

地黄属核基因与叶绿体基因 DNA 条形码比较研究

李建芳, 黄 静, 程芳婷, 杜 优, 刘占林*

西北大学生命科学学院, 西部资源生物与现代生物技术教育部重点实验室, 陕西 西安 710069

摘要: 目的 比较单拷贝核基因、叶绿体基因及 ITS 等 DNA 条形码在地黄属的物种分辨率, 探讨地黄属种间亲缘关系。方法 利用距离法与建树法分析了地黄属植物 13 个居群的 12 个单拷贝核基因、6 个叶绿体基因以及 ITS 片段物种鉴定成功率, 构建了地黄属系统发育树。结果 叶绿体 *matK* 和 *rbcL*、单拷贝核基因物种分辨率普遍较低, 多基因联合分析有助于物种鉴定成功率的提高。在 NJ 树中, 地黄-茄叶地黄聚为一支, 其他物种都是单系群, 裂叶地黄位于系统发育树基部。结论 ITS 是地黄属物种鉴定的核心条形码, *psbA-trnH*、*trnL-F*、*trnM-V*、*trnS-G* 可以作为备选方案, 核基因 *RR255* 与 *R257* 可用作地黄专用鉴定条形码。四倍体地黄起源方式有待进一步澄清。

关键词: 地黄属; DNA 条形码; 核基因; 叶绿体基因; 亲缘关系

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2017)01-0165-07

DOI: 10.7501/j.issn.0253-2670.2017.01.023

Comparison on DNA barcodes of nuclear and chloroplast gene fragments in *Rehmannia Libosch. ex Fisch. et Mey. (Rehmanniaceae)*

LI Jian-fang, HUANG Jing, CHENG Fang-ting, DU You, LIU Zhan-lin

Key Laboratory of Resource Biology and Biotechnology in Western China (Ministry of Education), College of Life Sciences, Northwest University, Xi'an 710069, China

Abstract: Objective To compare the usage of single-copy nuclear genes (SCNGs), chloroplast gene fragments and internal transcribed spacer (ITS) as DNA barcodes in plants of *Rehmannia Libosch. ex Fisch. et Mey.* and elucidate the interspecific relationships in the genus. **Methods** Distance and tree-based methods were performed to analyze 12 SCNGs, six cpDNA regions, and ITS fragment in 13 populations of *Rehmannia Libosch. ex Fisch. et Mey.* Neighbor-joining tree was constructed to investigate the relationships among the species. **Results** CpDNA regions (*matK* and *rbcL*) and SCNGs had lower identification rates than others, especially ITS. Multiple DNA barcodes combination would be helpful to improve the species identification rate. The NJ tree indicated that *R. glutinosa* and *R. solanifolia* were clustered together, other species in the genus were monophyletic, and *R. piasezkii* was the basal group. **Conclusion** ITS should be used as the core barcode in the genus while *psbA-trnH*, *trnL-F*, *trnM-V*, and *trnS-G* would be considered as the candidates, and nuclear genes *R255* and *R257* might be utilized as unique barcode for the identification of *R. glutinosa*. The origin of tetraploid species (*R. glutinosa* and *R. solanifolia*) still remains unknown and further works should be done to solve the question.

Key words: *Rehmannia Libosch. ex Fisch. et Mey.*; DNA barcoding; nuclear genes; chloroplast genes; phylogenetic relationship

物种多样性是生物多样性的的重要组成部分, 随着全球生物多样性程度的丧失, 物种多样性保护已经成为公众关注的热点问题, 也是当前生命科学领域研究的重要方向。对物种数目统计及其相互关系的研究是物种多样性保护的前提与基础, 传统依赖于形态学标准的物种鉴定往往需要专门知识, 只有少数分类学专家才能有效掌握。生物多样性的保护急需新的、标准化的、高效的物种鉴定手段。2003

年 Hebert 等^[1]提出将 DNA 条形码应用于分类学研究, 希望能够利用一个或多个 DNA 片段实现物种的快速、准确鉴定, 用来探讨区系组成、评估生物多样性水平。2009 年, CBOL Plant Working Group 提出将叶绿体 *rbcL*+*matK* 片段组合作为陆地植物核心条形码, 叶绿体基因 *psbA-trnH* 和核基因片段 ITS 作为补偿^[2]。2011 年中国植物条形码研究团队 (China Plant BOL Group) 建议将 ITS 纳入核心条形

收稿日期: 2016-08-13

基金项目: 国家自然科学基金项目 (31670219, 31370353); 陕西省自然科学基金研究计划 (2015JM3106); 陕西省教育厅项目 (11JS093)

作者简介: 李建芳 (1975—), 女, 助理工程师, 研究方向为生物资源遗传分析。E-mail: jianfangli2010@qq.com

*通信作者 刘占林, 男, 教授, 从事进化植物学研究。E-mail: liuzl@nwu.edu.cn

码当中^[3]。由于叶绿体 *rbcL* 和 *matK* 片段进化速率较慢, 适合于科属及以上水平类群的分析, 而对于经历快速辐射或最近分化的类群, 这 2 个条形码标记就往往缺乏有效性^[4]。ITS 现已广泛应用于系统发育分析中, 但是鉴于 ITS 是多拷贝基因, 在许多类群中存在致同进化不完全现象以及 PCR 扩增的偏好性, 容易导致系统发育关系的不确定性, 为物种鉴定带来很大的难度^[5]。

与多拷贝核基因相比, 单拷贝核基因优势在于可直接区分直系同源 (orthologous) 和旁系同源 (paralogous), 易建立物种间直系同源的关系, 其兼具核基因的双亲遗传以及序列保守性的特点, 进化速率比质体基因组序列更快而且能够选取的序列片段更加丰富。单拷贝核基因条形码开发所面临的最大困难是大多数植物缺乏基因组信息, DNA 条形码的通用性较差。随着新一代测序 (next generation sequencing) 技术的飞速发展, 大规模转录组测序变得简便, 核基因数据呈几何式增长, 使开发不同类群通用的单拷贝核基因 DNA 条形码成为可能。Duarte 等^[6]发现 4 个属植物 (拟南芥属 *Arabidopsis* Heynh.、杨属 *Populus* Linn.、稻属 *Oryza* Linn. 和葡萄属 *Vitis* Linn. APOV) 之间共享的 959 个单拷贝核基因, Zhang 等^[7]进一步开发出 1 083 个高度保守的低拷贝基因。这些单拷贝基因已经应用于多种植物类群系统关系研究中, 如杨柳科^[8]、山龙眼科^[9]等, 并为单拷贝核基因 DNA 条形码的筛选提供了有效工具。

地黄属 *Rehmannia* Libosch. ex Fischer & Mey. 是隶属于玄参科 (Scrophulariaceae), 亦有研究认为应该独立成地黄科 (Rehmanniaceae)^[10]。《中国植物志》记载地黄属有 6 个种, 包括中国著名的传统中药地黄 *R. glutinosa* Libosch. ex Fisch. & C. A. Mey.、天目地黄 *R. chingii* Li、裂叶地黄 *R. piasezkii* Maxim.、高地黄 *R. elata* N. E. Br. ex Prain、湖北地黄 *R. henryi* N. E. Br. 和茄叶地黄 *R. solanifolia* Tsoong et Chin, 其中地黄和茄叶地黄是四倍体^[11]。地黄属植物的研究主要集中在传统药用植物地黄的药理作用、物质基础、繁育体系、遗传变异等方面^[11], 而缺乏属内其他物种的研究, 加上该属植物分布范围广, 外观形态差异较大, 经常造成分类学上的争议。比如, 最近的形态、化学成分以及分子证据显示高地黄和裂叶地黄实为同一物种^[11-12], 茄叶地黄与地黄在遗传上高度一致, 也应该予以归并^[11]。当前已有研究探讨 DNA 条形码在地黄属的

适用性。程芳婷等^[13]利用 *trnL-F*、*trnMV*、*trnS-G* 3 个叶绿体片段与核基因 ITS 对地黄属物种进行分析后, 推荐 ITS+*trnS-G* 作为地黄属的标准条形码。该研究没有说明 *psbA-trnH*、*matK*、*rbcL* 等国际标准候选条形码在地黄属应用的可行性。最近, 夏至等^[14]补充了该标准条形码的数据, 认为 *psbA-trnH* 也可以作为地黄属的核心条形码。遗憾的是, 除地黄外这一研究没有开展其他物种种内遗传变异分析, 无法确定这些标记的种内最大变异是否显著小于种间最小变异范围, 其物种鉴定的可靠性受到质疑。此外, 采用单一个体代表物种所进行系统发育分析, 往往会对种间亲缘关系的确定构成负面影响^[14]。本研究目的在于通过大量的群体数据, 对比多个叶绿体与核基因 DNA 标记的差异性, 筛选出适合地黄属的 DNA 条形码, 确立地黄属各物种的亲缘关系, 并探讨单拷贝核基因作为物种鉴定 DNA 条形码的可行性。

1 材料与方法

1.1 材料

由于高地黄与裂叶地黄的遗传组成相一致, 本研究将二者作为一个物种处理^[11]。实验选取了地黄属 5 个物种 13 个居群共 65 份个体作为研究对象^[13] (表 1)。所用实验材料均为野外采集的叶片, 经硅胶快速干燥, 常温保存备用, 标本保存于西北大学生命科学学院。

利用 CTAB 法从叶片中提取基因组 DNA, 叶绿体基因片段引物参照文献方法^[13]设计, 单拷贝核基因由地黄转录组开发获得^[11], PCR 扩增程序按照相关文献方法^[11,13]操作, PCR 产物由上海生工生物工程技术服务有限公司完成测序工作。根据先前的研究成果, 本研究总共对 12 个单拷贝核基因 (*R10*、*R15*、*R56*、*R84*、*R118*、*R124*、*R146*、*R163*、*R254*、*R255*、*R256*、*R257*)、1 个多拷贝核基因 (ITS) 和 6 个叶绿体基因 (*matK*、*psbA-trnH*、*rbcL*、*trnL-F*、*trnM-V*、*trnS-G*) 进行对比分析。

1.2 数据处理

原始序列采用 BioEdit ver. 7.0.9.1 软件^[15]编辑并进行人工校对。利用 MEGA ver. 6.0 软件^[16]统计每个基因片段的长度、插入/缺失 (indels)、变异位点数、简约信息位点数等序列特征; 分析各 DNA 片段种内、种间的 Kimura-2-parameter (K2P) 遗传距离, 同时以崖白菜 *Triaenophora rupestris* (Hemsl.) Solereder 为外类群构建单一片段以及片段组合的邻接 (neighbor-joining, NJ) 树。采用建树法和距离法这 2 种不同的方法评估 DNA 条形码的物种鉴定效果。建

表 1 材料来源

Table 1 Information of materials

物种	产地	居群编号	个体数	E 经度/°	N 纬度/°	海拔/m
地黄 <i>R. glutinosa</i>	山西大同	DT1~5	5	113.0	40.1	1 098
	甘肃崇信	CC1~5	5	107.0	35.2	1 215
	山西长治	CZ1~5	5	113.0	36.1	929
	四川隘口	AK1~5	5	107.5	31.5	787
	山东泰安	TA1~5	5	117.1	36.2	203
茄叶地黄 <i>R. solanifolia</i>	重庆城口	CK1~5	5	108.1	31.7	950
天目地黄 <i>R. chingii</i>	安徽绩溪	AH1~5	5	118.5	30.1	222
	安徽棠溪	TX1~5	5	117.6	30.4	86
	浙江丽水	LS1~5	5	120.0	28.4	195
湖北地黄 <i>R. henryi</i>	湖北宜昌	YC1~5	5	110.6	31.3	299
裂叶地黄 <i>R. piasezkii</i>	陕西后柳	HL1~5	5	108.2	32.9	383
	湖北保康	BK1~5	5	111.5	31.9	251
	湖北竹山	ZS1~5	5	110.2	32.1	495

树法认为只有同一个物种的不同个体在构建的 NJ 系统树上形成单系分支才认为物种鉴定成功；距离法采用 PWG 距离 (PWG-distance) 法，认为种间最小的遗传距离大于种内最大的遗传距离，则物种鉴定成功^[3]；此外，还应用 TaxonDNAver.1.7.8 软件^[17-18]，通过最佳匹配 (Best match) 计算每个个体间的相似性，判断每个个体所属物种的准确率。只有当最佳匹配值大于 0.4 时，该标记才会用于条形码组合分析。由于本研究侧重于分析叶绿体基因与单拷贝核基因在地黄属物种鉴定成功率的差异性，因此仅以 ITS 为参照，构建了叶绿体基因与单拷贝核基因的独立数据矩阵，并没有将二者进行联合数据集分析。考虑到前期研究均表明茄叶地黄往往包含在地黄系统分支内，把地黄-茄叶地黄分支组成单系群确定为地黄鉴定成功的标准。

2 结果

2.1 序列特征

本研究所采用的 19 个基因片段在地黄属 5 个种 13 个居群 65 个个体均能够进行成功扩增与测序。除 R84 外，其余 11 个单拷贝核基因片段序列均仅出现位点的碱基替代，没有插入/缺失现象。变异位点比率与简约信息位点比率变化范围均在 0.77%~4.33%，平均值分别为 2.44% 和 2.34%，R146 变异最低，而 R84 最高 (表 2)。裂叶地黄 (除 R15、R56 基因外) 与天目地黄 (除 R56 基因外) 在绝大多数核基因中都表现出较高水平的种内居群间变异，湖北地黄只在 R118 和 R257 中存在变异，而地黄与茄叶地黄序列完全相同，在所有基因中均呈现单态现象。多拷贝核基因 ITS 则体现出种内较高的多态性，每个物种种内均

表 2 核基因片段遗传特征以及物种鉴定成功率

Table 2 Characteristics of each nuclear or chloroplast fragment and their successful identification rates for species in *Rehmannia Libosch. ex Fisch. et Mey.*

标记	序列长度/bp	排列后序列长度/bp	GC 量/%	变异位点/%	信息位点数/%	插入缺失数目/bp	建树方法物种鉴定率	PWG 距离	Best match 法	种内距离(平均值)	种间距离(平均值)
R10	320	320	49.0	7 (2.19)	7 (2.19)	0	0.25 (y)	0.50	0.10	0~0.001 4 (0.000 7)	0.000 8~0.013 5 (0.007 1)
R15	399	399	41.6	11 (2.76)	11 (2.76)	0	0.50 (y)	0.75	0.31	0~0.006 4 (0.001 6)	0.005 0~0.012 7 (0.009 4)
R56	254	254	42.8	4 (1.58)	4 (1.58)	0	0.25 (y)	0.00	0.00	0	0~0.015 9 (0.008 0)
R84	279~299	300	39.0	13 (4.33)	13 (4.33)	3 (3~12)	0.00 (n)	0.00	0.41	0~0.014 9 (0.004 5)	0~0.024 6 (0.012 2)
R118	350	350	37.9	10 (2.86)	9 (2.57)	0	0.25 (y)	0.25	0.31	0~0.001 8 (0.000 6)	0.001 2~0.021 1 (0.013 8)
R124	269	269	45.3	3 (1.12)	3 (1.12)	0	0.25 (n)	0.00	0.00	0~0.001 8 (0.000 7)	0~0.005 5 (0.002 7)
R146	260	260	43.1	2 (0.77)	2 (0.77)	0	0.00 (n)	0.00	0.00	0~0.001 8 (0.000 9)	0~0.003 9 (0.001 9)
R163	254	254	44.5	7 (2.76)	7 (2.76)	0	0.25 (n)	0.50	0.00	0~0.003 2 (0.001 2)	0.001 1~0.015 2 (0.008 1)
R254	429	429	42.6	15 (3.50)	14 (3.26)	0	0.50 (y)	0.50	0.41	0~0.008 4 (0.000 9)	0.006 3~0.019 8 (0.012 8)
R255	390	390	42.2	10 (2.56)	9 (2.31)	0	0.51 (y)	1.00	0.57	0~0.006 2 (0.001 8)	0.007 8~0.011 2 (0.009 9)
R256	330	330	41.4	8 (2.42)	7 (2.12)	0	0.25 (y)	1.00	0.23	0~0.004 4 (0.001 8)	0.006 1~0.013 5 (0.008 8)
R257	574	574	39.0	10 (1.74)	10 (1.74)	0	0.75 (y)	1.00	0.57	0~0.002 3 (0.000 8)	0.002 5~0.011 2 (0.007 6)
ITS	680~682	685	61.0	24 (3.50)	21 (3.07)	5 (1)	1.00 (y)	1.00	0.98	0~0.002 1 (0.001 3)	0.003 0~0.016 5 (0.013 3)

建树法物种鉴定成功率，括号内字母 y 表示是，n 表示否对地黄-茄叶地黄分支进行准确鉴定，下同

successful identification by tree-based method, letters in parenthesis indicate whether the barcode is suitable for *R. solanifolia*-*R. glutinosa*, same as below

存在变异, 变异位点、简约信息位点比率分别为 3.50% 和 3.07%。在 6 个叶绿体基因片段中, 除 *matK* 和 *rbcL*, 其他基因均存在碱基的插入/缺失, 其中 *psbA-trnH* 与 *trnS-G* 插入/缺失十分频繁, 各有 10 个和 8 个, 最长缺失分别长达 76 bp 和 30 bp (表 2)。叶绿体基因的变异位点比率、简约信息位点比率的变化幅度分别为 0.15%~4.02% 和 0.15%~3.62%, 平均值为 1.67% 和 1.39%, 稍低于单拷贝核基因。*rbcL* 变异率最低, 仅在天目地黄种内有 1 个变异位点, 其他物种均共享一个单倍型; *psbA-trnH* 基因有 20 个变异位点, 序列变异比率最高; 其次是 *trnS-G*, 而 *trnL-F* 与 *trnM-V* 在地黄属内变异水平相当 (表 3)。在所有地黄属物种中, 多数叶绿体基因片段表现出居群个体间的多态性。但也有例外, 比如 *rbcL* 个体间同源性近乎 100%, 湖北地黄也表现出种内低的变异水平, 仅在 *psbA-trnH* 基因上存在个体间遗传变异。

2.2 地黄属 DNA 条形码鉴定

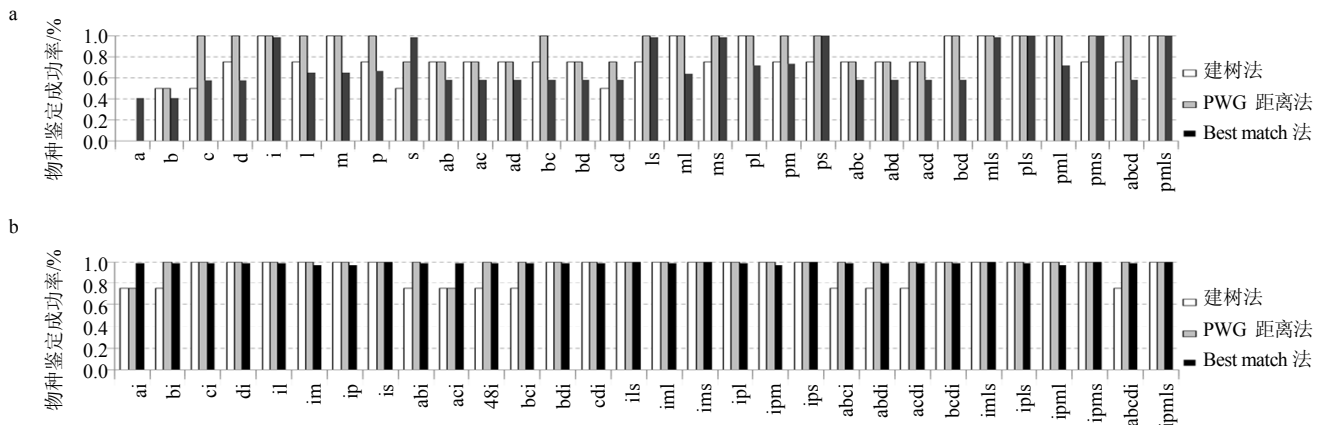
利用距离法 (PWG 距离法和 Best match 法) 和

建树法比较不同 DNA 条形码对地黄属物种的鉴定成功率。ITS 作为推荐的核心条形码在 3 种方法中都具有最高的分辨率。PWG 距离法分析表明, 多数单拷贝核基因片段存在种内最大遗传距离大于种间最小遗传距离, 单拷贝核基因使用有一定的限制性。只有 R255、R256 和 R257 能够与 ITS 媲美, 符合种内最大遗传距离小于种间最小遗传距离的 DNA 条形码的标准 (表 2)。而依据 Best match 法, 物种分辨率最高的 R255、R257 在地黄属也只有 57% 的鉴定成功率。建树法结果则表明单拷贝核基因不能单独地 100% 完成物种鉴定, 分辨率最高的 R257 也只能成功分辨 75% 的物种。在本研究涉及的 6 个叶绿体基因片段中, PWG 距离法、Best match 法和建树法都显示 *rbcL* 没有分辨率, *matK* 的物种分辨率也较低。虽然在不同条形码分析方法中得到的物种分辨率略有差异, 但是 *psbA-trnH*、*trnL-F*、*trnM-V* 与 *trnS-G* 都显示较高水平的物种鉴定成功率 (50%~100%), 见表 3 和图 1-a。

表 3 叶绿体基因片段遗传特征以及物种鉴定成功率

Table 3 Characteristics of chloroplast fragment and their successful identification rates for species in *Rehmannia* Libosch. ex Fisch. et Mey.

标记	序列长度 /bp	排列后序列长度/bp	GC 量/%	变异位点/%	信息位点 数/%	插入缺失数 目/bp	建树方法物种 鉴定率/%	PWG 距离 法	Best match 法	种内距离 (平均值)	种间距离 (平均值)
<i>matK</i>	814	814	34.1	7 (0.86)	5 (0.61)	0	0.25 (n)	0.75	0.25	0~0.000 6 (0.000 5)	0.001 0~0.002 2 (0.001 8)
<i>psbA-trnH</i>	388~479	497	26.3	20 (4.02)	18 (3.62)	10 (1~76)	0.75 (y)	1.00	0.65	0~0.003 4 (0.002 3)	0.006 6~0.022 1 (0.014 5)
<i>rbcL</i>	681	681	43.3	1 (0.15)	1 (0.15)	0	0.00 (n)	0.00	0.17	0~0.000 7 (0.000 2)	0.000 0~0.001 0 (0.000 5)
<i>trnL-F</i>	794~801	802	35.0	9 (1.12)	8 (1.00)	2 (1~7)	1.00 (y)	1.00	0.65	0~0.000 6 (0.000 2)	0.001 6~0.006 8 (0.004 1)
<i>trnM-V</i>	747~755	756	36.5	12 (1.59)	8 (1.06)	4 (1~5)	0.75 (y)	1.00	0.67	0~0.000 6 (0.000 5)	0.002 7~0.005 8 (0.004 8)
<i>trnS-G</i>	645~686	738	27.7	17 (2.30)	14 (1.90)	8 (1~30)	0.50 (n)	0.75	0.98	0~0.002 0 (0.001 4)	0.002 6~0.007 6 (0.005 6)



a=R84 b=R254 c=R255 d=R257 l=trnL-F m=trnM-V s=trnS-G p=psbA-trnH i=ITS

图 1 不包含 ITS 的条形码组合 (a) 和包括 ITS 的条形码组合 (b) 的物种鉴定成功率

Fig. 1 Successful identification rates of barcode combinations without ITS fragment (a) and barcode combinations containing ITS (b)

通过 Best match 法分辨率大于 0.4 的标准进一步筛选标记, 单拷贝核基因与叶绿体基因各自获得 4 个片段用于构建基因片段组合数据集, 分别是核基因的 R84、R254、R255、R257 和叶绿体基因的 psbA-trnH、trnL-F、trnM-V 与 trnS-G (表 1)。DNA 条形码分析表明, 随着基因片段组合数目的增加, 叶绿体基因物种分辨率快速增加, 按照 PWG 距离法 2 个叶绿体片段的组合就能达到 100% 的分辨率, 而按照建树法则需要 3 个叶绿体片段组合。与之相对, 2、3 与 4 个单拷贝核基因组合数据得到的物种分辨率没有显著区别, 只有 R255+R257+R84 组合能够依照 PWG 距离法和建树法 100% 将地黄属各物种区分开 (图 1-a)。当把 ITS 作为必需的 DNA 条形码组合时, 几乎所有的基因片段组合方式都有近乎 100% 的分辨率 (图 1-b), 这进一步证实了 ITS 作为核心条形码的必要性^[19]。

2.3 地黄属的系统发育关系

以崖白菜为外类群, 利用 12 个单拷贝核基因、ITS 和 6 个叶绿体基因片段构建了地黄属 5 个物种的 NJ 树。结果表明, 裂叶地黄、天目地黄、湖北地黄均为单系群, 地黄与茄叶地黄为并系群, 每个分支都有 98% 以上的支持率。湖北地黄与天目地黄是姊妹类群, 裂叶地黄处于系统树的基部 (图 2)。

3 讨论

3.1 DNA 条形码的比较

有研究认为, 叶绿体基因进化速率慢, 使用叶绿体基因片段对植物类群鉴定存在上限, 即增加片段组合也不能提高分辨率^[19]。研究结果与此稍有不同, 基因数目的增加有助于物种鉴定成功率的提高。以建树法为例, 2 个基因片段组合的物种分辨率稍高于单个基因 (组合前后平均值分别为 75% 与 83%), 但多于 2 个基因组合后, 物种鉴定成功率快速增加 (3、4 个基因组合的分辨率分别为 93.7%、100%, 图 1)。究其原因, 可能与所选叶绿体基因具体的进化规律有关。生命条码联盟植物工作组推荐的 DNA 条形码, 比如 matk 和 rbcL 进化速率较慢, 适合科属以上水平的系统学分析, 对属下水平的物种鉴定提供的信息非常有限^[20]。而本研究分析所用的 psbA-trnH、trnL-F、trnM-V、trnS-G 片段全部属于非编码片段, 进化速率快于基因编码区。特别是 trnS-G 基因片段, 序列中含有类似小卫星重复区域, 插入/缺失反复出现, 拥有这一特征的基因其进化速率可能比普通叶绿体基因进化速率高几个数量级^[21]。

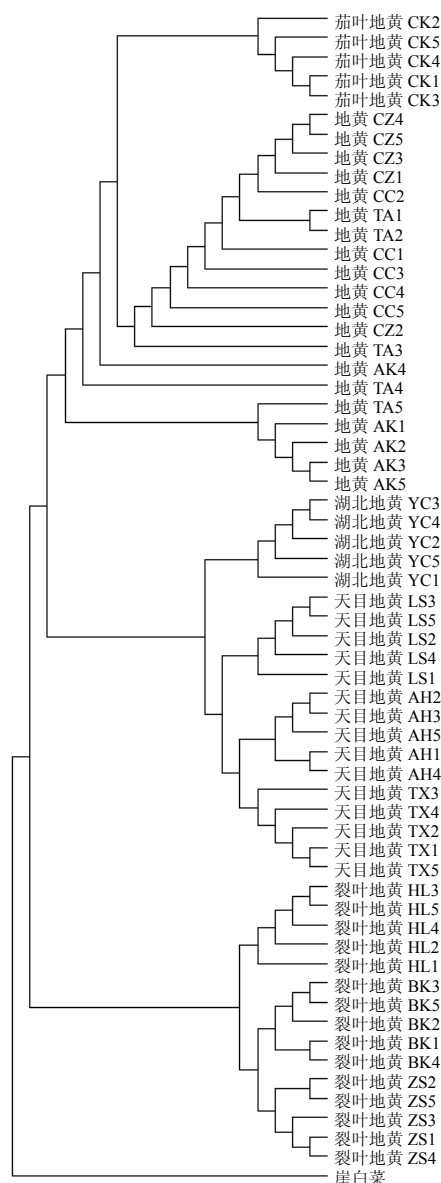


图 2 基于 6 个叶绿体基因与 13 个核基因构建的地黄属植物 NJ 树

Fig. 2 NJ tree constructed with six chloroplast genes and 13 nuclear gene fragments

与叶绿体基因结果不同, 单拷贝核基因片段数目增加并不能显著提高地黄属物种的鉴定成功率, 这主要有 2 个原因: 一是本研究所选单拷贝核基因是根据 APOV 等模式植物筛选的, 基因在类群间高度保守, 并且绝大部分区段都是基因编码区, 进化速率较慢^[22]; 另一个原因是不同基因在地黄属种内种间的变异程度不同, 进化规律的冲突也会导致片段组合后也会降低物种的鉴定效果。

本研究中, ITS 在 3 种分析方法中均表现出很高的物种分辨率 (98%~100%), 证明 ITS 确实应该作为地黄属的 DNA 核心条形码, 这与先前的研

究相一致^[3-4,13]。4 个叶绿体基因片段 (psbA-trnH、trnL-F、trnM-V、trnS-G) 在地黄属表现的分辨率相近, 种间最小距离都大于种内最大距离 (表 2), ITS 与其中任何一个序列组合均能将地黄属物种 100% 鉴定成功, 因此它们均可以作为备选 DNA 条形码。同时建议单拷贝核基因 R255 与 R257 也可以考虑作为备选项。虽然这些核基因在地黄属物种鉴定中表现逊于叶绿体备选条形码, 但依据 PWG 距离法仍能 100% 鉴定地黄属 5 个物种, 并且这 2 个基因能各自独立地把地黄与其他二倍体物种区分开来 (表 2), 这为中药地黄的分子鉴定提供了更多有效手段。随着基因组数据的增长与开发, 单拷贝核基因将会提供更多合适的单拷贝核基因 DNA 条形码。

3.2 地黄属种间亲缘关系

通过 19 个基因片段构建的地黄属 NJ 树强烈支持地黄属二倍体物种都是单系群 (图 2), 支持率茄叶地黄与地黄关系密切, 这 2 个四倍体, 在遗传上显著区别于本属二倍体物种, 这与前人研究相一致^[4,11,13,23]。茄叶地黄处于地黄分布区西南缘, 前者可能只是地黄地理扩张过程中新形成的生态型。先前叶绿体证据, 支持湖北地黄与天目地黄是并系群^[24]。本研究综合了大量的核基因与叶绿体基因数据, 不支持上述结论, NJ 树显示湖北地黄与天目地黄是姊妹类群, 二者组成单系群, 可能拥有最近共同祖先。以前研究均认为天目地黄是本属基本类群^[13,23-24], 但本研究认为裂叶地黄可能更近于祖先状态。裂叶地黄现今主要分布于大巴山东缘, 其分布范围与地黄属模拟得到的祖先分布区相一致^[14], 裂叶地黄可能是地黄属祖先类群原地分化的产物, 而天目地黄可能是祖先物种向东扩张过程中与祖先形成地理隔离产生的衍生种, 其具体机制有待深入研究。关于四倍体地黄-茄叶地黄的起源方式一直受到各项研究的关注, 前人研究推测地黄-茄叶地黄可能源于天目地黄与裂叶地黄的杂交再经过基因组加倍而形成的异源多倍体^[14,23]。本研究的 NJ 树显示地黄-茄叶地黄分支与天目地黄-湖北地黄祖先是姊妹类群。这一系统发育关系既可以解释为裂叶地黄与天目地黄-湖北地黄的祖先物种分化后再杂交形成四倍体的地黄-茄叶地黄, 也可以解释为地黄-茄叶地黄是裂叶地黄一部分群体直接加倍后而形成的同源四倍体。因此关于四倍体的起源方式仍需要更多的研究工作进行深入探讨。

4 结论

通过比较 12 个单拷贝核基因、ITS 与 6 个叶绿体基因片段作为 DNA 条形码的可行性, 确定了 ITS 作为核心条形码的地位, 同时推选叶绿体片段 psbA-trnH、trnL-F、trnM-V、trnS-G 作为备选项, 核基因 RR255 与 R257 可作为地黄专用鉴定条形码。这些 DNA 标记为探讨地黄属物种间的亲缘关系提供了新的资料。

参考文献

- [1] Hebert P D H, Cywinska A, Ball S L, *et al.* Biological identifications through DNA barcodes [J]. *Proc Roy Soc Lond, Ser B: Biol Sci*, 2003, 270(1512): 313-321.
- [2] Group CPW. A DNA barcode for land plants [J]. *Proc Nat Acad Sci USA*, 2009, 106: 12794-12797.
- [3] Li D Z, Gao L M, Li H T, *et al.* Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants [J]. *Proc Nat Acad Sci USA*, 2011, 108(49): 19641-19646.
- [4] 夏至, 王璐静, 黄勇, 等. 地黄属植物 DNA 条形码鉴定及地黄栽培起源研究 [J]. *中草药*, 2016, 47(4): 648-654.
- [5] 胡兆锋, 向秋云, 黄宏文. 四照花亚属的 nrDNA ITS 致同进化不完全 [J]. *热带亚热带植物学报*, 2013, 21(1): 29-37.
- [6] Duarte J M, Wall P K, Edger P P, *et al.* Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels [J]. *BMC Evol Biol*, 2010, 10: 61.
- [7] Zhang N, Zeng L, Shan H, *et al.* Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms [J]. *New Phytol*, 2012, 195(4): 923-937.
- [8] Du S, Wang Z, Zhang J. A novel set of single-copy nuclear DNA markers for the genetic study of Salicaceae [J]. *Gen Mol Res*, 2014, 13(3): 4911-4917.
- [9] Tonnabel J, Olivieri I, Mignot A, *et al.* Developing nuclear DNA phylogenetic markers in the angiosperm genus *Leucadendron* (Proteaceae): a next-generation sequencing transcriptomic approach [J]. *Mol Phylogenet Evol*, 2014, 70: 37-46.
- [10] Reveal J L. Summary of recent systems of angiosperm classification [J]. *Kew Bulletin*, 2011, 66(1): 5-48.
- [11] Huang J, Zeng S Y, Zhao J H, *et al.* Genetic variation and phylogenetic relationships among *Rehmannia* (Scrophulariaceae) species as revealed by a novel set of

- single-copy nuclear gene markers [J]. *Biochem Syst Ecol*, 2016, 66: 43-49.
- [12] 李宏庆, 刘国丽. 地黄属的一个新异名 [J]. 广西植物, 2012, 32(1): 23-26.
- [13] 程芳婷, 李忠虎, 刘春艳, 等. 地黄属植物的 DNA 条形码研究 [J]. 植物科学学报, 2015, 33(1): 25-32.
- [14] 刘占林, 李建芳. 地黄属分子系统学分析 [J]. 西北植物学报, 2014, 34(1): 77-82.
- [15] Hall T A. BioEdit: a user-friendly biological sequence alignment editor program for Windows 95/98/NT [J]. *Nuclear Acids Symposium Series*, 1999, 41(1): 95-98.
- [16] Tamura K, Stecher G, Peterson D, *et al.* MEGA6: Molecular evolutionary genetics analysis version 6.0 [J]. *Mol Biol Evol*, 2013, 30(12): 2725-2729.
- [17] Meier R, Shiyang K, Vaidya G, *et al.* DNA barcoding and taxonomy in *Diptera*: a tale of high intraspecific variability and low identification success [J]. *Syst Biol*, 2006, 55(5): 715-728.
- [18] Vaidya G, Lohman D J, Meier R. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information [J]. *Cladistics*, 2011, 27(2): 171-180.
- [19] Lahaye R, van der Bank M, Bogarin D, *et al.* DNA barcoding the floras of biodiversity hotspots [J]. *Proc Nat Acad Sci USA*, 2008, 105(8): 2923-2928.
- [20] Dong W, Liu J, Yu J, *et al.* Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding [J]. *PLoS One*, 2012, 7(4): e35071.
- [21] Cozzolino S, Noce M E, Musacchio A, *et al.* Variation at a chloroplast minisatellite locus reveals the signature of habitat fragmentation and genetic bottlenecks in the rare orchid *Anacamptis palustris* (Orchidaceae) [J]. *Am J Bot*, 2003, 90(12): 1681-1687.
- [22] Zeng L, Zhang Q, Sun R, *et al.* Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times [J]. *Nat Commun*, 2014, 5: 4956-4965.
- [23] Albach D C, Li H Q, Zhao N, *et al.* Molecular systematics and phytochemistry of *Rehmannia* (Scrophulariaceae) [J]. *Biochem Syst Ecol*, 2007, 35(5): 293-300.
- [24] Xia Z, Wang Y Z, Smith J F. Familial placement and relations of *Rehmannia* and *Triaenophora* (Scrophulariaceae) inferred from five gene regions [J]. *Am J Bot*, 2009, 96(2): 519-530.