

基于色木枫转录组数据的核低拷贝基因引物开发^{*}

邓雅文^{1,2,3)} 陈荣樟^{1,2,3)} 鲍 蕾^{1,2,3)} 管 宇^{1,2,3)} 王红芳^{1,2,3)†}

(1)生物多样性与生态工程教育部重点实验室, 100875, 北京;

2)东北虎豹国家公园保护生态学国家林草局重点实验室, 100875, 北京;

3)北京师范大学生命科学学院, 100875, 北京)

摘要 基于色木枫的转录组数据初步设计 166 对引物, 并根据琼脂糖凝胶电泳与 Sanger 测序结果筛选出其中的 26 对核低拷贝基因引物. 在东亚 63 个种群的 182 个个体中大批量扩增所开发引物, 单倍型多样性范围为 0.45~0.97, 核苷酸多样性为 $(1.90\sim 18.03)\times 10^{-3}$; 9~15 对引物可在金钱槭、茶条槭、青榨槭和鸡爪槭中扩增; STRUCTURE 聚类结果与之前核微卫星研究结果类似, 且展示了更多的谱系地理结构, 结果表明引物多态性高、拓展性强、具有可靠性. 本研究的转录组数据为槭属的研究提供重要遗传信息, 所用方法被证明可以有效开发大量核低拷贝基因引物, 已开发出的 26 对引物将为色木枫的进化历史和槭属系统发育研究提供丰富有效的分子标记.

关键词 引物开发; 转录组测序; 核低拷贝基因; 槭属

中图分类号 Q37

DOI: 10.12202/j.0476-0301.2022025

0 引言

色木枫 (*Acer pictum*) 属于无患子科 (Sapindaceae)、槭属 (*Acer*), 是第三纪孑遗的温带树种, 作为落叶林的主要组成部分在东亚广泛分布^[1]. 色木枫种群的核微卫星和叶绿体遗传结构表明, 北方种群和南方种群的遗传多样性具有显著差异, 南方种群的遗传多样性高, 遗传组成局域分布明显. 在北方种群和南方种群接壤的华北地区, 核微卫星和叶绿体遗传组成存在“核质分离”现象^[2-3]. 此外, 日本和东亚大陆的色木枫种群的遗传组成关系也无法被简单归纳^[4]. 由于色木枫和近缘种间, 以及亚种间存在广泛的基因流, 形态变异丰富, 缺乏断然可分的形态证据, 因此色木枫与元宝枫 (*A. truncatum*) 之间长期存在分类学争议. 《Flora of China》的编委 Chang Chinsung 认为元宝枫属于色木枫的亚种^[5], 而 Xu 等^[6] 则将二者视为 2 个独立的物种. 本研究以色木枫指代色木枫-元宝枫物种复合体. 为更清晰地揭示色木枫种水平和种下水平的进化历史和物种分类地位, 开发突变机制相对明确, 多样性高的核低拷贝基因引物十分必要.

核低拷贝基因指基因组中只有 1 个或少数几个拷贝的核基因, 主要功能为编码蛋白质, 因此序列相对保守, 主要以点突变的形式保存系统发育信息, 可以在较为广泛的分类群中扩增^[7-9], 同源关系较易确

定. 转录组指样本在某一功能状态下所转录出的所有产物的集合, 反映了该物种特定状态下几乎所有相关序列的信息^[10]. 转录组测序作为有力的工具, 被越来越广泛地用于筛选大量的核低拷贝基因引物^[11-13]. 本研究根据色木枫的转录组数据开发核低拷贝基因引物, 利用所开发引物对槭属其他物种以及金钱槭属 (*Dipteronia*) 物种进行交叉扩增, 根据扩增成功情况判断引物的拓展性, 此外还对东亚地区的色木枫种群进行聚类分析来验证引物的可靠性.

1 材料与方 法

1.1 试验材料 在陕西省宁陕县旬阳坝镇大寺沟 (33°34'N, 108°31'E) 与北京市北京师范大学校园 (39°57'N, 116°21'E) 中各选取 1 株色木枫作为试验材料, 采集树上的新鲜叶片并用蒸馏水洗净, 吸水纸吸干残余水分后剪碎装至 RNA 稳定液中, 交由诺禾致源公司进行 RNA 提取和转录组测序. 在东亚地区 63 个不同地点的种群中选取 182 株色木枫 (附表 A), 采集树上的新鲜叶片, 装袋并使用硅胶干燥保存, 用于后续全基因组 DNA 的提取.

1.2 试验方法 转录组原始测序数据通过了测序错误率分布检查、ATGC 含量分布检查的品质评估. 由于原始测序数据中含有低品质的序列片段, 因此去除

^{*} 国家自然科学基金面上项目 (32071494, 32171654)

[†] 通信作者: 王红芳 (1982—), 女, 博士, 副教授. 研究方向: 景观遗传学和谱系地理学. E-mail: wanghf@bnu.edu.cn

收稿日期: 2022-01-11

其中低品质碱基($Q \leq 20$)比例 $>50\%$ 的序列片段,以及无法确定碱基比例 $>0.1\%$ 的序列片段,再去掉带接头的序列片段,过滤后得到有效序列数据.采用 Trinity^[14]对有效序列进行拼接,得到转录本序列,选择每条基因中最长的转录本序列代表该基因.为了提高所设计核低拷贝基因引物的多态概率,在初始选择基因位点时,选择2个样品的转录组结果中具有 SNP 多态的基因位点进行设计,基因长度尽量控制在 300~1000 bp,基因中单核苷酸多态性(SNP)的数量超过2个,而且不位于基因首尾两端.将筛选后的基因在 NCBI 数据库中进行比对,找出最近的同源物种 mRNA 序列,通过 BioEdit 7.1.9^[15]的 Local BLAST 功能,以已知的功能基因作为数据库确定该基因的编码区,并验证外显子位置、内含子长度和假定功能.

确定基因编码区后进一步使用 Primer Premier 5 软件 (PREMIER Biosoft International, Palo Alto, California, USA)在该区域设计引物,原则如下:1)引物长度为 18~27 bp;2)退火温度为 45~64 °C;3)引物对之间的退火温度差异 <4 °C;4)3'端无连续碱基;5)3'端及近 3'端无二聚体;6)设计引物无错配情况、发卡结构以及交叉二聚体;7)引物对得分 >90 ;8)理论扩增产物长度 <1200 bp.

利用天根试剂盒提取色木枫的全基因组 DNA,提取过程严格按照说明书操作,获得的 DNA 产物 -20 °C 保存,用于后续引物测试.本研究使用 20 μ L PCR 扩增反应体系,其中包含 10~20 ng DNA, 0.8 μ L 正反向引物($10 \mu\text{mol} \cdot \text{L}^{-1}$), 15.4 μ L ddH₂O 和 10 μ L 2x TSINGKE Master Mix (Tsingke Biotechnology Co., Ltd.), 反应条件为 95 °C 预变性 5 min, 94 °C 变性 1 min, 对应最佳退火温度下退火 50 s, 72 °C 延伸 1 min, 以及 72 °C 最终再延伸 8 min (表 1). 琼脂糖凝胶电泳筛选出扩增产物只产生一条明亮清晰条带的引物.在我国东北 (大黑山、长白山、胜山)、西南 (香格里拉、海螺沟)、华东 (黄山) 的 6 个色木枫种群中各取 1 个个体,利用所筛引物小批量扩增,扩增产物的测序结果用软件 CodonCode Aligner 3.6.1 (CodonCode Corporation, Centerville, Massachusetts, USA; <http://www.codoncode.com/aligner/>)读取分析.去除 6 个个体中无多态位点或者可能为核多拷贝基因的引物后,再在东亚地区 63 个种群的 182 个个体中大批量扩增 (附表 A).

用 CodonCode Aligner 3.6.1 读取大批量扩增的数据并分型矫正,通过 DnaSP 5.10.01^[16]的 PHASE 功能确定杂合位点、多态位点和单倍型,计算每个位点的遗传多样性,包括核苷酸多样性 (π) 和单倍型多样性 (H_d).所有核低拷贝基因引物在金钱槭 (*Dipteronia*

sinensis)、茶条槭 (*Acer ginnala*)、青榨槭 (*A. davidii*)、鸡爪槭 (*A. palmatum*) 4 个近缘种中交叉扩增.此外,通过 STRUCTURE 2.3.4 软件^[17]对所有个体进行聚类分析,选择混合祖先模型和等位基因频率关联模型,种群数 K 设置为 1~20,重复运行 30 次, MCMC 运行长度设为 2000000, Burnin 长度设为 1000000,根据对数似然值 ($L(K)$) 以及对数似然值的变化速率 (ΔK)^[18] 综合确定出最有可能的种群组数,再通过 CLUMPP^[19] 和 DISTRICT 软件^[20] 作图分析聚类情况,与之前已有的核微卫星研究结果进行比较,验证所开发引物的可靠性.

2 结果与分析

2.1 基于转录组数据的引物设计结果

2 个样品的转录组数据品质见附表 B. 拼接有效序列后共获得 133506 条转录本序列,共计 101346 个基因,按材料方法所述标准筛选出 702 条基因序列用于引物设计.

本研究一共设计出 166 对测试引物,以 c80508_g1 的基因序列为例,在 NCBI 数据库中 blast 比对后得到最近的同源物种的 mRNA 序列 (图 1), 确定编码区域为 139~584 bp. 根据引物设计原则在 Primer Premier 5 软件中设计出测试引物 AM00500 (表 1).

2.2 设计引物的筛选结果

在 166 对测试引物中有 116 对未成功扩增或测序失败. 50 对测序成功的引物在小批量扩增结果中出现了 3 种情况: 1) 没有多态位点 (图 2), 尽管我们选择设计的转录本都有 SNP 多态, 以提高获得多态引物的可能性, 然而该 SNP 多态可能只是以极低频率出现在某个种群, 因此我们批量扩增情况下, 未检测到该多态. 在小批量扩增的样品来自核微卫星和叶绿体反映遗传差异显著的不同地理区域的情况下, 仍然没有发现多态, 表明该基因片段的遗传多样性可能非常低, 因此放弃该引物. 2) 多态位点只以杂合形式存在 (图 3), 因为无等位基因分离, 所以该基因片段极有可能存在多拷贝, 因此放弃该引物. 3) 多态位点既有杂合也有纯合 (图 4), 这类基因片段可能为核低拷贝基因引物, 保留用于后续分析.

2.3 核低拷贝基因引物的多态性和拓展性

保留的 30 对引物在 182 个个体的大批量扩增中又发现 4 对疑似多拷贝引物, 因此最终开发出 26 对核低拷贝基因引物 (表 1). 这些引物的长度范围为 143~652 bp, 多态位点和单倍型的数量分别为 6~69 和 7~74, 平均值分别为 37 和 42. 单倍型多样性范围为 0.45~0.97, 平均值为 0.80, 核苷酸多样性范围为 $(1.90\sim 18.03) \times 10^{-3}$, 平均值为 7.56×10^{-3} , 多态性较高. 13 对引物在金钱槭和青榨槭中成功扩增, 14 对引物在鸡爪槭中成功扩

表 1 26 对色木枫核低拷贝基因引物的位点信息

位点	引物序列(5'-3')	产物长度/ bp	退火温度/ °C	外显子/ bp	内含子/ bp	推测的功能
AM00033	F: CTCACCTCCAAGCAAT R: GGCCTTCCAGTCTCCC	327	62	1~236	237~327	Citrus sinensis protein kinase 2B,
AM00046	F: AAGAAGACCCTTATCAGTAT R: TCCTTTGGCTTGTAGTT	631	50	1~89 192~629	90~191	chloroplastic Ziziphus jujuba cultivar Dongzao chromosome 12, ZizJuj_1.1
AM00059	F: CATTTCGGCTTCACTATCT R: CGACTTCTCCACCTTTA	311	53	1~114 215~311	115~214	Herrania umbratica cultivar Fairchild unplaced genomic scaffold, ASM216827v2
AM00075	F: AGTCATCCAGCACTCTAC R: CTCACTTACACCTCCCTA	268	59	136~258	1~135 259~268	Malus x domestica cultivar Golden Delicious chromosome 12, MalDomGD1.0
AM00096	F: TGGGAACGAAGAAACAT R: AACTACAGGCTACATACAAT	320	56	1~291	292~320	Citrus sinensis serine/threonine-protein kinase KIPK-like
AM00107	F: GTCCTCGGTAGGATGTTTGG R: ATCGAGAGCAAGAGCCGATA	159	56	1~155		Citrus sinensis cultivar Valencia chromosome 4, Csi_Valencia_1.0
AM00126	F: TTCGGAGAAAGACAGGA R: GCTTACTAACAGGCACAGG	342	53	1~342		Citrus sinensis cultivar Valencia chromosome 9, Csi_Valencia_1.0
AM00134	F: GCAGCAGCAAACAAAAAGAA R: GCCAATGATTTAGAGCAAGGA	326	53	1~291	292~326	Bactrocera dorsalis strain Punador unplaced genomic scaffold
AM00150	F: TCAGAAGTCATGCCATTCA R: TCAAAATCCTCAACCCCAAG	259	56	1~158	159~259	Gossypium arboreum cultivar Shixiya1 chromosome 11, Gossypium_arboreum_v1.0
AM00410	F: ATCAAGCGAGGGAACGA R: GAGCCAACCTGAGCACC	375	62	1~356	357~375	Theobroma cacao cultivar B97-61/B2 chromosome 7, Criollo_cocoa_genome_v2
AM00478	F: CAGTTTCACGAATGCTCCTC R: TGCCTCCAGTAGAATCCAAG	385	65	267~385	1~266	Vitis vinifera cultivar PN40024 chromosome 18, 12X
AM00497	F: ATCCAACCCACATCACC R: GGAACAACGAGCATTAGAG	288	56	1~288		Durio zibethinus cultivar Musang King isolate D1 unplaced genomic scaffold, Duzib1.0 scaffold_27
AM00498	F: CCAGGCAGGAGTCAAGT R: AAATGGACGAGCAAATC	205	53	1~205		Populus euphratica unplaced genomic scaffold, PopEup_1.0 scaffold50.1
AM00500	F: CCTTATCTTTCTTTCCCTT R: TATGCGTGCCGAACTGG	143	59	1~143		Citrus sinensis cultivar Valencia chromosome 5, Csi_Valencia_1.0

续表

位点	引物序列(5'-3')	产物长度/ bp	退火温度/ °C	外显子/ bp	内含子/ bp	推测的功能
AM00525	F: CGAGCAAGAATGTCAAA R: TCAAGGGCGATAATAAG	272	53	1~181 182~272		Herrania umbratica cultivar Fairchild unplaced genomic scaffold, ASM216827v2 scaffold_4.1 Durio zibethinus
AM00527	F: GCACCATCGTGTCTTTC R: GCAGTGGGAGTTCTTGAT	310	59	145~310	1~144	cultivar Musang King isolate D1 unplaced genomic scaffold Citrus sinensis cultivar Valencia chromosome 7, Csi_Valencia_1.0
AM00549	F: CCACCATCGGAGAAACA R: AGCCCTTAGAGCCAAAA	397	56	1~397		Citrus sinensis cultivar Valencia chromosome 5, Csi_Valencia_1.0
AM00566	F: AAGGGTGGATAGTCTGC R: AAGTGGAATGAGCGTAG	643	59	1~643		Citrus sinensis cultivar Valencia chromosome 9, Csi_Valencia_1.0
AM00609	F: AGACAGGCACTCCAACAC R: AAGATTCTATCGGCACATT	593	59	1~593		Citrus sinensis cultivar Valencia chromosome 2, Csi_Valencia_1.0
AM00612	F: ATCATACAATCTACAGCCACAA R: TCCAAGGGACCAAGCAAT	297	56	1~293		Xanthoceras sorbifolium
AM00644	F: CATCAACCCAAGAATCCA R: GTCTGCCAATCAAGCCAC	229	59	1~229		microsatellite BI-H177 sequence Hevea brasiliensis cultivar reyan7-33-97
AM00645	F: GGCTATGGTAGATGGGTG R: ATAAATCGCAAGGAGGGA F: ATCAGAAGACCCTTACACTC	427	59		1~14, 125~203, 324~427	unplaced genomic scaffold, ASM165405v1 scaffold0645 Durio zibethinus
AM00646	R: GTTTATCAGCAATCCCTC F: CTCATCCAACGGGTCAA	652	59	1~652		cultivar Musang King isolate D1 unplaced genomic scaffold, Duzib1.0 scaffold_6 Citrus sinensis cultivar Valencia unplaced genomic scaffold, Csi_Valencia_1.0
AM00661	R: GTCGTGCTCTACTTACAAACA F: GAGCCAGCACTCCTAAC	416	53	1~416		scaffold_0096 Theobroma cacao cultivar B97-61/B2 chromosome 3, Criollo_cocoa_genome _v2
AM00687	R: CTCAAGAAGTATTTACCCC F: GAGGTTGGTGCGGTCTG	399	53	1~399		Citrus sinensis cultivar Valencia chromosome 8, Csi_Valencia_1.0
AM00699	R: TCGCTTCCCCTTGCTAT	406	59	1~406		

增, 15 对引物在茶条槭中成功扩增, 有 9 对引物在 4 个物种中都可以成功扩增, 拓展性较好(表 2)。

2.4 STRUCTURE 聚类分析结果 STRUCTURE 聚类分析结果显示, 当 $K=11$ 时, $L(K)$ 最大, ΔK 也最大(图 5), 在此基础上绘制了 K 为 2~11 的分组情况(图 6), 当 $K=2$ 时, 核低拷贝基因的分组结果与核微卫星的分

组结果非常类似, 分为南北 2 个分支, 在华北地区存在交错, 因为核低拷贝基因数量更多, 多态信息更丰富, 所以核低拷贝基因的最佳分组是 11 组。不同组之间存在明显地理分离, 东北、华中、华北以及云南地区种群分别属于不同的独立单倍型, 其中华北与东北地区的种群存在基因交流, 而华东等地区单倍型组成

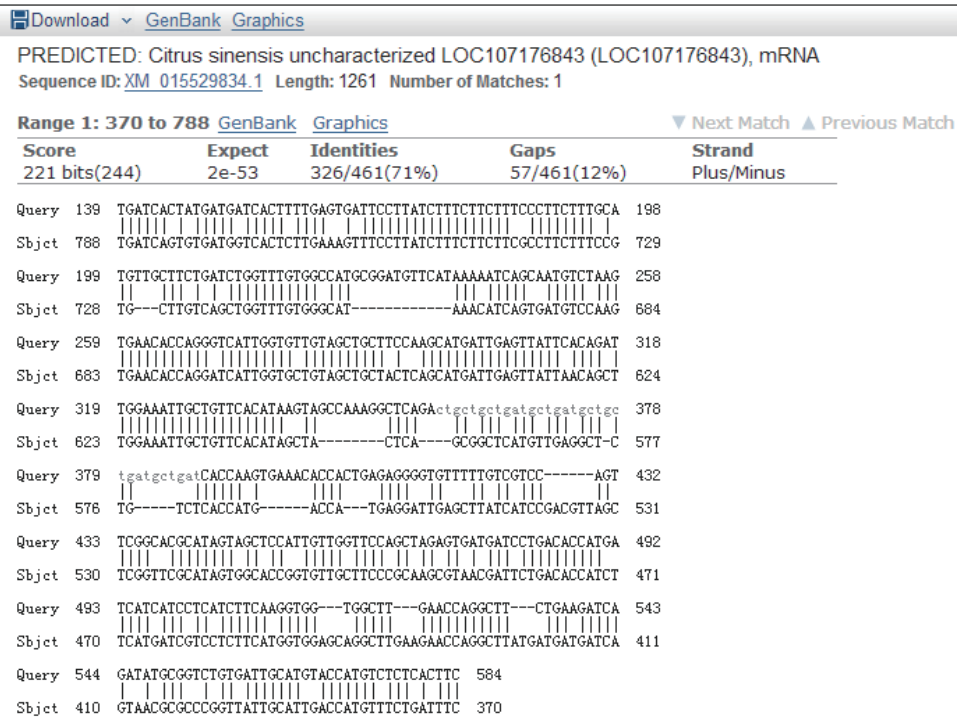


图 1 c80508_g1 序列与最近的同源物种 mRNA 序列比对结果



图 2 无任何多态位点的情况

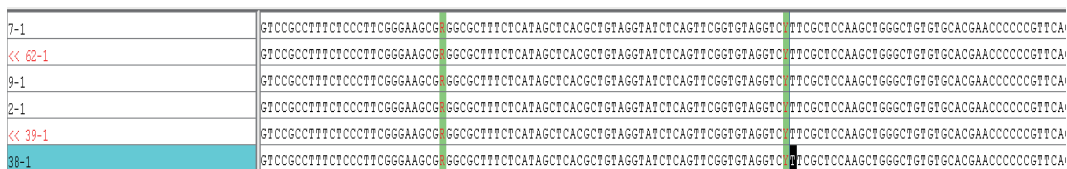


图 3 多态位点只以杂合形式存在的情况

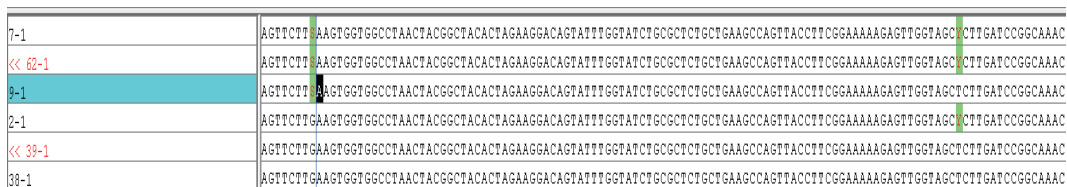


图 4 多态位点既有杂合也有纯合的情况

较复杂,可能还存在更细致的地理结构.核低拷贝基因展示了更多的谱系地理结构,所以为研究色木枞物种水平和种下水平的进化历史提供了理想的分子标记.

3 结论与讨论

转录组测序作为功能强大、成本低廉的手段已经

普及^[21],转录组序列中包含大量的 SNP 信息,且大多为编码基因,以此开发的分子标记相对保守,因此被越来越多地用于筛选核低拷贝基因引物^[13],而本研究也再一次证明了利用转录组测序筛选大量具有多态的核低拷贝基因是可行的.由于转录组测序结果数据量极大,且没有明确的外显子位置和内含子序列的

表 2 26 个色木枫核低拷贝基因位点的遗传多样性和交叉扩增情况

位点	样本量	变异位点	单变异位点	简约信息位点	单倍型 (H)	单倍型多样性(H _d)	核苷酸多样性 (π×10 ⁻³)	<i>Dipteronia sinensis</i>	<i>Acer ginnala</i>	<i>Acer davidii</i>	<i>Acer palmatum</i>
AM00033	182	17	5	12	19	0.67	3.09	-	+	-	-
AM00046	182	50	13	37	56	0.90	4.31	+	+	-	-
AM00059	181	16	4	12	17	0.50	2.11	-	-	-	-
AM00075	182	35	9	26	50	0.87	7.81	-	+	-	-
AM00096	182	27	11	16	30	0.60	2.86	-	-	+	+
AM00107	182	33	13	20	57	0.90	14.47	+	+	+	+
AM00126	182	37	11	26	46	0.88	8.87	+	-	-	-
AM00134	173	74	19	55	61	0.93	18.03	-	-	-	-
AM00150	182	20	8	12	23	0.83	6.76	+	+	+	+
AM00410	177	37	8	29	57	0.92	11.90	+	+	+	+
AM00478	182	34	4	30	27	0.79	10.64	-	-	-	-
AM00497	182	17	3	14	19	0.63	3.48	+	+	+	+
AM00498	182	23	8	15	29	0.69	5.48	+	+	+	+
AM00500	182	6	1	5	7	0.62	5.23	-	+	+	+
AM00525	172	36	11	25	37	0.83	9.36	-	+	-	+
AM00527	182	28	7	21	32	0.45	1.90	-	+	-	+
AM00549	182	43	13	30	56	0.94	9.05	+	+	+	+
AM00566	173	26	8	18	20	0.74	3.03	-	-	-	-
AM00609	182	65	20	45	74	0.94	5.03	-	-	+	+
AM00612	182	45	18	27	51	0.89	13.28	+	-	+	-
AM00644	182	33	13	20	42	0.91	14.53	+	+	+	+
AM00645	178	65	27	38	57	0.78	6.76	-	-	-	-
AM00646	182	42	13	29	52	0.88	3.67	+	-	-	-
AM00661	182	69	19	50	63	0.94	12.19	+	+	+	+
AM00687	182	43	15	28	67	0.97	7.64	+	+	+	+
AM00699	172	43	13	30	49	0.83	5.20	-	-	-	-

“+”表示该行引物可在该列物种中成功扩增;“-”表示该行引物无法在该列物种中成功扩增。

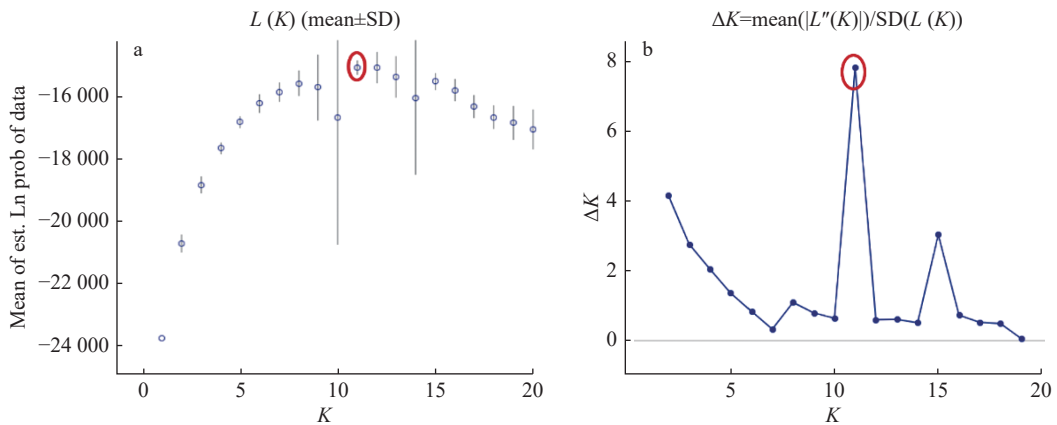


图 5 STRUCTURE 分析中 L(K)(a) 与 ΔK(b) 随 K 值的变化

信息, 因此使用转录组数据直接开发引物的成功率并不高^[11, 22]。如果先确定基因的编码区域、筛选具有

SNP 的位点后再设计引物, 可以大大提高成功率。本研究从 166 对引物中开发出 26 对引物, 比例达到了

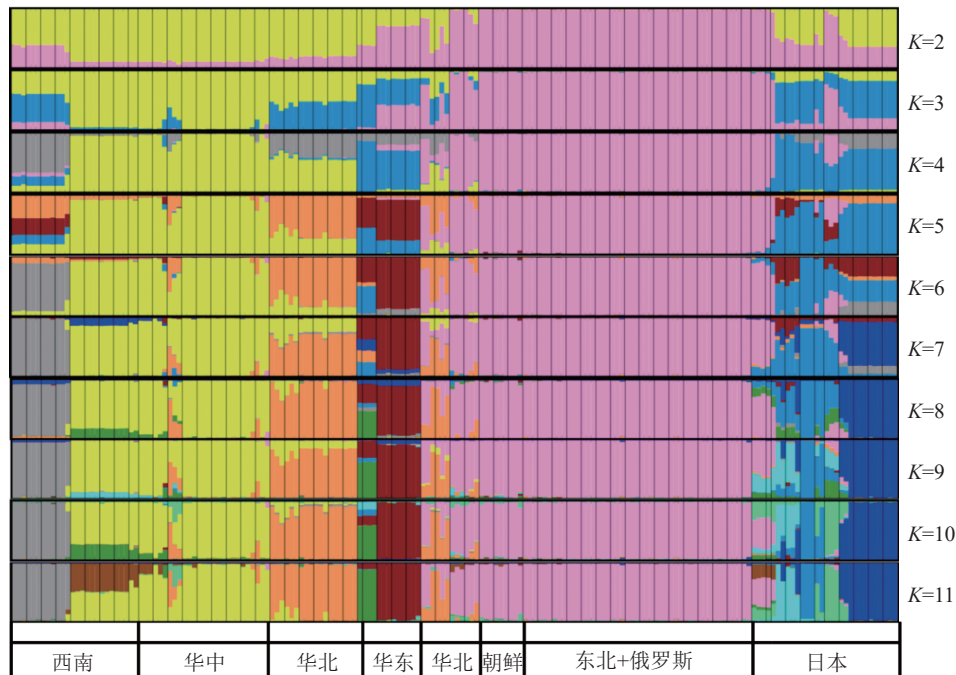


图 6 $K=2\sim 11$ 的遗传分组结果以及 $K=11$ 时分组对应的地理位置

15.6%, 与同样使用该方法的 Ye 等^[23] 的成功率相近, 比陈君^[24]、陈吉雯^[25] 等直接设计引物的成功率提高 1 倍。核低拷贝基因多态性高、直系同源、双亲遗传, 可以提供大量信息位点, 能够根据序列的差异清晰地地区分出不同的物种^[24, 26-29]。本研究开发的 26 对核低拷贝基因引物, 在地理跨度极大的 63 个色木枫种群中批量扩增时成功率非常高, 而且检测出的遗传多样性较高、交叉扩增时拓展性较好, 反映了更精细的谱系地理结构, 表明这些引物位点包含了丰富的关键进化信息, 将为解译色木枫复杂的物种进化历史, 以及建立种以上水平的系统发育关系, 提供除基因组外性价比比较高的途径。

尽管转录组筛选核低拷贝基因具有广泛的应用前景, 然而开发效率仍然较低。未来可以在转录组水平进行生物信息分析, 预筛选多态性高的核低拷贝基因, 而不需要通过实验判断是否为低拷贝位点, 如将色木枫的转录组基因序列与已知的其他物种核低拷贝基因进行同源比较, 通过 HaMSTR 软件预测核低拷贝直系同源基因^[30-32], 或是通过 OrthoMCL 软件将各同源直系基因进行聚类, 区分出核多拷贝基因与核低拷贝基因^[33]。此外, 转录组存在可变剪切^[34], 容易导致根据转录本设计的引物无法在基因组中成功扩增, 或者扩增的片段与预期片段不一致。在我们设计引物过程中, 50 对可以成功测序的引物中部分未发现多态, 我们猜测至少部分原因可能是转录组中的可变剪切造成的, 因此未来设计引物时, 可以评估可变剪切对引物设计的影响^[35]。虽然本研究中筛选的 26 个核

低拷贝基因引物交叉扩增的成功率非常高, 但只选用了 4 个物种, 代表性可能不够, 因此这些引物的在其他槭属物种中的扩增拓展性还需要进一步验证。

本研究基于色木枫的转录组数据开发了 26 对核低拷贝基因引物, 证明了该方法的有效性, 并且为后续应用于色木枫的进化历史和槭属系统发育研究提供丰富而有效的分子标记。

数据分享: 附表 A 和附表 B 见期刊网站。数据链接: <http://www.bnujournal.com/article/doi/10.12202/j.0476-0301.2022025>

4 参考文献

- [1] 徐廷志. 槭树科的地理分布[J]. 云南植物研究, 1996, 18(1): 43
- [2] LIU C P, TSUDA Y, SHEN H L, et al. Genetic structure and hierarchical population divergence history of *Acer mono* var. *mono* in South and Northeast China[J]. PLoS One, 2014, 9(1): e87187
- [3] GUO X D, WANG H F, BAO L, et al. Evolutionary history of a widespread tree species *Acer mono* in East Asia[J]. Ecology and Evolution, 2014, 4(22): 4332
- [4] YE J W, GUO X D, WANG S H, et al. Molecular evidence reveals a closer relationship between Japanese and mainland subtropical specimens of a widespread tree species, *Acer mono*[J]. Biochemical Systematics and Ecology, 2015, 60: 143
- [5] WU Z Y, RAVEN P, ZHANG L. Flora of China Illustrations, Volume 11[M]. Beijing: Science Publishing, 2009

- [6] XU T Z, CHEN Y S, DE JONG P C, et al. Flora of China Volume. 11[M]. Beijing: Science Press, 2008
- [7] 马兰, 黄原. 单拷贝核基因在昆虫分子系统学中的应用[J]. 昆虫知识, 2006, 43(1): 6
- [8] ZIMMER E A, WEN J. Using nuclear gene data for plant phylogenetics: Progress and prospects II. Next-gen approaches[J]. Journal of Systematics and Evolution, 2015, 53(5): 371; 379
- [9] LI Z, DE LA TORRE A R, STERCK L, et al. Single-copy genes as molecular markers for phylogenomic studies in seed plants[J]. Genome Biology and Evolution, 2017, 9(5): 1130
- [10] 肖永. 基于转录组学桫欏科单拷贝分子标记开发与应用[D]. 广州: 仲恺农业工程学院, 2019
- [11] BAI W N, ZHANG D Y. Current status and future directions in plant phylogeography[J]. Chinese Bulletin of Life Sciences, 2014, 26(2): 125
- [12] SHEN H, JIN D M, SHU J P, et al. Large-scale phylogenomic analysis resolves a backbone phylogeny in ferns[J]. Giga Science, 2017, 7(2): 116
- [13] QI X P, KUO L Y, GUO C C, et al. A well-resolved fern nuclear phylogeny reveals the evolution history of numerous transcription factor families[J]. Molecular Phylogenetics and Evolution, 2018, 127: 961
- [14] GRABHERR M G, HAAS B J, YASSOUR M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome[J]. Nature Biotechnology, 2011, 29(7): 644
- [15] HALL T. BioEdit: a user-friendly biological sequence alignment program for Windows 95/98/NT[J]. Nucleic Acids Symposium Series, 1999, 41: 95
- [16] ROZAS J, SÁNCHEZ-DELBARRIO J C, MESSEGUER X, et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods[J]. Bioinformatics, 2003, 19(18): 2496
- [17] GILBERT K J, ANDREW R L, BOCK D G, et al. Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program structure[J]. Molecular Ecology, 2012, 21(20): 4925
- [18] EVANNO G, REGNAUT S, GOUDET J. Detecting the number of clusters of individuals using the software structure: a simulation study[J]. Molecular Ecology, 2005, 14(8): 2611
- [19] JAKOBSSON M, ROSENBERG N A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure[J]. Bioinformatics, 2007, 23(14): 1801
- [20] ROSENBERG N A. Distruct: a program for the graphical display of population structure[J]. Molecular Ecology Notes, 2004, 4(1): 137
- [21] 张春兰, 秦孜娟, 王桂芝, 等. 转录组与RNA-Seq技术[J]. 生物技术通报, 2012(12): 51
- [22] CURTO M A, PUPPO P, FERREIRA D, et al. Development of phylogenetic markers from single-copy nuclear genes for multi locus, species level analyses in the mint family (Lamiaceae)[J]. Molecular Phylogenetics and Evolution, 2012, 63(3): 758
- [23] YE J W, LI Q, TIAN X Y, et al. Twenty-seven low-copy nuclear primers for *Lindera obtusiloba* (Lauraceae): a tertiary relict species in East Asia[J]. Applications in Plant Sciences, 2017, 5(12): 1700120
- [24] 陈君. 中国白栎组单拷贝核基因标记的开发及系统发育关系研究[D]. 北京: 中国林业科学研究院, 2017
- [25] 陈吉雯. 宽叶香蒲(*Typha latifolia*)单拷贝核基因标记的开发及谱系地理学研究[D]. 武汉: 武汉大学, 2020
- [26] WANG Q, YU Q S, LIU J Q. Are nuclear loci ideal for barcoding plants? A case study of genetic delimitation of two sister species using multiple loci and multiple intraspecific individuals[J]. Journal of Systematics and Evolution, 2011, 49(3): 182
- [27] 赵奉彬. 中国杨属物种DNA条形码及分子系统学初步研究[D]. 北京: 北京林业大学, 2016
- [28] 杨晨阳, 于超, 马玉杰, 等. 基于SSR标记和单拷贝核基因的蔷薇属植物系统发生分析[J]. 北京林业大学学报, 2018, 40(12): 85
- [29] JOLY S, STARR J R, LEWIS W H, et al. Polyploid and hybrid evolution in roses east of the Rocky Mountains[J]. American Journal of Botany, 2006, 93(3): 412
- [30] EBERSBERGER I, STRAUSS S, VON HAESELER A. HaMStR: profile hidden Markov model based search for orthologs in ESTs[J]. BMC Evolutionary Biology, 2009, 9: 157
- [31] 刘勉, 张彩飞, 黄建勋, 等. 利用低拷贝核基因重建菊科紫菀亚科族间系统发育关系[J]. 植物学报, 2015, 50(5): 549
- [32] WANG Q, LIU J Q, ALLEN G A, et al. Arctic plant origins and early formation of circumarctic distributions: a case study of the mountain sorrel, *Oxyria digyna*[J]. New Phytologist, 2016, 209(1): 343
- [33] HUANG C H, SUN R R, HU Y, et al. Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution[J]. Molecular Biology and Evolution, 2016, 33(2): 394
- [34] LI Y P, DAI C, HU C G, et al. Global identification of alternative splicing via comparative analysis of SMRT- and Illumina-based RNA-seq in strawberry[J]. The Plant Journal, 2017, 90(1): 164
- [35] 马素平. 利用转录组测序数据分析可变剪接的方法[J]. 科学与信息化, 2020(8): 2

Low-copy nuclear gene primers based on transcriptome data of *Acer pictum*

DENG Yawen^{1, 2, 3)} CHEN Rongzhang^{1, 2, 3)} BAO Lei^{1, 2, 3)}
GUAN Yu^{1, 2, 3)} WANG Hongfang^{1, 2, 3)}

(1) Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, 100875, Beijing, China;

2) National Forestry and Grassland Administration Key Laboratory for Conservation Ecology in the Northeast Tiger and Leopard National Park, 100875, Beijing, China;

3) College of Life Sciences, Beijing Normal University, 100875, Beijing, China)

Abstract Tertiary relic temperate tree species *Acer pictum* is excellent to study evolutionary history in East Asia. Abundant morphological variations exist in this species, contributing to controversial classification at the species and subspecies level. Therefore low-copy primers with clear mutation mechanism and high diversity are of great significance in the evolution and taxonomy of *A. pictum*. In this study, we designed 166 pairs of primers based on transcriptome data, and screened for low-copy genes by agarose gel electrophoresis and Sanger sequencing. A total of 26 pairs of low-copy primers were used. Effectiveness of primers was verified by amplifications in 182 individuals of 63 populations in East Asia. Nucleotide diversity was found to range from 1.90×10^{-3} to 18.03×10^{-3} , haplotype diversity from 0.45 to 0.97. Nine to fifteen pairs of primers were used for successful amplifications in *Dipteronia sinensis*, *Acer ginnala*, *A. davidii* and *A. palmatum*. STRUCTURE clustering data were found similar to previous nuclear microsatellite studies and showed more phylogenetic geography. Transcriptome data provided important genetic information for *Acer*, the method was effective. The 26 pairs of primers will provide rich and effective molecular markers for the study of evolutionary history of *A. pictum* and the phylogeny of *Acer*.

Keywords primer exploitation; transcriptome sequencing; low-copy nuclear gene; *Acer*

【责任编辑: 武 佳】