

基于桉木属转录组测序的 SSR 分子标记的开发

饶龙兵¹, 杨汉波¹, 郭洪英^{2*}, 段红平³, 陈益泰¹

(1. 中国林业科学研究院亚热带林业研究所, 浙江 杭州 311400; 2. 四川省林业科学研究院, 四川 成都 610081;
3. 云南农业大学资源与环境学院, 云南 昆明 650201)

摘要: [目的] 基于转录组数据开发适用于桉木属树种的 SSR 标记, 揭示其在转录组序列中的分布类型及特征, 为桉木属树种分子标记辅助育种提供有利工具。 [方法] 利用 MicroSATellite (MISA) 软件对所有转录组序列进行 SSRs 搜索, 并对 SSR 位点的数量、分布特征进行统计分析。设计 100 对 SSR 引物, 采用琼脂糖凝胶电泳和毛细管电泳分离检测方法对 3 种不同倍性桉木属植物 (12 份材料) 进行遗传多样性检测, 确定引物多态性及通用性。 [结果] 85 769 条 Unigenes 序列中发现 8 678 个 SSR 位点, 分布在 8 298 条 Unigenes 中发生频率为 9. 67%, 转录组序列中平均每 14. 04 kb 长度就有一个 SSR 位点分布。其中, 二核苷酸重复类型数量最多, 占 65. 87%。根据转录组 Unigenes 序列, 利用 Primer 3 软件共设计出 4 531 对符合要求的引物, 挑选出的 100 对 SSR 引物中, 获得 18 对多态性高、稳定性好的 SSR 引物。 [结论] 本研究可扩增出多态性位点的引物重复单元以二、三核苷酸重复为主。基于桉木属转录组序列的 SSR 标记开发是可行的, 开发的引物为桉木属遗传多样性分析、分子育种、遗传图谱构建和功能基因的挖掘提供了丰富的候选分子标记。

关键词: 桉木属; 转录组; SSR; 引物

中图分类号: S718. 46

文献标识码: A

Development of SSR Molecular Markers Based on Transcriptome Sequences of *Alnus*

RAO Long-bing¹, YANG Han-bo¹, GUO Hong-ying², DUAN Hong-ping³, CHEN Yi-tai¹

(1. Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Hangzhou 311400, Zhejiang, China; 2. Sichuan Academy of Forestry, Chengdu 610081, Sichuan, China; 3. College of Resources and Environment, Yunnan Agricultural University, Kunming 650201, Yunnan, China)

Abstract: [Objective] To develop SSR markers of gene transcriptional areas from *Alnus* based on the results of transcriptome sequencing. [Methods] Distribution patterns of the markers in the transcriptome sequences and their characteristics were analyzed, in order to provide more powerful tools for molecular marker-assisted breeding *Alnus*. The SSR locus from transcriptome sequences were searched by MicroSATellite (MISA), and statistical analyses were conducted for the amount, distribution and characteristics of SSR loci. 100 pairs of SSR primers were designed and synthesized. Agarose electrophoresis was used for initial check and capillary electrophoresis for separation and detection of the polymorphism of the primers. [Results] A total of 8 298 Unigenes containing 8 678 SSR locus were searched from 85 769 Unigenes by MicroSATellite (MISA) sequence analysis software, accounting for 9. 67% of the transcriptome sequences, with an average of one SSR per 14. 04kb. The dinucleotide repeat is the most abundant repeat type, accounting for 65. 87% of the total number of SSRs. Additionally, a small amount of GC repeats was found. A total of 4 531 pairs of primers were designed by Primer 3 software according to the Unigene sequences of

收稿日期: 2016-03-25

基金项目: “十二五”国家科技支撑计划项目桉木育种专题“抗逆生态树种桉木新品种选育技术研究”(2012BAD01B0604)。

作者简介: 饶龙兵(1973—), 男, 湖北天门人, 助理研究员, 博士, 主要从事植物遗传育种研究。

* 通讯作者: 郭洪英(1973—), 女, 四川越西人, 研究员, 博士, 主要从事林木遗传育种研究。 ghy0607@163.com.

transcriptional. 18 pairs of primers produced bands with expected sizes were selected from 100 pairs of primers. [Conclusion] The amplified primers of the polymorphism loci were mainly dinucleotide and trinucleotide repeats. This study has an important value to develop SSR molecular markers based on the transcriptome sequencing analysis. This study is also important for analyzing genetic diversity, marker assisted selection, genetic linkage mapping and functional gene mining of *Alnus* by using SSR molecular markers.

Keywords: *Alnus*; transcriptome; SSR; primer

桤木属 (*Alnus* Mill.) 为桦木科 (Betulaceae) 植物, 主要分布于北半球寒温带、温带和亚热带地区, 美洲最南达秘鲁。自 20 世纪 80 年代桤木属树种在工业生产上的发现与应用, 该属树种在生态、经济及遗传改良等方面的价值引起研究人员的极大关注^[1-2]。桤木属树种根系发达, 可固沙保土, 大多数树种含根瘤, 能固定空气中游离的氮素, 增加土壤肥力^[3]。通过近几十年来的不断努力, 研究者在桤木属树种形态学、细胞学和生理生化等研究中投入了大量的精力, 积累了宝贵的成果^[4-6]。随着分子生物学的发展, 对桤木属树种的研究热点逐渐转入分子水平, RAPD、AFLP、ISSR 等现代分子生物学手段为桤木属遗传多样性、种质创新和品种改良等研究带来很大便利^[7-9]。

简单重复序列 (simple sequence repeat, SSR), 又称微卫星, 是广泛存在于真核生物和部分原核生物基因组中的 1~6 个核苷酸串联重复单元^[10-11]。相对于其它标记, SSR 标记具有多态性高、重复性高、共显性、覆盖面积广等优点, 在 DNA 指纹图谱的构建、遗传多样性分析、基因定位、分子标记辅助育种等方面得到广泛应用^[12-13]。传统微卫星标记开发以基因文库构建法 (包括 SSR 富集文库) 为主, 其实验过程繁杂、费时费力、效率较低 (Zane 等, 2002)^[14]。随着高通量测序技术的快速发展和测序成本的降低, 利用新一代高通量测序技术对植物进行转录组测序, 并产生丰富的转录组数据, 其中包含了大量的 SSR 序列^[15]。通过转录组信息开发 SSR 标记的方法已广泛应用于非模式植物的研究, 在南方红豆杉 (*Taxus chinensis* (Pilger) Rehd. var. *mairei* (Lemee et Levl.) Cheng et L. K. Fu)^[16]、杜仲 (*Eucommia ulmoides* Oliver)^[17]、蜡梅 (*Chimonanthus praecox* (Linn.) Link)^[18]、红松 (*Pinus koraiensis* Sieb. et Zucc.)^[19] 等多年生木本植物中均有报道。关于桤木属植物 SSR 标记, Drašnarová 等^[20] 开发了 17 对具有多态性的 SSR 引物; Lepais 等^[21] 采用焦磷酸测序法开发 12 对多态性的 SSR 引物, 但桤木属树

种为多倍性木本植物, 具有复杂和丰富的遗传背景, 这些引物还难以满足研究需要。为此, 本研究利用高通量测序技术成功获得不同倍性桤木属树种混合样本的转录组数据, 并对这些数据进行了系统分析, 最终找到了适用于桤木属树种 SSR 标记研究的引物, 为进行桤木属树种遗传多样性分析、分子标记辅助育种等提供丰富的候选标记。

1 材料与方法

1.1 试验材料

转录组测序样本采自中国林业科学研究院亚热带林业研究所试验大棚欧洲桤木 (*A. glutinosa* (L.) Gaertn., 2n=28)、四川桤木 (*A. cremastogyne* Burk., 2n=56)、硬桤木 (*A. firma* Sieb. et Zucc., 2n=112) 的幼嫩新叶, 液氮速冻后送华大基因公司 (北京) 进行转录组测序, 共包含 85 769 条 Unigenes。

SSR 引物筛选所用的 12 份不同倍性桤木属植物样品于 2015 年 7 月取自中国林业科学研究院亚热带林业研究所试验大棚 (表 1)。样品取健壮、无病虫害的当年生嫩叶, 液氮速冻后保存于 -80℃ 冰箱中备用。

表 1 引物筛选试验材料

树种	倍性	数量	编号	引种地区	引种年份
欧洲桤木	28	4	AGL201005、AGL201011、 AGL201026、AGL201047	捷克	2010
四川桤木	56	4	ACR201007、ACR201015、 ACR201025、ACR201033	中国	2010
硬桤木	112	4	AFI201203、AFI201209、 AFI201218、AFI201221	日本	2012

1.2 试验方法

1.2.1 RNA 提取和转录组测序 采用 Trizol 法提取桤木属植物材料总 RNA, 0.8% 琼脂糖凝胶电泳检测 DNA 纯度和完整性, NanoDrop-2000 超微量分光光度计 (Thermo, 美国) 检测其浓度。利用 Oligo (dT) 吸附纯化 mRNA, 将 mRNA 打断成短片段 (Thermomixer), 以打断后的 mRNA 为模板合成第一链 cDNA, 然后配制二链合成反应体系合成二链 cD-

NA,并使用试剂盒纯化回收、粘性末端修复,获 cDNA 文库;在双链 cDNA 片段的 3' 末端加上碱基“A”并连接接头,并进行片段大小选择,最后进行 PCR 扩增得到测序文库,利用 ABI StepOnePlus Real-Time PCR System 和 Agilent 2100 Bioanalyzer 对构建好的测序文库进行质检,质检合格后,采用 Illumina HiSeq™2000 测序仪对构建的文库进行测序。

1.2.2 SSR 位点的鉴别及 SSR 引物设计 对柾木属转录组序列的 85 769 条 Unigenes 利用 SSR 软件 MicroSatellite (MISA) 查找 SSR 位点,查找标准:核苷酸的最小重复次数为 5 次,并且 SSR 位点侧翼序列长度 ≥ 50 bp。

用 Primer 3 软件进行引物设计和评价。设计引物时设置的主要参数为:GC 含量 40%~70%,退火温度 55~65℃,引物长 18~28 bp,预期扩增产物长度 80~300 bp,且无二级结构和二聚体。引物命名为 AR 加序号,如 AR1。

1.2.3 SSR 引物筛选 采用鼎国生物公司植物基因组提取试剂盒提取基因组 DNA,用 1% 的琼脂糖凝胶电泳检测 DNA 的质量,Nanodrop2000 测定 DAN 浓度,稀释至 $25 \text{ ng} \cdot \mu\text{L}^{-1}$, -20°C 保存备用。

挑选重复单元重复数较高的 100 对引物由生工生物技术公司(上海)合成以用于扩增, $2 \times \text{Taq PCR MasterMix}$ 购自 TAKARA 生物公司(大连)。PCR 反应体系为 $20 \mu\text{L}$:DNA 模板 25 ng , $2 \times \text{Taq PCR MasterMix}$ $9 \mu\text{L}$,上下游引物浓度 $0.4 \mu\text{mol} \cdot \text{L}^{-1}$,去离子水补足 $20 \mu\text{L}$ 。PCR 反应在 Takara PCR Thermal Cycler Dice Touch 上进行扩增: 94°C 预变性 4min;再进行 35 个循环,每个循环包括 94°C 变性 30 s, 60°C 复性 30 s, 72°C 延伸 1 min;最后 70°C 延伸 10 min。PCR 扩增产物首先用 2% 的琼脂糖凝胶电泳检测有无特异性条带,有特异性条带的再用 Qsep100™ 进行毛细管电泳分离检测和片段大小的测定。

1.2.4 数据分析 对 SSR 检测结果进行峰图及等位基因的读取,然后采用 Cervus 3.0.7 软件^[22]计算多态信息含量(PIC)。利用 TASSEL 2.1 软件^[23]衡量 SSR 标记两两间的连锁不平衡水平。利用 GenAlEx 6.502 软件^[24]进行遗传多样性分析;分别计算等位基因数(Na)、有效等位基因数(Ne)、固定指数(Fis),并进行哈迪温伯格平衡检测。GenAlEx 6.502 软件计算群体间的 Nei's 遗传相似系数和遗传距离,并利用 NTSYS pc 2.0 软件^[25]绘制基于 UPGMA 法的树状聚类图。

2 结果与分析

2.1 SSR 位点的分布特点

在柾木属转录组 85 769 条 Unigenes 序列中发现 8 678 个 SSR 位点,分布在 8 298 条 Unigenes 中,发生频率(含有 SSR 的 Unigenes 数量与总 Unigenes 数量之比)为 9.67%。其中有 6 971 条 Unigenes 序列中只含 1 个 SSR 位点,含 2 个及 2 个以上 SSR 位点的 Unigenes 序列有 1 327 条,SSR 的分布频率(SSR 的个数与总 Unigenes 的数量比)为 10.13%,柾木属转录组序列中平均 14.04 kb 就能发现一个 SSR 位点(表 2)。

柾木属转录组中 SSR 的主要重复类型是二核苷酸重复,占 SSR 总数的 65.87%;其次是三核苷酸重复,占 SSR 总数的 28.36%;四、五、六核苷酸重复类型的数量很少,总计 5.77%(图 1)。重复序列长度 12~17 bp 的最多,占 60.99%;其次是长度 18~24 bp 的序列重复,占 38.55%;大于 40 bp 的重复序列仅 0.02%(图 2)。SSR 重复单元的重复次数分布在 5~13 次之间,其中 6 次重复的最多,有 2 393 个 SSR,占 28.74%;其次为 7 和 5 次重复,分别有 1 573、1 498 个 SSR,分别占 18.89% 和 17.99%;12 次以上重复的仅有 13 个 SSR,占 0.15%(图 3)。

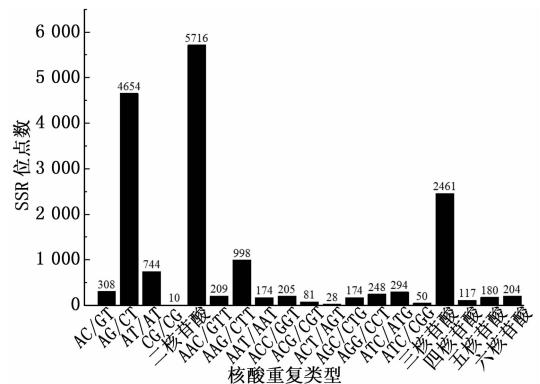


图 1 SSR 不同重复单元

2.2 SSR 分布特征

8 678 个 SSR 位点共包含 149 种重复单元,二、三、四、五、六核苷酸重复各有 12、63、48、11、15 种。出现频率最高的重复类型是 AG/CT,占总 SSR 位点的 53.63%;其次是 AAG/CTT 和 AT/AT,分别占总数 SSR 位点的 11.50% 和 8.57%。在二核苷酸重复中,AG/CT 出现的次数最多,占二核苷酸 SSR 位点的 81.42%;在三核苷酸重复中,AAG/CTT 出现的次数最多,占三核苷酸 SSR 位点的 40.55%;同时在二

表2 桉木转录组中 SSR 重复单元的分布特征

重复类型	数量	频率/ %	平均距 离/kb	平均长 度/bp	主要重复单元
二核苷酸	5 716	6.67	1.42	15	AC/GT, AG/CT, AT/AT, CG/CG AAC/GTT, AAG/CTT, AAT/ATT, ACC/GGT, ACG/CGT, ACT/AGT, AGC/CTG, AGG/CCT, ATC/ATG, CCG/CGG
三核苷酸	2 461	2.87	2.24	11	AAAC/GTTT, AAAG/CTTT, AAAT/ ATTT, AACG/GGTT, AAGC/CTTG, AAGG/CCTT, AATC/ATTG, AATG/ ATTC, AATT/AATT, ACAG/CTGT, ACAT/ATGT, ACGC/CGTG, ACTC/ AGTG, AGAT/ATCT, AGCC/CTGG, AGCG/CGCT, AGCT/AGCT, AGGC/ CCTG, AGGG/CGCT, ATCC/ATGG, ATCG/ATCG
四核苷酸	117	0.14	30.71	10	AAAAT/ATTTT, AAAC/AGTTT, AAAG/CCTTT, AAAT/ATTTG, AACTC/AGTTG, AAAAC/GTTT, AAGCT/AGCTT, AAGTG/ACTTC, ACCTC/AGGTG, ACTGG/AGTCC, AGCCC/CTGGG
五核苷酸	180	0.21	19.96	11	AACGCC/CGTTGG, AATCCG/AT- TCCG, AATGGC/ATTGCC, AAT- TCC/AATTGG, ACATAT/ATATGT, ACCACT/AGTGGT, ACCGAG/CG- GTCT, ACCGGC/CCGGTG, AGC- CCG/CGGGCT, AGCCGC/CG- GCTG, AGCCTG/AGGCTC, AG- GATG/ATCCTC, AGGGCC/CCT- GCG, ATCATG/ATCATG
六核苷酸	204	0.24	15.87	12	
合计	8 678	-	-	-	
均值	-	2.03	14.04	12	

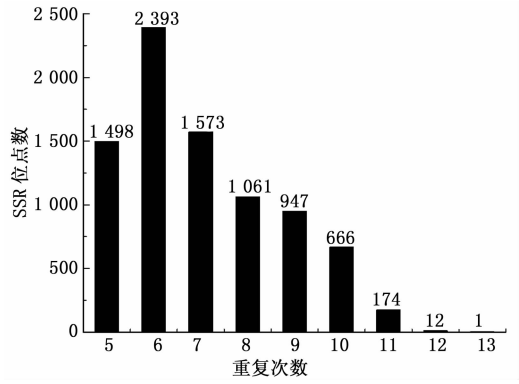
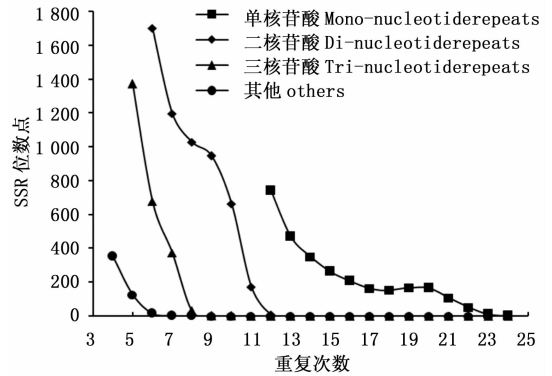


图3 SSR 重复次数分布

次、其它碱基重复次数达到9次时, SSR 的下降速率降低, 进入平台期。



其它:表示四、五、六碱基重复 SSR 的总和

图4 桉木转录组中 SSR 数量随重复次数的变化曲线

2.3 SSR 引物的有效性检测

对含 SSR 位点的转录组序列进行引物设计, 挑选 100 对符合筛选条件的引物进行合成, SSR 位点包括 2~6 核苷酸重复单元。选用 3 种不同倍性的桉木属植物(12 份材料)基因组为模板进行引物筛选。100 对引物中有 39 对引物扩增出特异性条带, 扩增效率 39%; 其中有 2 对引物扩增片段与预期的扩增产物片段大小不符, 剩余 37 对引物中有 18 对引物的扩增产物具有多态性, 占设计引物总数 18%, 18 对 SSR 引物的信息及遗传参数见表 3、表 4。利用 R^2 评价 SSR 位点的连锁不平衡(LD)情况, 结果显示大部分 SSR 位点间不存在连锁, 处于完全独立的状态($R^2 < 0.3$), 但也发现部分位点间(AR5 与 AR62、AR76)具有 LD 水平($R^2 \geq 0.3$)的, 但不显著。引物 AR5、AR8、AR36 和 AR41 在 12 份材料之间的扩增效果见图 6, 4 对引物的目的片段长度分别在 145~170、130~157、102~220、110~130 bp 之间。通过聚类分析, 18 对引物将 12 个不同桉木

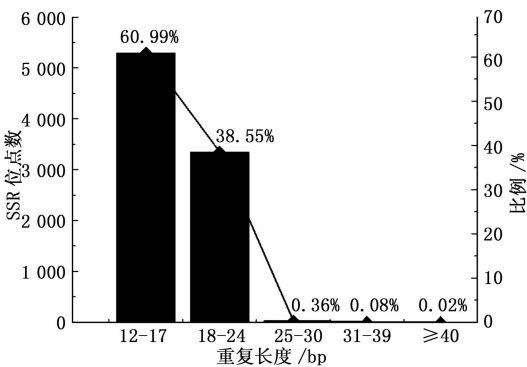


图2 SSR 重复长度分布

核苷酸重复中还发现了少量的 CG 重复, 占二核苷酸 SSR 位点的 0.17% (图 1)。

图 4 表明, SSR 数量随着重复次数增加呈明显下降趋势, 而且碱基重复次数越少, SSR 数量下降的速率就越快。单核苷酸重复次数达到 23 次、二核苷酸重复次数达到 13 次、三核苷酸重复次数达到 11

材料按物种水平分别聚在一起(图7),这与形态学分类相符合,说明这18对SSR引物可以用于桉木属

植物遗传多样性分析和亲缘关系研究中,基于桉木转录组序列的SSR标记开发是可行的。

表3 18对SSR引物信息

引物编号	重复单元	引物序列(5' - 3')	退火温度/°C	预期片段大小/bp	片段范围/bp
AR5	(AT) ₈	F: TTTGTGGAGATGTGTACGTGAAG R: TAGAGGGGAAAAGGATAAAAACCA	60	147	145 ~ 170
AR8	(TC) ₈	F: CTTTTAAGCTCCAATCCTTTCC R: GTCAGATCCATTCAGGAAGTGAG	60	135	130 ~ 157
AR12	(ATC) ₈	F: GAAACAATCTCTAGAATTCCCGC R: GATGATCTGAGCAATGATGATGA	60	113	100 ~ 137
AR32	(ATG) ₇	F: CATCAAAATATAATGCGGGGTT R: CAAGTGCTTTGGCTCAGTAGAAT	60	160	145 ~ 168
AR36	(TA) ₁₀	F: AGTTCTAGCATGGTGTCTGTGATT R: TGTCTTCCCAGCATTACATAAG	60	144	102 ~ 220
AR37	(GA) ₁₀	F: TCAGCTTTGGACAGAGAAAAGAAAT R: ACAAGGCAGTAGATTCAACATGG	60	138	130 ~ 174
AR41	(ACA) ₇	F: CACCTGGGTCTCTATGTATGGT R: TTAACCCTTGCCTTCAACACTA	60	118	110 ~ 130
AR62	(CCT) ₇	F: CAGTTGGTCTACCTCTGGAGTTG R: TCGAAAAGATACAGCTAAGCAC	60	142	128 ~ 300
AR66	(TCA) ₇	F: TTCTCTGATGCAAACACTCTCA R: ACAAATTAGTCCAAGTTGCTCCA	60	158	145 ~ 277
AR68	(ATCC) ₅	F: ACCCTTATTTTCCCATCAACAAT R: CCGGATAATGGAAAGATTGATAG	60	151	143 ~ 158
AR73	(TATG) ₅	F: CAGATTGTGAAAATGATTGTGGA R: CGAACCCCTTCTCTTCTACCTTC	60	114	110 ~ 142
AR75	(CCAG) ₆	F: ATAAGGAGATGTGTCTCATGCT R: ATTFACACTCCACCATGAAAAGGA	60	155	150 ~ 172
AR76	(CCTC) ₆	F: AAGACACATTTGGTTTTGGAAGA R: CCCAAAGTTTCTGCTCACTCTTA	60	110	104 ~ 120
AR77	(TC) ₉	F: TCTCTCAGGCAAGCATCTCTATC R: ACTTAGCCGTTGGTATCTTCTCC	60	145	145 ~ 176
AR80	(TA) ₉	F: CGCAAATCTTAAAGGACATGACT R: GCTGAGGCTTCGATTATTAGGA	60	156	150 ~ 188
AR81	(GA) ₉	F: AACTCGCACACAGAGAAAAGAAAG R: ATCTTCTCTCCTTCAGCCTCTT	60	127	127 ~ 141
AR83	(TA) ₉	F: AGAAGCTTGTAAACACGTCCTCAG R: TGACCGGAGTGTGTAACCTTTTCT	60	150	149 ~ 164
AR98	(AG) ₉	F: GGGTTTCTGCTTTAGTCTCTCT R: TTTAGCTTCCAGTTCCAGACTTG	60	121	120 ~ 132

3 讨论

随着高通量测序技术的发展,测序数据几乎涵盖了特定组织或时期的所有转录本,特别是那些缺乏基因组信息的物种,基于转录组建立的SSR标记是根据基因本身的差异,具有操作简便、经济、信息量大的优势,得到了广泛的应用^[26]。桉木属转录组序列中SSR位点出现的频率为9.67%,高于杜仲的

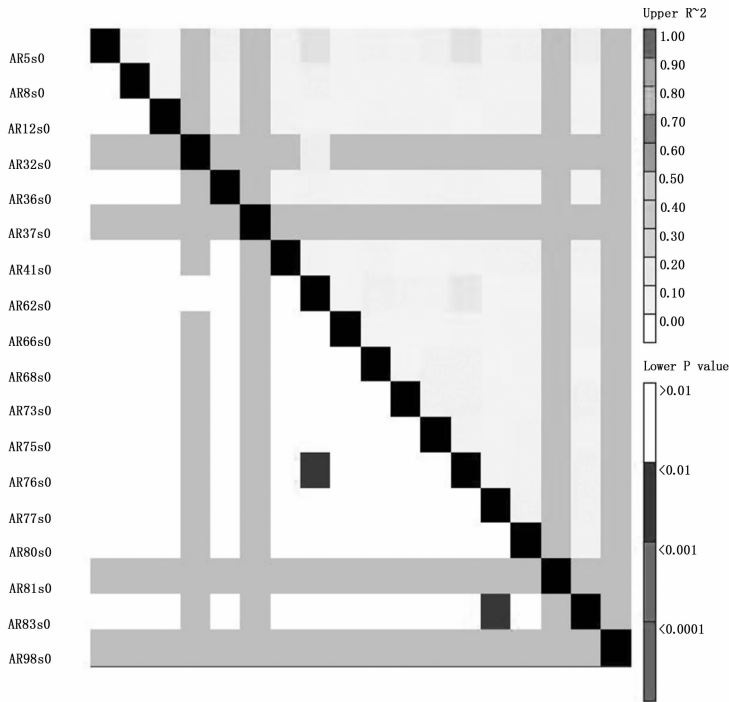
2.9%^[17]、南方红豆杉的2.07%^[16],但显著低于柑橘(*Citrus reticulata* Blanco)21.74%^[27]、萝卜(*Raphanus sativus* L.)23.79%^[28]。目前,在大多数植物中,SSR类型以二、三核苷酸重复为主,但不同的物种间优势重复单元类型也存在很大差异^[29],刺梨(*Rosa roxburghii* Tratt)^[30]、桑树(*Morus alba* L.)^[31]等以三核苷酸重复为最广泛的重复单元,桉木属转录组主要重复类型为二、三核苷酸重复,占SSR总数的

表4 18对SSR引物在12份材料中的遗传多样性参数

引物编号	等位基因数(Na)	有效等位基因数(Ne)	多态信息含量(PIC)	固定指数(Fis)
AR5*	7	2.618	0.589	0.294
AR8ns	12	8.471	0.871	0.123
AR12ns	11	5.878	0.814	-0.257
AR32*	6	4.255	0.728	-0.108
AR36ns	11	7.579	0.855	0.500
AR37ns	8	6.737	0.835	0.244
AR41ns	10	7.784	0.858	-0.247
AR62**	6	4.000	0.719	1.000
AR66**	11	9.000	0.879	0.765
AR68*	9	4.966	0.774	0.724
AR73ns	6	3.236	0.659	-0.021
AR75*	7	5.333	0.791	-0.103
AR76ns	6	3.600	0.677	-0.026
AR77ns	10	8.471	0.870	-0.315
AR80*	12	9.290	0.883	0.188
AR81ns	2	1.600	0.305	1.000
AR83**	8	4.721	0.764	0.564
AR98ns	3	2.667	0.555	1.000

65.87%和28.36%，四、五、六核苷酸重复类型的数量很少，总计5.73%。在二核苷酸重复单元中AG/CT占总SSR位点53.63%，这与报道的红掌(*Anthurium andraeanum* Linden)、刺梨(*Rosa roxburghii* Tratt)等多数植物相同^[30-32]。另外，CG/CG重复单元在桤木属转录组SSR中检测到了10个，这个重复单元在二核苷酸重复中是十分罕见的，其它报道如腊梅^[18]中检测到6次，甘蓝(*Brassica oleracea* L.)和杜仲中都仅检测到1次^[17, 33]，表现出了明显的偏倚性。桤木属转录组SSRs出现的频率较低，但重复单元的类型较为丰富，平均分布距离14.04 kb，出现频率为2.03%。

在合成的100对引物中有2对引物扩增片段与预期产物片段大小不符，显著大于预期产物片段长度，这可能与扩增片段中插入内含子有关，属于无效扩增^[34]。合成的100对SSR引物的扩增效率为39%，部分引物扩增失败的原因可能与引物位置或



SSR位点排列:黑色对角线上方的每个像素格使用右侧色差代码表征成对位点间R2值大小,黑色对角线下方为成对位点间LD的支持率

图5 18个SSR位点间LD分布

者质量有关。不同物种,检测出多态性SSR引物的差异较大,本研究中39对引物中18对具有多态性,占可扩增引物46.15%,这个比例低于刺梨52.17%^[26]、三七(*Panax pseudoginseng* Wall. var. *notoginseng* (Burkill) Hoo et Tseng)50%^[11],但高于杜仲44.44%^[17]和南方红豆杉38.71%^[16]。本研究检测的18对具有多态性的引物重复单元多为二、三

核苷酸重复,进一步验证了低级单元SSR的多态性普遍高于高级单元的推断^[35],因此,在设计SSR引物时应侧重含低级重复单元的序列。前人在桦木科植物SSR通用性方面作了研究,Drašnarová等^[20]根据桦木属(*Betula* L.)、桤木属和榛属(*Corylus* L.)的SSR序列设计的23个SSR引物在欧洲桤木和灰桤木(*A. incana* (L.) Moench)中的扩增效率分别为

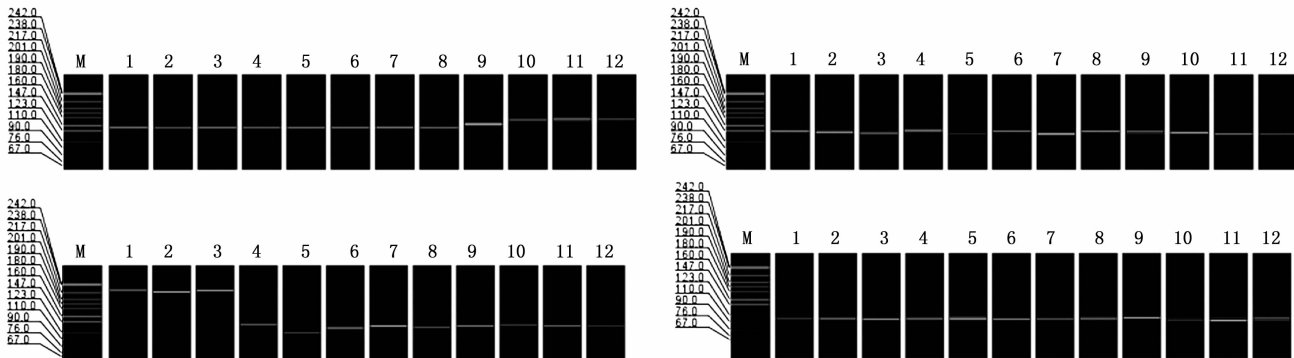


图6 引物 AR5、AR8、AR36 和 AR41 在 12 份桉木材料间的多态性

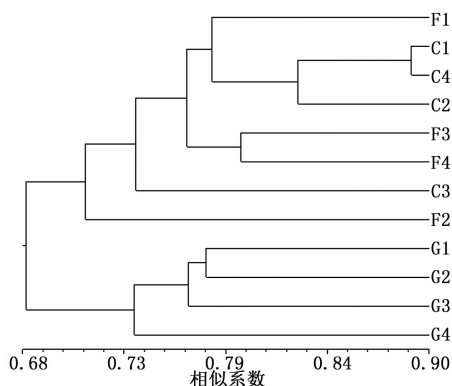


图7 12 份桉木材料的 UPGMA 聚类

37.78% 和 39.13%。本研究检测的 18 对 SSR 引物,在欧洲桉木、四川桉木和硬桉木中的扩增效率分别为 100%、88.89% 和 88.89%。植物 SSR 侧翼序列的保守程度决定着 SSR 的通用率,说明所选取的桉木属 SSR 侧翼序列的保守性较好。

第二代高通量测序技术显著提高了 SSR 开发效率和通量,但如何进一步提高多态性 SSR 位点鉴别、筛选效率是面临的又一挑战。可根据研究材料、研究目的以及所需位点数量的不同,灵活运用重复单元长度判断法^[36]和电子预测法^[37],借助多重 PCR 技术^[38]进一步提高多态性位点检测效率,降低试验成本。随着第三代测序技术的发展与应用,能有效解决大规模微卫星标记开发中的难题,不需要经过 PCR 扩增,实现对每一条 DNA 分子的单独测序会使 SSR 标记开发、功能基因定位更为便捷与准确。同时,高精度的多样本平行测序使大规模多态性位点电子检测成为可能,真正实现多态性位点选择的规模化和高效化^[36]。

4 结论

本研究以 3 种不同倍性桉木属树种转录组序列

信息进行桉木属 SSR 标记的开发,共搜索出 8 678 个 SSR 位点,分布在 8 298 条 Unigenes 中,SSR 检出率为 9.67%,重复单元中二核苷酸重复数量最多(65.87%),其次为三核苷酸(28.36%)。随机挑选 100 对 SSR 引物进行扩增筛选,扩增效率 39%,进一步筛选得到 18 对多态性、通用性均较强的 SSR 引物。本研究表明利用转录组数据开发桉木属树种 SSR 标记是可行的,开发的引物为桉木属树种遗传多样性分析、分子标记辅助育种和育种群体的建立等提供了丰富的候选标记。

参考文献:

- [1] 于东阳,梅芳,王军辉,等.不同种源桉木制浆造纸性能的遗传变异[J].河南农业大学学报,2013,47(6):703-709.
- [2] 饶龙兵,杨汉波,郭洪英,等.17种桉木属植物的亲缘关系研究及模糊种鉴定[J].植物研究,2015(4):528-534.
- [3] 杨汉波,饶龙兵,郭洪英,等.5种桉木属植物的核型分析[J].植物遗传资源学报,2013,14(6):1203-1207.
- [4] 刘军.桉木和四川桂花的形态、细胞及分子分类学研究[D].雅安:四川农业大学,2006.
- [5] 饶龙兵,杨汉波,郭洪英,等.桉木属7种植物的核型分析[J].西北植物学报,2013,33(7):1333-1338.
- [6] 李秀媛,刘西平,Hang DUONG,等.美国海滨桉木和薄叶桉木水分生理特性的比较[J].植物生态学报,2011,35(1):73-81.
- [7] 饶龙兵,杨汉波,郭洪英,等.桉木属植物 AFLP 反应体系的建立与优化[J].分子植物育种,2014(3):547-553.
- [8] 饶龙兵,杨汉波,郭洪英,等.不同倍性桉木属植物遗传差异的 AFLP 分析[J].植物研究,2014(6):803-809.
- [9] 夏勇.江南桉木遗传多样性研究[D].南京:南京林业大学,2012.
- [10] Kalia R K, Rai M K, Kalia S, et al. Microsatellite markers: An overview of the recent progress in plants [J]. Euphytica, 2011, 177(3):309-334.
- [11] 李翠婷,张广辉,马春花,等.野三七转录组中 SSR 位点信息分析及其多态性研究[J].中草药,2014,45(10):1468

- 1472.
- [12] Morgante M, Olivieri A M. PCR-amplified microsatellites as markers in plant genetics [J]. *Plant Journal for Cell & Molecular Biology*, 1993, 3(1):175-182.
- [13] Milee A, Neeta S, Harish P. Advances in molecular marker techniques and their applications in plant sciences[J]. *Plant Cell Reports*, 2008, 27(4):617-631.
- [14] 程小毛, 黄晓霞. SSR 标记开发及其在植物中的应用[J]. *中国农学通报*, 2011, 27(5):304-307.
- [15] Simon S A, Zhai J, Nandety R S, *et al.* Short-read sequencing technologies for transcriptional analyses [J]. *Annual Review of Plant Biology*, 2009, 60(1):305-333.
- [16] 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.
- [17] 黄海燕, 杜红岩, 乌云塔娜, 等. 基于杜仲转录组序列的 SSR 分子标记的开发[J]. *林业科学*, 2013, 49(5):176-181.
- [18] 赵凯歌, 陈玉星, 唐锐君, 等. 蜡梅转录组 EST-SSR 标记开发与引物筛选[J]. *北京林业大学学报*, 2013(S1):25-32.
- [19] 张 振, 张含国, 莫 迟, 等. 红松转录组 SSR 分析及 EST-SSR 标记开发[J]. *林业科学*, 2015(8):114-120.
- [20] Drašnarová A, Krak K, Vít P, *et al.* Cross-amplification and multiplexing of SSR markers for *Alnus glutinosa* and *A. incana* [J]. *Tree Genetics & Genomes*, 2014, 10(4):865-873.
- [21] Lepais O. De novo discovery and multiplexed amplification of microsatellite markers for black alder (*Alnus glutinosa*) and related species using SSR-enriched shotgun pyrosequencing [J]. *Journal of Heredity*, 2011, 102(5):627-632.
- [22] Kalinowski S T, Taper M L, Marshall T C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment [J]. *Molecular Breeding*, 2007, 16(5):1099-1106.
- [23] Bradbury P J, Zhang Z, Kroon D E, *et al.* TASSEL: software for association mapping of complex traits in diverse samples [J]. *Bioinform*, 2007, 23(19):2633-2635.
- [24] Peakall R, Smouse P E. GenALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update [J]. *Bioinformatics*, 2011, 28(19):2537-2539.
- [25] Rohlf F J. NYSYS pc, numerical taxonomy and multivariate analysis system, version 2.1 [CP]. Setauket, NY: Exeter Publishing, Ltd., 2000
- [26] 李小白, 向 林, 罗 洁, 等. 转录组测序 (RNA-seq) 策略及其数据在分子标记开发上的应用[J]. *中国细胞生物学学报*, 2013(5):720-726.
- [27] Jiang D, Zhong G Y, Hong Q B. Analysis of Microsatellites in *Citrus Unigenes* [J]. *Investigación Agraria Sistemas Y Recursos Forestales*, 2000, 9(1):233-254.
- [28] Wang S, Wang X, He Q, *et al.* Transcriptome analysis of the roots at early and late seedling stages using Illumina paired-end sequencing and development of EST-SSR markers in radish [J]. *Plant Cell Reports*, 2012, 31(8):1437-1447.
- [29] Vendramin E, Dettori M T, Giovannazi J, *et al.* A set of EST-SSRs isolated from peach fruit transcriptome and their transportability across *Prunus* species [J]. *Molecular Ecology Notes*, 2007, 7(2):307-310.
- [30] 鄢秀芹, 鲁 敏, 安华明. 刺梨转录组 SSR 信息分析及其分子标记开发[J]. *园艺学报*, 2015, 42(2):341-349.
- [31] 姚刘慧, 唐翠明, 戴凡炜, 等. 基于桑树转录组测序的 SSR 标记开发和引物筛选[J]. *蚕业科学*, 2015(1):18-27.
- [32] 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(6):1349-1354.
- [33] 陈 琛, 庄 木, 李康宁, 等. 甘蓝 EST-SSR 标记的开发与应用[J]. *园艺学报*, 2010, 37(2):221-228.
- [34] Saha M C, Mian M A R, Imad E, *et al.* Tall fescue EST-SSR markers with transferability across several grass species [J]. *Tag-theoretical & Applied Geneticstheoretische Und Angewandte Genetik*, 2004, 109(4):783-791.
- [35] Dreisigacker S, Zhang P, Warburton M L, *et al.* SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different megaenvironments [J]. *Crop Science*, 2004, 44(2):381-388.
- [36] 文亚峰, 韩文军, 周 宏, 等. 杉木转录组 SSR 挖掘及 EST-SSR 标记规模化开发[J]. *林业科学*, 2015, 51(11):40-49.
- [37] Duran C, Singhania R, Raman H, *et al.* Predicting polymorphic EST-SSRs in silico [J]. *Molecular Ecology Resources*, 2013, 13(3):538-545.
- [38] Blacket M J, Robin C, Good R T, *et al.* Universal primers for fluorescent labelling of PCR fragments-an efficient and cost-effective approach to genotyping by fluorescence [J]. *Molecular Ecology Resources*, 2012, 12(3):456-463.

(责任编辑:金立新)