

基于叶绿体 *ndhA* 基因内含子序列的 DNA 条形码在秋海棠属物种鉴定中的应用

赵 博¹, 李景剑², 毛世忠¹, 唐文秀¹

(1. 中国科学院 广西植物研究所, 广西 桂林 541006; 2. 华南农业大学 林学与风景园林学院, 广东 广州 510642)

摘 要:以 18 种秋海棠属植物为试材, 基于叶绿体 *ndhA* 基因内含子片段, 利用 Codon Code Aligner 软件拼接序列后, 通过 MEGA 软件进行比对并计算 19 个秋海棠属植物种间及种内遗传距离, 最后利用邻接法构建分子系统树, 研究了该片段作为 DNA 条形码对秋海棠属植物进行物种鉴定的可行性。结果表明: 秋海棠属 19 个种类共 59 个个体的 *ndhA* 基因内含子序列长度为 1 182 bp, 在所考察的候选秋海棠属植物中具有最大的种间变异和较小的种内变异, 且二者存在极显著差异。在系统树中, 秋海棠属植物每一物种能够分别形成各自独立的分支, 表现出良好单系性。基于 *ndhA* 基因内含子的 DNA 条形码在识别秋海棠属植物物种方面和传统形态学基本一致。该研究表明, 以 *ndhA* 基因内含子作为秋海棠属植物 DNA 条形码进行物种鉴定具有一定的可行性。

关键词:秋海棠属; *ndhA* 基因内含子; DNA 条形码; 物种鉴定

中图分类号:Q 948 **文献标识码:**A **文章编号:**1001-0009(2016)16-0103-05

秋海棠属(*Begonia* L.)为秋海棠科(Begoniaceae)一年或多年生草本、灌木或小树枝状植物, 全世界约 1 500~1 600 种^[1]。在中国约有 173 种, 分布于长江以南的地区, 常见于云南、广西和贵州^[2-3]。由于秋海棠属植物在习性、生境、性别、形态等方面均表现出很高的多样性, 品种繁多, 其中大部分具有极高的观赏价值, 在国内外花卉市场颇受人们的喜爱^[4]。另外, 秋海棠属植物也是我国少数民族中常见的民族药, 具有清热解毒、散瘀消肿的功效^[5-6]。由此可见, 秋海棠属植物中蕴藏着丰富的观赏、药用资源有待进一步的开发利用。然而,

秋海棠属植物性状受环境影响较大, 即使同一种类的不同居群, 在某些性状上的变异也比较显著, 很多近缘种类难以鉴定^[7]。同时该属大多是多汁草本植物, 具有肉质的茎、娇嫩的花和果实, 传统方法压制成标本后, 在叶片颜色、花纹、毛被颜色和花被片颜色等方面都有很大程度的失真, 果实和子房的形态解剖学特征等是属下分组不可缺少的重要鉴定特征, 在多汁的果实被压干以后容易变形, 给传统的分类鉴定带来较大的困难^[8-9]。因此, 为秋海棠属植物寻找快速、高效的分类鉴定方法一直都是研究的热点。

DNA 条形码(DNA barcoding)是通过使用短的标准 DNA 片段, 对物种进行快速、准确的识别和鉴定^[10-13]。DNA 条形码不仅是传统植物分类与鉴定的强有力补充, 而且能突破对经验的过度依赖, 可以在较短时间内建成易于利用的应用系统, 在物种分类和鉴定方面展示出了强大的生命力^[14]。植物 DNA 条形码片段主要在叶

第一作者简介:赵博(1981-), 女, 河南南阳人, 博士, 副研究员, 研究方向为植物细胞学与分子系统发育。E-mail: calljone@163.com.

责任作者:唐文秀(1973-), 女, 本科, 副研究员, 现主要从事珍稀濒危植物保护等研究工作。E-mail: tangwx99@163.com.

基金项目:广西自然科学基金资助项目(2012GXNSFBA053075)。

收稿日期:2016-04-26

Abstract: Taking chestnut as test material, GBSS gene was cloned from the first strand of *Castanea mollissima* seed's cDNA through RT-PCR methods, and the bioinformatics characters were analysed. The results showed that a full length cDNA of *Cgbss* (GenBank accession number: KU162945) was obtained. The length of *Cgbss* gene was 2 085 bp, which contained an open reading frame encoding 611 amino acids. The bioinformatics analysis indicated that the pI and MW of amino acids encoded by *Cgbss* were predicted to be 8.36 and 67.580 9 kDa, respectively, and the protein had one GT1_glycogen_synthase domain. And the fragment had 85% homology identical to the already reported sequence of GBSS I.

Keywords: *Castanea mollissima*; granule-bound starch synthase; clone; sequence analysis

绿体基因组上进行选择,此外还有核基因 ITS^[15-17]。因为用于 DNA 条形码研究的标准基因,一方面应该足够保守,能够利用通用引物进行大范围的扩增;另一方面应该有足够的变异来区分不同物种 DNA 序列,从而进行鉴定^[18]。而目前为止,植物 DNA 条形码尚处于对所提议的各片段比较和评价阶段,还未获得统一的标准^[19]。

目前已有关于利用 DNA 条形码对秋海棠属植物进行物种鉴定的相关研究报道,但是这些研究结果表明,叶绿体基因 *rbcL*、*matK* 和 *psbA-trnH* 种内和种间变异都较小,都不足以区分秋海棠属植物,只有 ITS/ITS2 种内和种间变异大,可考虑作为秋海棠属 DNA 条形码鉴定的候选片段^[8,20-21]。因此,在鉴定秋海棠属植物时还需选择更多的有效基因。该试验应用植物 DNA 条形码片段 *ndhA* 基因内含子区对中国秋海棠属植物 19 个种类共 59 个个体进行了研究,评价其 PCR 扩增和测序成功率,种内种间差异及物种鉴定能力,探讨该基因片段是否适合作为秋海棠属的候选 DNA 条形码。

表 1 供试 18 种秋海棠样品情况

Table 1 *Begonia oxyloba* samples for test

样品名称 Samples name	采集地点 Location	GenBank 号 GenBank No.	样品名称 Samples name	采集地点 Location	GenBank 号 GenBank No.
		KT599066			KT599047
<i>Begonia jingxiensis</i>	广西桂林植物园、广西靖西	KT599067	<i>Begonia pseudodaxinensis</i>	广西桂林植物园、广西大新	KT599048
		KT599068			KT599049
		KT599077			KT599072
<i>Begonia filiformis</i>	广西桂林植物园、广西龙州	KT599078	<i>Begonia grandis</i>	广西桂林植物园	KT599073
		KT599079			KT599091
		KT599045	<i>Begonia chingii</i>	广西桂林植物园	KT599092
<i>Begonia umbraculifolia</i>	广西桂林植物园、广西大新	KT599046			KT599093
					KT599039
		KT599063			KT599040
<i>Begonia leprosa</i>	广西桂林植物园、广西大尧山	KT599064	<i>Begonia fimbriatipula</i>	广西桂林、广东鼎湖山、云南昆明植物研究所	KT599041
		KT599065			KT599042
		KT599088			KT599043
<i>Begonia daxinensis</i>	广西桂林植物园、广西大新	KT599089	<i>Begonia palmata</i>	广西金秀	KT599054
		KT599090			KT599055
		KT599050			KT599056
		KT599051	<i>Begonia longifolia</i>	云南	KT599060
<i>Begonia picturata</i>	广西桂林植物园、广西靖西	KT599052			KT599061
		KT599053			KT599062
		KT599057			KT599069
<i>Begonia luochengensis</i>	广西桂林植物园、广西罗成	KT599058	<i>Begonia handelii</i>	广西桂林植物园、广西金秀	KT599070
		KT599059			KT599071
		KT599080			KT599094
<i>Begonia fangii</i>	广西桂林植物园、广西龙州	KT599081	<i>Begonia cathayana</i>	广西桂林植物园、广西上思	KT599095
		KT599082			KT599096
					KT599083
		KT599074			KT599084
<i>Begonia gigaphylla</i>	广西桂林植物园、广西龙州	KT599075	<i>Begonia edulis</i>	广西桂林植物园、广西靖西	KT599085
		KT599076			KT599086
					KT599087

1 材料与方法

1.1 试验材料

采集分布于中国广西、云南和广东地区的秋海棠属植物 18 种共 58 个个体(表 1),标本存放于中国科学院广西植物研究所,同时从 GenBank 数据库中下载 *Begonia oxyloba* 序列 1 条(GenBank ID:JF756335),共计 19 种 59 个个体。

1.2 试验方法

1.2.1 秋海棠 DNA 的提取 由于秋海棠属植物的 DNA 较难提取,在提取总 DNA 时,用液氮将叶片彻底磨成粉末状后,加入 CTAB-free+2%PVP 溶液将粉末混匀,然后放入 -20 ℃ 冰冻 10 min 后,离心弃上清液,再采用植物基因 DNA 提取试剂盒(Tiangen Biotech Co., China)按说明书操作步骤提取样品总 DNA。通过紫外分光光度计和 1% 琼脂糖凝胶电泳检测 DNA 的质量和浓度,最后将浓度调整为 30~50 ng · μL⁻¹,于 -20 ℃ 贮存备用。

1.2.2 序列扩增 选用 *ndhA* 基因内含子区通用引物 *ndhAx1* (GCY CAA TCW ATT AGT TATG AAA TACC) 和 *ndhAx2* (GGT TGA CGC CAM ARA TTC CA)。PCR 扩增反应体系 (25 μ L): $2 \times Taq$ PCR Mix 12.5 μ L, 引物 (2.5 μ mol \cdot L⁻¹) 各 1.0 μ L, DNA 模板 2 μ L, 最后用 ddH₂O 补足至 25 μ L。PCR 扩增程序: 94 $^{\circ}$ C 预变性 5 min, 94 $^{\circ}$ C 变性 30 s, 57 $^{\circ}$ C 退火 45 s, 68 $^{\circ}$ C 延伸 2 min, 循环 40 次, 72 $^{\circ}$ C 延伸 7 min, 4 $^{\circ}$ C 保存。PCR 产物用 1% 琼脂糖凝胶电泳检测后由上海生工生物公司完成测序工作。

1.2.3 序列分析 测序所获得的 *ndhA* 基因内含子区域峰图经序列拼接软件 Codon Code Aligner (Codon Code Co., USA) 校对拼接, 并依据 NCBI 数据库上已测 *ndhA* 基因内含子序列对所测序列去除低质量序列并结合人工校对, 确定每一位点准确无误。利用 MEGA 5.1 软件进行多序列比对、查错并计算变异位点。利用 Kimura 两参数模型 (Kimura-2-parameter, K2P) 计算样品间两两遗传距离 (pairwise distances)、种内距离 (distances within species) 和种间距离 (distances between species)。采用邻接距离法 (Neighbor-Joining) 构建系统发育树, 选用 K2P 替代模型, 进行 1 000 次自检。遗传距离的计算和系统发育树的构建均采用 MEGA 5.0 软件。

2 结果与分析

分子生物学研究需要制备用于 PCR 反应的核酸, 前期的研究工作发现, 秋海棠属植物的核酸非常难以提取, 推测可能是该属植物富含多糖、多酚、酯类和酸类等次生代谢产物, 导致传统的 CTAB 法和试剂盒法无法提取到可用于扩增的 DNA。课题组应用改良试剂盒法提取的秋海棠属植物 DNA 可以满足于 SSR 引物开发、系统发育和 DNA 条形码等分子标记研究。DNA 条形码技术的核心为 PCR 扩增, 而延伸温度是影响 PCR 扩增的重要参数, 在 DNA 聚合酶允许的温度范围内, 选择合适的延伸温度可相对提高 DNA 聚合酶的活性, 从而促进引物和模板间的结合, 提高 PCR 扩增的效率。该试验对 *ndhA* 基因内含子序列的扩增过程中发现, 当延伸温度为 72 $^{\circ}$ C 时, *ndhA* 基因内含子的引物对于秋海棠属植物的扩增能力较差, 有些物种未能成功扩增出条带; 延伸温度改为 68 $^{\circ}$ C 时扩增时, 58 个个体的电泳条带明亮单一, 且测序成功率为 100% (图 1), 所以确定大 DNA 条形码引物 PCR 反应的最适延伸温度为 68 $^{\circ}$ C。

58 个个体中, *ndhA* 基因内含子的 PCR 扩增成功率和测序成功率 (测序后获得高质量的序列且正反向序列顺利拼接即判定为成功) 为 100%。所得的 58 条序列比对后, *ndhA* 基因内含子序列长度为 1 113~1 130 bp, 平均长度为 1 182 bp, 种间存在 101 个变异位点。种间和种内遗传距离的大小是进行物种鉴别的主要标准。结

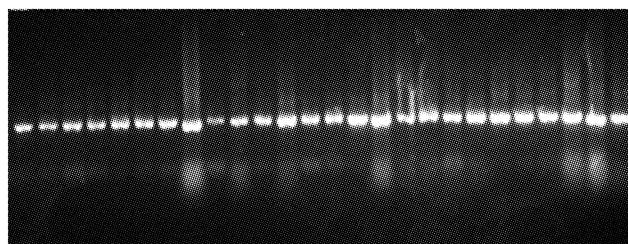


图 1 *ndhA* 基因 PCR 产物电泳结果

果显示, 在所考察的候选秋海棠属植物中具有最大的种间变异和较小的种内变异, 种间遗传距离为 0.01940 ± 0.00019 , 种内遗传距离为 0.00020 ± 0.00006 , 二者存在极显著差异。

通过对全部共 59 条 *ndhA* 基因内含子序列构建邻接 (NJ) 树 (图 2), 不同种之间能够明显分开, 18 个种类

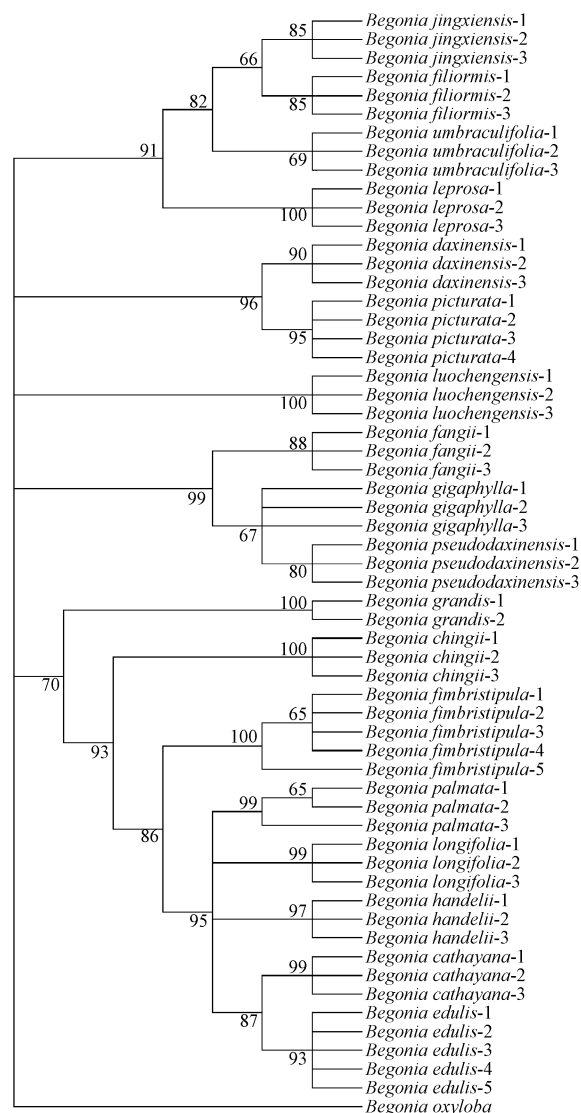


图 2 基于 *ndhA* 基因内含子序列的秋海棠属聚类分析

Fig. 2 Cluster analysis of *Begonia* based on *ndhA* sequence

的所有个体都形成单系,除了 *B. gigaphylla* 的分支支持率没有高于 50%,其它秋海棠的支持率均高于 50%,物种分辨率较高。

3 讨论

该试验采用改良的植物试剂盒方法提取秋海棠属植物 DNA,相较于常规的试剂盒法,仅在试验前期增加了一个步骤,即在叶片彻底磨成粉末状后,加入 CTAB-free+2%PVP 溶液预处理,再按常规的试剂盒提取流程提取。该方法提取的 DNA 纯度较高,浓度也较大,说明该方法适合于提取秋海棠属植物的 DNA。

生物条形码联盟(CBOL)认为理想的 DNA 条形码应该符合以下 3 个标准:1)在种内的变异应足够小,同时在种间有明显的遗传变异和分化可将物种区分开来;2)所用的片段尽量短,便于 PCR 扩增,而且一个反应就能完成测序,尤其是对存在 DNA 降解的材料(如:炮制过的药材、存放已久的蜡叶标本);3)两端连接相对保守的区域,便于设计通用引物。与动物学领域相比,在植物中的 DNA 条形码的研究进展相对缓慢,主要有 2 个方面原因:1)植物线粒体基因组进化速率较慢,遗传分化小,因此动物中的标准片段 COI 不适用于植物;2)系统学研究中常用的片段(如 *trnL-F*)变异较小,不适合用作条形码片段^[19]。由于有些植物的染色体是多倍体,核基因组通常具有多拷贝的特性,引物通用性差。因此,植物中最可能的条形码还是从叶绿体基因组中选择^[22-23]。虽然叶绿体基因相对保守,但仍然包含许多变异区域,同时叶绿体基因组有其自身的优势:单亲遗传避免了基因重组;植物个体中均有大量的叶绿体,即使 DNA 高度降解也容易扩增。生物条形码联盟建议的植物叶绿体条形码片段有 *matK*、*psbA-trnH*、*rbcL*。

在有关于秋海棠属植物的 DNA 条形码研究中,焦丽娟等^[8]的研究结果表明,3 个叶绿体片段(*rbcL*、*matK* 和 *psbA-trnH*)可能是由于进化速率慢,谱系分化不完全或是秋海棠属自身物种分化时间短等生物学原因导致种内、种间距离较小,不存在遗传差异鸿沟(barcoding gap),物种正确鉴定率较低,不适合作为秋海棠属的标准 DNA 鉴定条形码。在秋海棠属中,只有核基因 ITS 片段种内变异和种间变异比较大,存在遗传差异鸿沟,单片段的物种正确鉴定率达到 100%,可考虑作为秋海棠属 DNA 条形码鉴定的标准条码。但是秋海棠属植物含有不规则的多倍体,ITS 片段有多拷贝现象。因此,在鉴定秋海棠属植物时还需开发有效的叶绿体基因片段。

近年来,*ndhA* 基因内含子片段在一些类群中已被证实具有很好的鉴定效果。该研究中,基于 *ndhA* 基因

内含子区的遗传距离分析表明,种内距离小于种间距离,存在遗传差异鸿沟,物种分辨率较高,只有 *B. gigaphylla* 物种的分支支持率没有高于 50%(图 2)。虽然 *ndhA* 基因内含子的各种支持率与 ITS 相比稍有降低,但 *ndhA* 基因内含子和其它 3 个片段(*rbcL*、*matK* 和 *psbA-trnH*)联合分析的物种正确鉴定率可能会提高,那么 *ndhA* 基因内含子可以作为秋海棠属植物 DNA 条形码的一个有效的补充片段。

综上所述,该研究通过对秋海棠属 19 个物种 59 个个体的 *ndhA* 基因内含子序列分析,发现这些类群中 *ndhA* 基因内含子序列能够较好的区分各物种,18 个种类的所有个体都形成单系类群,对秋海棠属物种而言,是一个有效的 DNA 条形码序列。该研究为探讨以 *ndhA* 基因内含子序列作为 DNA 条形码对秋海棠属植物进行物种识别的可行性提供了一定的参考。

参考文献

- [1] HUGHES M, HOLLINGSWORTH P M. Population genetic divergence corresponds with species level biodiversity patterns in the large genus *Begonia* [J]. *Molecular Ecology*, 2008, 17(11): 2643-2651.
- [2] KU S M, LIU Y, PENG C. Four new species of *Begonia* sect. *Coelocentrum* (Begoniaceae) from limestone areas in Guangxi, China [J]. *Bot Stud* (Taipei, Taiwan), 2006, 47: 207-222.
- [3] de WILDE J J F E, van VALKENBURG J L C H. *Begonia sosefiana* (Begoniaceae): A new species in section *Loasibegonia* from Gabon [J]. *Blumea*, 2005, 50: 467-471.
- [4] THOMAS D C, HUGHES M, PHUTTHAI T, et al. A non-coding plastid DNA phylogeny of Asian *Begonia* (Begoniaceae): Evidence for morphological homoplasy and sectional polyphyly [J]. *Molecular Phylogenetics and Evolution*, 2011, 60(3): 428-444.
- [5] 张少平, 赖正锋, 吴水金, 等. 药食同源植物紫背天葵研究现状与展望 [J]. *中国农学通报*, 2014, 30(4): 58-61.
- [6] 王红珊, 曹毅敏, 李国家, 等. 紫背天葵提取物对糖尿病肾病大鼠的作用 [J]. *中国生化药物杂志*, 2012, 33(3): 272-274.
- [7] FORREST L L, HUGHES M, HOLLINGSWORTH P M. Phylogeny of *Begonia* using nuclear ribosomal sequence data and morphological characters [J]. *Systematic Botany*, 2005, 30(3): 671-682.
- [8] 焦丽娟, 税玉民. 中国秋海棠属(秋海棠科)植物的 DNA 条形码评价 [J]. *植物分类与资源学报*, 2013, 35(6): 715-724.
- [9] 张嵘梅, 陈文红, 税玉民, 等. 中国秋海棠属植物的叶表皮特征及其分类学意义 [J]. *云南植物研究*, 2008, 30(6): 665-678.
- [10] HEBERT P D N, CYWINSKA A, BALL S L, et al. Biological identifications through DNA barcodes [J]. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 2003, 270(1512): 313-321.
- [11] CHEN S L, YAO H, HAN J P, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species [J]. *PloS One*, 2010, 5(1): e8613.
- [12] 彭居刚, 王绪祯, 王丁, 等. 基于线粒体 COI 基因序列的 DNA 条形码在鲤科鲃属鱼类物种鉴定中的应用 [J]. *水生生物学报*, 2009, 33(2): 271-276.

- [13] 张辉,姚辉,崔丽娜,等. 基于 COI 条形码序列的《中国药典》动物药材鉴定研究[J]. 世界科学技术-中医药现代化,2013,15(3):371-380.
- [14] LI D Z, LIU J Q, CHEN Z D, et al. Plant DNA barcoding in China[J]. Journal of Systematics and Evolution, 2011, 49(3):165-168.
- [15] CHASE M W, SALAMIN N, WILKINSON M, et al. Land plants and DNA barcodes: Short-term and long-term goals[J]. Philos Trans R Soc Lond B Biol Sci, 2005, 360:1889-1895.
- [16] KRESS W J, WURDACK K J, ZIMMER E A, et al. Use of DNA barcodes to identify flowering plants[J]. Proc Natl Acad Sci, USA, 2005, 102: 8369-8374.
- [17] KRESS W J, ERICKSON D L. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region[J]. PLoS One, 2007, 2(6):e508.
- [18] FAZEKAS A J, BURGESS K S, KESANGAKURTI P R, et al. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well[J]. PLoS One, 2008, 3:e2802.
- [19] 宁淑萍,颜海飞,郝刚,等. 植物 DNA 条形码研究进展[J]. 生物多样性, 2008, 16(5):417-425.
- [20] THOMAS D C, HUGHES M, PHUTTHAI T, et al. A non-coding plastid DNA phylogeny of Asian *Begonia* (Begoniaceae): Evidence for morphological homoplasy and sectional polyphyly[J]. Molecular Phylogenetics and Evolution, 2011, 60:428-444.
- [21] CHUNG K F, LEONG W C, RUBITE R R, et al. Phylogenetic analyses of *Begonia* sect. *Coelocentrum* and allied limestone species of China shed light on the evolution of Sino-Vietnamese karst flora[J]. Botanical Studies, 2014, 55:1.
- [22] COWAN R S, CHASE M W, KRESS W J, et al. 300000 species to identify: Problems, progress, and prospects in DNA barcoding of land plants[J]. Taxon, 2006, 55, 611-616.
- [23] NEWMASER S G, FAZEKAS A J, RAGUPATHY S. DNA barcoding in land plants: Evaluation of *rbcL* in a multigene tiered approach[J]. Canadian Journal of Botany, 2006, 84:335-341.

Application of DNA Barcoding Based on the Chloroplast *ndhA* Intron Region in Classification of *Begonia* (Begoniaceae)

ZHAO Bo¹, LI Jingjian², MAO Shizhong¹, TANG Wenxiu¹

(1. Guangxi Institute of Botany, Chinese Academy of Sciences, Guilin, Guangxi 541006; 2. College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou, Guangdong 510642)

Abstract: Taking 58 samples from 18 species of *Begonia* as materials, chloroplast *ndhA* intron region was used as a popular marker of DNA barcoding for species identification to evaluate its species resolution rates in the genus *Begonia*. The sequences of the *ndhA* intron region were assembled using the Codon Code Aligner, and then aligned and the genetic distances were computed using MEGA 5.0 in accordance with the kimura 2-parameter (K2P) model. The neighbor-joining (NJ) phylogenetic trees were constructed. The results showed that the *ndhA* intron region was about 1 182 bp in length, and exhibited the highest inter-specific divergence, and this was significantly higher than the intra-specific variation. The *ndhA* intron region had reasonable high identification efficiency among all tested samples. The NJ trees revealed that all individuals of each species formed a strong monophyletic group. Species identification with *ndhA* intron region in cluster supported the taxonomy based on morphological characters. The conclusion suggested that chloroplast *ndhA* intron region was a valid DNA barcoding gene for species identification in genus *Begonia*.

Keywords: *Begonia*; *ndhA* intron region; DNA barcoding; species identification