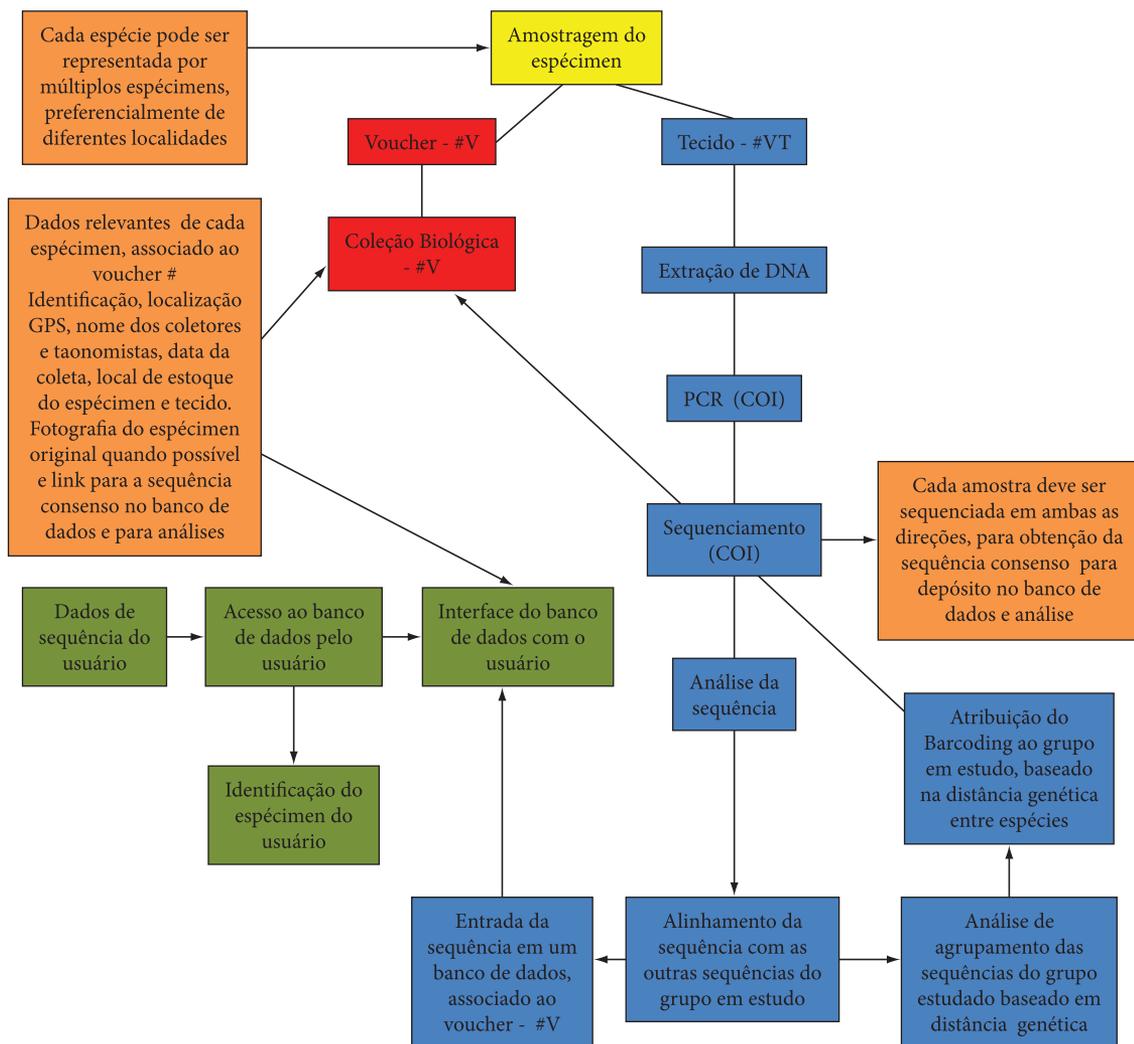


Figura 1. Estrutura do desenvolvimento de *DNA Barcoding*. Quadrados em azul representam as análises moleculares desde a amostra de tecidos para extração de DNA até a análise e entrada das seqüências no banco de dados. Quadrados vermelhos representam o depósito do *voucher* na coleção biológica com a entrada de todos os dados relevantes no banco de dados. Quadrados verdes representam o acesso público aos dados através da interface com o usuário. Quadrados laranja representam as informações importantes sobre os dados e o espécimen.

universal de *DNA Barcoding*, o “*Consortium for the Barcode of Life*” (www.barcoding.si.edu), para comparações de taxa não identificados com o banco de dados de *Barcoding*.

Muitas promessas foram feitas sobre a aplicação do *DNA Barcoding*. Por exemplo, Hebert & Gregory (2005, p. 853) observaram que a aplicação do *DNA Barcoding* será uma ferramenta poderosa como facilitador da taxonomia alfa:

“[...] we are confident that DNA barcoding will play an increasingly important role as a taxonomic screening tool because of its ability to rapidly reveal the genetic discontinuities that ordinarily separate distinct species [...] Its application in this fashion will allow an inversion of standard taxonomic approaches that operate in an a priori fashion—seeking the morphological discontinuities that signal reproductive isolation among unsorted assemblages of organisms. By contrast, DNA barcoding allows a more efficient a posteriori approach where predefined, genetically divergent groups are examined for trait variation. In this sense, DNA barcoding will clearly be a powerful enabler of alpha taxonomy.”

Entretanto, para a maioria das espécies de plantas o gene *coI* não é aplicável como *DNA Barcoding* devido a taxa evolutiva lenta, e também à instabilidade da estrutura do genoma mitocondrial (e.g. Wolfe *et al.* 1987; Mackenzie 2007). Além disso, a grande variação nas taxas evolutivas no genoma de cloroplasto entre e dentro das famílias de plantas (e.g. Shaw *et al.* 2007), dificulta a determinação de uma sequência de *DNA Barcoding* universal para as plantas. Assim, o debate sobre qual a sequência que irá identificar as plantas em nível de espécie ainda está em aberto (e.g. Chase *et al.* 2005; Kress *et al.* 2005; Taberlet *et al.* 2006; Fazekas *et al.* 2008; *CBOL Plant Working Group* 2009).

Olhando dentro da “caixa-preta”, o *DNA Barcoding* segue o mesmo princípio do conceito de espécie de Carl Linnaeus (1707-1778). A abordagem utilizada de associar uma “etiqueta” a uma espécie (uma sequência padronizada, chamada de *Barcoding*) sugere a idéia que cada espécie existe como uma “forma ideal”, sem variação, ou com variação mínima, entre indivíduos e populações. De outra forma seria impossível “etiquetar” uma espécie com um único identificador, como se faz com o código de barras dos produtos nas prateleiras dos supermercados. A teoria de evolução das espécies por meio da seleção natural de Charles Darwin (1809-1882) tem duas implicações importantes para esta discussão, que fundamentalmente desafia as suposições ou idéias por trás do *DNA Barcoding*. Primeiro, ela estabelece que as espécies não são só similares, mas elas são realmente relacionadas por descendência. Segundo, ela supõe que as espécies não são homogêneas, fixas – membros de uma espécie são diferentes e as espécies devem mudar ao longo do tempo:

“From these remarks it will be seen that I look at the term species, as one arbitrarily given for the sake of convenience to a set of individuals closely resembling each other, and it does not essentially differ from the term

variety, which is given to less distinct and more fluctuating forms. The term variety, again, in comparison with mere individual differences, is also applied arbitrarily and for mere convenience sake” (Darwin 1859, p. 52, Chapter II).

Van Valen (1973) também destacou a variação genética natural nas populações. Ele sugeriu que as espécies estão sob constante seleção natural de forma a seguir as mudanças no ambiente e se elas não tiverem variação genética adequada elas extinguem porque elas não conseguirão seguir a evolução dos parasitas e predadores. Esta evolução do “armamento de luta” (“*arm race*” evolution) foi chamada de “Hipótese da Rainha Vermelha” (the “*Red Queen Hypothesis*”), em referência ao diálogo da Rainha Vermelha com Alice na estória “*Through the Looking Glass*” de Lewis Carroll: “*Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!*” (Carroll, 1871, p. 136), “*Though the looking-glass and what Alice found there*”).

Este paralelo feito por Van Vallen sugere que as espécies não tem delimitação clara, mas ao contrário são produto das populações onde as frequências gênicas estão constantemente mudando ao longo do tempo e do espaço (De Queiroz 2005). Assim, embora o conceito Linneano para identificar as espécies seja útil e corriqueiramente utilizado, nós não podemos pensar nas espécies como entidades imutáveis e invariáveis. Desta forma, a pergunta que temos que fazer é: “quão útil é a técnica de *DNA Barcoding* em um mundo onde as espécies são variáveis?”

A um primeiro olhar, seria o melhor dos mundos – o uso de uma ferramenta nova, rápida e de baixo custo para resolver o velho problema de identificar a biodiversidade com precisão. Mas, além dos problemas conceituais apontados acima, o uso de métodos fenéticos para atribuir um espécimen a um ponto no “espaço n-dimensional de sequências”, e então definir a espécie, não leva em conta a variação dentro e entre espécies e as relações filogenéticas entre as espécies (para uma discussão detalhada sobre este aspecto veja Will & Rubinoff 2004; Moritz & Cicero 2004; DeSalle *et al.* 2005; Rubinoff *et al.* 2006). Mesmo que Herbert e colaboradores usem o termo “*profile*” (perfil) em vez de filogenia, provavelmente para evitar interpretações erradas obtidas a partir de resultados fenéticos como relações filogenéticas, o problema ainda permanece.

Então, Nós Realmente Precisamos do *DNA Barcoding*?

Não é surpreendente que a abordagem de *DNA Barcoding*, como proposta por Herbert *et al.* (2003), funcione bem em taxa bem diferenciados. Como comentado por Will & Rubinoff (2004), em taxa bem diferenciados as distâncias entre as espécies são suficientemente grandes para não serem afetadas por taxa ou indivíduos com diferenças intermediárias. O mecanismo proposto de usar um limiar de distância genética para separar a variação intraespecífica da interespecífica para a avaliação de um espécimen não

identificado dentro de um agrupamento (Hebert *et al.* 2003, 2004) é vulnerável e passível de erros pela indicação tanto de falsos positivos quanto de falsos negativos (Meyer & Paulay 2005), porque a acurácia do método depende do nível de sobreposição na variação ao longo de toda a filogenia (Will & Rubinoff 2004; Moritz & Cicero 2004; Meyer & Paulay 2005). Falsos positivos são taxa novos identificados erroneamente dentro de uma espécie na qual a variação intraespecífica é mais profunda que o limiar escolhido. Por outro lado, falsos negativos são taxa identificados erroneamente dentro de um agrupamento no qual as divergências interespecíficas são mais superficiais do que o limiar proposto (para detalhes veja Meyer & Paulay 2005).

Assim, a resposta à questão acima “Nós realmente precisamos de *DNA Barcoding*?” também é uma questão: Nós precisamos de *DNA Barcoding* para identificar espécies bem definidas? Ou: “O *DNA Barcoding* fornecerá alguma vantagem adicional em relação a identificação morfológica para espécies bem definidas?” Além disso, e mais importante, “A abordagem de *DNA Barcoding*, conforme proposta por Herbert e colaboradores, fornece identificação de espécies rápida, precisa e automatizável para *taxa* taxonomicamente problemáticos?” Eu temo que a resposta a todas estas questões seja “não”.

Espécies bem definidas não são um problema real. A identificação baseada em caracteres morfológicos é tipicamente fácil e, em muitos casos, os indivíduos podem ser identificados no campo pelo pesquisador ou pelo taxonomista para posterior checagem. Assim, para este caso, qualquer vantagem adicional dada pelo sequenciamento para a identificação é negligenciável (Will & Rubinoff 2004; Rubinoff *et al.* 2006). Para plantas, uma solução baseada em uma única sequência (Fazekas *et al.* 2008), ou mesmo a utilidade do *DNA Barcoding* está longe do ideal, não só pela variação, mas também devido a hibridização e retenção de polimorfismos ancestrais (*e.g.* Fazekas *et al.* 2009; Gonzalez *et al.* 2009). Além disso, muitas espécies são polifiléticas ou parafiléticas, o que implica que não podem ser diferenciadas pelo método de *DNA Barcoding* (Funk & Omland 2003). Desta forma, para espécies estreitamente relacionadas, como espécies crípticas, grupos complexos altamente variáveis, geralmente amplamente distribuídos e taxonomicamente desafiadores, a diferenciação morfológica e genética é difícil e intrincada e a atribuição de uma espécie é ambígua (Will & Rubinoff 2004; Meyer & Paulay 2005). Neste caso, a contribuição do *DNA Barcoding*, conforme proposto por Hebert *et al.* (2003), para identificar as espécies é nula porque a ambiguidade permanecerá.

Então, do que Realmente Precisamos?

É notório que a abertura da “caixa preta” do “impedimento taxonômico à conservação da biodiversidade” nos revela que as causas principais do nosso desconhecimento da biodiversidade são: falta de taxonomistas e infraestrutura inadequada. O fomento para formação de recursos humanos

na área de taxonomia ainda é tímido para o enorme desafio de identificar e descrever novas espécies. Especificamente no Brasil, a falta de um plano de carreira específico para os curadores de coleções também desestimula a formação de novos taxonomistas. Somado a isto, a falta de técnicos de nível superior para dar apoio em campo e nas coleções, e a falta de infraestrutura adequada para estoque, triagem, sistematização, gerenciamento e formação de banco de dados das coleções comprometem a conservação, velocidade de identificação e descrição de novas espécies. Por fim, o excesso de burocracias que precisam ser atendidas e superadas, tanto em relação às obrigações dentro das universidades, onde estão grande parte das coleções no Brasil, como em relação às várias normas e leis dos diversos órgãos do governo que devem ser atendidas para a realização de coletas, transporte de amostras, permuta entre coleções, etc., contribuem para que o processo de identificação e descrição de espécies seja uma máquina paquidérmica.

Agradecimentos

Ao Editor-Chefe, Prof. José Alexandre F. Diniz-Filho pelas sugestões e correções durante a elaboração do manuscrito.

Referências

- Avice JC, 2004. *Molecular markers, natural history, and evolution*. 2nd ed. Sunderland, Massachusetts: Sinauer Associates, Inc. 684 p.
- Carroll L, 1871. Through the looking-glass and what Alice found there. In: Green RL, editor. *The works of Lewis Carroll*. London: The Hamlyn Publ. Group Ltda. Chap. 2, p. 111-220.
- Chase MW *et al.*, 2005. Land plants and DNA barcodes: short-term and long-term goals. *Philosophical Transactions of the Royal Society B*, 360:1889-1895. PMID:16214746. PMCID:1609218. <http://dx.doi.org/10.1098/rstb.2005.1720>
- CBOL Plant Working Group, 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Science USA*, 106:12794-12797. PMID:19666622. PMCID:2722355. <http://dx.doi.org/10.1073/pnas.0905845106>
- Darwin C, 1859. *On the origin of species by means of natural selection*. London: John Murray.
- DeSalle R, Egan MG & Siddall M, 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transaction of the Royal Society B*, 360:1905-1916. PMID:16214748. PMCID:1609226. <http://dx.doi.org/10.1098/rstb.2005.1722>
- Fazekas AJ *et al.* 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE*, 3:e2802. PMID:18665273. PMCID:2475660. <http://dx.doi.org/10.1371/journal.pone.0002802>
- Fazekas AJ *et al.* 2009. Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Molecular Ecology Resources*, 9:130-139. <http://dx.doi.org/10.1111/j.1755-0998.2009.02652.x>
- Funk DJ & Omland KE, 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with

- insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution and Systematics*, 34:397-423. <http://dx.doi.org/10.1146/annurev.ecolsys.34.011802.132421>
- Gonzalez MA *et al.* 2009. Identification of Amazonian trees with DNA barcodes. *PLoS ONE*, 4:e7483. PMID:19834612. PMCid:2759516. <http://dx.doi.org/10.1371/journal.pone.0007483>
- Graham CH *et al.* 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, 19:487-503. <http://dx.doi.org/10.1016/j.tree.2004.07.006>
- Hebert PDN *et al.*, 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B*, 270:313-321. PMID:12614582. PMCid:1691236. <http://dx.doi.org/10.1098/rspb.2002.2218>
- Hebert PDN *et al.*, 2004. Identification of birds through DNA barcodes. *PLoS Biology*, 2:e312. PMID:15455034. PMCid:518999. <http://dx.doi.org/10.1371/journal.pbio.0020312>
- Hebert PDN & Gregory TR, 2005. The Promise of DNA Barcoding for Taxonomy. *Systematic Biology*, 54:852-859. PMID:16243770. <http://dx.doi.org/10.1080/10635150500354886>
- Hopkins GW & Freckleton RP, 2002. Declines in the numbers of amateur and professional taxonomists: implications for conservation. *Animal Conservation*, 5:245-249. <http://dx.doi.org/10.1017/S1367943002002299>
- Kress WJ *et al.*, 2005. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences USA*, 102:8369-8374. PMID:15928076. PMCid:1142120. <http://dx.doi.org/10.1073/pnas.0503123102>
- Kress WJ & Erickson DL, 2008. DNA barcodes: Genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences USA*, 105:2761-2762. PMID:18287050. PMCid:2268532. <http://dx.doi.org/10.1073/pnas.0800476105>
- Lee MSY, 2000. A worrying systematic decline. *Trends in Ecology and Evolution*, 15:346. [http://dx.doi.org/10.1016/S0169-5347\(00\)01907-8](http://dx.doi.org/10.1016/S0169-5347(00)01907-8)
- Mackenzie SA, 2007. The unique biology of mitochondrial genome instability in plants. *Annual Plant Reviews*, 31:36-49.
- Meyer CP & Paulay G, 2005. DNA Barcoding: error rates based on comprehensive sampling. *PLoS Biology*, 3:e422. PMID:16336051. PMCid:1287506. <http://dx.doi.org/10.1371/journal.pbio.0030422>
- Moritz C & Cicero C, 2004. DNA barcoding: Promise and pitfalls. *PLoS Biology*, 2:e354. PMID:15486587. PMCid:519004. <http://dx.doi.org/10.1371/journal.pbio.0020354>
- De Queiroz K, 2005. Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences USA*, 102:6600-6607. PMID:15851674. PMCid:1131873. <http://dx.doi.org/10.1073/pnas.0502030102>
- Rubinoff D, Cameron S & Will K, 2006. Are plant DNA barcodes a search for the Holy Grail? *Trends in Ecology and Evolution*, 21:1-2. <http://dx.doi.org/10.1016/j.tree.2005.10.019>
- Shaw J *et al.*, 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, 94:275-288. <http://dx.doi.org/10.3732/ajb.94.3.275>
- Taberlet P *et al.* 2006. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, 35(3):e14. <http://dx.doi.org/10.1093/nar/gkl938>
- Van Valen L, 1973. A new evolutionary law. *Evolutionary Theory*, 1:1-30.
- Wheeler QD, Raven PH & Wilson EO, 2004. Taxonomy: Impediment or Expedient? *Science*, 303:285. PMID:14726557. <http://dx.doi.org/10.1126/science.303.5656.285>
- Whittaker RJ *et al.* 2005. Conservation biogeography: assessment and prospect. *Diversity and Distributions*, 11:3-23. <http://dx.doi.org/10.1111/j.1366-9516.2005.00143.x>
- Will K & Rubinoff D, 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics*, 20:47-55. <http://dx.doi.org/10.1111/j.1096-0031.2003.00008.x>
- Wolfe KH, Li W & Sharp PM, 1987. Evolution rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceeding of the National Academy of Science USA*, 84: 9054-9058. <http://dx.doi.org/10.1073/pnas.84.24.9054>

Recebido: Dezembro 2010
Primeira Decisão: Fevereiro 2011
Aceito: Fevereiro 2011

