

## 普通油茶转录组EST-SSR分子标记开发

李海波<sup>1</sup>, 王珊<sup>2</sup>, 丁红梅<sup>3</sup>, 陈友吾<sup>1</sup>, 胡传久<sup>1</sup>, 李楠<sup>1</sup>, 魏海龙<sup>1,\*</sup>

<sup>1</sup>浙江省林业科学研究院, 杭州310023; <sup>2</sup>浙江农林大学林业与生物技术学院, 浙江临安311300; <sup>3</sup>浙江中医药大学分析测试中心, 杭州310053

**摘要:** 通过转录组测序技术鉴别普通油茶(*Camellia oleifera*) EST-SSR位点, 开发筛选适用于普通油茶种质资源评价与利用的多态性SSR标记, 对普通油茶的遗传变异研究和油茶良种的分子标记辅助育种具有重要意义。利用 MISA 工具从普通油茶转录组67 434条Unigenes (63.39 MB)中筛选得到26 751个包含1~6个核苷酸重复类型的SSR位点, SSR的分布频率为39.67%, 发生频率为1/2.33 kb; SSR位点中的主导类型是二核苷酸重复, 占总SSR的44.28%, 其次是单核苷酸(35.28%)和三核苷酸重复(18.27%)。通过SSR标记的琼脂糖凝胶电泳和荧光标记毛细管电泳, 从143对SSR引物中遴选出了20对扩增效率高、稳定性好的多态性SSR引物, 并对来自浙江产区的57个普通油茶品种进行遗传关系分析。结果显示: 20对SSR引物共检测出129个等位基因, 每对引物2~17个, 平均6.45个; 多态性信息量(PIC)为0.0499~0.8635, 平均0.4355。57个普通油茶品种在Dice遗传相似系数大约0.47时聚为一类, 在大约0.60时可分为七个大类。本研究印证了利用普通油茶转录组数据开发SSR标记的可行性, 也证明了茶属树种SSR标记的可转移性; 不仅为普通油茶遗传多样性分析和遗传图谱构建提供了有价值的候选标记, 也为普通油茶良种的鉴别与分子选育提供了分子技术手段。

**关键词:** 油茶; 转录组; EST-SSR; 遗传关系; 多态性

油茶是山茶属(*Camellia*)种子含油量高、有栽培价值物种的统称(陈永忠等2005a)。普通油茶(*Camellia oleifera*)是油茶中分布面积最广、总产量最多的一个物种, 常用于榨取高质量的食用茶油。普通油茶是世界四大木本油料之一, 在我国的栽培历史悠久, 主要分布于我国南方18个省市, 是我国南方重要的木本油料树种, 在林业生产和生态经济建设中占有重要地位。自上个世纪六十年代以来, 我国林业育种工作者就开展了油茶良种的常规选育工作, 并取得了重大突破(林萍等2010)。进入九十年代以来, 分子生物学的飞速发展成为油茶分子育种提供了新的技术手段, 获得含油率高、品质优良、适应性强的优质高产新品种是油茶分子育种的基本方向。

鉴定试验材料的遗传多样性是研究物种起源进化、发现新的基因资源、改良现有育种材料的基础工作。DNA分子标记是研究种质资源遗传多样性的重要手段, 也是一种重要的分子辅助育种方法。自上个世纪以来, 分子标记技术包括随机扩增多态性DNA (random amplified polymorphic DNA, RAPD)、相关序列扩增多态性(sequence-related amplified polymorphism, SRAP)、简单重复序列间区(inter simple sequence repeat, ISSR)和扩增片段长度多态性(amplified fragment length polymorphism, AFLP)陆续应用于普通油茶无性系的鉴别和油茶种质资源的遗传多样性分析(陈永忠等

2005b; 黄永芳等2006; 张国武等2007; 王保明等2008; 温强等2008; 林萍等2010; 彭绍峰等2011; 张婷等2011a, b; 于小玉等2013; 刘扬等2016)。这些在分子水平上开展的研究也为普通油茶优良无性系的鉴别与优良品系的分子选育提供了可参考的技术手段。

随着高通量测序技术的快速发展与测序成本的不断降低, 在许多物种的研究上, 利用新一代测序技术产生了丰富的转录组数据。基于Illumina转录组测序(RNA-Seq)获得的EST数据进一步挖掘有价值的简单重复序列(simple sequence repeats, SSR)位点已在水生植物蕃菜以及杜仲、洋葱、红松、印度南瓜等物种上得到了成功验证(袁阳阳等2013; 黄海燕等2013; 李满堂等2015; 张振等2015; 王洋洋等2016)。简单重复序列又称微卫星(microsatellite)标记, 主要是以1~6个核苷酸为基本重复单位的串联重复序列(Kalia等2011)。与RAPD、ISSR、SRAP和AFLP四种显性分子标记相比, SSR为共显性标记, 能区分纯合子和杂合子, 能检测复等位基因, 且具有多态性丰富、操作简单、结果可靠、重复性好等优点, 在遗传多样性检测中更具优势。SSR标记在DNA指纹图谱的构建、遗传

收稿 2017-04-17 修定 2017-06-08

资助 浙江省科研院所扶持专项(2014F30002和2013F50012)和浙江省重大科技专项(2010C02005-1)。

\* 通讯作者(E-mail: 1178646251@qq.com)。

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(罗冉等2010)。但在普通油茶SSR标记的开发与利用上,迄今为止,有价值的SSR标记数量与利用尚十分有限,远远不能满足油茶丰富的遗传多样性研究与加快分子育种工作的需要,高质量的SSR分子标记亟待大量开发。温强等(2013)挖掘比较了普通油茶、浙江红山茶(*C. chekiangoleosa*)和短柱茶(*C. brevistyla*)转录组和基因组中微卫星重复序列的组成与分布特征,但尚未开发出多态性高的油茶微卫星标记。彭婵等(2013)和张恩慧等(2016)利用相同的10个SSR标记分别对湖北和广西油茶的种质资源进行了分析评价; Jia等(2014)开发出了15个多态性的油茶SSR标记。此外,虽然也有山茶属内红花油茶(*C. chekiangoleosa*)的SSR标记开发(Wen等2012)、SSR标记用于小果油茶(*C. meiocarpa*)与普通油茶居群的遗传多样性、遗传结构、以及种间杂交渐渗的研究报道(黄勇2013),但这些山茶属内不同种之间SSR标记的通用性、多态性还有待进一步验证。浙江省具有丰富的油茶种质资源,油茶栽培与利用历史悠久。现有油茶栽培面积16.33万 $\text{hm}^2$ ,位居全国第4位(娄浩峰2014)。迄今为止,基于SSR标记对浙江产区油茶种质资源的评价与利用也尚未有报道。

利用Illumina转录组测序技术,课题组前期已获得了普通油茶幼叶的大量转录组数据。基于序列数据组装拼接后得到的Unigenes,本研究将进行普通油茶EST-SSR位点的鉴别,并对SSR位点的组成、分布及特征进行分析;将利用SSR荧光标记的毛细管电泳检测技术,以浙江产区内的57个普通油茶品种作为研究材料,筛选出适用于普通油茶种质资源评价的多态性SSR标记,并对不同SSR引物所揭示的油茶SSR位点的多态性水平进行分析评价。本研究工作旨在印证基于普通油茶转录组数据开发SSR标记的可行性,以期普通油茶遗传多样性分析和遗传图谱构建提供有价值的候选标记,为普通油茶优良无性系的鉴别、优良品系的分子选育提供分子技术手段。

## 材料与方法

### 1 材料与试剂

#### 1.1 植物材料

用于本研究的57个普通油茶(*Camellia oleifera* Abel.)品种样品于2014年4~5月采集。其中12个长林系列油茶样品来自浙江省林业科学研究院在浙江省武义县俞源镇钟蓬村建立的百灵谷油茶幼林基地,45个龙林系列油茶样品来自浙江省龙游县龙

表1 用于本研究的普通油茶品种  
Table 1 *Camellia oleifera* cultivars used in this study

品种编号	油茶品种	品种编号	油茶品种	品种编号	油茶品种
CL3	‘长林3号’	LL8	‘龙林8号’	LL27	‘武10’
CL4	‘长林4号’	LL9	‘龙林9号’	LL28	‘丽水12’
CL18	‘长林18号’	LL10	‘龙林10号’	LL29	‘越南’
CL21	‘长林21号’	LL11	‘龙林11号’	LL30	‘江西小果’
CL23	‘长林23号’	LL12	‘龙林12号’	LL31	‘丽水50’
CL26	‘长林26号’	LL13	‘龙林13号’	LL32	‘油茶WM’
CL27	‘长林27号’	LL14	‘龙林14号’	LL33	‘木茶’
CL40	‘长林40号’	LL15	‘龙林15号’	LL34	‘青田6’
CL53	‘长林53号’	LL16	‘龙林16号’	LL35	‘巨40’
CL55	‘长林55号’	LL17	‘龙林17号’	LL36	‘陆川’
CL56	‘长林56号’	LL18	‘普油’	LL37	‘常德’
CL166	‘长林166号’	LL19	‘衡东大桃’	LL38	‘茶王’
LL1	‘龙林1号’	LL20	‘青田24’	LL39	‘衡大’
LL2	‘龙林2号’	LL21	‘永兴1’	LL40	‘遂78-18’
LL3	‘龙林3号’	LL22	‘武7’	LL41	‘遂15’
LL4	‘龙林4号’	LL23	‘秀山76-22’	LL42	‘秀山茶王3’
LL5	‘龙林5号’	LL24	‘青田4’	LL43	‘四川茶王’
LL6	‘龙林6号’	LL25	‘普茶’	LL44	‘攸县油茶’
LL7	‘龙林7号’	LL26	‘丽水13’	LL45	‘芳6’

游林场(表1)。在本实验中, 每个油茶品种均采集3株的幼嫩叶片作为提取基因组DNA的供试材料, 放置于硅胶密封保存。

转录组测序样本为取自油茶良种‘长林18号’的幼嫩叶片, 液氮速冻后进行RNA提取、文库构建和Illumina转录组测序, 拼接组装后得到67 434条Unigenes, 总长为63.39 MB。

## 1.2 主要试剂及引物

研究所用主要分子生物学试剂包括新型快速植物基因组DNA提取试剂(BioTeke, 北京百泰克)、PCR扩增试剂2×TSINGKE Master Mix (TSINGKE, 北京擎科), 在正向SSR引物上加注的荧光染料是FAM (蓝), 荧光SSR引物和普通SSR引物合成(TSINGKE, 北京)。用于毛细管电泳检测的内标试剂是Hi-Di™ Formamide (Applied Biosystems)、GeneScan™-500 LIZ Size Standard (Applied Biosystems)。

## 2 仪器与设备

PCR仪: Life ECO (Bioer, 杭州博日); 凝胶电泳仪: PowerPac™ Universal Power Supply, 美国伯乐(Bio-Rad, Hercules, CA, USA); 凝胶成像系统: Chemi Doc XRS imaging system, 美国伯乐(Bio-Rad, Hercules, CA, USA); 分光光度计: NanoDrop 2000, 美国热电(Thermo Scientific, USA); DNA分析仪: 96-well plate, 3730 XL, DNA Analyzer (Applied Biosystems, USA)。

## 3 实验方法

### 3.1 油茶基因组DNA的提取及浓度测定

每个油茶品种任取3株经硅胶干燥后的幼嫩叶片粉碎, 等量混合后用于提取基因组DNA。DNA提取方法参照新型快速植物基因组DNA提取试剂盒的操作说明, 提取后的DNA经NanoDrop 2000测定浓度, 并经1.5%琼脂糖凝胶电泳检测, 于-20℃保存备用。

### 3.2 油茶转录组SSR位点的鉴别及SSR引物设计

利用MicroSAteLLite (MISA, <http://pgrc.ipk-gatersleben.de/misa/>)工具对拼接组装得到的Unigenes进行SSR位点的识别定位。筛选标准: 单核苷酸重复的次数在10次或10次以上, 二核苷酸重复的次数在6次或6次以上, 三至六核苷酸重复的次数在5次或5次以上。同时, 也筛选中间被少数碱基(间隔小于100或等于100)打断的不完全重复

的SSR。利用Primer 3.0对筛选得到的SSR位点设计引物, 引物设计原则: 引物序列长度18~22 bp, PCR产物长度100~300 bp, GC含量40%~60%, 退火温度50~63℃, 尽量避免出现发夹结构、二聚体、错配和引物二聚体。

### 3.3 SSR-PCR分析

20 μL SSR-PCR的扩增体系为: 2×TSINGKE Master Mix 10 μL, 荧光标记(FAM)的正向SSR引物(10 μmol·L<sup>-1</sup>) 1.5 μL, 反向引物(10 μmol·L<sup>-1</sup>) 1.5 μL, DNA模板(20 ng·μL<sup>-1</sup>) 3 μL, 加ddH<sub>2</sub>O补齐到20 μL。PCR反应条件为: 94℃预变性7 min后; 94℃变性45 s, 退火45 s (不同SSR引物的最佳退火温度不同), 72℃延伸2 min, 共30个循环; 最后于72℃补平7 min, 终止温度为4℃。PCR扩增后的产物用1.5%的琼脂糖凝胶(EB染色)电泳, 扩增图谱的摄取采用Chemi Doc XRS imaging system。SSR标记扩增产物的测序验证交付北京擎科(TSINGKE, 北京)完成。

### 3.4 毛细管电泳检测

内标配制: 取10 mL Hi-Di和80 μL 500 LIZ混匀, 离心, 以每孔10 μL分装于96孔内标板, 离心。将SSR-PCR产物电泳, 根据电泳胶图做一定稀释(最低可检测标准为0.1 ng·μL<sup>-1</sup>), 离心。取稀释产物0.5 μL加入到分配好的内标板中, 混匀, 离心, 放入PCR仪, 于96℃变性5 min。20℃下迅速冷冻2 min, 离心。置入DNA分析仪进行毛细管电泳检测。用Data Collection 3.0软件收集数据。

### 3.5 数据统计与分析

用GeneMapper 4.1软件对Data Collection软件收集的原始数据进行分析。软件系统将根据目标峰的位置与同一泳道中的内标GeneScan™ 500 LIZ进行比较, 直接给出目标SSR片段的准确数值。将SSR原始数据转换为0/1原始数据矩阵格式, 即对于每个SSR检测位点, 具有相同数值的SSR条带记为1, 无该数值的SSR条带记为0。利用NTSYS-pc 2.10e软件进行遗传关系分析, 使用软件中的SimQual程序计算样品间的Dice遗传相似系数, 用SAHN程序进行UPGMA (unweighted pair group method with arithmetic mean, 非加权成对算术平均法)聚类分析, 通过Tree plot模块生成聚类图。利用SSR数据格式转换软件DataFormater (樊文强等



2016)将NTSYS软件的0/1数据矩阵转换为POP-GENE 1.32 (Yeh等2000)软件的数据格式,统计每对引物在57个油茶品种中扩增的等位基因数(number of alleles),计算等位基因频率(allele frequency)。利用PIC-CALC 0.6软件计算SSR位点的多态性信息量(polymorphism information content, PIC) (Botstein等1980)。

## 实验结果

### 1 油茶转录组中的SSR位点的数量与分布

利用MISA工具对普通油茶转录组67 434条Unigenes (63.39 MB)的cDNA序列进行识别,在

18 040条Unigene中找到26 751个符合条件的SSR位点,发生频率(含有SSR的Unigene数/总Unigene数)为26.75%。其中,5 834条unigenes含1个以上SSR位点,12 206条unigenes含单个SSR位点,SSR的分布频率(SSR的个数/总Unigene数)为39.67%。油茶转录组中平均2.33 kb发现1个SSR位点。

油茶转录组SSR种类较为丰富,各重复类型均有,但各种类型的出现频率有较大差异(表2),二核苷酸重复类型的数量最多,占44.28%,其次是单核苷酸和三核苷酸重复类型,分别占35.28%和18.27%。四、五、六核苷酸重复类型的数量很少,只占到总SSR的2.16%。

表2 普通油茶转录组SSR类型和出现频率

Table 2 Types and distribution frequency of the SSR motifs in transcriptome of *Camellia oleifera*

重复类型	数量/个	比例/%	出现频率/%	总长度/bp	平均长度/bp	平均距离/kb
单核苷酸	9 439	35.28	14.00	131 883	13.97	6.61
二核苷酸	11 846	44.28	17.57	206 560	17.44	5.27
三核苷酸	4 887	18.27	7.25	84 570	17.31	12.77
四核苷酸	292	1.09	0.43	6 192	21.21	213.65
五核苷酸	111	0.41	0.16	3 155	28.42	562.04
六核苷酸	176	0.66	0.26	5 718	32.49	354.47
总计	26 751	100	39.67	438 078	16.38	2.33

油茶转录组中SSR位点的序列总长度为438 078 bp,位点平均长度为16.38 bp,其中,单、二、三、四、五和六核苷酸类型SSR位点的平均长度分别为13.97、17.44、17.31、21.21、28.42以及32.49 bp。油茶转录组中SSR位点的重复次数以13次及以上最多(5 896个),占总SSR的20.31%,其中以单核苷酸最多(4 391个)。其次为重复6次(4 421个,15.23%)、重复12次(3 539个,12.19%)、重复10次

(3 222个,11.10%)和重复5次(3 142个,10.82%)。重复次数相对较少的SSR位点为重复7、8、9和11次,依次为9.01%、7.95%、6.79%和6.59%(表3)。

### 2 油茶转录组SSR位点的分布特征

从油茶SSR核苷酸基序类型(表4)来看,SSR以二核苷酸(11 846)重复基序为主要类型,其次是单核苷酸(9 439)和三核苷酸(4 887)重复类型。在二核苷酸重复基序中以AG/CT为主(9 629,81.28%),

表3 普通油茶转录组SSR位点重复次数分布

Table 3 Distribution of the number of repeats in SSR loci in transcriptome of *Camellia oleifera*

重复类型	重复数/次								
	5	6	7	8	9	10	11	12	≥13
单核苷酸	—	—	—	—	—	2 468	1 510	1 070	4 391
二核苷酸	—	2 920	2 091	2 198	1 903	706	378	314	1 336
三核苷酸	2 703	1 405	497	101	63	48	25	15	30
四核苷酸	235	37	16	1	2	0	0	0	17
五核苷酸	83	16	4	4	2	0	1	0	122
六核苷酸	121	43	7	4	1	0	0	43	0
总计	3 142	4 421	2 615	2 308	1 971	3 222	1 914	3 539	5 896

表4 普通油茶转录组中的SSR类型分布  
Table 4 Distribution of SSR types in transcriptome of *Camellia oleifera*

重复类型	基序	重复数/次	合计/次
单核苷酸	A/T	8 941	9 439
	C/G	498	
二核苷酸	AC/GT	969	11 846
	AG/CT	9 629	
	AT/AT	1 222	
三核苷酸	CG/CG	26	4 887
	AAC/GTT	449	
	AAG/CTT	1 210	
	AAT/ATT	294	
	ACC/GGT	915	
	ACG/CGT	93	
	ACT/AGT	104	
	AGC/CTG	349	
	AGG/CCT	596	
	ATC/ATG	623	
	CCG/CGG	254	
四单核苷酸			292
五核苷酸			111
六核苷酸			176

且以6~9次重复为主; 在单核苷酸重复基序中以A/T为主(8 941, 94.72%), 且以13次以上重复为主; 在三核苷酸重复基序中以AAG/CTT为主(1 210, 24.76%), 且以5~6次重复为主。

### 3 SSR引物多态性分析

为了进一步利用基于转录组测序鉴别出的普通油茶EST-SSR引物, 对含SSR位点的Unigene序列进行引物设计, 通过对引物的评价优化, 最终选出89对SSR引物进行合成, 这些SSR引物扩增的位点含2~6个核苷酸重复单元。以普通油茶品种‘长林26号’的DNA作为模板, 利用89对SSR引物进行PCR扩增, 以验证这些引物的实际有效性。琼脂糖凝胶电泳检测结果表明, 89对引物中的32对能有效扩增出预期大小片段。此外, 为了进一步丰富普通油茶的多态性SSR标记, 对Wen等(2012)、彭婵等(2013)、黄勇(2013)、Jia等(2014)文献报道的25对来自普通油茶、29对来自茶属内红花油茶和茶树(*C. sinensis*)的SSR引物也进行验证, 发现其中的38对能有效扩增出预期大小片段。对初步选出的这70对SSR引物的正向引物进行荧光标记, 先从本研究收集的57个不同油茶品种中随机挑选15个进行

PCR扩增及毛细管电泳检测, 共筛选出了23对扩增效率高、稳定性好的多态性SSR引物。收集PCR扩增产物, 经进一步测序验证, 发现其中的20对引物能够扩增出目标重复序列(表5)。

利用20对多态性SSR引物对57个油茶品种进行检测(图1)。表6显示, 20对多态性SSR引物在57个油茶品种中共检测出129个等位基因, 长度大小为96~277 bp, 每对引物检测出2~17个等位基因, 平均为6.45个等位基因。同时, 在129个等位基因中, 每个等位基因出现的频率差异也很大, 从最低的0.0088到最高的0.9737不等。这同时也反映了20对SSR引物对57个油茶品种的多态性识别存在很大差异, 其中以Co3、Co12、C037和Co65四个引物的多态性最低, TUGMS17的多态性最高。

多态信息量(PIC)是等位基因频率和等位基因数变化的函数, 是衡量引物适合程度的重要指标, 也反映在微卫星座位上的遗传变异程度。PIC高表明引物揭示的多态性丰富, 微卫星座位的一致性差, 变异大; PIC低表明引物揭示的多态性匮乏, 该座位变异小。20个引物的PIC为0.0499~0.8635, 平均为0.4355。根据Botstein等(1980)的理论, 6个标记为低度多态位点( $0 < \text{PIC} < 0.25$ ), 6个标记为中度多态位点( $0.25 < \text{PIC} < 0.5$ ), 8个标记为高度多态位点( $\text{PIC} > 0.5$ ), 以Co3、Co12、C037和Co65四个位点的PIC最低(0.0499), TUGMS17位点的PIC最高(0.8635)(表6)。

### 4 基于SSR的遗传关系分析

基于20对多态性SSR引物扩增出的129个条带对57个普通油茶品种进行聚类分析(图2)。UPGMA聚类图总体上显示57个油茶品种并没有形成明显的具有规律性的成组分类, 在Dice遗传相似系数大约为0.47时才聚在一起, 在大约0.60时可由上至下分为七个大类。第一大类包含5个品种(8.8%), 分别是‘长林3号’、‘长林56号’、‘龙林4号’、‘龙林15号’和‘武7’; 第二大类包含了57个品种中的43个(77.2%), 在Dice遗传相似系数大约为0.645时, 第二大类可以进一步细分为八个组, 分别包含了8、5、11、7、5、3、2和2个品种。第三大类包含5个品种(8.8%), 分别是‘长林26号’、‘越南’、‘青田6’、‘油茶WM’和‘巨40’; 第四、五、六、七大类均只有一个品种(各占1.7%), 分别是‘丽水50’、‘龙林7号’、‘武10’和‘攸县油茶’。

表5 用于本研究的普通油茶多态性SSR引物

Table 5 Polymorphic SSR primers used in this study for *Camellia oleifera*

SSR标记	重复序列	SSR引物序列(5'→3')	最适退火温度/°C	目标条带扩增大小/bp	来源
Co3	(CTG) <sub>n</sub>	CL3F: GGGATCTGTGATGGCAAAGT CL3R: CCCTTGATACACCTGCCTGT	54.0	138~141	转录组测序开发
Co11	(CCT) <sub>n</sub>	Co11F: GAAGGAACCACCTTCAACCA Co11R: TTGTCGAGGAGGAGGGATT	55.0	208~211	转录组测序开发
Co12	(CCT) <sub>n</sub>	Co12F: TCACAATTCTGGCACCTCAA Co12R: CCACCAATTCCATCAATTCC	54.0	246~249	转录组测序开发
Co37	(GT) <sub>n</sub>	Co37F: AATTCCAACCATCCCAACAA Co37R: CAAGGCAAAAAGAACTCACCA	55.0	273~277	转录组测序开发
Co41	(GGA) <sub>n</sub>	Co41F: GCTAGCTGATCATGGGACCT Co41R: CAAAAGGCACCAAAGAAACC	54.5	226~262	转录组测序开发
Co51	(AG) <sub>n</sub>	Co51F: TTGGAGGAGGCCATAGTTG Co51R: CACAATCCAATCTAACCGCC	54.5	148~150	转录组测序开发
Co53	(TCG) <sub>n</sub>	CL53F: AGTGCTTGAAGGCGTTGATT CL53R: CACTACGAAGACGACGACGA	55.0	267~273	转录组测序开发
Co62	(AAC) <sub>n</sub>	CL62F: CATCCATAGCATCATCACCG CL62R: GCGGAGAGTTTCGAGTTTTG	55.0	206~215	转录组测序开发
Co65	(GGT) <sub>n</sub>	Co65F: AGGACGTGAAGGGTGAATTG Co65R: TTCCTTCATGATTGCTCC	55.5	274~277	转录组测序开发
Co66	(GAA) <sub>n</sub>	Co66F: CGTCCAATCCTCAACCAACT Co66R: AGAGGCTCGTCTTCCACAAA	55.0	96~99	转录组测序开发
Co68	(TC) <sub>n</sub>	CL68F: TCGATTCAATCACCTCTCC CL68R: TTGTCACAAAAGCCCATCA	54.5	258~268	转录组测序开发
Co74	(TC) <sub>n</sub>	CL74F: CGCTTTTGAACAGCACTCTG CL74R: AAGTAAGGCGTCCGGAATTT	56.0	167~197	转录组测序开发
Co76	(AAC) <sub>n</sub>	Co76F: ACACAACCAATCGACGAA Co76R: TTGGGTTTTGGTTTTCTTG	55.0	156~159	转录组测序开发
Co81	(GAT) <sub>n</sub>	Co81F: GGTCAAAACGAAGAAGAAGATCA Co81R: GGGATTCCCAATAGAGAGCC	54.5	146~161	转录组测序开发
Co84	(GAT) <sub>n</sub>	Co84F: TGATTGGAGAACAACCGA Co84R: GACAGTCAAGCCTCCATTCC	55.0	100~112	转录组测序开发
CoUg10161	(CT) <sub>n</sub>	Co10161F: ATGCTATTTGAGGGTCTTG Co10161R: ACTTGGAGTTGGAGTTGTC	50.0	186~208	Jia等(2014)
Ck03	(CCT) <sub>n</sub>	Ck03F: GTGGCGTCTTTCTTTACT Ck03R: ATTGACAAGCCTCATCGT	55.0	126~150	Wen等(2012)
Ck55	(AG) <sub>n</sub>	Ck55F: TGAACCAAGTAGCGTAGACC Ck55R: TGCAAAGCATCCACATT	55.0	174~208	Wen等(2012)
TUGMS17	(TC) <sub>n</sub>	TU17F: GGGGAATTTAGACAGACAC TU17R: GCCGTTCAAGTGTAGTAGATCG	55.0	216~262	黄勇(2013)
TUGMS28	(TG) <sub>n</sub> /(GA) <sub>n</sub>	TU28F: GTCCCCATTGCTCTTAGTTT TU28R: GACAATCATTGCCACCACAT	55.0	160~208	黄勇(2013)

从12个长林系列油茶在聚类图上的总体分布来看, 2个品种位于第一大类, 9个品种分别位于第二大类的四个组内, 1个品种位于第三大类。从17个龙林系列油茶在聚类图上的分布来看, 2个品种位于第一大类, 14个品种分别位于第二大类的五个组内, 1个品种位于第五大类。这表明无论长林系

列还是龙林系列油茶均具有较为复杂的遗传背景, 品种间的遗传距离较大, 并没有形成明显的具有规律性的成组分类, 尤其象‘龙林7号’、‘长林26号’与其他长林、龙林系列品种的遗传距离更大, 应该具有更为复杂的遗传背景。此外, 聚类图也显示, 在个别大类或组内, 也有个别品种成组聚在一

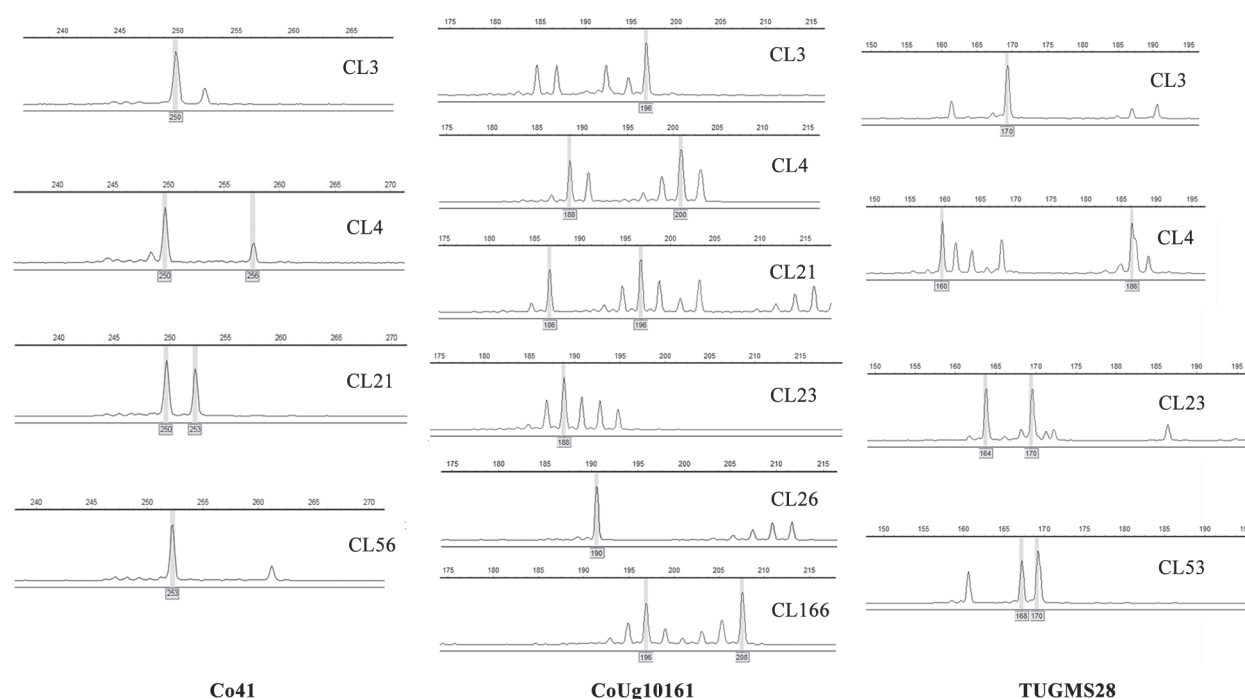


图1 部分油茶品种在3个SSR位点的基因型差异

Fig.1 Genotypic difference among some *Camellia oleifera* cultivars at three SSR loci

表6 20对SSR引物在57个油茶品种中多态性扩增结果  
Table 6 Amplification information of 20 SSR primer pairs in 57 *Camellia oleifera* cultivars

SSR位点	等位基因数/个	等位基因频率	多态信息含量
Co3	2	0.0263~0.9737	0.0499
Co11	2	0.0526~0.9474	0.0947
Co12	2	0.0263~0.9737	0.0499
C037	2	0.0263~0.9737	0.0499
Co41	9	0.0175~0.4912	0.6819
Co51	2	0.2368~0.7632	0.2961
Co53	3	0.0263~0.9474	0.0986
Co62	4	0.0175~0.6754	0.4513
Co65	2	0.0263~0.9737	0.0499
Co66	4	0.0088~0.7543	0.3470
C068	5	0.0088~0.5877	0.4607
Co74	15	0.0088~0.1842	0.8206
Co76	2	0.4825~0.5175	0.3747
Co81	6	0.0351~0.5088	0.6514
Co84	5	0.0088~0.6490	0.3842
CoUg10161	10	0.0088~0.2281	0.7897
Ck03	9	0.0088~0.3684	0.7797
Ck55	15	0.0088~0.3947	0.7752
TUGMS17	17	0.0088~0.2018	0.8635
TUGMS28	13	0.0088~0.5263	0.6407
平均值	6.45		0.4355

起, *Dice*遗传相似系数均达到了0.90以上, 这些包括‘长林4号’、‘长林27号’与‘长林40号’、‘长林18号’与‘龙林2号’、‘龙林8号’与‘龙林11号’、‘龙林6号’与‘龙林12号’, 这表明这些油茶品种间的遗传差异很小, 亲缘关系非常接近, 也可能存在同种异名现象。

## 讨 论

在我们当前设定的MISA筛选条件下, 共从普通油茶转录组67 434条unigenes中鉴定出26 751个含1~6个核苷酸重复类型的SSR位点, 平均分布频率(SSR的个数/总Unigene数)为39.67%, 平均出现频率为1/2.33 kb。这与温强等(2013)报道的普通油茶的出现频率(1/7.09 kb)、红花油茶(1/6.25 kb)、短柱茶(1/6.62 kb)以及在其他树种上报道的平均分布频率, 例如杨树(14.83%)、红松(4.24%)、马尾松(3.62%)、杜仲(2.9%)等差异较大(张振等2015), 其原因首先在于物种的特异性, 其次也与MISA软件检索设置的标准有很大关系。本研究利用MISA筛选的SSR位点包含1~6个核苷酸重复类型, 而其他树种报道的SSR位点包含2~6个核苷酸重复类

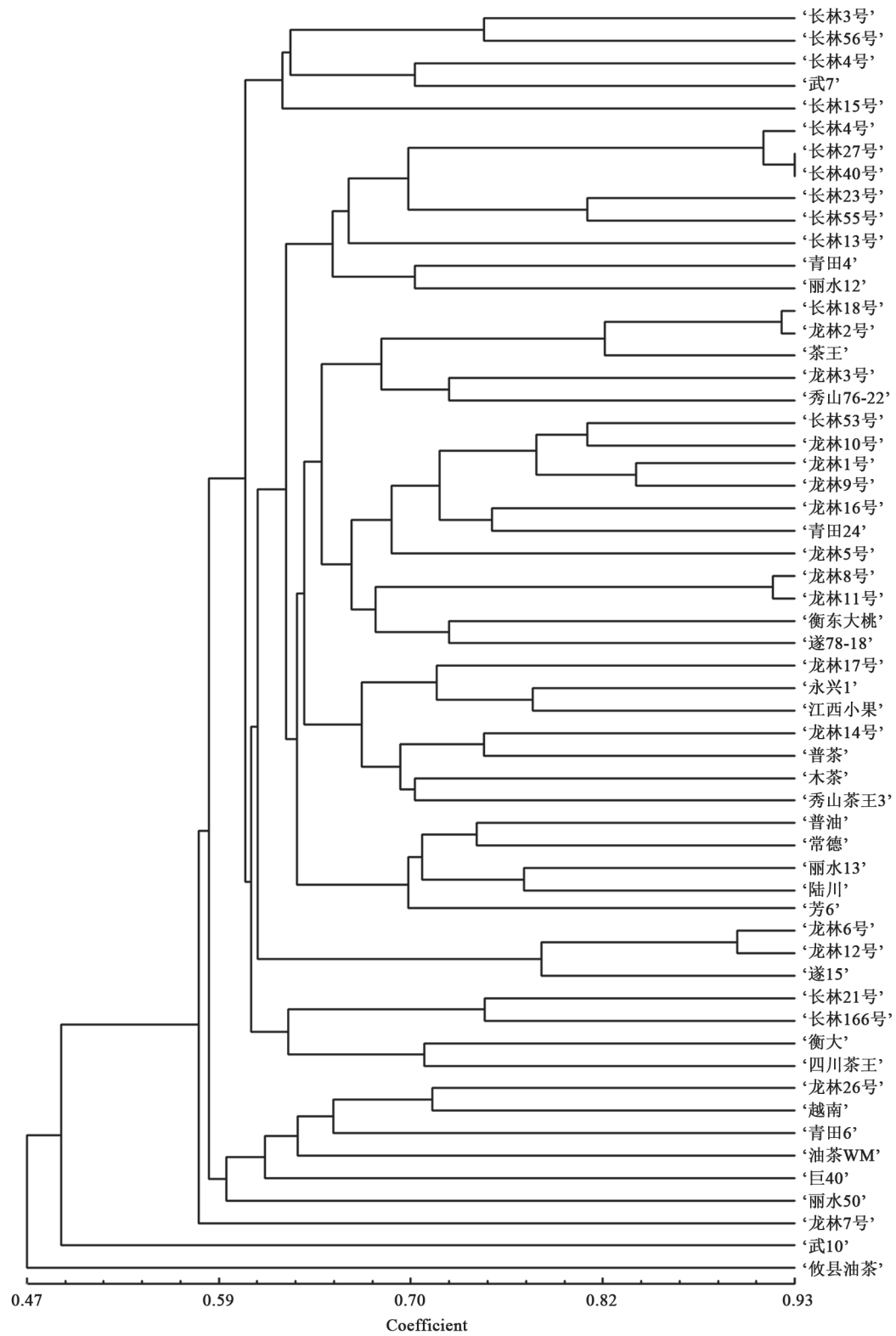


图2 基于SSR标记构建的57个油茶品种聚类树状图

Fig.2 UPGMA cluster diagram of 57 studied *Camellia oleifera* cultivars based on SSR markers



型,且核苷酸重复次数不同。可见在进行物种间SSRs频率对比时,不同筛选参数设定下的检测结果其可比性很差,这也意味着亟需一个统一的SSR检测参数设置标准,以便于更客观有效地在不同物种的研究中进行SSR分布类型与特点的比较。从SSR位点的重复类型与重复基序来看,大多数植物的SSRs主要以二、三核苷酸重复为主,但优势的重复基序有差别。我们鉴定出的普通油茶EST-SSR,除以A/T基序为主、重复12次以上的单核苷酸外,以二核苷酸重复基序最多,其次是三核苷酸。二核苷酸重复基序以AG/CT为主,三核苷酸重复基序以AAG/CTT为主。这与温强等(2013)报道的普通油茶、红花油茶和短柱茶基本一致,但与其他树种相比存在一定差别,这可能与不同物种中相应的编码蛋白的使用频率有关。

尽管来自转录组数据开发的EST-SSR标记用于种质资源遗传多样性评价的优势非常明显,但由于EST-SSR是来自相对保守的功能基因表达序列,因此与来自基因组数据的Genomic-SSR标记相比较,EST-SSR标记的多态性较低(Eujayl等2001)。本研究从143对SSR引物中仅遴选出了20对(14%)扩增效率高、稳定性好的多态性SSR引物也印证了这一点。因此,为了弥补EST-SSR标记本身多态性较低的不足,还需要继续从本研究高通量测序获得的EST序列,以及公共数据库中不断增加的EST序列进一步开发出更多高质量的SSR标记,以满足油茶遗传育种工作的需要。

对许多植物的研究业已表明,SSR标记在属内种间乃至属间的基因组具有一定的通用性(可转移性)(Yasodha等2005;陈金金和彭俊华2011;Lesser等2012;邱道寿等2013;邓绍勇等2014;卢玉飞等2015;张华丽等2016)。在本研究中,我们从29对来自茶属内的红花油茶和茶树的SSR引物中,筛选出了17对可扩增出预期大小片段的引物,其中4对引物(Ck03、Ck55、TUGMS17和TUGMS28)具有高度多态性(PIC>0.5)。这表明茶属内不同种间的SSR标记也具有较高的可转移性,可以相互验证使用。茶属SSR标记的通用性极大地减轻了SSR引物开发的工作量,为廉价、快捷获得SSR标记提供了可用资源,同时也为研究茶属种间进化关系提供了有效途径。

用于SSR引物设计的转录组EST序列来自于并不包含内含子的基因表达数据RNA,但在对SSR位点扩增的DNA序列中可能包含内含子,这样就可能出现跨越内含子的无效扩增,或出现与目标重复序列长度不一致的情况(Saha等2004)。此外,DNA序列中也会存在没有表达的基因序列,在这些SSR位点也会导致无效扩增,或者在同源性很高的非SSR位点处扩增。我们对于普通油茶89对EST-SSR引物的验证也很好说明了这一点。在自主设计的89对SSR引物中,有高达57对扩增失败。在23对扩增效率高、稳定性好的多态性SSR引物中,3对扩增的PCR产物测序表明非目标重复序列,而且20对成功的均为二、三核苷酸重复。因此,对于SSR标记的有效性验证尤为关键。首先要排除无效扩增,其次是多态性扩增验证,最后要排除非目标重复序列。此外也提示对于普通油茶SSR引物设计,要侧重于含二、三核苷酸重复的低级单元序列,同时验证了低级单元SSR的多态性普遍高于高级单元(Dreisigacker等2004)的推断。

本研究克服了聚丙烯酰胺凝胶电泳方法对SSR片段人工判别的误差,利用SSR荧光标记毛细管电泳技术,印证了利用普通油茶转录组数据开发SSR标记的可行性,以及茶属树种SSR标记的可转移性。本研究开发的SSR标记不仅为普通油茶遗传多样性分析和遗传图谱构建提供了有价值的候选标记,也为普通油茶良种的鉴别与分子选育提供了分子技术手段。

#### 参考文献

- Botstein D, White RL, Skolnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment polymorphisms. *Am J Human Genet*, 32 (3): 314-331
- Chen JJ, Peng JH (2011). Transferability of wheat (*Triticum aestivum*) EST-derived SSR markers to several potential energy plants in Gramineae. *Plant Sci J*, 29 (6): 696-703 (in Chinese with English abstract) [陈金金, 彭俊华(2011). 小麦EST-SSR标记对几种有潜力的禾本科能源植物的可转移性研究. *植物科学学报*, 29 (6): 696-703]
- Chen YZ, Yang XH, Peng SF, Li DX, Wang XN, Duan W (2005a). The current status and development strategy of fine varieties of *Camellia oleifera* breeding in China. *J For Engin*, 19 (4): 1-4 (in Chinese with English abstract) [陈永忠, 杨小胡, 彭邵锋, 李党训, 王湘南, 段玮(2005a). 我国油茶良种选育研究现状及发展策略. *林业科技开发*, 19 (4): 1-4]
- Chen YZ, Zhang ZJ, Tan XF (2005b). Identification of oil tea (*Camellia oleifera*) superior clones by RAPD molecular marker. *J Cent*

- South For Univ, 25 (4): 40–45 (in Chinese with English abstract) [陈永忠, 张智俊, 谭晓风(2005b). 油茶优良无性系的RAPD分子鉴别. 中南林学院学报, 25 (4): 40–45]
- Deng SY, Wen Q, Li KQ, Ye JS, Zhu PL (2014). Screening and application of SSR primers in plants of *Clerodendrum* L. Chin Tradit Herbal Drugs, 45 (22): 3317–3322 (in Chinese with English abstract) [邓绍勇, 温强, 李康琴, 叶金山, 朱培林(2014). 大青属植物通用性SSR引物筛选及应用. 中草药, 45 (22): 3317–3322]
- Dreisigacker S, Zhang P, Warburton ML, Van Ginkel M, Hoisington D, Bohn M, Melchinger AE (2004). SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted different megaenvironments. Crop Sci, 44 (2): 381–388
- Eujayl I, Sorrells M, Baum M, Wolters P, Powell W (2001). Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. Euphytica, 119 (12): 39–43
- Fan WQ, Gai HM, Sun X, Yang AG, Zhang ZF, Ren M (2016). DataFormater, a software for SSR data formatting to develop population genetics analysis. Mol Plant Breeding (online), 14 (5): 1029–1034 (in Chinese with English abstract) [樊文强, 盖红梅, 孙鑫, 杨爱国, 张忠锋, 任民(2016). SSR数据格式转换软件DataFormater. 分子植物育种(网络版), 14 (5): 1029–1034]
- Huang HY, Du HY, Wuyun TN, Liu PF (2013). Development of SSR molecular markers based on transcriptome sequencing of *Eucommia ulmoides*. Sci Silv Sin, 49 (5): 176–181 (in Chinese with English abstract) [黄海燕, 杜红岩, 乌云塔娜, 刘攀峰(2013). 基于杜仲转录组序列的SSR分子标记的开发. 林业科学, 49 (5): 176–181]
- Huang Y (2013). Population genetic structure and interspecific introgressive hybridization between *Camellia meiocarpa* and *C. oleifera*. Chin J Appl Ecol, 24 (8): 2345–2352 (in Chinese with English abstract) [黄勇(2013). 小果油茶与普通油茶居群遗传结构及种间杂交渐渗. 应用生态学报, 24 (8): 2345–2352]
- Huang YF, Chen XM, Zhuang XY, Lei ZG, Chen YZ, Peng SF (2006). Analysis of genetic diversity in *Camellia oleifera* germplasm. Sci Silv Sin, 42 (4): 38–43 (in Chinese with English abstract) [黄永芳, 陈锡沐, 庄雪影, 雷治国, 陈永忠, 彭邵峰(2006). 油茶种质资源遗传多样性分析. 林业科学, 42 (4): 38–43]
- Jia B, Lin Q, Zhang L, Tan X, Lei X, Hu X, Shao F (2014). Development of 15 genic-SSR markers in oil-tea tree (*Camellia oleifera*) based on transcriptome sequencing. Genetika, 46 (3): 789–797
- Kalia R, Rai M, Kalia S, Singh R, Dhawan AK (2011). Microsatellite markers: an overview of the recent progress in plants. Euphytica, 177 (3): 309–334
- Lesser MR, Parchman TL, Buerkle CA (2012). Cross-species transferability of SSR loci developed from transcriptome sequencing in lodgepole pine. Mol Ecol Resour, 12 (3): 448–455
- Li MT, Zhang SL, Deng P, Hou XL, Wang JJ (2015). Analysis on SSR information in transcriptome of onion and the polymorphism. Act Hort Sin, 42 (6): 1103–1111 (in Chinese with English abstract) [李满堂, 张仕林, 邓鹏, 侯喜林, 王建军(2015). 洋葱转录组SSR信息分析及其多态性研究. 园艺学报, 42 (6): 1103–1111]
- Lin P, Yao XH, Wang KL, Zheng TT, Teng JH (2010). Identification and genetic analysis of *Camellia oleifera* changlin series superior clones by SRAP molecular marker. J Agr Biotech, 18 (2): 272–279 (in Chinese with English abstract) [林萍, 姚小华, 王开良, 郑婷婷, 滕建华(2010). 油茶长林系列优良无性系的SRAP分子鉴别及遗传分析. 农业生物技术学报, 18 (2): 272–279]
- Liu Y, Peng Y, He R, Li LS, Tang GH, Wang CY, Xu M, Zeng CB, Tang H (2016). Genetic diversity of *Camellia oleifera* in Hainan by ISSR. Mol Plant Breeding, (2): 517–523 (in Chinese with English abstract) [刘扬, 彭赞, 贺瑞, 李连胜, 唐光华, 王辰尧, 徐敏, 曾灿彬, 汤华(2016). 基于ISSR标记的海南油茶种质资源遗传多样性分析. 分子植物育种, (2): 517–523]
- Lou HF (2014). Analysis and strategy of development of *Camellia Oleifera* industry in Zhejiang Province based on SWOT Method. Mod Agr Sci Tech, 5: 328–330 (in Chinese) [娄浩峰(2014). 浙江省油茶产业发展的SWOT分析及其策略. 现代农业科技, 5: 328–330]
- Lu YF, Jiang JX, Yi ZL (2015). Study on the transferability of maize SSR and sugarcane EST-SSR markers to *Miscanthus* (Poaceae). Acta Pratacult Sin, 21 (5): 86–95 (in Chinese with English abstract) [卢玉飞, 蒋建雄, 易自力(2015). 玉米SSR引物和甘蔗EST-SSR引物在芒属中的通用性研究. 草业学报, 21 (5): 86–95]
- Luo R, Wu WL, Zhang Y, Li YH (2010). SSR marker and its application to crop genetics and breeding. Gen Appl Biol, 29 (1): 137–143 (in Chinese with English abstract) [罗冉, 吴委林, 张昉, 李玉花(2010). SSR分子标记在作物遗传育种中的应用. 基因组学与应用生物学, 29 (1): 137–143]
- Peng C, Li ZF, Chen HL, Li AH, Zhang XY (2013). Genetic analysis of the germplasm of *Camellia oleifera* from Hubei Province based on SSR marker. Hubei For Sci Tech, 42 (5): 1–5 (in Chinese with English abstract) [彭婵, 李振芳, 陈慧玲, 李爱华, 张新叶(2013). 湖北油茶种质资源SSR分析. 湖北林业科技, 42 (5): 1–5]
- Peng SF, Zhang DQ, Chen YZ, Song ZD, Han X, Zhang HY (2011). Genetic diversity analysis of elite cultivars of *Camellia oleifera* by SRAP. J Cent South Forestry Univ, 31 (1): 80–85 (in Chinese with English abstract) [彭邵峰, 张党权, 陈永忠, 宋志丹, 韩欣, 章怀云(2011). 14个油茶良种遗传多样性的SRAP分析. 中南林业科技大学学报, 31 (1): 80–85]
- Qiu DS, Zheng XL, Cai SK, Zheng JR, Luo HM, Zhang L, Deng RY, Li W, Liu XJ (2013). Development and transfer analysis of SSR in *Dendrobium*. Plant Sci J, 31 (5): 500–509 (in Chinese with English abstract) [邱道寿, 郑希龙, 蔡时可, 郑锦荣, 罗焕明, 张蕾, 邓瑞云, 李武, 刘晓津(2013). 石斛SSR标记的开发及可转移性分析. 植物科学学报, 31 (5): 500–509]
- Saha MC, Mian MA, Eujayl I, Zwonitzer JC, Wang L, May GD (2004). Tall fescue EST-SSR markers with transferability across several grass species. Theor Appl Genet, 109 (4): 783–791
- Wang BM, Chen YZ, Tan XF, Peng SF, Shi MW (2008). Genetic diversity of elite clones of *Camellia oleifera* by ISSR. J Northeast Forestry Univ, 36 (6): 19–23 (in Chinese with English abstract) [王保明, 陈永忠, 谭晓风, 彭邵峰, 石明旺(2008). 应用ISSR分析油茶无性系的遗传多样性. 东北林业大学学报, 36 (6): 19–23]

- Wang YY, Shan WQ, Xu WL, Cui CS, Qu SP (2016). Analysis on SSR information in transcriptome and the polymorphism of *Cucurbita maxima*. *Act Hort Sin*, 43 (3): 578–586 (in Chinese with English abstract) [王洋洋, 单文琪, 徐文龙, 崔崇士, 屈淑平 (2016). 印度南瓜转录组SSR信息分析及其多态性研究. *园艺学报*, 43 (3): 578–586]
- Wen Q, Lei XL, Ye JS, Jiang M, Zuo JL, Huang LL, Jiang XM, Xu LC (2008). Identification of *Camellia oleifera* superior clones by ISSR molecular markers. *J Cent South For Univ*, 28 (1): 39–43 (in Chinese with English abstract) [温强, 雷小林, 叶金山, 江梅, 左继林, 黄丽莉, 江香梅, 徐林初(2008). 油茶高产无性系的ISSR分子鉴别. *中南林业科技大学学报*, 28 (1): 39–43]
- Wen Q, Xu L, Gu Y, Huang M, Xu L (2012). Development of polymorphic microsatellite markers in *Camellia chekiangoleosa* (Theaceae) using 454-ESTs. *Am J Bot*, 99 (5): e203–e205
- Wen Q, Xu LC, Jiang XM, Li J, Gu YC, Xu LA, Huang MR (2013). Survey and analysis of microsatellites from DNA sequences in *Camellia* species using 454 pyrosequencing. *Sci Silv Sin*, 49 (8): 43–50 (in Chinese with English abstract) [温强, 徐林初, 江香梅, 李江, 顾胤聪, 徐立安, 黄敏仁(2013). 基于454测序的油茶DNA序列微卫星观察与分析. *林业科学*, 49 (8): 43–50]
- Yasodha R, Ghosh M, Sumathi R, Gurumurthi K (2005). Cross-species amplification of eucalyptus SSR markers in Casuarinaceae. *Acta Bot Croat*, 64 (1): 115–120
- Yeh FC, Yang R, Boyle TJ, Ye Z, Xiyang JM (2000). PopGene32, Microsoft Windows-based freeware for population genetic analysis, version 1.32. Molecular Biology and Biotechnology Centre, Edmonton, Alberta, Canada: University of Alberta. Available at [http://www.ualberta.ca/~fyeh/popgene\\_download.html](http://www.ualberta.ca/~fyeh/popgene_download.html)
- Yuan YY, Wang QF, Chen JM (2013). Development of SSR markers in aquatic plant *Nymphoides peltata* (Menyanthaceae) based on information from transcriptome sequencing. *Plant Sci J*, 31 (5): 485–492 (in Chinese with English abstract) [袁阳阳, 王青锋, 陈进明(2013). 基于转录组测序信息的水生植物蓴菜SSR标记开发. *植物科学学报*, 31 (5): 485–492]
- Yu XY, Yu FY, Liu JB, Chen JY (2013). Identification and genetic diversity analysis of *Camellia oleifera* varieties using ISSR marker. *J Nanjing For Uni (Nat Sci)*, 37 (1): 61–66 (in Chinese with English abstract) [于小玉, 喻方圆, 刘建兵, 陈军勇(2013). ISSR在油茶品种鉴别和遗传多样性分析中的应用. *南京林业大学学报*, 37 (1): 61–66]
- Zhang EH, Wang XY, Qin ZH, Zhao WD, Wei CJ, Wang PL (2016). Genetic diversity analysis of *Camellia oleifera* in Guangxi using SSR markers. *Guihaia*, 36 (7): 806–811 (in Chinese with English abstract) [张恩慧, 王晓云, 覃子海, 赵文东, 韦长江, 王鹏良 (2016). 广西普通油茶种质资源遗传多样性的SSR分析. *广西植物*, 36 (7): 806–811]
- Zhang GW, Zhong WB, Wuyun TN, Tan XF, Du TZ (2007). Identification of oil tea (*Camellia oleifera*) superior clones by ISSR molecular marker. *Fore Res*, 20 (2): 278–282 (in Chinese with English abstract) [张国武, 钟文斌, 乌云塔娜, 谭晓风, 杜天真 (2007). 油茶优良无性系ISSR分子鉴别. *林业科学研究*, 20 (2): 278–282]
- Zhang HL, Wang T, Song LN, Dong AX, Xin HB, Yi MF, Zhao JH (2016). Study on EST-derived SSR marker development of *Chrysanthemum morifolium* and its transferability to *Tagetes erecta* L. *J Nucl Agric Sci*, 30 (1): 42–49 (in Chinese with English abstract) [张华丽, 王涛, 宋利娜, 董爱香, 辛海波, 义鸣放, 赵金华(2016). 菊花EST-SSR标记开发及其对万寿菊的可转移性研究. *核农学报*, 30 (1): 42–49]
- Zhang T, Liu SQ, Mei H, Dong YL (2011a). Genetic diversity of *Camellia oleifera* from Hubei Province revealed by SRAP analysis. *J Cent China Nor Uni (Nat Sci)*, 45 (3): 477–484 (in Chinese with English abstract) [张婷, 刘双青, 梅辉, 董妍玲(2011a). 湖北省油茶资源遗传多样性的SRAP分析. *华中师范大学学报*, 45 (3): 477–484]
- Zhang T, Liu SQ, Mei H, Dong YL, Lü MZ (2011b). AFLP analysis on the genetic diversity of *Camellia oleifera* from different regions of Hubei Province. *J Anhui Agr Sci*, 39 (23): 14070–14071, 14075 (in Chinese with English abstract) [张婷, 刘双青, 梅辉, 董妍玲, 吕明治(2011b). 湖北省不同地区油茶遗传多样性的AFLP分析. *安徽农业科学*, 39 (23): 14070–14071, 14075]
- Zhang Z, Zhang HG, Mo C, Zhang L (2015). Transcriptome sequencing analysis and development of EST-SSR markers for *Pinus koraiensis*. *Sci Silv Sin*, 51 (8): 114–120 (in Chinese with English abstract) [张振, 张含国, 莫迟, 张磊(2015). 红松转录组SSR分析及EST-SSR标记开发. *林业科学*, 51 (8): 114–120]

## Development of EST-SSR molecular markers based on transcriptome sequencing of *Camellia oleifera*

LI Hai-Bo<sup>1</sup>, WANG Shan<sup>2</sup>, DING Hong-Mei<sup>3</sup>, CHEN You-Wu<sup>1</sup>, HU Chuan-Jiu<sup>1</sup>, LI Nan<sup>1</sup>, WEI Hai-Long<sup>1,\*</sup>

<sup>1</sup>Zhejiang Academy of Forestry, Hangzhou 310023, China; <sup>2</sup>School of Forestry and Bio-technology, Zhejiang Agriculture and Forestry University, Lin'an, Zhejiang 311300, China; <sup>3</sup>Analysis and Test Center, Zhejiang Chinese Medical University, Hangzhou 310053, China

**Abstract:** Identification of EST-SSR loci based on transcriptome data and development of polymorphic SSR markers suitable for evaluation and application of germplasm resources on *Camellia oleifera* are critical for the study on genetic variation of *C. oleifera* and molecular marker assisted breeding in superior cultivars of *C. oleifera*. A total of 26 751 SSR loci, including the types of 1–6 nucleotide repeats with occurring frequency of 1/2.33 kb, were identified from 67 434 unigenes (63.39 MB) in *C. oleifera* transcriptome by using MISA software. The distribution frequency of SSRs was 39.67%. Dinucleotide repeat was the main type, accounted for as much as 44.28% of all SSRs, followed by mononucleotide repeat (35.28%) and trinucleotide repeat (18.27%). Using agarose electrophoresis and capillary electrophoresis, 143 SSR markers were detected, among which 20 can produce clear, reproducible and polymorphic bands. Genetic relationship analysis among 57 *C. oleifera* cultivars from Zhejiang production areas was performed with the 20 fluorescent labeled SSR primer pairs. The results showed 129 alleles were detected from the 20 SSR loci, the number of alleles detected by each SSR primer pairs ranged from 2 to 17, with an average of 6.45. Polymorphism information content (PIC) ranged from 0.0499–0.8635, with an average of 0.4355. Fifty-seven *C. oleifera* cultivars were clustered into one group at a genetic similarity coefficients (GSC) level by *Dice* of 0.47, seven groups at 0.60. Our present study verified that using *C. oleifera* transcriptome data to develop SSR markers was feasible, also proved that the transferability of SSR markers across species of *Camellia*. The SSR markers obtained in the study could be valuable candidate markers for genetic diversity analysis and genetic mapping construction in *C. oleifera*, also be used as a molecular technical tool for identification and molecular marker assisted breeding in superior cultivars of *C. oleifera*.

**Key words:** *Camellia oleifera*; transcriptome; EST-SSR; polymorphism; genetic relationship

Received 2017-04-17 Accepted 2017-06-08

This work was supported by the Key Project of Science and Technology Programme of Zhejiang Science Technology Department, China (Grant Nos. 2014F30002 and 2013F50012), and the Key Science and Technology Special Project of Zhejiang Province (Grant No. 2010C02005-1).

\*Corresponding author (E-mail: 1178646251@qq.com).