

南方红豆杉转录组 SSR 位点分析及其分子标记开发

蒋路园^{1,2}, 吴文丽¹, 张恺恺¹, 张林凤¹, 邵芬娟¹, 陈段芬², 邱德有¹, 杨艳芳^{1*}

1. 中国林业科学研究院林业研究所 林木遗传育种全国重点实验室, 国家林业局林木培育重点实验室, 北京 100091

2. 河北农业大学园艺学院, 河北 保定 071001

摘要: **目的** 基于赤霉素 (gibberellin A₃, GA₃) 处理下的南方红豆杉 *Taxus chinensis* var. *mairei* 转录组数据鉴定和分析 SSR 位点, 开发可用于南方红豆杉种质资源鉴定和遗传多样性分析的 SSR 标记。**方法** 使用 MISA 软件对南方红豆杉转录组数据进行 SSR 分析, 搜索 SSR 位点。可搜索单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸, 最小重复数分别设置为 10、6、5、5、5、5 次。采用 Primer 3 (2.3.5 版, 默认参数) 进行 SSR 引物设计。通过毛细管荧光电泳检测, 最终选定多态性好, 扩增成功率高的 SSR 引物用于后续数据分析; 通过 MEGA 进行 UPGMA 聚类分析, 构建聚类树。**结果** 转录组测序共获得 202 361 条 Unigene, 采用 MISA 软件搜索出 12 008 个 SSR 位点, 分布在 10 894 条 Unigenes 上, 发生频率为 5.38%, 平均距离为 1/7.64 kb。单核苷酸重复类型最丰富, 占总 SSR 位点的 51.74%, 其次为三核苷酸 (20.92%) 和二核苷酸重复类型 (20.88%)。A/T、AG/CT 是优势重复基序, 分别占总 SSR 重复类型的 49.60% 和 10.93%。随机抽选 96 对 SSR 引物, 以 4 个地区 12 份南方红豆杉种质进行引物有效性和多态性验证, 筛选出 13 对具多态性的 SSR 引物, 多态性引物比率为 13.54%。13 对引物在东北红豆杉 *Taxus cuspidata* 扩增效率高达 100%, 但在曼地亚红豆杉 *Taxus × media* 扩增效率仅为 23.08%。聚类分析表明太行山地区南方红豆杉聚为一类, 浙江、福建和湖南等南方来源的南方红豆杉聚为一类, 后三者彼此之间遗传距离更近。**结论** 南方红豆杉转录组测序产生的 Unigene 信息可以用来开发 SSR 分子标记, 开发的 13 个标记将有助于南方红豆杉物种遗传多样性、分子育种、资源保护等方面的研究。

关键词: 南方红豆杉; 转录组; SSR; 东北红豆杉; 曼地亚红豆杉

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2024)03-0928-09

DOI: 10.7501/j.issn.0253-2670.2024.03.023

Transcriptome sequencing and development of SSR molecular markers of *Taxus chinensis* var. *mairei*

JIANG Luyuan^{1,2}, WU Wenli¹, ZHANG Kaikai¹, ZHANG Linfeng¹, SHAO Fenjuan¹, CHEN Duanfen², QIU Deyou¹, YANG Yanfang¹

1 State Key Laboratory of Tree Genetics and Breeding, State Forestry Administration Key Laboratory of Tree Breeding and Cultivation, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China

2 College of Horticulture, Hebei Agricultural University, Baoding 071001, China

Abstract: **Objective** To identify and analyze SSR loci based on the transcriptome data of *Taxus chinensis* var. *mairei*, Lemee et levl S. Y. Hu et Liu treated with gibberellin A₃ (GA₃), and to develop SSR markers that can be used for the identification of the germplasm resources and genetic diversity analysis of *T. chinensis* var. *mairei*. **Method** Perform SSR analysis on transcriptome data of southern Chinese yew using MISA software and search for SSR loci. Single nucleotide, dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide can be searched, with the minimum number of repetitions set to 10, 6, 5, 5, and 5, respectively. Primer 3 (version 2.3.5, default parameters) was used for SSR primer design. Through capillary fluorescence electrophoresis detection, SSR primers with good polymorphism and high amplification success rate were ultimately selected for subsequent data analysis; Perform UPGMA clustering analysis using MEGA and construct a clustering tree. **Results** A total of 202 361 Unigenes were obtained from the transcriptome database of the needle of *T. chinensis* var. *mairei* treated with GA₃, and 12 008 SSR were found by MISA software, which were distributed on 10 894 Unigenes with an occurrence frequency of 5.38% and a mean distance of 7.64 kb per one. Mono-nucleotide repeats were major types, accounting for 51.74% of the total number of SSR, followed by trinucleotide repeats

收稿日期: 2023-06-09

基金项目: 国家林业和草原局科技发展中心生物安全与遗传资源管理项目 (KJZXSA202037); 中央级公益性科研院所基本科研业务费专项资金 (CAFYBB2019SY006)

作者简介: 蒋路园 (1994—), 女, 博士研究生, 主要从事林木遗传育种研究。E-mail: bachelorjly@163.com

*通信作者: 杨艳芳, 研究员, 主要从事林木次生代谢工程与调控研究。Tel: (010)62889638 E-mail: echoyyf@caf.ac.cn

(20.92%) and dinucleotide repeats (20.88%). A/T and AG/CT were dominant repeat motifs, accounting for 49.60% and 10.93% of the total number of SSR repeat types, respectively. A total of 96 SSR primers was selected randomly, and verified with 12 germplasms of *T. chinensis* var. *mairei*. 13 pairs of SSR primers with polymorphism were screened out, and the ratio of polymorphic primers was 13.54%. The amplification rate of 13 pairs of primers was as high as 100% in *T. cuspidata*, but only 23.08% in *T. × media*. Cluster analysis showed that *T. chinensis* var. *mairei* in Taihang Mountain area was clustered into one group, and *T. chinensis* var. *mairei* from southern China such as Zhejiang, Fujian and Hunan were clustered into one group, and the genetic distance between the latter three groups was closer. **Conclusion** The Unigenes information from the transcriptome of *T. chinensis* var. *mairei* can be used to identify SSR markers. The SSR markers screened in this study are helpful to the study of genetic diversity, molecular breeding and resource protection of *T. chinensis* var. *mairei*.

Key words: *Taxus chinensis* Smith var. *mairei*, Lemee et Lévl S. Y. Hu et Liu; transcriptome; SSR; *Taxus cuspidata* Siebold & Zucc.; *Taxus × media* Rehder

红豆杉属 *Taxus* spp. 植物, 俗称紫杉, 不仅具有较高的观赏价值, 更是抗癌药物紫杉醇的重要原材料, 具有极高的药用价值。红豆杉属植物主要分布在北半球, 全球有 11 个种^[1]。南方红豆杉 *Taxus chinensis* Smith var. *mairei*, Lemee et Lévl S. Y. Hu et Liu 又称美丽红豆杉, 是中国红豆杉 *T. chinensis* (Pilg.) Rehd. 的变种, 在我国分布范围最广。自 1985 年美国国立癌症研究院发现红豆杉属植物含有多种抗肿瘤活性成分后, 人类为谋取暴利随意采伐, 导致南方红豆杉自然资源日益减少^[2]。红豆杉资源保护及可持续开发利用是影响红豆杉产业发展的关键问题。

简单重复序列标记 (simple sequence repeats, SSR), 又称短串联重复序列或微卫星标记, 是一类由几个核苷酸 (1~6 个) 为重复单位组成的长达几十个核苷酸的重复序列, 长度较短, 且广泛分布于真核生物基因组中^[3]。由于重复单位的核苷酸不同以及重复次数不完全相同, 造成 SSR 长度的高度变异性。SSR 遵循孟德尔遗传方式, 呈共显性遗传, 具有数量丰富、高度多态性、可靠性好、检测快速等特征, 已成为群体遗传学分析中的一个有用的标记系统, 被广泛应用于遗传图谱的构建、种质资源遗传多样性研究、遗传关系与分化及群体结构分析等领域。此外, 近年来, 有研究发现 SSR 还可能对植物的逆境胁迫响应有重要作用^[4]。

目前, 已有研究者在红豆杉中开发出了多个具有多态性的 SSR 分子标记^[5-11], 并利用 SSR 分子标记技术在红豆杉中开展相关研究, 如国外研究者利用 SSR 分子标记技术研究了生境断裂对欧洲红豆杉 *T. baccata* L. 的遗传结构的影响^[12], 国内研究者也通过 SSR 技术发现生境断裂是导致云南红豆杉 *T. yunnanensis* Cheng et L. K. Fu 显著遗传瓶颈、近亲繁殖和种群分化主要原因^[13]。开发南方红豆杉 SSR

分子标记, 了解南方红豆杉的遗传多样性, 对南方红豆杉资源保护、育种和遗传改良非常重要。

赤霉素 (gibberellin A₃, GA₃) 和紫杉醇同为二萜类化合物, 在生物合成途径中存在底物竞争关系。前人研究表明, GA₃ 处理可以提高红豆杉中紫杉醇产量^[14], 紫杉醇合成关键酶基因的启动子序列含有 GA₃ 应答成分^[15]。因此, 基于 GA₃ 处理南方红豆杉转录组数据开发 SSR 标记, 有可能获得紫杉醇合成相关的 SSR 位点。

随着科技发展, 下一代测序技术 (next-generation sequencing, NGS) 的出现使得 SSR 标记快速经济开发成为现实。目前, 人们已经基于组学数据对包含红豆杉在内的多种植物开发获得了一些具有多态性的标记^[5,8-10], 然而, 目前可检索到的南方红豆杉 SSR 标记数量仅有少量具有多态性, 远远不能满足南方红豆杉遗传多样性等研究。本课题组前期利用 RNA-seq 技术对 GA₃ 处理下的南方红豆杉的叶片进行了转录组测序, 本研究基于此转录组数据对 SSR 分子标记进行了鉴定和分析, 并开发出了 13 对具有多态性的 SSR 标记, 丰富南方红豆杉 SSR 标记数量, 为红豆杉品种鉴定、分子辅助标记育种和遗传多样性分析提供理论基础。

1 材料、仪器及试剂

南方红豆杉 *T. chinensis* Smith var. *mairei*, Lemee et Lévl S. Y. Hu et Liu、东北红豆杉 *T. cuspidata* Siebold & Zucc. 和曼地亚红豆杉 *T. × media* Rehder 材料经中国林业科学研究院林业研究所杨艳芳研究员鉴定后, 分别种植于中国林业科学研究院科研温室和廊坊种植基地, 样品信息见表 1。

3730XL 型基因测序仪 (Applied Biosystems)、veriti™ 96 型 PCR 仪 (Thermo Fisher Scientific)、5430R 型离心机 (Eppendorf)、NanoDrop 分光光度计 (Gene Company Limited) 等。本实验所用主要试

表 1 样本信息

Table 1 Sample information

序号	编号	种	个数	产地	纬度 (N)	经度 (E)
1	FJNP	南方红豆杉	4	福建南平	26°63'	118°17'
2	HNYZ	南方红豆杉	2	湖南永州	26°43'	111°60'
3	ZJLS	南方红豆杉	2	浙江丽水	28°45'	119°92'
4	TH	南方红豆杉	4	陕西长治	36°19'	113°11'
5	DD	东北红豆杉	1	辽宁丹东	40°07'	124°23'
6	MDY	曼地亚红豆杉	1	山东济南	36°40'	117°00'

剂包括: CTAB (北京酷来搏科技有限公司)、TaqDNA 聚合酶 (康为世纪生物科技股份有限公司)、琼脂糖 (南京森贝伽生物科技有限公司)、TAE buffer (北京酷来搏科技有限公司)、ddH₂O (北京酷来搏科技有限公司) 等。

2 方法

2.1 样品的处理

5 年生南方红豆杉种植于中国林业科学研究院科研温室, 利用 200 μmol/L GA₃ 分别在上午 9:00 和下午 15:00 喷施于红豆杉叶片, 处理 2 d 后, 于第 3 天上午 9:00 收集叶片, 液氮速冻后保存在-80 °C 冰箱, 用于后续转录组测序。纯水处理作为对照。样品经过 RNA 提取、cDNA 文库构建质检合格后, 采用 Illumina 测序平台进行转录组测序, 过滤原始数据后, 采用 Trinity 软件进行组装。

2.2 红豆杉基因组 DNA 提取

温室中的红豆杉叶片采集后立刻置于液氮中, 放置-80 °C 冰箱。廊坊基地红豆杉叶片于采集时装入含有硅胶的自封袋中进行干燥。样品通过液氮研磨后用于基因组 DNA 提取。提取方法采用 mCTAB 法 (modified CTAB method) [16], 在 1% 的琼脂糖凝胶电泳检测完整度, 并根据 Nanodrop 测定将 DNA 原液稀释至 50 ng/μL 保存于-20 °C 条件下用于后续实验。

2.3 红豆杉转录组 SSR 位点的筛选及引物设计

使用 MISA 软件对南方红豆杉转录组数据进行 SSR 分析, 搜索 SSR 位点。可搜索单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸, 最小重复数分别设置为 10、6、5、5、5、5 次。采用 Primer 3 (2.3.5 版, 默认参数) 进行 SSR 引物设计。

基于转录组数据开发的一批 SSR 引物中, 选择重复单元碱基数 2~3 bp, 重复单元 7~10 次的 96 对 SSR 引物进行筛选实验。所选择的引物长度在 20~25 bp, 理论退火温度 52~56 °C, PCR 产物长度处于 150~350 bp。

2.4 SSR 引物的 PCR 扩增

PCR 扩增为 2 轮扩增。第 1 轮 (获取相应片段) PCR 扩增反应体系: 双蒸水 (ddH₂O) 5.9 μL, 10×Taq Buffer 1 μL, 2 mmol/L dNTP 1 μL, 5 μmol/L 正向 SSR 引物 0.5 μL, 5 μmol/L 反向 SSR 引物 0.5 μL, TaqDNA 聚合酶 0.1 μL, 50 ng/μL 的 DNA 模板 1 μL。反应程序: 94 °C 预变性 4 min; 20 个循环的 3 步 PCR (94 °C 变性 0.5 min, 退火时间 0.5 min, 退火温度为 52 °C, 72 °C 延伸 1 min); 72 °C 延伸 10 min, 10 °C 保存。第 2 轮 (添加荧光标记) PCR 扩增反应体系: ddH₂O 5.9 μL, 10×Taq Buffer 1 μL, 2 mmol/L dNTP 1 μL, 5 μmol/L 带荧光的 M13 正向接头引物 0.5 μL, 5 μmol/L 反向 SSR 引物 0.5 μL, TaqDNA 聚合酶 0.1 μL, 50 ng/μL 的第 1 轮模板 1 μL。反应程序: 94 °C 预变性 4 min; 30 个循环的 3 步 PCR (94 °C 变性 0.5 min, 退火时间 0.5 min, 退火温度为 50 °C, 72 °C 延伸 1 min); 72 °C 延伸 10 min, 10 °C 保存。

2.5 毛细管电泳检测

将 PCR 产物根据不同荧光标记进行混板, 在生工生物工程 (上海) 股份有限公司通过 3730XL 测序仪检测对 PCR 产物进行毛细管荧光电泳检测。通过毛细管荧光电泳检测, 最终选定多态性好, 扩增成功率高的 SSR 引物用于后续数据分析。

2.6 数据分析

将下机的原始数据 “.fsa” 文件导入到分析软件 GeneMarker 4.0 中进行峰图判读, 具有多态峰图的则表示该引物具有多态性, 记录等位基因大小, 得到等位基因矩阵。根据判读所得等位基因位点矩阵, 通过 GenALEx 来计算每个个体间的遗传距离和每对引物的等位基因数 (number of alleles, N_a)、有效等位基因数 (effective number of allele, N_e)、期望杂合度 (expected heterozygosity, H_e)、观测杂合度 (observed heterozygosity, H_o)、Shannon's 指数 (Shannon's information index, I) 等遗传多样性

数据。利用 PIC-CALC 0.6 软件计算 SSR 位点的多态性信息数据 (polymorphism information content, PIC)。基于 GenALEx 计算所得的个体遗传距离矩阵, 通过 MEGA 进行 UPGMA 聚类分析, 构建聚类树。

3 结果与分析

3.1 转录组 SSR 位点数量与分布

利用 Trinity 拼接从 GA₃ 处理下的南方红豆杉转录组中共获得 202 361 条 Unigene, 通过鉴定, 共获得了 12 008 个 SSR 位点, 出现频率 (SSR 数与总 Unigene 数之比) 为 5.93%, 这些 SSR 分布在 10 894 条 Unigene 中, 发生频率 (含 SSR 的 Unigene 数与总 Unigene 数之比) 为 5.38%, SSR 平均分布距离为 1/7.64 kb。含有 1 个以上 SSR 位点的 Unigene 有 975 条, 包含复合形 SSR 的

Unigene 有 489 条。

南方红豆杉 SSR 类型丰富, 单核苷酸至六核苷酸重复类型均存在。12 008 个 SSR 位点包括 6 213 个单核苷酸 (51.74%)、2 507 个二核苷酸 (20.88%)、2 992 个三核苷酸 (24.92%)、226 个四核苷酸 (1.88%)、27 个五核苷酸 (0.22%)、43 个六核苷酸 (0.36%)。其中单核苷酸最丰富, 其次是三核苷酸和二核苷酸类型, 五核苷酸类型最少。所有 SSR 中以 10 次重复的 SSR 最多, 占 24.48%, 5 次重复占 16.73%, 6 次重复占 15.30%。南方红豆杉转录组 SSR, 基序的重复次数在 5~72, 以 5~10 次重复的 SSR 最多, 共 8 413 个, 占总 SSR 的 70.06%; 其次是 10~14 次重复, 共 2 566 个, 占总 SSR 的 21.37%; 14 次以上的重复较少, 共 1 029 个, 占总 SSR 的 8.57% (表 2)。

表 2 南方红豆杉 SSR 不同基序长度和重复次数的分布

Table 2 Distribution of SSR with different motif types and repeat numbers in *T. chinensis* var. *mairei*

重复类型	SSR 位点重复次数											总计	占比/%
	5	6	7	8	9	10	11	12	13	14	>15		
单核苷酸	0	0	0	0	0	2 733	1 237	591	381	247	1 024	6 213	51.74
二核苷酸	0	1 039	559	346	251	206	99	6	0	0	1	2 507	20.88
三核苷酸	1 780	752	421	36	1	0	1	0	0	0	1	2 992	24.92
四核苷酸	185	34	0	4	1	0	0	0	0	0	2	226	1.88
五核苷酸	19	3	2	0	0	0	0	1	1	0	1	27	0.22
六核苷酸	25	9	3	3	0	1	0	2	0	0	0	43	0.36
总计	2 009	1 837	985	389	253	2 940	1 337	600	382	247	1 029	12 008	
比例/%	16.73	15.30	8.2	3.24	2.11	24.48	11.13	5.00	3.18	2.06	8.57		

3.2 转录组 SSR 基序重复类型和频率特征

研究发现, 在 12 008 个 SSR 位点中包含 114 种重复基元, 其中单、二、三、四、五、六核苷酸重复分别有 2、4、10、29、26、43 种。从出现频率来看 (表 3), 频率较高的重复基元分别为 A/T, 占总位点数的 49.60%。单核苷酸重复类型中, A/T 出现频率明显高于 C/G, 分别占总位点数的 49.60% 和 2.14%; 二核苷酸重复类型中, AG/CT 最多, 约占总位点数的 10.93%, 其次是 AT/AT 和 AC/GT, 分别占总位点数的 5.14%、4.65%, CG/CG 最少, 约占总位点数的 0.17%; 三核苷酸重复类型中, AAG/CTT 最多, 约占总位点数的 5.49%, ACT/AGT 最少, 约占总位点数的 0.44%; 四核苷酸重复类型中, AAAG/CTTT 最多, 约占总位点数的 0.23%, AACG/CGTT、AACT/AGTT、ACCT/AGGT、ATCG/ATCG 最少, 占总位点数不足 0.001%; 五、六核苷酸约占总位点数 0.58%。

表 3 南方红豆杉 SSR 类型分布

Table 3 Distribution of SSR repeat types in *T. chinensis* var. *mairei*

重复类型	基序	重复数	占比/%	合计
单核苷酸	A/T	5 956	49.60	6 213
	G/C	257	2.14	
二核苷酸	AG/CT	1 312	10.93	2 507
	AT/AT	617	5.14	
	AC/GT	558	3.92	
	CG/CG	20	0.17	
三核苷酸	AAG/CTT	659	5.49	2 992
	CCG/CGG	471	3.92	
四核苷酸	ACT/AGT	53	0.44	226
	AAAG/CTTT	28	0.23	
	AAAT/ATTT	21	0.17	
	ATCG/ATCG	1	0.008 3	

3.3 SSR 引物有效性验证

选择重复单元碱基数 2~3 bp, 重复单元 7~10 次的 96 对 SSR 引物进行筛选实验。结果显示, 以我国多个地区的南方红豆杉 DNA 为模板进行 SSR-PCR, 13 对 (表 4) 引物具有多态性, 多态性引物比率为 13.54%。引物多态峰图举例见图 1。

13 个 SSR 位点的多态性信息见表 5。12 个样本的 N_a 为 2~10, 13 个 SSR 位点上总的 N_a 为 55, 平均值为 4.23。 N_e 为 1.53~5.43, 平均值为 2.76。 I 为 0.53~1.85, 平均值为 1.10。 H_o 为 0.14~0.82, 平均值为 0.39。 H_e 为 0.35~0.82, 平均值为 0.58 个。 PIC 为 0.29~0.79, HDS5、HDS14、HDS52、HDS54

和 HDS93 为中多态性位点 ($0.25 < PIC < 0.5$), HDS1、HDS10、HDS56、HDS76、HDS82、HDS84、HDS86 和 HDS88 则为高多态性位点 ($PIC > 0.5$)。其中, HDS 84 的多态性最高, 其 PIC 值为 0.79。

3.4 SSR 位点种间扩增

将上述 13 对具有多态性的 SSR 引物在东北红豆杉和曼地亚红豆杉 DNA 样品中进行扩增, 结果发现有 13 对引物在东北红豆杉 DNA 样品中有扩增, 而在曼地亚红豆杉 DNA 样品中仅有 3 对引物有扩增 (表 6)。两者相比较, 13 对引物在东北红豆杉 DNA 样品中的有效扩增率 (100%), 明显高于在曼地亚红豆杉 DNA 样品中的有效扩增率 (23.08%)。

表 4 SSR 引物信息
Table 4 Information of primers

编号	引物序列 (5'-3')	重复单元	退火温度/℃	荧光标签
HDS1	F: <u>CACGACGTTGTAAAACGACG</u> GAACCTGGCCTCCTTCATC R: GAAGGTGCTGTTCCCGAAG	(GGC) ₇	55	FAM
HDS5	F: <u>CACGACGTTGTAAAACGACTCT</u> CCTGTTGTGAGCATGG R: GTTCTTGCTCCTGCTTTTGC	(CAG) ₇	55	FAM
HDS10	F: <u>CACGACGTTGTAAAACGACGTGG</u> CGTGAAGAAAGTTGGT R: CTCAGCAACAGACACAGGA	(CTG) ₇	55	FAM
HDS14	F: <u>CACGACGTTGTAAAACGACAAC</u> AGAGAGGAAGAACGCCA R: CTTCTTCGACTTTTCCTGCG	(AAG) ₇	55	FAM
HDS52	F: <u>CACGACGTTGTAAAACGACAAC</u> ACAACACCCTCGTCTCC R: GTAAACATCGCCTTCCTTGC	(GAA) ₉	55	VIC
HDS54	F: <u>CACGACGTTGTAAAACGACCAAT</u> TCCGGAAACTTGCTGT R: TTCACCTGCCAATCTAAGGG	(ATA) ₉	55	VIC
HDS56	F: <u>CACGACGTTGTAAAACGACGG</u> AGCTCAGGATCGAGTCAC R: AGAATTGCAGCATTGCACAC	(AG) ₉	54	VIC
HDS76	F: <u>CACGACGTTGTAAAACGACTGCA</u> AATTCCTCCCCATTA R: GCTTCAACATTGCAAAACGA	(GA) ₁₀	55	NED
HDS82	F: <u>CACGACGTTGTAAAACGACATTC</u> ACCGTGACTTCCAAGG R: GTGGATTTCGGTCACTTTGGT	(TA) ₁₀	54	NED
HDS84	F: <u>CACGACGTTGTAAAACGACTGAT</u> GGCTTTTGTGCCACA R: AAGCATAAAAGGCGAGAGCA	(CT) ₁₀	55	NED
HDS86	F: <u>CACGACGTTGTAAAACGACGCT</u> CTCTCTGGATAGGCGTG R: CGTTTCAAGAAAAACGGGAA	(AC) ₁₀	55	NED
HDS88	F: <u>CACGACGTTGTAAAACGACTGCC</u> ATCTGCTCTATTGGATT R: TGCCCAATAAACTATCATCTCC	(TC) ₁₀	53	NED
HDS93	F: <u>CACGACGTTGTAAAACGACCCAT</u> TTCCGTCGATAAAGGA R: GTGCACCACCACACCTACAC	(AT) ₉	54	NED

下划线部分为 M13 接头序列。
The underlined part is the sequence of M13 connectors.

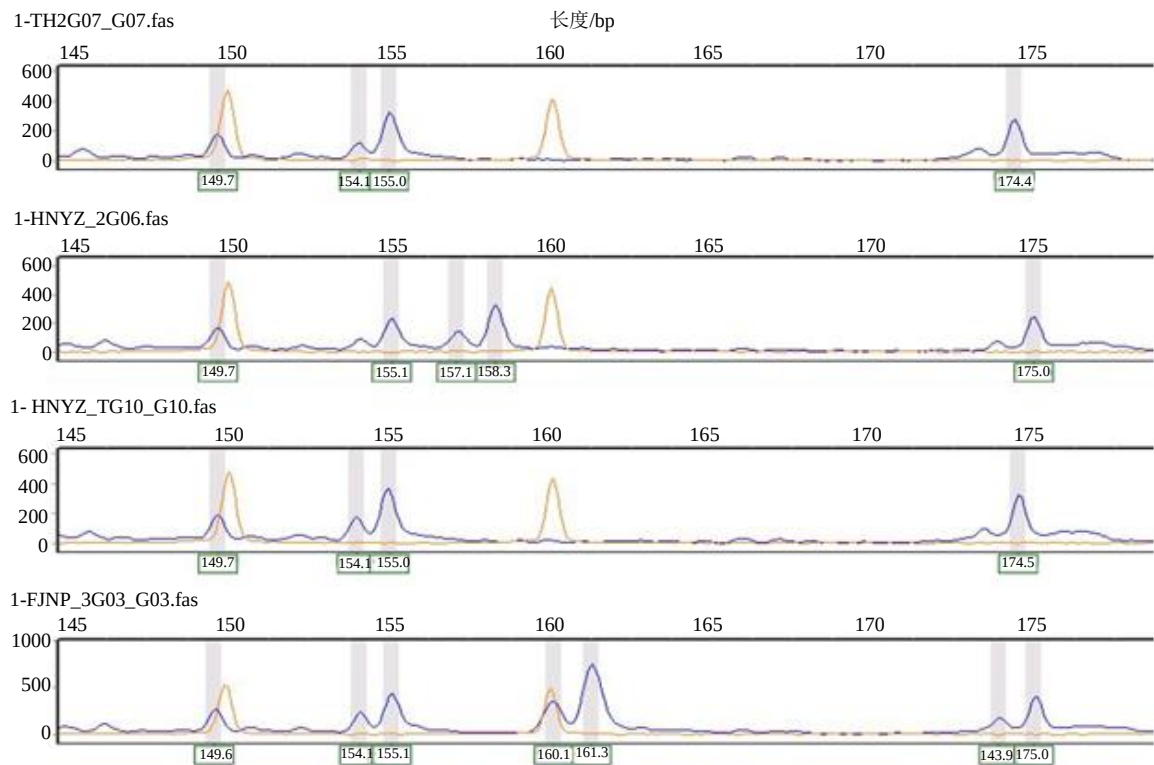


图 1 HDS10 标记在太行山、湖南和福建地区红豆杉样本中的基因分型

Fig. 1 Genotyping by HDS10 marker for *T. chinensis* var. *mairei* from Taihang mountain, Hunan, and Fujian Provinces

表 5 SSR 引物位点多样性

Table 5 SSR loci diversity of 13 SSR primers						
基因位点	N_a	N_e	I	H_o	H_e	PIC
HDS 1	4	3.38	1.30	0.43	0.70	0.65
HDS 5	3	1.94	0.79	0.50	0.49	0.41
HDS 10	5	2.42	1.13	0.33	0.59	0.54
HDS 14	2	1.53	0.53	0.44	0.35	0.29
HDS 52	2	1.54	0.54	0.09	0.35	0.29
HDS 54	2	1.80	0.64	0.22	0.44	0.35
HDS 56	6	3.92	1.55	0.40	0.75	0.71
HDS 76	8	4.26	1.77	0.57	0.77	0.74
HDS 82	4	2.39	1.05	0.14	0.58	0.52
HDS 84	8	5.43	1.85	0.75	0.82	0.79
HDS 86	4	2.69	1.17	0.36	0.63	0.58
HDS 88	3	2.78	1.06	0.45	0.64	0.57
HDS 93	4	1.85	0.90	0.43	0.46	0.43
平均值	4.23	2.76	1.10	0.39	0.58	0.53

3.5 聚类分析

UPGMA 聚类分析表明，来自浙江、福建和湖南纬度较低地区的南方红豆杉聚类在一起，而来自太行山地区的 4 份样品聚为一类(图 2)。可以看出，浙江、福建和湖南地区的南方红豆杉亲缘关系更近，而太行山地区的南方红豆杉与上述纬度更低南方来源的南方红豆杉亲缘关系较远。此外，在 3 种纬度

表 6 种间扩增结果统计

Table 6 Statistics of interspecific amplification results				
编号	扩增数			
	东北红豆杉	曼地亚红豆杉		
HDS1	170	170	0	0
HDS5	260	266	272	272
HDS10	155	155	155	155
HDS14	161	161	161	161
HDS52	225	225	0	0
HDS54	277	277	0	0
HDS56	266	266	0	0
HDS76	128	130	0	0
HDS82	210	210	0	0
HDS84	210	212	0	0
HDS85	189	189	0	0
HDS86	291	307	0	0
HDS88	180	180	0	0
HDS93	221	221	0	0

较低的南方地区的南方红豆杉中，也可以看出，浙江地区和湖南地区的距离相较福建地区更近。

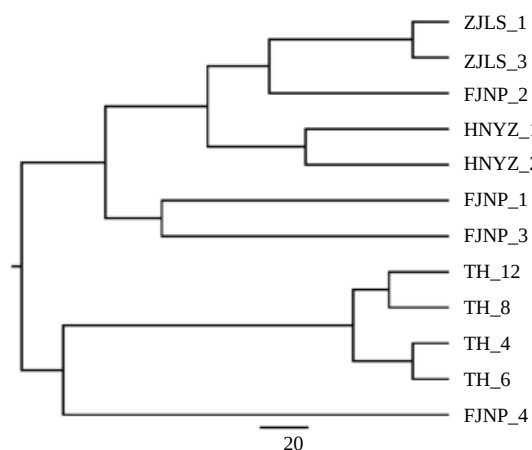


图 2 12 份南方红豆杉聚类分析

Fig. 2 Cluster analysis of 12 samples of *T. chinensis* var. *mairei*

4 讨论

随着科技的飞速发展,高通量测序技术的出现以及更新迭代,使得利用基因组和转录组数据鉴定获取 SSR 标记更加方便,数量更加丰富,更加有助于进行植物种质资源鉴定及遗传多样性评价等方面研究。目前,已经有多种植物基于基因组和转录组数据挖掘鉴定出 SSR 标记^[7,17-19]。相较于基因组 SSR,转录组 EST-SSR 标记能更准确地鉴别植物基因型,在种间通用性更好,能准确评估种间亲缘关系^[20-23]。在南方红豆杉中,也有研究人员通过转录组数据开发鉴定出了多个 SSR 标记,并用于连锁图谱构建等研究^[4,7]。但是,已开发的 SSR 标记还不能满足连锁图谱构建和遗传分析的需要,红豆杉 SSR 分子标记相关研究仍需要进一步加强。

本研究利用 GA₃ 处理的南方红豆杉叶片转录组数据搜索到二核苷酸至六核苷酸重复类型的 SSR 位点有 6 213 个,高于李炎林等^[10]利用公共数据库中南方红豆杉根、茎、叶转录组数据搜索到的 SSR 位点数(2 160 个)。差异可能是与转录组数据来源、MISA 软件检索设置参数等不同导致,也有可能与 GA₃ 处理有关。研究表明,二、三核苷酸重复类型为植物的 SSR 位点的主要类型^[24-26]。易官美等^[9]研究报道在南方红豆杉中主要以二核苷酸重复类型为主,而李炎林等^[10]报道则是以三核苷酸重复类型为主。此外,在武星彤^[5]的研究中,南方红豆杉 SSR 重复基序是以六核苷酸重复为主,其次为三核苷酸。本研究南方红豆杉转录组数据分析结果显示,三核苷酸重复基序比二核苷酸重复基序更丰富,与李炎

林等^[10]结果更为相似。分析原因,可能与南方红豆杉的产地不同以及转录组数据和软件设置参数等有关。从优势重复基序来看,单核苷酸至五核苷酸分别为 A/T、AG/CT、AAG/CTT、AAAG/CTTT、ATCCG/ATCCG。二、三核苷酸的优势重复基序与前人结果较为一致^[5,9-10]。

遗传多样性能够体现生物遗传信息变异的程度,群体遗传多样性主要由等位基因数、杂合度和多态信息含量等 3 个方面体现^[27]。本研究中 13 对 SSR 引物共有 55 个位点具有多态,在 12 个样品中的 H_o 平均值低于 H_e ,与武星彤的结果相似,而 PIC 平均值为 0.53,则高于武星彤^[5]报道的 7 个省份的南方红豆杉 PIC 数值 0.453,表明本研究所检测的 4 个南方红豆杉群体的遗传多样性要高于武星彤所采用的南方红豆杉群体。究其原因,推测是由于本研究中所检测的太行山南方红豆杉属于目前发现的最北端的南方红豆杉,其与福建、浙江、广西等南方地区的生长的南方红豆杉生长环境且表型特征都具有较大不同。此外,研究中所涉及的 SSR 位点不同也有可能造成结果的差异。尽管本研究和武星彤^[5]研究中采用的群体以及样品数量不多,但是两者结果高度相似,也能充分表明当前人为因素对南方红豆杉的遗传多样性产生不可忽略的影响。

种间扩增法被认为能有效开发近源物种新 SSR 位点。武星彤开发的 107 个 SSR 位点在 5 种红豆杉种间平均扩增率为 96.07%^[5]。本研究获得的 13 对引物在东北红豆杉 DNA 样品中扩增率达 100%,但在曼地亚红豆杉 DNA 样品中扩增率仅达 23.08%。标记通用性的高低一定程度上反映物种基因组差异的大小和亲缘关系的远近^[28-29]。本研究的结果说明相较于曼地亚红豆杉,南方红豆杉与东北红豆杉亲缘关系更近,基因组差异更小。分析原因可能有 2 个:一是因为曼地亚红豆杉是欧洲红豆杉 *T. baccata* 和东北红豆杉的天然杂交种^[30],遗传背景复杂;二是因为曼地亚红豆杉不是我国本土树种,是从外国引种而来,与南方红豆杉的亲缘关系、自然分布距离远。因此,本研究开发的 SSR 引物可用于红豆杉种间亲缘关系鉴定、系统发育等研究。

研究中所获得 13 对 SSR 多态性引物,将 12 个不同地区的南方红豆杉进行了很好的聚类,聚类结果与其实际来源地较为一致。来自太行山的 4 个样品明显聚为一类,与其他南方地区的红豆杉可以得以显著区分。可见,本研究所开发的 SSR 引物能够

用于后续南方红豆杉遗传多样性分析及遗传结构分析等研究。

综上所述, 本研究结果表明通过南方红豆杉转录组数据开发 SSR 标记切实可行, 鉴定筛选出的 13 对 SSR 引物为南方红豆杉的遗传多样性分析提供了候选标记。随着红豆杉基因组的发表^[31-32]以及更多转录组数据的公布, 将会有更多的与功能基因相关的 SSR 位点被开发, 例如与紫杉醇生物合成相关的 SSR 位点, 希望今后研究人员可以基于此, 在今后的研究中充分利用 SSR 分子标记技术, 并对红豆杉的群体遗传、遗传图谱构建和谱系地理等诸多方面进行更加深入的研究, 这些都将为未来的红豆杉分子育种提供帮助。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Yang L, Ruan X A, Gao X X, *et al.* Morphological characteristics, geographical distribution, secondary metabolites, and biological activities of *Taxus* [J]. *Mini Rev Org Chem*, 2017, 14(3): 217-226.
- [2] 王亚飞, 王强, 阮晓, 等. 红豆杉属植物资源的研究现状与开发利用对策 [J]. *林业科学*, 2012, 48(5): 116-125.
- [3] Ueno S, Moriguchi Y, Uchiyama K, *et al.* A second generation framework for the analysis of microsatellites in expressed sequence tags and the development of EST-SSR markers for a conifer, *Cryptomeria japonica* [J]. *BMC Genomics*, 2012, 13: 136.
- [4] Song X M, Yang Q H, Bai Y, *et al.* Comprehensive analysis of SSRs and database construction using all complete gene-coding sequences in major horticultural and representative plants [J]. *Hortic Res*, 2021, 8(1): 122.
- [5] 武星彤. 南方红豆杉 EST-SSR 标记的开发及其在父本分析中的应用 [D]. 长沙: 中南林业科技大学, 2017.
- [6] Qin H T, Yang G Q, Provan J, *et al.* Using MiddRAD-seq data to develop polymorphic microsatellite markers for an endangered yew species [J]. *Plant Divers*, 2017, 39(5): 294-299.
- [7] Majeed A, Singh A, Choudhary S, *et al.* Transcriptome characterization and development of functional polymorphic SSR marker resource for Himalayan endangered species, *Taxus contorta* (Griff) [J]. *Ind Crops Prod*, 2019, 140: 111600.
- [8] Wu X, Wen Y, Ueno S, *et al.* Development and characterization of EST-SSR markers for *Taxus mairei* (Taxaceae) and their transferability across species [J]. *Silvae Genet*, 2016, 65(1): 67-70.
- [9] 易官美, 黎建辉, 王冬梅, 等. 南方红豆杉 SSR 分布特征分析及分子标记的开发 [J]. *园艺学报*, 2013, 40(3): 571-578.
- [10] 李炎林, 杨星星, 张家银, 等. 南方红豆杉转录组 SSR 挖掘及分子标记的研究 [J]. *园艺学报*, 2014, 41(4): 735-745.
- [11] 吴琼, 段小群, 陈旭, 等. 基于高通量测序的红豆杉 EST-SSRs 标记研究 [J]. *中国中药杂志*, 2012, 37(24): 3728-3733.
- [12] Chybicki I J, Oleksa A, Burczyk J. Increased inbreeding and strong kinship structure in *Taxus baccata* estimated from both AFLP and SSR data [J]. *Heredity*, 2011, 107(6): 589-600.
- [13] 缪迎春. 濒危植物云南红豆杉保护遗传学研究 [D]. 昆明: 云南大学, 2013.
- [14] 李娟花. 红豆杉内生真菌代谢产紫杉烷类物质的研究 [D]. 汉中: 陕西理工学院, 2011.
- [15] Jiang L Y, Zhang K K, Lü X, *et al.* Characterization and expression analysis of genes encoding Taxol biosynthetic enzymes in *Taxus* spp. [J]. *J For Res*, 2021, 32(6): 2507-2515.
- [16] 李金璐, 王硕, 于婧, 等. 一种改良的植物 DNA 提取方法 [J]. *植物学报*, 2013, 48(1): 72-78.
- [17] Zhao T M, Zheng S G, Hu Y D, *et al.* Classification of interspecific and intraspecific species by genome-wide SSR markers on *Dendrobium* [J]. *S Afr N J Bot*, 2019, 127: 136-146.
- [18] Xiang X Y, Zhang Z X, Wang Z G, *et al.* Transcriptome sequencing and development of EST-SSR markers in *Pinus dabeshanensis*, an endangered conifer endemic to China [J]. *Mol Breeding*, 2015, 35(8): 158.
- [19] 张淑文, 梁森苗, 郑锡良, 等. 杨梅基因组 SSR 引物的开发与应用 [J]. *园艺学报*, 2019, 46(1): 149-156.
- [20] 刘果, 谢耀坚, 陈鸿鹏, 等. 桉树 Genomic-SSR 和 EST-SSR 分子标记的遗传差异性分析 [J]. *桉树科技*, 2017, 34(3): 1-8.
- [21] 宋跃朋, 江锡兵, 张曼, 等. 杨树 Genomic-SSR 与 EST-SSR 分子标记遗传差异性分析 [J]. *北京林业大学学报*, 2010, 32(5): 1-7.
- [22] 陈和明, 肖文芳, 吕复兵. 蝴蝶兰 genomic-SSR 与 EST-SSR 分子标记遗传差异性分析 [A] // 中国园艺学会 2015 年学术年会论文摘要集 [C]. 厦门: 中国园艺学会, 2015: 215.
- [23] 张亚东, 彭婵, 李振芳, 等. 基因组 SSR 与 EST-SSR 标记在杨树不同种间的遗传差异 [J]. *东北林业大学学报*, 2011, 39(12): 8-11.
- [24] 郁永明, 田丹青, 潘晓韵, 等. 基于红掌转录组序列

- 的 SSR 标记分析与开发 [J]. 分子植物育种, 2015, 13(6): 1349-1354.
- [25] 郭俊, 朱婕, 谢尚潜, 等. 油梨转录组 SSR 分子标记开发与种质资源亲缘关系分析 [J]. 园艺学报, 2020, 47(8): 1552-1564.
- [26] 徐海滨, 钱俊, 汪波, 等. 丹参基因组 SSR 位点的特征分析及引物开发 [J]. 世界科学技术—中医药现代化, 2013, 15(3): 367-370.
- [27] 陈友吾, 廖荣俊, 叶碧欢, 等. 多花黄精转录组 SSR 位点分析及分子标记开发 [J]. 中草药, 2020, 51(1): 182-189.
- [28] 倪州献, 白天道, 蔡恒, 等. 马尾松基因组 SSR 标记在松属其他树种中的通用性分析 [J]. 分子植物育种, 2015, 13(12): 2811-2817.
- [29] 蔡年辉, 许玉兰, 王大玮, 等. 思茅松 EST-SSR 标记在几种松属植物中的通用性分析 [J]. 西南林业大学学报: 自然科学, 2017, 37(4): 8-14.
- [30] 汪群. 曼地亚红豆杉年生长节律研究 [J]. 安徽农学通报, 2016, 22(16): 92-94.
- [31] Xiong X Y, Gou J B, Liao Q G, *et al.* The *Taxus* genome provides insights into paclitaxel biosynthesis [J]. *Nat Plants*, 2021, 7(8): 1026-1036.
- [32] Cheng J, Wang X, Liu X N, *et al.* Chromosome-level genome of Himalayan yew provides insights into the origin and evolution of the paclitaxel biosynthetic pathway [J]. *Molecular Plant*, 2021, 21: 00160-X.

[责任编辑 时圣明]