

· 研究报告 ·

DNA条形码在热带龙脑香科树种鉴定中的应用

胡建霖^{1,2}, 刘志芳^{1,2}, 慈秀芹¹, 李捷^{1*}

¹中国科学院西双版纳热带植物园综合保护中心, 植物系统发育与保护生物学实验室, 昆明 650223

²中国科学院大学, 北京 100049

摘要 龙脑香科植物是东南亚地区重要的热带木材来源树种, 对其开展DNA条形码评估在林业监管及森林资源保护等方面具有非常重要的实际应用价值。通过对龙脑香科植物样品进行*rbcl*、*matK*、*trnL-trnF*和*ITS2*四个片段的扩增和测序, 结合GenBank下载的数据, 共获得龙脑香科树种14属244种共计899条序列。通过比较4个片段的通用性、序列特征、种内和种间的遗传变异, 基于Best Match (BM)、Best Close Match (BCM)、相似性搜索算法(BLAST)和邻接树(NJ) 4种方法评估DNA条形码对于龙脑香科树种的鉴定能力。结果表明, *ITS2*在龙脑香科树种中鉴定效率最高, 通过优化的扩增体系能够从该科植物叶片中获得较高质量的*ITS2*片段; 叶绿体*matK*片段扩增和测序效率为100%, 且种内及种间遗传变异明显, 鉴定成功率高于其它叶绿体片段, 并据此提出*ITS2*和*matK*适合作为龙脑香科树种的DNA条形码。

关键词 龙脑香科, DNA条形码, *ITS2*, *matK*, 热带地区

胡建霖, 刘志芳, 慈秀芹, 李捷 (2019). DNA条形码在热带龙脑香科树种鉴定中的应用. 植物学报 54, 350–359.

自林奈时期, 人们就利用生物的形态学特征对动、植物进行分类与鉴别, 但是分类学发展了200多年, 地球上仅有约20%的物种被正式鉴定命名(Wilson, 2016)。面对传统形态学鉴定的局限性和不断缩减的分类学研究队伍, 分类学的发展面临巨大的挑战(Hebert et al., 2003), 尤其是在生物多样性极为丰富的热带东南亚地区, 如何对这些区域的热带树种进行分类鉴定是植物分类学和生态学工作者面临的一个重大挑战。

龙脑香科(Dipterocarpaceae)是典型的泛热带分布科, 作为东南亚低地湿性雨林的优势科, 是热带雨林特征性树种(李锡文等, 2002)。龙脑香科被划分为龙脑香亚科(Dipterocarpaceae)、非洲香亚科(Monotoideae)及美洲香亚科(Pakaraimoideae) 3个亚科, 共有17属约529种(Ashton, 1982)。中国有龙脑香科植物5属13种, 均局域分布在西藏东南部、云南西南至东南部、广西西南部及海南, 目前已有10种被列入国家重点保护野生植物名录, 占中国分布总数的76.9%, 其中一级保护种类1种, 二级保护种类6种,

三级保护种类3种(童绍全和陶国达, 1990; 于永福, 1999)。龙脑香科植物具有速生、丰产和材质优良等特点, 是东南亚国家传统商品木材的重要来源(罗良才, 1988)。在东南亚地区, 龙脑香科木材产量约为其它各科之总和(杨家驹等, 1989)。据统计, 云南进口龙脑香科树种的木材有7属, 25–30种, 且进口比例呈逐年上升的趋势(罗良才, 2008)。云南婆罗双(*Shorea assamica* Dyer)树干挺直, 分枝高, 径级大, 出材量高, 其木材为散孔材, 边材黄白色, 心材黄褐色, 比重及硬度中等, 木材纹理直, 结构中至略粗, 加工性能良好, 可供旋切、工业、建筑及装修等多种用途(杨家驹等, 1991; 周亮等, 2013)。东京龙脑香(*Dipterocarpus retusus* Bl.), 其商品材名称为Holong (印度、缅甸)、Hullung (印度)、Kanyin (缅甸)、Keruing gunong (马来西亚), 主要生长于潮湿的沟谷雨林及石灰山密林中, 树形高大, 木材暗灰色, 材质略硬重, 是稀有珍贵的大径级用材树种, 被列为国家I级重点保护物种(成俊卿, 1980; 童绍全和陶国达, 1990)。冰片是龙脑香科植物中树脂凝结形成的一种近于白色

收稿日期: 2018-09-27; 接受日期: 2019-02-11

基金项目: 中国科学院战略生物资源服务网络计划生物多样性保护策略(No.ZSSD-013)和中国科学院东南亚生物多样性中心资助(No.Y4ZK111B01)

* 通讯作者。E-mail: jieli@xtbg.ac.cn

的结晶体,是一种名贵香料,古代称之为龙脑,以示其珍贵。另外,龙脑香科植物的树脂和木油是化工和医药原料(王锦亮等, 1992; 戴文君等, 2017)。由于历史上龙脑香科的有关资料缺乏,加之该科植物又都是热带雨林的上层乔木树种,分类和鉴定存在较大困难(朱华和王洪, 1992)。龙脑香科树种多为高大乔木,难以获取花和果实等具明显特征的植物器官用于形态鉴定,因此对其进行分类鉴定往往依赖于经验丰富的分类专家。在实际工作中,针对龙脑香科树种、幼苗、植物碎片及木材产品等的准确鉴定尤为困难。

DNA条形码技术首先由Paul Hebert提出,是利用1个或少数几个DNA片段对地球上所有物种进行识别和鉴定(Hebert et al., 2003; Kress et al., 2005)。自DNA条形码技术的概念被提出以来,研究者制订了DNA条形码的选择标准:(1) 尽量短的片段,片段两端连接相对保守的区域,易于开发通用引物;(2) 要有足够的变异可将物种区分开来(CBOL Plant Working Group et al., 2009; 任保青和陈之端, 2010)。利用以上标准,研究者陆续提出将*matK*、*rbcL*、*trnH-psbA*和*ITS/ITS2*作为陆地植物的候选DNA条形码(Kress et al., 2005; China Plant BOL Group et al., 2011)。近些年, DNA条形码技术的发展十分迅速,已经应用于一些热带珍贵植物树种的鉴定。例如, Hassold等(2016)对马达加斯加的红木类树种(*red wood*)进行DNA条形码研究,评估3个DNA条形码片段(*matK*、*rbcL*和*trnL*)对于不同产地及不同种类红木的鉴定效率; Jiao等(2018)利用6种紫檀属(*Pterocarpus* Jacq.)木材标本和干燥叶片样品材料,评估*ITS2*、*matK*、*ndhF-rpl32*和*rbcL*四个DNA条形码鉴定效果,并构建6种紫檀属DNA条形码数据库。在龙脑香科树种中也开展了不少相关研究。例如, Tsumura等(2011)通过比较*trnL*、*trnL-trnF*、*trnH-trnK*和*psbC-trnS*片段构建了84种娑罗双属(*Shorea* Roxb.)植物分子数据库; Trang等(2015)评估*rbcL*、*matK*和*trnH-psbA*片段在东南亚地区4种坡垒属(*Hopea* Roxb.)近缘树种中的鉴定效果,发现*matK*鉴定效果优于*rbcL*和*trnH-psbA*。目前,龙脑香科植物的DNA条形码研究仅见于对娑罗双属和坡垒属等少数类群,而整个龙脑香科的条形码尚未形成统一的标准(Kamiya et al., 2005; Tsumura et al., 2011; Trang et al., 2015)。

本研究选择目前常用于龙脑香科植物分子系统学研究的叶绿体*rbcL*、*matK*和*trnL-trnF*片段以及*ITS2*,综合评估其在热带龙脑香科树种中的通用性和鉴定能力,筛选出适合龙脑香科木材树种鉴定的DNA条形码及组合片段,以期为热带地区森林树种鉴定与保护提供科学依据。

1 材料与方法

1.1 实验材料

研究样品采集于中国科学院西双版纳热带植物园龙脑香科专类园区,每个树种采集2–3个个体的新鲜叶片,共采集龙脑香科5属9种共计19份样品(表1)。凭证材料均保存于中国科学院西双版纳热带植物园标本馆(HITBC)。为了在更大范围内考察条形码片段对龙脑香科树种的鉴别能力,我们整理了来自GenBank中常用于龙脑香科植物系统发育分析的叶绿体*rbcL*、*matK*和*trnL-trnF*片段以及进化速率较快的*ITS2*,最终共获得龙脑香科14属244种共计899条序列,包括娑罗双属(*Shorea* Roxb.) 119种、坡垒属(*Hopea* Roxb.) 40种、龙脑香属(*Dipterocarpus* Gaertn. f.) 26种、青梅属(*Vatica* L.) 20种以及柳安属(*Parashorea* Kurz) 7种。

1.2 方法

1.2.1 DNA提取和扩增

使用变色硅胶快速干燥叶片材料。每次称取30 mg叶片加入液氮研碎,采用改进后的CTAB法(Doyle and Doyle, 1987)提取植物总DNA,使用1%琼脂糖凝胶进行电泳检测。以*rbcL*、*matK*、*trnL-trnF*和*ITS2*作为扩增目的片段,扩增体系(20 μ L)为: 2 μ L DNA模板, 2.0 μ L 10 \times PCR Buffer, 2.0 μ L 25 mmol \cdot L $^{-1}$ MgCl $_2$, 1.6 μ L 2.5 mmol \cdot L $^{-1}$ dNTPs, 0.2 μ L 5 U \cdot μ L $^{-1}$ rTaq, 12.2 μ L ddH $_2$ O, 10 μ mol \cdot L $^{-1}$ 正反扩增引物各1 μ L。对扩增或测序失败的样品使用高保真聚合酶(*TransTaq* HiFi DNA Polymerase)进行重复实验,并调整扩增体系为: 2 μ L DNA模板, 2.0 μ L HiFi Buffer I, 1.6 μ L 2.5 mmol \cdot L $^{-1}$ dNTPs, 0.2 μ L 5 U \cdot μ L $^{-1}$ HiFi DNA Polymerase, 14.2 μ L ddH $_2$ O。由于*matK*片段引物通用性较低,实验中使用*matK3-F/matK1R*与*matK-390F/matK-1326R*两对引物

(Chen et al., 2010)进行扩增。另外, 在*matK*和*ITS2*扩增体系中添加BSA和DMSO以提高扩增效率。除此之外, 由于*ITS2*片段有较高的GC/AT含量, 在扩增

失败的样品反应体系中添加5% GC Enhancer。扩增产物送华大基因(广州华大公司)进行双向测序, 扩增引物及反应程序见表2。

表1 实验材料
Table 1 Information of experimental materials

No.	Species	Voucher numbers	GenBank accession number			
			<i>rbcL</i>	<i>matK</i>	<i>trnL-trnF</i>	<i>ITS2</i>
1	<i>Dipterocarpus turbinatus</i> C. F. Gaertner	JL5901	MK030561	MK051080	MK051054	MK051069
2	<i>D. turbinatus</i> C. F. Gaertner	JL5902	MK030562	MK051081	MK051055	MK051070
3	<i>D. retusus</i> Blume	JL9601	MK030559	MK051078	MK051052	MK193032
4	<i>D. retusus</i> Blume	JL9602	MK030560	MK051079	MK051053	MK193031
5	<i>Hopea chinensis</i> (Merr.) Hand.-Mazz.	JL110A	MK030563	MK051082	MK193024	MK193026
6	<i>H. chinensis</i> (Merr.) Hand.-Mazz.	JL110B	MK030564	MK051083	MK051059	MK193025
7	<i>H. hainanensis</i> Merrill & Chun	JL2201	MK030565	MK051084	MK051060	MK193029
8	<i>H. hainanensis</i> Merrill & Chun	JL2202	MK030566	MK051085	MK051061	MK193028
9	<i>H. hainanensis</i> Merrill & Chun	JL2203	MK030567	MK051086	MK051062	MK193030
10	<i>Parashorea chinensis</i> H. Wang	JL6101	MK030568	MK051087	MK051056	MK051071
11	<i>P. chinensis</i> H. Wang	JL6102	MK030569	MK051088	MK051057	MK051072
12	<i>Shorea assamica</i> Dyer	JL3901	MK030571	MK051090	MK051063	—
13	<i>S. assamica</i> Dyer	JL3902	MK030570	MK051089	MK051058	MK051073
14	<i>Vatica diospyroides</i> Symington	JL7601	MK030572	MK051091	MK051064	MK051074
15	<i>V. diospyroides</i> Symington	JL7602	MK030573	MK051092	MK193023	MK051075
16	<i>V. guangxiensis</i> X. L. Mo	JL4801	MK030574	MK051093	MK051065	MK051076
17	<i>V. guangxiensis</i> X. L. Mo	JL4802	MK030575	MK051094	MK051066	MK051077
18	<i>V. mangachapoi</i> Blanco	JL111A	MK030576	MK051095	MK051067	—
19	<i>V. mangachapoi</i> Blanco	JL111B	MK030577	MK051096	MK051068	MK193027

表2 PCR扩增引物及程序
Table 2 Primer pairs used for PCR amplification and amplification protocol

Barcodes	Primers sequence (5'–3')	Amplification protocol	References
<i>rbcL</i>	1F: ATGTCACCACAAACAGAGACTAAAGC 724R: TCGCATGTACCTGCAGTAGC	94°C 4 min; 94°C 30 s, 55°C 45 s, 72°C 1 min, 5 cycles; 94°C 30 s, 54°C 45 s, 72°C 10 min, 30 cycles; 72°C 10 min	Fay et al., 1997 Olmstead et al., 1992
<i>matK</i>	3F: CGTACAGTACTTTTGTGTTTACGA 1R: ACCCAGTCCATCTGGAAATCTTGG 390F: CGATCTATTCAATCAATATTTTC 1326R: TCTAGCACACGAAAGAAGT	94°C 4 min; 94°C 30 s, 51°C 50 s, 72°C 50 s, 35 cycles; 72°C 10 min 94°C 4 min; 94°C 30 s, 50°C 50 s, 72°C 50 s, 35 cycles; 72°C 10 min	Unpublished Cuénoud et al., 2002
<i>trnL-trnF</i>	F: GGTTC AAGTCCCTCTATCCC R: ATTTGA ACTGGTGACACGAG	94°C 4 min; 94°C 30 s, 55°C 1 min, 72°C 1 min, 35 cycles; 72°C 5 min	Taberlet et al., 1991
<i>ITS2</i>	<i>ITS2</i> 2F: ATGCGATACTTGGTGTGAAT <i>ITS2</i> 3R: GACGCTTCTCCAGACTACAAT <i>ITS3</i> : GCATCGATGAAGAACGCAGC 26SE: TAGAATTCCCCGGTTCGCTCGCCGTTAC	94°C 4 min; 94°C 30 s, 55°C 45 s, 72°C 45 s, 35 cycles; 72°C 10 min 94°C 4 min; 94°C 30 s, 57°C 45 s, 72°C 45 s, 35 cycles; 72°C 10 min	Chen et al., 2010 Chen et al., 2010 White et al., 1990; Sun et al., 1994

1.2.2 数据处理

使用Sequencher 4.14软件(GeneCodes Corp., Ann Arbor, Michigan, USA)对双向测序结果进行序列拼接和人工校对, 去除低质量序列和两端引物区。使用MEGA 6.0软件(Tamura et al., 2013)对获得的序列进行比对, 采用K2P (Kimura-2-parameter)模型计算种内及种间平均遗传距离。种内、种间遗传变异通过Taxton DNA软件(Meier et al., 2006)进行计算, 随后使用SPSS V21.0软件(IBM Corp., Armonk, New York, USA)对各片段的种内、种间遗传距离差异进行Wilcoxon Signed Rank检验(Lahaye et al., 2008)。采用Taxon DNA软件中的BM (Best Match)和BCM (Best Close Match)两种方法分析鉴定成功率。利用Geneious 6.15软件(Biomatters Ltd.)中Custom Blast功能实现本地BLAST, 以相同位点的百分比(identical sites)作为量化标准, 如果相同物种的identical sites值都大于与其它所有物种个体间的identical sites值, 即认为这个物种被准确鉴定(卢孟孟等, 2013)。使用Geneious 6.15软件构建邻接(NJ)树, 设置bootstrap值为1 000, 检验各支点支持率。如果同一个物种的所有个体形成1个单系, 且节点支持率大于50%, 则认为该物种被准确鉴定。

2 结果与讨论

2.1 候选条形码扩增和测序成功率

我们统计了4个片段的扩增和测序成功率, 将获得清晰的目的条带判定为扩增成功, 测序后获得高质量目的片段序列判定为测序成功。有效片段获得率计算公式为:

$$\text{有效片段获得率} = \text{扩增成功率} \times \text{测序成功率}$$

计算结果见表3。叶绿体片段*rbcL*、*matK*和*trnL-trnF*均获得100%的扩增成功率和测序成功率。通过采用引物*ITS2F/ITS3R* (Chen et al., 2010)进行扩增, *ITS2*的扩增成功率为100%, 而有效序列获得率仅为47.4%; 通过调整引物*ITS3/26SE* (White et al., 1990; Sun et al., 1994)并优化扩增体系, *ITS2*的有效序列获得率提高到89.5%。19个龙脑香科植物样品材料的扩增和测序结果表明, 叶绿体片段的有效序列获得率明显高于*ITS2*, 但采用2对*ITS2*引物进行扩增能够获得较高质量的*ITS2*序列。

表3 条形码扩增和测序成功率

Table 3 The success rate of PCR amplification and sequencing

Bar-codes	Amplification efficiency (%)	Success rate of sequencing (%)	Effective sequence ratio (%)
<i>rbcL</i>	100.0	100.0	100.0
<i>matK</i>	100.0	100.0	100.0
<i>trnL-trnF</i>	100.0	100.0	100.0
<i>ITS2</i>	100.0	89.5	89.5

2.2 序列特征分析及种内种间遗传变异比较

我们共获得899条序列, 包括*rbcL*片段259条、*matK*片段207条、*trnL-trnF*基因间隔区365条和*ITS2*片段68条。经过序列比对后得到*rbcL*、*matK*、*trnL-trnF*和*ITS2*四个片段的序列矩阵, 其长度分别为684、969、454和473 bp, GC含量分别为46.0%、35.8%、27.2%和72.5% (表4)。由于获得的序列包含*ITS2*片段数目较少, 因此没有对*ITS2*片段进行种内种间的遗传变异统计。数据分析结果显示, 3个叶绿体片段*rbcL*、*matK*和*trnL-trnF*的种内和种间变异分布有较多的重叠区域, 没有形成明显的间隔区; *trnL-trnF*和*matK*片段的种内遗传变异大于*rbcL*片段($P < 0.05$), 而*trnL-trnF*和*matK*片段的种内遗传变异差异不显著, 3个叶绿体片段的种内遗传变异的大小关系为*trnL-trnF* > *rbcL*, *matK* > *rbcL*, *trnL-trnF*和*matK*种内变异不显著 (表5); 在种间遗传距离的Wilcoxon Signed Rank检验中, 3个叶绿体片段的种间遗传变异差异显著, 其变异大小为*trnL-trnF* > *matK* > *rbcL* (表6)。

通过对种内和种间的K2P遗传距离分布进行分析, 表明3条叶绿体片段*rbcL*、*matK*和*trnL-trnF*的种内和种间遗传距离分布没有显示清晰的Barcoding Gap (图1)。

2.3 候选条形码鉴定能力评估

利用BM、BCM、BLAST和NJ四种方法分别对4个候选条形码物种水平的鉴定能力进行分析(图2), 结果获得*ITS2* 68条, 用BM/BCM方法获得100%的鉴定成功率; 基于BLAST方法, 鉴定成功率为93.33%; 基于NJ法, 鉴定成功率为85.71%。*matK*片段207条, 基于BM/BCM方法, 鉴定成功率为31.45%; 基于

表4 序列特征信息

Table 4 Information of sequences characteristic

Barcodes	No. of sequences	Aligned length (bp)	GC content (%)	No. of variable sites	Intraspecific distance (mean)	Interspecific distance (mean)
<i>rbcL</i>	259	684	46.0	23	0.0006	0.0161
<i>matK</i>	207	969	35.8	27	0.0032	0.0273
<i>trnL-trnF</i>	365	454	27.2	46	0.0041	0.0273
<i>ITS2</i>	68	473	72.5	51	0.0017	0.1311

表5 叶绿体条形码片段在龙脑香科物种种内变异Wilcoxon检验

Table 5 Wilcoxon signed rank tests of intraspecific divergence among chloroplast markers in the species of Dipterocarpaceae

W+	W-	Intra relative ranks	N	P value	Result
<i>trnL-trnF</i>	<i>rbcL</i>	W+=1258, W-=120	96	2.20E-07	$P<0.05$, <i>trnL-trnF</i> > <i>rbcL</i>
<i>trnL-trnF</i>	<i>matK</i>	W+=298, W-=297	43	9.93E-01	$P>0.05$, <i>matK</i> = <i>trnL-trnF</i>
<i>matK</i>	<i>rbcL</i>	W+=408, W-=57	44	9.18E-04	$P<0.05$, <i>matK</i> > <i>rbcL</i>

表6 叶绿体条形码片段在龙脑香科物种种间变异Wilcoxon检验

Table 6 Wilcoxon signed rank tests of interspecific divergence among chloroplast markers in the species of Dipterocarpaceae

W+	W-	Inter relative ranks	N	P value	Result
<i>trnL-trnF</i>	<i>rbcL</i>	W+=8136415.5, W-=1811614.5	4560	3.78E-296	$P<0.05$, <i>trnL-trnF</i> > <i>rbcL</i>
<i>trnL-trnF</i>	<i>matK</i>	W+=236579.5, W-=156361.5	903	1.41E-07	$P<0.05$, <i>trnL-trnF</i> > <i>matK</i>
<i>matK</i>	<i>rbcL</i>	W+=411099, W-=36832	946	1.00E-13	$P<0.05$, <i>matK</i> > <i>rbcL</i>

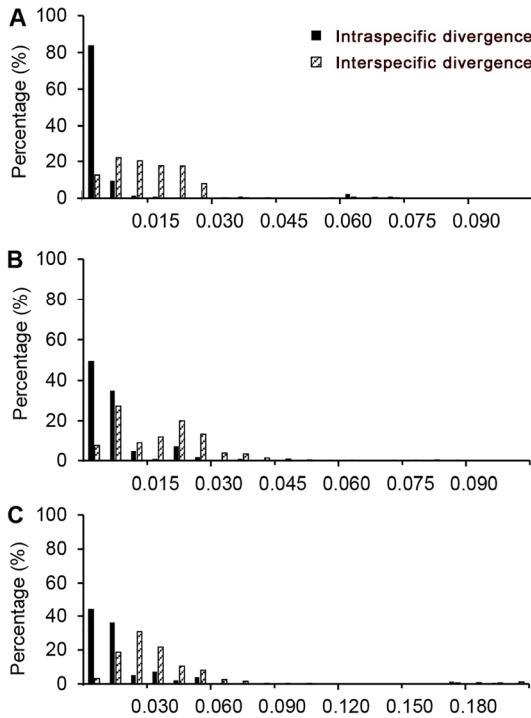


图1 条形码片段在龙脑香科物种种内和种间遗传变异分布 (A) *rbcL*序列; (B) *matK*序列; (C) *trnL-trnF*序列

Figure 1 Distribution for intra and inter-specific variation of markers in the species of Dipterocarpaceae (A) *rbcL* sequence; (B) *matK* sequence; (C) *trnL-trnF* sequence

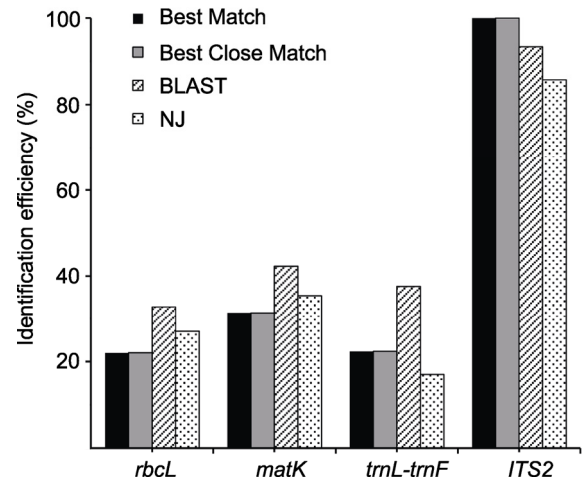


图2 基于4种方法的条形码鉴定效率比较

Figure 2 Comparison of the identification efficiency of bar-codes base on four methods

BLAST方法, 鉴定成功率为42.31%;基于NJ法, 成功率为35.42%。获得 *trnL-trnF* 片段365条, 基于BM/BCM、BLAST和NJ方法分别获得22.34%、37.61%和17.04%的鉴定成功率。获得 *rbcL* 片段259条, 基于BM/BCM、BLAST和NJ方法, 分别获得22.03%、32.79%和27%的鉴定成功率。总体来看, 基于4种评估方法获得了较为一致的分析结果。

2.4 讨论

2.4.1 龙脑香科树种DNA条形码综合评价

理想的DNA条形码需具备片段较短、种间变异较大且鉴定效率高等特点(Kress et al., 2005; Taberlet et al., 2007; 任保青和陈之端, 2010)。本研究, 我们发现*ITS2*片段在龙脑香科植物树种中呈现较高的扩增率和测序成功率(分别为100%和89.5%), 且基于BM、BCM、BLAST和NJ树4种条形码评估方法均获得最高的鉴定成功率, 因此我们认为*ITS2*是龙脑香科树种首选的DNA条形码。能否成功获得高质量条形码序列决定着条形码技术可否应用到实际物种鉴定工作中, 因此条形码扩增率和测序成功率是条形码筛选时的重要指标。*ITS2*是*ITS*序列的1个片段, 位于5.8S和26S rRNA之间, 其序列长度较短, 易于扩增和测序, 且其二级结构能够提供更多的信息(Schultz and Wolf, 2009; China Plant BOL Group et al., 2011)。Chen等(2010)通过大规模的取样分析, 认为可将*ITS2*作为药用植物鉴定的核心条形码; 同时, *ITS2*在山葡萄(*Vitis amurensis*)、兰科及红景天属等植物中也显示较好的鉴定效果(宋慧芳等, 2017; De Boer et al., 2017; Zhu et al., 2017)。本研究采用引物*ITS3/26SE*并优化扩增体系, 通过在扩增体系中添加DMSO及BSA提高PCR的产物特异性和产量。此外, 根据*ITS2*具有较高的GC/AT含量, 使用高保真DNA聚合酶和GC Enhancer优化PCR过程, 提高了*ITS2*在龙脑香科树种的获得率, 我们建议将20 μ L体系作为龙脑香科植物树种*ITS2*片段扩增的标准体系。

以往的研究表明, 叶绿体基因组进化速率相对较慢, 能够提供鉴定分析的变异位点较少(Newmaster et al., 2006; Chase et al., 2007; Kress and Erickson, 2007)。根据Wilcoxon Signed Rank检验结果, *matK*和*trnL-trnF*片段的种间遗传变异较高, 且差异显著, 显示出较高的应用潜力。*matK*基因序列是植物叶绿体基因中进化较快的编码基因, 基于样品序列*matK*片段在龙脑香科树种中进行物种鉴定的成功率仅次于*ITS2*片段, 表明*matK*用于龙脑香科树种鉴定效果较好。叶绿体基因间隔区*trnL-trnF*常用于龙脑香科树种的分子系统学研究, 在GenBank数据库中有较多的数据, 但在本研究中基于4种评估方法获得的鉴定成功率较低。*rbcL*片段是DNA条形码的核心之

一, 由于其引物通用性较高而被广泛使用, 但因变异较小不适合用于种级水平的物种鉴定(Kress and Erickson, 2007; Lahaye et al., 2008)。本研究结果显示, 叶绿体*rbcL*片段种内种间遗传变异不明显, 且鉴定成功率低于其它片段, 因此不适合作为热带龙脑香科树种的DNA条形码。综上, 我们建议将*ITS2*和*matK*片段作为热带龙脑香科树种的DNA条形码。

2.4.2 条形码技术在热带龙脑香科树种鉴定中的应用展望

随着世界各国对森林资源保护的重视, 使用合法认证的木材和木制品已成为国际共识, 因此对木材种类进行准确鉴定的重要性日趋凸显(Dormontt et al., 2015)。基于宏观和微观特征的木材鉴定方法往往依赖于经验丰富的技术专家, 且大多数解剖学特征也只能将木材产品鉴定到属级水平(Gasson et al., 2010; Dormontt et al., 2015)。近年来, 植物幼苗和木材等非标准材料DNA提取方法的改进大大提高了条形码技术在木材树种及其产品鉴定中的可行性, 研究者在树木形成层(Huang et al., 2015)、植物幼苗(Gonzalez et al., 2009)和木材标本(Yu et al., 2017; Jiao et al., 2018; 乔梦吉等, 2018)等样品材料中均成功获得条形码片段, 尤其是Tsumura等(2011)在龙脑香科木材样品中成功扩增到*trnL*、*trnL-trnF*、*trnH-trnK*和*psbC-trnS*片段。这些研究表明以植物叶片为材料扩增到的植物条形码片段也能用于龙脑香科木材产品的鉴定。本研究以龙脑香科树种的干燥叶片为样品材料, 结合GenBank数据分析评估4种条形码片段在龙脑香科植物鉴定方面的应用潜力, 能够为龙脑香科树种及木材产品的条形码鉴定提供借鉴。

龙脑香科植物是热带雨林中具有重要生态价值和经济意义的树种, 其木材产品在东南亚地区热带木材市场中有重要地位。本研究涉及的条形码序列来自龙脑香科14属244种, 共计899条, 涵盖了龙脑香科树种82.4%的属和46.1%的种, 序列数据大多来源于东南亚地区龙脑香科树种的相关研究, 在较大范围内代表了龙脑香科树种的遗传变异。目前, 龙脑香科植物的物种亲缘关系还未得到很好的解决与澄清, 这可能在一定程度上影响了DNA条形码在此类群中的应用。分子系统学研究结果不支持将龙脑香亚科(Dipterocarpoideae)分为Dipterocarpeae和Shoreaeae

2个族,理由是龙脑香属(*Dipterocarpus* Gaertn. f.)被认为是冰片香属(*Dryobalanops* C.F. Gaertn.)和Shoreeae的姐妹类群,且娑罗双属(*Shorea* Roxb.)为非单系分支(Heckenhauer et al., 2017; Ng et al., 2017)。利用DNA条形码进行物种鉴定需要较为完善的物种系统发育框架,因此龙脑香科条形码的研究有赖于其系统发育关系树的构建。随着高通量测序技术(next generation sequencing)的发展, DNA条形码技术衍生出新的发展趋势。例如,基因组浅层测序(genome skimming)通过获得全基因组较低测序深度基因组数据,借助生物信息学手段可获得叶绿体全基因组以及核基因组中高度重复区域(nrDNA)等序列(Hollingsworth et al., 2016; Zeng et al., 2018)。利用高通量测序技术能够获得物种的基因组水平信息,为快速分化的近缘物种鉴定提供更加有效的解决办法(李德铎和曾春霞, 2015)。随着条形码技术的不断成熟,以及龙脑香科木材树种DNA条形码数据库的不断完善和龙脑香科系统发育关系的日趋明确,相信DNA条形码技术将成为热带地区龙脑香科木材树种鉴定的有力工具。

致谢 本研究的样品采集承蒙中国科学院西双版纳热带植物园园林园艺部工作人员和实习生刘志毅的大力帮助;实验处理得到了蔡超男和张灿瑜的帮助。在此一并表示感谢!

参考文献

- 成俊卿 (1980). 中国热带及亚热带木材——识别、材性和利用. 北京: 科学出版社. pp. 413–418.
- 戴文君, 周磊, 杨梅 (2017). 中国龙脑香科植物研究及利用现状. 世界林业研究 30(6), 46–51.
- 李德铎, 曾春霞 (2015). 植物DNA条形码研究展望. 生物多样性 23, 297–298.
- 李锡文, 李捷, Ashton P (2002). 中国龙脑香料植物纪要. 云南植物研究 24, 409–420.
- 卢孟孟, 慈秀芹, 杨国平, 李捷 (2013). 亚热带森林乔木树种DNA条形码研究——以哀牢山自然保护区为例. 植物分类与资源学报 35, 733–741.
- 罗良才 (1988). 云南龙脑香科树种木材识别与利用的研究. 云南林业科技 (1), 1–13.
- 罗良才 (2008). 云南进口龙脑香科树种的木材识别与利用.

- 西部林业科学 37, 99–107.
- 乔梦吉, 陈柏旭, 符韵林 (2018). 基于DNA条形码技术的楠属和润楠属5种木材的识别. 基因组学与应用生物学 37, 4013–4021.
- 任保青, 陈之端 (2010). 植物DNA条形码技术. 植物学报 45, 1–12.
- 宋慧芳, 刘海双, 杨义明, 范书田, 李昌禹, 艾军 (2017). 山葡萄种质资源DNA条形码通用序列的筛选. 植物学报 52, 723–732.
- 童绍全, 陶国达 (1990). 龙脑香科. 见: 钱崇澍, 陈焕镛, 主编. 中国植物志, 第50卷2分册. 北京: 科学出版社. pp. 113–131.
- 王锦亮, 丁靖培, 程治英, 杨崇仁 (1992). 两种云南龙脑香属植物树脂精油的倍半萜成分及其季节性变化. 云南植物研究 14, 337–342.
- 杨家驹, 刘鹏, 卢鸿俊 (1989). 龙脑香亚科木材. 木材工业 3(3), 43–49.
- 杨家驹, 刘鹏, 卢鸿俊 (1991). 娑罗双属商品材. 木材工业 5, 45–50, 55.
- 于永福 (1999). 中国野生植物保护工作的里程碑——《国家重点保护野生植物名录(第一批)》出台. 植物杂志 (5), 3.
- 周亮, 黄建平, 黄白云 (2013). 热带植物云南娑罗双. 园林 (6), 62–63.
- 朱华, 王洪 (1992). 西双版纳龙脑香科植物纪要. 云南植物研究 14, 21–26.
- Ashton PS (1982). Flora Malesiana. Series I—Spermatophyta. Flowering Plants, Vol. 9, Part 2. Dipterocarpaceae. The Netherlands: Martinus Nijhoff Pub.
- CBOL Plant Working Group, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS, Erickson DL, Fazekas AJ, Graham SW, James KE, Kim KJ, Kress WJ, Schneider H, van AlphenStahl J, Barrett SCH, van den Berg C, Bogarin D, Burgess KS, Cameron KM, Carine M, Chacón J, Clark A, Clarkson JJ, Conrad F, Devey DS, Ford CS, Hedderson TAJ, Hollingsworth ML, Husband BC, Kelly LJ, Kesanakurti PR, Kim JS, Kim YD, Lahaye R, Lee HL, Long DG, Madriñán S, Maurin O, Meusnier I, Newmaster SG, Park CW, Percy DM, Petersen G, Richardson JE, Salazar GA, Savolainen V, Seberg O, Wilkinson MJ, Yi DK, Little DP (2009). A DNA barcode for land plants. *Proc Natl Acad Sci USA* 106, 12794–12797.
- Chase MW, Cowan RS, Hollingsworth PM, van den Berg

- C, Madriñán S, Petersen G, Seberg O, Jørgensen 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.
- Chen SL, Yao H, Han JP, Liu C, Song JY, Shi LC, Zhu YJ, Ma XY, Gao T, Pang XH, Luo K, Li Y, Li XW, Jia XC, Lin YL, Leon C (2010). Validation of the *ITS2* region as a novel DNA barcode for identifying medicinal plant species. *PLoS One* **5**, e8613.
- China Plant BOL Group, Li DZ, Gao LM, Li HT, Wang H, Ge XJ, Liu JQ, Chen ZD, Zhou SL, Chen SL, Yang JB, Fu CX, Zeng CX, Yan HF, Zhu YJ, Sun YS, Chen SY, Zhao L, Wang K, Yang T, Duan GW (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proc Natl Acad Sci USA* **108**, 19641–19646.
- Cuénoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW (2002). Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. *Am J Bot* **89**, 132–144.
- de Boer HJ, Ghorbani A, Manzanilla V, Raclariu AC, Kreziou A, Ounjai S, Osathanunkul M, Gravendeel B (2017). DNA metabarcoding of orchid-derived products reveals widespread illegal orchid trade. *Proc Roy Soc B* **284**, 20171182.
- Dormontt EE, Boner M, Braun B, Breulmann G, Degen B, Espinoza E, Gardner S, Guillery P, Hermanson JC, Koch G, Lee SL, Kanashiro M, Rimbawanto A, Thomas D, Wiedenhoeft AC, Yin YF, Zahnen J, Lowe AJ (2015). Forensic timber identification: it's time to integrate disciplines to combat illegal logging. *Biol Conserv* **191**, 790–798.
- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* **19**, 11–15.
- Fay MF, Cameron KM, Prance GT, Lledó MD, Chase MW (1997). Familial relationships of *Rhabdodendron* (Rhabdodendraceae): plastid *rbcL* sequences indicate a caryophyllid placement. *Kew Bull* **52**, 923–932.
- Gasson P, Miller R, Stekel DJ, Whinder F, Ziemińska K (2010). Wood identification of *Dalbergia nigra* (CITES Appendix I) using quantitative wood anatomy, principal components analysis and naïve Bayes classification. *Ann Bot* **105**, 45–56.
- Gonzalez MA, Baraloto C, Engel J, Mori SA, Pétronelli P, Riéra B, Roger A, Thébaud C, Chave J (2009). Identification of Amazonian trees with DNA barcodes. *PLoS One* **4**, e7483.
- Hassold S, Lowry II PP, Bauert MR, Razafintsalama A, Ramamonjisoa L, Widmer A (2016). DNA barcoding of Malagasy rosewoods: towards a molecular identification of CITES-listed *Dalbergia* species. *PLoS One* **11**, e0157881.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003). Biological identifications through DNA barcodes. *Proc Roy Soc B* **270**, 313–321.
- Heckenhauer J, Samuel R, Ashton PS, Turner B, Barfuss MHJ, Jang TS, Temsch EM, Mccann J, Salim KA, Attanayake AMAS, Chase MW (2017). Phylogenetic analyses of plastid DNA suggest a different interpretation of morphological evolution than those used as the basis for previous classifications of Dipterocarpaceae (Malvales). *Bot J Linn Soc* **185**, 1–26.
- Hollingsworth PM, Li DZ, van der Bank M, Twyford AD (2016). Telling plant species apart with DNA: from barcodes to genomes. *Philos Trans Roy Soc B* **371**, 2015–0338.
- Huang XC, Ci XQ, Conran JG, Li J (2015). Application of DNA barcodes in Asian tropical trees—a case study from Xishuangbanna Nature Reserve, Southwest China. *PLoS One* **10**, e0129295.
- Jiao LC, Yu M, Wiedenhoeft AC, He T, Li JN, Liu B, Jiang XM, Yin YF (2018). DNA barcode authentication and library development for the wood of six commercial *Pterocarpus* species: the critical role of xylarium specimens. *Sci Rep* **8**, 1945.
- Kamiya K, Harada K, Tachida H, Ashton PS (2005). Phylogeny of *PgiC* gene in *Shorea* and its closely related genera (Dipterocarpaceae), the dominant trees in south-east Asian tropical rain forests. *Am J Bot* **92**, 775–788.
- 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, 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, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V (2008). DNA barcoding the floras of biodi-

- versity hotspots. *Proc Natl Acad Sci USA* **105**, 2923–2928.
- Meier R, Shiyang K, Vaidya G, Ng PKL** (2006). DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst Biol* **55**, 715–728.
- Newmaster SG, Fazekas AJ, Ragupathy S** (2006). DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. *Can J Bot* **84**, 335–341.
- Ng CH, Lee SL, Tnah LH, Ng KKS, Lee CT, Maria M** (2016). Genome size variation and evolution in Diptero-
carpaceae. *Plant Ecol Divers* **9**, 437–446.
- Olmstead RG, Michaels HJ, Scott KM, Palmer JD** (1992). Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbcL*. *Ann Mo Bot Gard* **79**, 249–265.
- Schultz J, Wolf M** (2009). ITS2 sequence-structure analysis in phylogenetics: a how-to manual for molecular systematics. *Mol Phylogenet Evol* **52**, 520–523.
- Sun Y, Skinner DZ, Liang GH, Hulbert SH** (1994). Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theor Appl Genet* **89**, 26–32.
- Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermet T, Corthier G, Brochmann C, Willerslev E** (2007). Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. *Nucleic Acids Res* **35**, e14.
- Taberlet P, Gielly L, Pautou G, Bouvet J** (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol Biol* **17**, 1105–1109.
- Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S** (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
- Trang NTP, Duc NM, Sinh NV, Triest L** (2015). Application of DNA barcoding markers to the identification of *Hopea* species. *Genet Mol Res* **14**, 9181–9190.
- Tsumura Y, Kado T, Yoshida K, Abe H, Ohtani M, Taguchi Y, Fukue Y, Tani N, Ueno S, Yoshimura K, Kamiya K, Harada K, Takeuchi Y, Diway B, Finkeldey R, Na'iem M, Indrioko S, Ng KKS, Muhammad N, Lee SL** (2011). Molecular database for classifying *Shorea* species (Dipterocarpaceae) and techniques for checking the legitimacy of timber and wood products. *J Plant Res* **124**, 35–48.
- White TJ, Bruns TD, Lee S, Taylor J** (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. *PCR Protocols: a Guide to Methods and Applications*. San Diego: Academic Press. pp. 315–322.
- Wilson EO** (2016). *Half-Earth: Our Planet's Fight for Life*. New York: Liveright Publishing Corporation.
- Yu M, Jiao LC, Guo J, Wiedenhoef AC, He T, Jiang XM, Yin YF** (2017). DNA barcoding of vouchered xylarium wood specimens of nine endangered *Dalbergia* species. *Planta* **246**, 1165–1176.
- Zeng CX, Hollingsworth PM, Yang J, He ZS, Zhang ZR, Li DZ, Yang JB** (2018). Genome skimming herbarium specimens for DNA barcoding and phylogenomics. *Plant Methods* **14**, 43.
- Zhu RW, Li YC, Zhong DL, Zhang JQ** (2017). Establishment of the most comprehensive *ITS2* barcode database to date of the traditional medicinal plant *Rhodiola* (Crassulaceae). *Sci Rep* **7**, 10051.

Use of DNA Barcoding in Identifying Tropical Trees from Dipterocarpaceae

Jianlin Hu^{1, 2}, Zhifang Liu^{1, 2}, Xiuqin Ci¹, Jie Li^{1*}

¹Laboratory of Plant Phylogenetics and Conservation, Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Kunming 650223, China; ²University of Chinese Academy of Sciences, Beijing 100049, China

Abstract Plants from Dipterocarpaceae are vital tropical timber trees in Southeast Asia. They have high potential in forest supervision, forest resource protection and other aspects to evaluate the efficiency of DNA barcodes. We obtained 899 sequences from 244 species for 14 genera combined with sequences from the GenBank. We amplified and sequenced four DNA regions (*rbcL*, *matK*, *trnL-trnF* and *ITS2*) in samples from the family Dipterocarpaceae. The discrimination ability of DNA barcodes was evaluated by using effective sequence ratios, characteristics of sequences, divergence of intra- and inter-varieties, and identification success rates (Best Match, Best Close Match, BLAST and neighbor-joining methods). The *ITS2* fragment showed the highest identification efficiency. High-quality *ITS2* sequence can be obtained from the leaves of Dipterocarpaceae plants by using an optimized amplification system. The *matK* fragment showed obvious genetic variation and 100% amplification and sequencing efficiency, and its identification ability was higher than with other chloroplast markers. Therefore, *ITS2* and *matK* are suitable markers for tropical trees from Dipterocarpaceae.

Key words Dipterocarpaceae, DNA barcoding, *ITS2*, *matK*, tropical region

Hu JL, Liu ZF, Ci XQ, Li J (2019). Use of DNA barcoding in identifying tropical trees from Dipterocarpaceae. *Chin Bull Bot* 54, 350–359.

* Author for correspondence. E-mail: jieli@xtbg.ac.cn

(责任编辑: 白羽红)