植物学报 *Chinese Bulletin of Botany* 2010, **45** (1): 102–108, www.chinbullbotany.com
doi: 10.3969/j.issn.1674-3466.2010.01.015

·专题论坛·

DNA条形码技术在植物中的研究现状

闫化学*, 于杰

西南大学园艺园林学院, 重庆 400716

摘要 DNA条形码技术(DNA barcoding)是用短的DNA片段对物种进行识别和鉴定的分子生物学技术。在动物研究中该技术已经成功应用于利用线粒体细胞色素c氧化酶亚基I(COI)进行物种鉴定和发现隐种或新物种。相对于动物, *COI*基因在高等植物中进化速率较慢, 因此植物条形码研究以叶绿体基因组作为重点, 但目前还处于寻找合适的基因片段阶段。许多学者对此进行了积极的探索, 报道了多种植物条形码的候选片段或组合, 但还没有获得满足所有标准的特征位点片段。该文介绍了DNA条形码的标准、优点、工作流程及数据分析方法, 总结了DNA条形码在植物中的研究现状。

关键词 *COI*基因, DNA条形码, 植物

闫化学, 于杰 (2010). DNA条形码技术在植物中的研究现状. 植物学报 **45**, 102–108.

自卡尔·林奈对生物物种进行系统分类以来, 生物学家利用各种各样的性状——颜色、外形和行为等来鉴定动物和植物。最近数十年, 研究者开始利用DNA中携带的遗传信息来完成这个任务。DNA条形码(DNA barcoding)技术是一种利用短的DNA片段对物种进行识别和鉴定的新的分子生物学技术, 是生物学近期研究的热点之一。2003年, 加拿大圭尔夫大学(University of Guelph) Paul Hebert教授提出了“DNA条形码”概念, 将条形码技术引入生物界。其思想产生于现代商品零售业的条形编码系统, 就像以超市条形码识别产品一样, 利用A、T、C和G 4个碱基在基因中的排列顺序识别物种。Paul Hebert教授率先于2003年选取线粒体细胞色素c氧化酶亚基 I (cytochrome *c* oxidase subunit 1, COI)作为动物中通用的物种鉴定标记(Hebert et al., 2003a, 2003b), 并提出DNA条形码的定义: 通过使用短的标准DNA片段, 对物种进行快速、准确的识别和鉴定。目前, DNA条形码技术在很多动物分类群中得到了成功应用(Hebert et al., 2004a, 2004b; Ward et al., 2005; Hajibabaei et al., 2006; Yoo et al., 2006; Kerr et al., 2007), 有关植物条形码的工作也在进行中, 重点是选择合适的片段并对其进行评价(Pennisi, 2007; Kress and Erickson, 2008)。本文就有关研究工作的进展综述如下。

1 DNA条形码的标准及优点

理论上, DNA条形码利用A、T、C和G 4个碱基, 只需要15个碱基位点就可以得到4^{15}个排列方式, 远大于现存物种的数量(肖金花等, 2004)。Meyer和Paulay(2005)根据每百万年2%的物种进化速率, 推测对于一个有1百万年生殖隔离历史的物种类群, 一段长度为600 bp的DNA序列平均就有12个特征信号位点可用于识别, 即可以包括所有物种。即使在亲缘关系很近的类群中, 大多数物种的进化历史都超过了1百万年, 所以600 bp的DNA片段足够分析当前绝大多数的物种。

Kress等(2005)和Taberlet等(2007)提出了理想的DNA条形码标准, 综合起来有以下几点: (1)具有可以区分物种的足够变异和分化, 同时种内变异必须足够小; (2)有高度保守的引物设计区以便于设计通用引物; (3)片段足够短, 以便于DNA提取和PCR扩增, 尤其是对部分降解的DNA的扩增。

生命条形码联盟(consortium for the barcode of life, CBOL)阐述了DNA条形码的优点(http://phe.rockefeller.edu/barcode/), 可以概括为以下几点。(1)以DNA序列为检测对象, 其在个体发育过程中不会

收稿日期: 2008-12-15; 接受日期: 2009-03-13
基金项目: 国家自然科学基金(No. 30671450)
* 通讯作者。E-mail: ylp840203@163.com

改变。同种生物不同生长时期的DNA序列信息是相同的，即使经过加工，形态发生变化，而DNA序列信息不会改变，较之传统的方法，扩大了检测样本的范围；同时样本部分受损也不会影响识别结果。(2)可进行非专家物种鉴定。该技术是可机械重复的，只要设计一套简单的实验方案，经过简单培训的技术员即可操作。(3)准确性高。特定的物种具有特定的DNA序列信息，而形态学鉴别特征会因趋同和变异导致物种的鉴定误差。(4)通过建立DNA条形码数据库，可一次性快速鉴定大量样本。分类学家新的研究成果将不断地加入数据库，成为永久性资料，从而推动分类学科更加快速深入地发展。

2 DNA条形码的操作及分析方法

DNA条形码的操作过程与分子生物学实验类似，包括采集材料并提取DNA、利用通用引物PCR扩增目的片段、纯化PCR产物、序列测定与分析以及提交结果到相关数据库(Hajibabaei et al., 2006)。在寻找植物条形码的过程中，生物信息学的应用可能会起到非常重要的作用(Newmaster et al., 2006; Kress and Erickson, 2008)，条形码的应用目标是所有物种，并且每个物种需要多份材料，地球上的陆生植物大约需要200万–300万份材料(Newmaster et al., 2006)。采集材料时应以传统的形态分类学知识为依据，尽可能地涵盖传统分类学中的变异式样，材料来源范围越广泛，得到的结果就越具说服力。通常认为每个物种至少需要10份材料，并最好包括5个不同居群，这样既可以发掘出变异程度较低的种内变异，又可以更准确地计算出种内遗传距离(Stoeckle and Hebert, 2008)。

生物条形码工程的首要目标是建立可用来作为鉴定标本工具的基因序列数据库(Hebert et al., 2004a)。目前只建立了动物条形码数据库，植物条形码尚处于评估阶段，只有该技术在植物中进一步完善后才有可能建立相应的数据库。建立植物条形码数据库时，可参照名为生命条形码系统(barcode of life data system, BOLD, http://www.barcodinglife. org)的动物数据库，该数据库目前已经收录46多万条记录，涵盖了动物界46 000多个物种。上传的每一条记录都应包括物种名称、条形码区序列、采集地点、对照样本链接、图像信息以及其它生物学数据(Stoeckle and Hebert, 2008)。

序列数据分析是DNA条形码探索的最重要环节。动物条形码的分析方法成熟且简单，首先进行序列比对和人工校正，通过MEGA或PAUP计算种内和种间的K2P距离(Kimura-2-parameter distance)用于表示不同分类阶元之间的序列变异程度，比较种、属和科3个水平上的序列差异，然后根据计算结果建立NJ树(neighbour-joining tree)，最后依据DNA条形码遗传距离就能对未知标本进行分类和鉴定。然而由于植物的种间杂交现象比较普遍，因此植物条形码的分析方法也处于不断的摸索中，目前报道的分析方法基本可分为以下几步。

(1) 序列比对和人工校正。一般采用Clustal W，生命条形码数据库中使用Hidden Markov Models进行序列比对(Newmaster et al., 2006)，也可以用BLAST search在Genbank中搜索相似的基因片段，对比片段信息的可靠性。

(2) 遗传分析。计算遗传距离是条形码分析的重要手段，当前植物条形码探索还处于对不同片段的评价阶段，需要对不同片段的种间和种内变异进行对比，以选择最佳的片段或片段组合。种间距离基本采用Kimura-2-parameter distance (K2P)模型计算(Meyer and Paulay, 2005; Newmaster et al., 2006; Lahaye et al., 2008a, 2008b)，该模型也是生命条形码联盟(CBOL)建议使用的遗传距离计算模型(http://www.barcoding.si.edu)。不同学者计算遗传距离的方法也有差异。Newmaster等(2008)使用mega3.1计算P距离(p-distance)评价肉豆蔻科植物的种间和种内变异。Kress和Erickson(2007)使用Wilcoxon Signed Rank Tests比较片段的种间和种内差异；Meyer和Paulay(2005)则报道了一个评价DNA条形码的重要指标barcoding gap，即用柱形图表示种间和种内的遗传距离分布频率，对所研究的片段进行检验。理想的DNA条形码检测到的种间遗传变异应明显大于种内遗传变异，且会形成一个明显的间隔区，在理想状况下，柱形图上的种内变异会集中在数值较小的一侧，而种间变异则集中在数值较高的一侧。Median and Wilcoxon Two-Sample Tests可以用来评估种内和种间变异的分布是否存在重叠(Lahaye et al., 2008a)。

(3) 系统学分析。采用标准的系统学分析方法，建立系统发生树，检验同一个物种的不同个体能否聚

在一起。Lahaye等(2008a)使用PAUP4b10*建树分析，结果发现MP树和UPGMA树可以得到较高的物种正确识别率，并进行了聚并分析(coalescence analyses)，检验每个条形码，验证那些从一个独立的聚并过程得到的聚类群是否与以前识别的物种分类相匹配。

另外，序列特征分析将加快植物条形码研究的进程。DNA序列包涵丰富的生物和化学信息，不同物种的DNA序列存在明显差异。分析目标片段时，首先要对同一DNA片段进行统计分析，这种方法已经用于蛋白质序列的研究分析中(戴晓峰等，2007)。植物叶绿体基因组相对比较保守，得到某一片段的全部序列后，用CLUSTAL进行序列对齐，再用MEGA统计GC含量、可变位点和信息位点。形态学上属于同一物种的类群，则它们的序列会呈现特有的固定特征，可以为片段选择提供依据。

3 DNA条形码在植物中的研究现状

虽然DNA条形码技术在动物分类中已经体现了其使用价值，但是在植物中的研究进展则比较缓慢。研究结果显示，*COI*基因在植物中的进化速率远慢于在动物中的进化速率，因此只适用于一些藻类(Saunders, 2005; Evans et al., 2007)。为了寻求一种通用的植物条形码，许多学者进行了积极的探索，已经提出了多种条形码片段或组合方案，但至今仍没有达成明确的共识(Kress et al., 2005; Newmaster et al., 2006; Chase et al., 2007; Kress and Erickson, 2008; Lahaye et al., 2008a, 2008b)。

生命条形码联盟(CBOL)最初建议了6个片段作为候选片段：*matK*、*rpoC1*、*rpoB*、*accD*、*nhdJ*和*YCF5*。随着研究的不断扩大，不同研究小组提出了自己的观点，在所有涉及的片段中，还没有一个片段能够满足DNA条形码的3个标准。Kress等(2005)通过对53科88属99个物种的9个质体基因片段(*trnK-rps16*、*trnH-psbA*、*rp136-rps8*、*atpB-rbcL*、*ycf6-psbM*、*trnV-atpE*、*trnC-ycf6*、*psbM-trnD*和*trnL-trnF*)和核基因ITS序列进行分析，认为被子植物中普遍存在杂交和多倍体现象，某一个片段不太可能对所有的植物物种进行准确鉴定，因此提出了片段组合方案，并预测ITS+*trnH-psbA*组合将在有花植物(flowering plants)中有较广泛的应用价值。Rubinoff等(2006a)认为可以用rDNA作为质体基因的补充，作为未来使用的低拷贝核基因标记。但由于核基因本身具有多拷贝的特性，扩增时对DNA要求较高，部分降解的DNA材料扩增效果较差，并且长度变异大，存在长的poly-G、poly-C和poly-A结构等(Kress and Erickson, 2007; Sass et al., 2007)，其作为植物条形码候选片段存在较多的限制。而全球大部分植物都包含叶绿体，且叶绿体基因组是单亲遗传，避免了基因重组的影响。因此，应在叶绿体基因组中寻找植物条形码片段，片段组合方案也是目前学者研究的重点。

现有的研究报道了在2种组合方案中选择片段的方法，分别是Chase等(2005)提出的交通灯法(traffic light approach)和Newmaster等(2006)提出的等级(tier)分类法。2007年9月，在台北举行的第二届国际生命条形码大会上，总结了5种片段组合，分别是：Chase等(2007)提出的*rpoC1*+*rpoB*+*matK*或*rpoC1*+*matK*+*trnH-psbA*；Kress和Erickson(2007)提出的*rbcL*+*trnH-psbA*，其可以对陆生植物进行识别和鉴定；韩国植物学家Ki-Joong Kim等提出的*matK*+*atpF-atpH*+*psbK-psbI*或*matK*+*atpF-atpH*+*trnH-psbA*(Pennisi, 2007)。片段组合方案为植物条形码研究指出了一个新方向，但现有的研究结果并不能满足DNA条形码标准，仍需要大量的实验数据去验证它们的物种识别和鉴定能力。

首先，作为研究重点的片段组合还需要更深入的研究。Chase等(2007)提出的2个组合在Sass等(2007)和Fazekas等(2008)的研究中识别能力并不很理想。对于Ki-Joong Kim等提出的2个组合，Lahaye等(2008a)通过对克鲁格国家公园从单子叶植物到真细菌类的18科38种101份材料进行分析，结果显示*psbK-psbI*和*atpF-atpH*片段的识别成功率较高，分别为98%和93.1%，当2个片段组合时，成功率均达到100%；采用UPGMA聚类法，*matK*+*atpF-atpH*+*psbK-psbI*和*matK*+*atpF-atpH*+*trnH-psbA*组合的物种单系性分辨率分别为93.1%和89.3%。Fazekas等(2008)对来自32属92种251份材料的9个片段(*rpoB*、*rpoC1*、*matK*、*trnH-psbA*、*rbcL*、ITS1、*cox1*、*atpF-atpH*和*psbK-psbI*)进行了分析，结果发现*psbK-psbI*的扩增成功率仅为79%，而*atpF-atpH*则达到88%。在成功扩增的个体中，单独使用*psbK-psbI*和*atpF-atpH*识别物种的正确率分别仅为44%和45%。*matK*+*atpF-atpH*+*psbK-psbI*组合的识别正确

率最高，也只达到了69%。实验结果中没有出现特别理想的片段或组合，从扩增成功率和物种识别率的综合表现看，Fazekas等(2008)建议，应在编码片段*rbcL*、*rpoB*、*matK*和非编码片段*trnH-psbA*和*atpF-atpH*中选择最佳的片段组合。

其次，目前研究的片段也都有一定的局限性，主要体现在如下方面。

(1) 关于*matK*最主要的争议在于其引物的通用性差(Chase et al., 2007)。Sass等(2007)在寻找苏铁目植物条形码时，对*matK*片段未能得到理想的扩增效果，在所选的其它片段中，也没有发现单一片段(或组合)能正确识别和鉴定这类植物。Kress和Erickson(2007)对48属96个物种*matK*片段的扩增成功率仅为39.3%；Newmaster等(2008)对肉豆蔻科内毛楠属(*Compsoneura* Warb)8个物种的*matK*片段扩增也均失败。但是在Lahaye等(2008a, 2008b)的结果中，*matK*片段扩增成功率分别达到了100%和99%，且有较高的物种识别和鉴定能力，他们认为*matK*可以作为单一的植物DNA条形码片段，对于不能适用的类群可配合其它的片段组合使用。针对该结果，Hollingsworth(2008)指出，在Lahaye的实验材料中物种缺乏重复个体，而且没有包括同属内关系最近的姊妹种，这样则可能提高了识别率。Fazekas等(2008)使用Lahaye报道的*matK*引物，对32属92种陆生植物物种进行分析，*matK*扩增成功率低于50%，并且在双向测序上也存在问题。

(2) *rbcL*被大部分学者作为研究对象。Chase等(2005)提出了一些较为典型的系统发育标记，其中就包括*rbcL*片段。但是, Kress等(2005)认为在有花植物中，*rbcL*进化速度较慢，不适合作为种水平上的物种鉴定标记。Newmaster等(2006)通过对GenBank中10 000多个长度大于1 000 bp的*rbcL*片段进行分析，发现其能识别同属85%的物种，并认为*rbcL*不仅可以作为苔藓类植物的单片段DNA条形码，还可以作为其它植物DNA条形码片段组合方案的1个候选片段。*rbcL*在序列测定上也存在问题，完整的*rbcL*片段长度约为1 400 bp，需要设计测序引物才能完成测序(Kress et al., 2005)。而且在随后的研究中，很多研究者都提出，*rbcL*的变异只存在于种以上的分类水平，在物种水平上变异不够大。

(3) *trnH-psbA*间隔区是叶绿体基因进化速率最快的片段之一，并且易于设计引物(Shaw et al., 2005)。Fazekas等(2008)建议*trnH-psbA*间隔区应作为条形码片段组合中重要的非编码区片段之一。Kress等(2005)认为*trnH-psbA*间隔区长度较短，适合于DNA提取和扩增，在所用的材料中，92%的物种扩增所得的片段长度一般为340–660 bp，具有物种水平的遗传变异和分化，且片段两端具有保守区易于设计引物，可以作为植物条形码片段之一。Vischi等(2006)通过对种内和种间序列变异的评价，认为*trnH-psbA*间隔区的变异足以正确地鉴定4个向日葵物种，双向测序拥有较高的准确性。Kress和Erickson(2007)比较分析了来自39目43科48属的96个种的9个候选片段(*rpoB*、*rpoC1*、*matK*、*trnH-psbA*、*rbcL*、ITS、*accD*、*nhdJ*和*YCF5*)，结果表明所有片段中，*trnH-psbA*在扩增成功率和物种识别率方面表现最好。Newmaster等(2008)对肉豆蔻科内毛楠属的研究表明，*trnH-psbA*可以对每个种产生特异片段，并能成功识别70%的种。Lahaye等(2008a, 2008b)的研究结果也显示，*trnH-psbA*有较高的扩增成功率，特别是对兰科植物的识别率达到90%以上，尤其对近缘种有较好的识别能力。DNA条形码应用的一个难题就是难于识别近缘种和近期分化的物种，*trnH-psbA*在这方面表现较好，但由于有过多的插入/缺失使其难以用于非同属物种间的比对。Logacheva等(2008)分析了代表所有独活属(*Heracleum* L.)及其近缘属主要分支的物种，结果显示*trnH-psbA*序列变异较小，不足以识别物种，不适合作为这个种群体的条形码片段。就现有的研究结果来说，即使*trnH-psbA*不能单独作为植物条形码使用，也仍将是片段组合中的重要组成部分。

(4) 其它非重点研究的片段也存在一些缺点。*YCF5*、*accD*和*ndhJ*等片段在一些重要植物中缺失；Taberlet等(2007)和Chase等(2007)对叶绿体*trnL*(UAA)内含子全序列(254–767 bp)及其P6环(10–143 bp)进行分析后认为，虽然该片段的引物高度保守，扩增体系稳定，P6环在高度降解的DNA样本中仍可以扩增出来，但由于该片段进化速率太慢，作为DNA条形码片段应用存在困难；UPA(universal plastid amplicon)片段在陆生植物中没有显著变异，因此也失去了研究价值(Sass et al., 2007; Fazekas et al., 2008; Newmaster et al., 2008)。

植物基因组的进化与动物基因组的完全不同。动物存在生殖隔离，不同物种的动物无法繁殖杂交后

代，可以此为标准鉴定动物物种，但许多植物物种间可以互相杂交，这就模糊了它们之间的遗传边界(genetic boundary)，造成了植物在种的水平上有较大的差异，不同物种同一片段的进化速率是不同的。因此在寻求整个植物界通用的条形码过程中，可首先找出适合于大部分植物的片段或组合，其它少部分则作为特例，再针对不同的科属选择不同的条形码。编码基因受选择的压力大，通常变异较小，特别是叶绿体基因组相对比较保守，使用编码基因和非编码基因的组合，可能会更好地识别和鉴定物种。多片段组合将是植物条形码选择的最终方案。

4 存在的争议和前景

经过近几年的发展，DNA条形码技术已经取得了较大进步，也得到了许多生物学家的支持(Gregory, 2005; Schindel and Miller, 2005)，但同时也有部分专家学者对此持怀疑态度(Ebach and Holdrege, 2005)，对该技术的价值存在一些争议，认为这个新技术会削弱或者取代以形态学为基础的传统分类方法。Rubinoff等(2006b)认为，从目前可用的数据量和对基因组结构的了解来看，找到像动物DNA条形码COI那样的、易于使用的单一植物基因片段的可能性非常渺茫。Dasmahapatra和Mallet(2006)则认为，目前的研究方法存在偏颇，很多研究只对每个物种的1或2个个体进行了分析，或只在一个受限制的地理区域内采样，这就低估了种内变异；有的分析中没有包括姊妹类群，从而可能导致种间差异被高估，造成实验的准确率提高。

尽管还存在上述诸多问题，但目前DNA条形码技术已经在动物中得到了广泛应用，植物条形码的研究正在开展，相信通过大规模的分析和整体评价，最终将会找出通用且适合的植物条形码。未来可以利用形态学、生物地理学和DNA序列数据分析，揭示隐存种(van Velzen et al., 2007)。隐存种不是新物种，是指在传统分类法中，没有被划分出来，被归属为同一个物种的不同物种。例如，Hebert等(2004a)研究了哥斯达黎加森林中的弄蝶(skipper butterflies)，它们之前被认为属于同一个种类(*Astraptes fulgerator* Walch)，后来经分析发现这些蝴蝶的DNA条形码被很清楚地归入了10个不同的组中，表明这些蝴蝶属于10个不同的种类。此外，还可以联合使用形态学和分子生物学方法显著改善生物安全的问题(Armstrong and Ball, 2005)。例如，我们可能会使用条形码工具(例如条形码仪)进行例行鉴定，发现新物种，解决生态问题，控制外来物种入侵，以及将之应用于食品和中草药产业的质量控制中(Newmaster et al., 2006)。目前的研究也绝不能仅仅限于现行通用的分子标记，应大胆尝试一些新的有解决问题潜力的基因。

致谢 衷心感谢西南大学园艺园林学院周志钦教授在资料收集上提供的帮助以及在论文撰写和修改过程中给予的建议！

参考文献

戴晓峰，肖玲，武玉花，吴刚，卢长明 (2007). 植物脂肪酸去饱和酶及其编码基因研究进展. 植物学通报 **24**, 105–113.

肖金花，肖晖，黄大卫 (2004). 生物分类学的新动向———DNA条形编码. 动物学报 **50**, 852–855.

Armstrong KF, Ball SL (2005). DNA barcodes for biosecurity: invasive species identification. *Philos Trans R Soc Lond B Biol Sci* **360**, 1813–1823.

Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madriñán S, Petersen G, Seberg O, Jørgsensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M (2007). A proposal for a standardised protocol to barcode all land plants. *Taxon* **56**, 295–299.

Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, Savolainen V (2005). Land plants and DNA barcodes: short-term and long-term goals. *Philos Trans R Soc Lond B Biol Sci* **360**, 1889–1895.

Dasmahapatra KK, Mallet J (2006). DNA barcodes: recent successes and future prospects. *Heredity* **97**, 254–255.

Ebach MC, Holdrege C (2005). DNA barcoding is no substitute for taxonomy. *Nature* **434**, 697–697.

Evans KM, Wortley AH, Mann DG (2007). An assessment of potential diatom ''barcode'' genes (*cox*1, *rbc*L, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist* **158**, 349–364.

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SC (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One* **3**, e2802.

Gregory TR (2005). DNA barcoding does not compete with taxonomy. *Nature* **434**, 1067–1067.

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006). DNA barcodes distinguish species of tropical Lepidoptera. *Proc Natl Acad Sci USA* **103**, 968–971.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003a). Biological identifications through DNA barcodes. *Proc Biol Sci* **270**, 313–321.

Hebert PDN, Ratnasingham S, deWaard JR (2003b). Barcoding animal life: cytochrome coxidase subunit 1 divergences among closely related species. *Proc Biol Sci* **270** (suppl.), S96–S99.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004a). 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 PDN, Stoeckle MY, Zemlak TS, Francis CM (2004b). Identification of birds through DNA barcodes. *PLoS Biol* **2**, e312.

Hollingsworth PM (2008). DNA barcoding plants in biodiversity hot spots: progress and outstanding questions. *Heredity* **101**, 1–2.

Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, Hebert PDN (2007). Comprehensive DNA barcode coverage of North American birds. *Mol Ecol Notes* **7**, 535–543.

Kress WJ, Erickson DL (2007). A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One* **2**, e508.

Kress WJ, Erickson DL (2008). DNA barcodes: genes, genomics, and bioinformatics. *Proc Natl Acad Sci USA* **105**, 2761–2762.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005). Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci USA* **102**, 8369–8374.

Lahaye R, Savolainen V, Duthoit S, Maurin O, van der Bank M (2008a). A test of *psbK-psbI* and *atpF-atpH* as potential plant DNA barcodes using the flora of the Kruger National Park (South Africa) as a model system. Available from Nature Precedings <http://hdl.handle.net/10101/npre.2008.1896.1>.

Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V (2008b). DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci USA* **105**, 2923–2928.

Logacheva MD, Valiejo-Roman CM, Pimenov MG (2008). ITS phylogeny of West Asian *Heracleum* species and related taxa of Umbelliferae–Tordylieae W.D.J.Koch, with notes on evolution of their *psb*A-*trn*H sequences. *Plant Syst Evol* **270**, 139–157.

Meyer CP, Paulay G (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol* **3**, e422.

Newmaster SG, Fazekas AJ, Ragupathy S (2006). DNA barcoding in land plants: evaluation of *rbc*L in a multigene tiered approach. *Can J Bot* **84**, 335–341.

Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J (2008). Testing candidate plant barcode regions in the Myristicaceae. *Mol Ecol Resour* **8**, 480–490.

Pennisi E (2007). Wanted: a barcode for plants. *Science* **318**, 190–191.

Rubinoff D, Cameron S, Will K (2006a). A genomic perspective on the shortcomings of mitochondrial DNA for "Barcoding" identification. *J Hered* **97**, 581–594.

Rubinoff D, Cameron S, Will K (2006b). Are plant DNA barcodes a search for the Holy Grail? *Trends Ecol Evol* **21**, 1–2.

Sass C, Little DP, Stevenson DW, Specht CD (2007). DNA barcoding in the Cycadales: testing the potential of proposed barcoding markers for species identification of Cycads. *PLoS One* **2**, e1154.

Saunders GW (2005). Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philos Trans R Soc Lond B Biol Sci* **360**, 1879–1888.

Schindel DE, Miller SE (2005). DNA barcoding a useful tool for taxonomists. *Nature* **435**, 17–17.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu WS, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005). The tortoise and the hare. II. Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am J Bot* **92**, 142–166.

Stoeckle MY, Hebert PDN (2008). Bar code of life: DNA tags help classify animals: inspired by commercial bar codes, DNA tags could provide a quick, inexpensive way to identify species. *Sci Am* **299**(4), 82–88.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E (2007). Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding. *Nucleic Acids Res* **35**, e14.

van Velzen R, Bakker FT, van Loon JJA (2007). DNA barcoding reveals hidden species diversity in *Cymothoe* (Nymphalidae). *Proc Neth Entomol Soc Meet* **18**, 95–103.

Vischi M, Arzenton F, De Paoli E, Paselli S, Tomat E, Olivieri AM (2006). Identification of wild species of sun-

flower by a specific plastid DNA sequence. *Helia* **29**(45), 11–18.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005). DNA barcoding Australia's fish species. *Philos Trans R Soc Lond B Biol Sci* **360**, 1847–1857.

Yoo HS, Eah JY, Kim JS, Kim YJ, Min MS, Paek WK, Lee H, Kim CB (2006). DNA barcoding Korean birds. *Mol Cells* **22**, 323–327.

Current Status of the Study of DNA Barcoding in Plants

Huaxue Yan*, Jie Yu

College of Horticulture and Landscape Architecture, Southwest University, Chongqing 400716, China

Abstract DNA barcoding with a short DNA region is used to identify and characterize species of organisms. The mitochondrial cytochrome *c* oxidase subunit 1 (COI or *cox1*) sequence is used in DNA barcoding for diverse groups of animals, with success in species identification and revealing cryptic species or new species. However, the sequence is not appropriate for most plant species because of its slower rate of evolution in higher plants than in animals. Plant barcoding is still in the stage of searching for a suitable locus and kinds of candidates. Many researchers propose a single barcoding region or combinations of regions from the chloroplast genome. At present, a well-characterized plant locus that meets all necessary criteria is lacking. This paper briefly discusses the advantages, standards, workflow, and analysis of barcoding, as well as the status and problems of current studies in plants.

Key words *COI*, DNA barcoding, plants

Yan HX, Yu J (2010). Current status of the study of DNA barcoding in plants. *Chin Bull Bot* **45**, 102–108.

---

* Author for correspondence. E-mail: ylp840203@163.com

(责任编辑: 刘慧君)