

DOI:10.13332/j.1000-1522.20190101

基于山茶转录组的 SSR 标记开发及亲缘关系分析

潘丽芹^{1,2} 李纪元¹ 李绍翠¹ 范正琪¹ 殷恒福¹ 何丽波^{1,3}

(1. 中国林业科学研究院亚热带林业研究所, 浙江 富阳 311400; 2. 台州科技职业学院园艺植物研究所, 浙江 台州 318020;
3. 湖南农业大学园艺园林学院, 湖南 长沙 410128)

摘要:【目的】开发基于山茶转录组的 EST-SSR 分子标记, 并应用该标记进行山茶种质的亲缘关系分析。【方法】本研究以山茶叶片转录组测序所获得的 50 518 条 Unigenes 为背景数据, 利用 MISA 软件搜索 SSR 标记, 设计并筛选 SSR 多态性引物, 对 8 份山茶种质进行了 UPGMA 聚类。【结果】共检索到 13 197 个 SSR 位点, SSR 的发生频率和分布频率分别为 19.52%、26.12%, 平均分布距离为 4.33 kb; 在 6 种 SSR 重复类型中, 以单、二、三核苷酸这 3 种重复类型为主, 分别占 48.420%、34.917% 和 15.473%; 出现频率高的基序类型为 A/T、AG/CT 和 AT/TA, 占总 SSR 位点的 79.61%; 三核苷酸中以 AAG/CTT、ACC/GGT、AAT/ATT 类型居多; 在设计出的 10 974 对引物中, 随机选用其中 90 对引物进行多态筛选, 73 对引物扩增产物与预期大小一致, 有效扩增率为 81.11%, 29 对引物存在扩增多态性, 占 39.73%。扩增得到 72 个等位基因, 有效等位基因数平均达到 3.264; 观察杂合度和期望杂合度的平均值分别为 0.208 和 0.638, 引物多态信息含量平均为 0.496; 8 份山茶种质在相似系数为 0.58 时, 可以分为 4 类: 山茶原种为 I 类, ‘黑蛋石’、‘红叶黑魔法’和‘黑骑士’为 II 类, ‘黑魔法’和‘金华美女’分别为 III 类和 IV 类。【结论】山茶转录组测序的 Unigenes 信息可作为 SSR 标记开发的有效来源, 为山茶的遗传多样性研究、亲缘关系鉴定以及分子标记辅助育种等奠定了一定的理论基础。

关键词: 山茶; 转录组; SSR; 标记; 亲缘关系

中图分类号: S794.9 文献标志码: A 文章编号: 1000-1522(2019)07-0111-10

引文格式: 潘丽芹, 李纪元, 李绍翠, 等. 基于山茶转录组的 SSR 标记开发及亲缘关系分析 [J]. 北京林业大学学报, 2019, 41(7):111-120. Pan Liqin, Li Jiyuan, Li Shaocui, et al. Development of SSR markers based on transcriptome of *Camellia japonica* and analysis of genetic relationship [J]. Journal of Beijing Forestry University, 2019, 41(7): 111-120.

Development of SSR markers based on transcriptome of *Camellia japonica* and analysis of genetic relationship

Pan Liqin^{1,2} Li Jiyuan¹ Li Shaocui¹ Fan Zhengqi¹ Yin Hengfu¹ He Libo^{1,3}

(1. Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Fuayang 311400, Zhejiang, China;
2. Research Institute of Horticultural Plants, Taizhou Vocational College of Science & Technology, Taizhou 318020, Zhejiang, China;
3. Horticulture & Landscape College, Hunan Agricultural University, Changsha 410128, Hunan, China)

Abstract: [Objective] Develop the EST-SSR markers based on transcriptome sequencing and apply these markers to analyze the genetic relationships of *Camellia* germplasms. [Method] Using 50 518 Unigenes obtained from the transcriptome sequencing of *C. 'Red Leaved Black Magic'* leaves as background data, SSR markers were searched out, and polymorphic primers were then designed and screened to construct the UPGMA evolutionary tree for 8 *Camellia* germplasms. [Result] 13 197 SSR sites were founded out. The

收稿日期: 2019-03-26 修回日期: 2019-05-26

基金项目: 浙江省农业(林木)新品种选育重大科技专项(2016C02056-12), 政府间国际科技创新合作重点专项(2016YFE0126100)。

第一作者: 潘丽芹, 副教授。主要研究方向: 园林植物分子育种。Email: 826673874@qq.com 地址: 318020 浙江省台州市黄岩区嘉木路 288 号台州科技职业学院。

责任作者: 李纪元, 研究员, 博士生导师。主要研究方向: 园林植物分子育种。Email: jiyuan_li@126.com 地址: 311400 浙江省杭州市富阳区大桥路 73 号中国林业科学研究院亚热带林业研究所。何丽波, 副教授。主要研究方向: 风景园林。Email: helibo@hunau.edu.cn 地址: 410128 湖南省长沙市芙蓉区农大路 1 号湖南农业大学园艺园林学院。

本刊网址: <http://j.bjfu.edu.cn>; <http://journal.bjfu.edu.cn>

occurrence frequency and the distribution frequency were estimated to be 19.52% and 26.12%, respectively, with the average distribution distance of 4.33 kb. Among the 6 types of SSR repeat motifs, mono-nucleotide, di-nucleotide and tri-nucleotide were the frequent dominant motifs, accounting for 48.420%, 34.917% and 15.473%, respectively. The types of motifs with high occurrence frequency included A/T, AG/CT and AT/TA, accounting for 79.61% of the total SSR loci; AAG/CTT, ACC/GGT, AAT/ATT were in the majority among the tri-nucleotides; 10 974 pairs of primers were designed, and 73 pairs of primers were finally used to amplify effectively based on 90 pairs selected randomly, resulting in the rate of effective amplification of 81.11%; 29 primers showed some amplification polymorphism, accounting for 39.73%. The average number of effective alleles was estimated to be 3.264 with a total of 72 alleles. The mean values of observed heterozygosity (H_o) and expected heterozygosity (H_e) were measured to 0.208 and 0.638 respectively, as well as the average polymorphic information content (PIC) of 0.496; genetic relationship and cluster analysis indicated that eight *Camellia* germplasms could be classified into four categories at a similarity coefficient of 0.58: wild *C. japonica* in group I; *C. 'Black Opal'*, *C. 'Red Leaved Black Magic'* and *C. 'Night Rider'* together in group II; *C. japonica 'Black Magic'* and *C. japonica 'Jinhua Meinü'* in group III and group IV respectively. **[Conclusion]** The result showed that *Camellia* transcriptome Unigenes can be used as an effective source for SSR markers exploiting, which lays a theoretical foundation for studies on genetic diversity, identification of genetic relationship and marker-assisted breeding of *Camellia*.

Key words: *Camellia japonica*; transcriptome; SSR; marker; genetic relationship

随着高通量测序技术的飞速发展,通过转录组数据挖掘 EST-SSR(Expressed sequence tag SSR, EST-SSR)信息,能开发出大量而有效的 SSR 标记。由于该技术从功能基因入手,可直接反映功能基因的多样性,具有技术简便、共显性表达、重复性好、省时快速、引物开发成本低廉^[1]等优点,已被广泛应用于植物种质遗传关系分析与鉴定^[2]、遗传多样性分析^[3]和遗传连锁图谱构建等研究领域^[4-7]。

山茶(*Camellia japonica*)是山茶科(Theaceae)山茶属中以观赏为主要栽培目的的物种、变种和品种的统称。据记载世界茶花品种已达 2.5 万余种,但很多品种间形态性状相似,遗传背景不明,尤其是生产中,同名异物,同物异名现象导致品种混淆,给产业开发带来不利影响。因此利用高效的分子标记技术,对厘清种质遗传背景,鉴别品种真伪,保护品种权益则显得尤为重要。目前,SSR 标记在山茶中的应用已逐步展开。如张亚利等^[8]利用 20 对 SSR 引物对 33 份山茶属植物种质的遗传多样性进行了研究;胡兴华等^[9]进行了茶花品种 SSR 指纹图谱构建的初步研究;李琳琳等^[10]利用 SSR 标记技术鉴定了茶花品种杂交 F1 代的真伪。基于转录组测序的 SSR 标记在山茶中的应用研究则未见报道。本文在山茶品种‘红叶黑魔法’(*C. 'Red Leaved Black Magic'*)叶片转录组测序的基础上,利用丰富的转录组数据,深入挖掘 SSR 标记,同时选取花色为红、黑红,且新叶均为红叶的几份近缘种质,进行亲缘关系分析,探讨该类标记在山茶属植物研究中的应用价

值,以期对山茶种质资源遗传多样性、品种鉴定、山茶属植物亲缘关系等研究提供一定的理论参考。

1 材料与方法

1.1 试验材料

转录组测序所用材料为山茶品种‘红叶黑魔法’叶片,样品经液氮处理后用于 RNA 提取,送交北京百迈客公司进行 RNA-Seq 转录组测序,以此为背景数据进行 SSR 标记的搜索与开发。选择 8 份山茶种质用于遗传多样性分析。于 2018 年 3—4 月间采集每份种质的幼嫩叶片,每份种质样品从 3 个不同植株上采样,作为 3 次生物学重复。经液氮速冻后带回实验室,用于基因组 DNA 的提取。SSR 扩增所用的材料见表 1。

1.2 试验方法

1.2.1 基因组 DNA 的提取

采用北京艾德莱生物科技有限公司(Aidlab Biotechnologies Co., Ltd)植物基因组快速提取试剂盒提取植物 DNA。用 NanoDrop2000 紫外分光光度计(Thermo Fisher 公司)检测 DNA 浓度和纯度;1% 琼脂糖凝胶电泳检测 DNA 质量,用于 SSR 标记分析。

1.2.2 转录组 SSR 位点的鉴别及 SSR 引物设计

本研究利用 MISA(<http://pgrc.jp/gatersleben.de/misa/>)对拼接组装得到的 Unigenes 进行 SSR 位点识别,搜索重复单元长度为 1~6 bp,搜索标准为一、二、三、四、五、六核苷酸,最少重复次数分别为

表 1 山茶材料及采样地点

Tab. 1 List of *Camellia* collections and locations

编号 No.	名称 Genotype	种源地 Origin	编号 No.	名称 Genotype	种源地 Origin
1	山茶 <i>C. japonica</i>	浙江杭州 Zhejiang Hangzhou	5	黑魔法 <i>C. japonica</i> 'Black Magic'	浙江金华 Zhejiang Jinhua
2	黑蛋石 <i>C.</i> 'Black Opal' ^[11-13]	浙江杭州 Zhejiang Hangzhou	6	黑魔法 <i>C. japonica</i> 'Black Magic' ^[11-13]	浙江杭州 Zhejiang Hangzhou
3	红叶黑魔法 <i>C.</i> 'Red Leaved Black Magic'	浙江杭州 Zhejiang Hangzhou	7	黑骑士 <i>C.</i> 'Night Rider' ^[11-13]	浙江杭州 Zhejiang Hangzhou
4	金华美女 <i>C. japonica</i> 'Jinhua Meinü' ^[14]	浙江杭州 Zhejiang Hangzhou	8	黑骑士 <i>C.</i> 'Night Rider'	浙江金华 Zhejiang Jinhua

10、6、5、5、5、5 次, 复合微卫星位点之间最大间隔碱基数为 100 bp。

利用 Primer 3.0 软件, 以 SSR 重复单元前后的序列进行引物设计, 每个 SSR 位点设计出 3 条引物, 引物序列长度 18~27 bp, GC 含量 40%~60%, 退火温度 55~65 °C, 预期扩增产物长度 100~280 bp, 尽量避免出现发夹结构和二聚体。引物设计完成后, 随机挑选 90 对引物送擎科生物工程(杭州)有限公司合成。

1.2.3 SSR-PCR 分析

扩增反应在 Thermal Cycler Dice PCR 仪(日本)上进行。25 μL 反应体系含有: 50~90 ng 模板 DNA 2 μL, Mix 12.5 μL, 引物(10 mmol/L)2 μL, 蒸馏水 9 μL。PCR 反应程序: 94 °C 预变性 5 min, 94 °C 变性 30 s, 55 °C 退火 30 s, 72 °C 延伸 30 s, 共 35 个循环, 最后在 72 °C 延伸 9 min。采用 Bioptic Qsep100 全自动核酸分析系统(Bioptic, Inc., 台湾)进行 PCR 产物检测, 并利用该系统自带的 Bioptic Qseq100 软件对 SSR 条带进行判读和分析。

1.2.4 数据统计与分析

利用 Bioptic Qseq100 的分析和判读结果, 建立原始数据矩阵。利用 GenAlEx 6.501 软件对山茶不同材料间的等位基因数、观察杂合度、期望杂合度及多态信息含量等进行统计; 利用 NTSYS-pc2.10e 软件计算遗传相似系数, 并构建 UPGMA 进化树。

2 结果与分析

2.1 山茶转录组中 SSR 位点的数量及分布特点

山茶‘红叶黑魔法’叶片转录组测序共获得 213.66 Gb 的 Clean Data, 经组装后共获得 50 518 条 Unigenes(序列总长度为 57 120 592 bp)。利用 MISA 软件对 1 kb 以上的 Unigenes 进行 SSR 搜索与分析, 发现有 9 859 条 Unigenes 序列中出现了 13 197 个 SSR 位点。SSR 的发生频率为 19.52%, 分布频率为 26.12%, 平均 4.33 kb 具有一个 SSR 位点。其中, 6 560 条 Unigenes 含有单个 SSR 位点, 占含有 SSR 位点 Unigenes 总数的 66.54%, 3 299 条 Unigenes 上出现了含有 2 个或 2 个以上的 SSR 位点, 占 33.46%。

山茶叶片转录组序列中的 SSR 类型较为丰富, 单核苷酸至六核苷酸重复类型均有出现, 但以单、二、三核苷酸这 3 种重复类型为主, 占 SSR 位点总数的 98.81%(表 2)。单核苷酸重复类型数量最大, 为 6 390 个, 占 SSR 位点总数的 48.420%; 其次是二核苷酸和三核苷酸重复类型, 分别为 4 608 个和 2 042 个, 占总数的 34.917% 和 15.473%; 而四、五、六核苷酸重复类型所占比例极低, 仅为 1.190%。出现频率最高的是单核苷酸重复序列, 为 12.649%, 其次是二、三核苷酸重复序列, 分别为 9.121% 和 4.042%, 而五、六核苷酸重复序列的出现频率最低, 仅为 0.028% 和 0.029%。山茶叶片转录组中 SSR 位

表 2 山茶叶片转录组中 SSR 的分布情况

Tab. 2 Distribution of SSR loci in the transcriptome of *Camellia* leaves

SSR 类型 SSR type	数量 Number	比例 Percentage/%	出现频率 Distribution frequency/%	平均距离 Average distance/kb
单核苷酸 Mono-nucleotide	6 390	48.420	12.649	8.94
二核苷酸 Di-nucleotide	4 608	34.917	9.121	12.40
三核苷酸 Tri-nucleotide	2 042	15.473	4.042	27.97
四核苷酸 Tetra-nucleotide	128	0.970	0.253	446.25
五核苷酸 Penta-nucleotide	14	0.106	0.028	4 080.04
六核苷酸 Hexa-nucleotide	15	0.114	0.029	3 808.04
总计 Total	13 197	100	26.123	4.33

点间的平均距离为 4.33 kb。

从 SSR 的重复次数来看,以 6 次和 10 次重复数量最多,分别为 2 359 和 2 358 个,占总位点的 17.88%, ≥ 17 次重复的数量仅有 181 个。从图 1 可以看出,山茶转录组中 SSR 的重复次数以 5~14 次重复为主,占总位点数的 93.33%,15~21 次的重复次数仅有 880 个,仅占 6.67%。

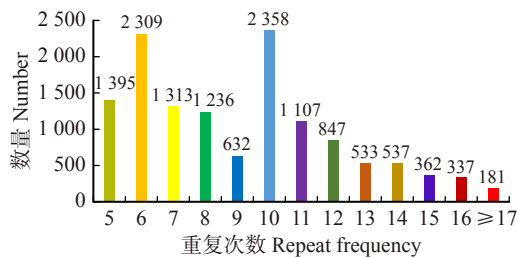


图 1 山茶‘红叶黑魔法’转录组中 SSR 重复次数分布
Fig. 1 Distribution of SSR repeats frequency in *C. 'Red Leaved Black Magic'* transcriptome

2.2 转录组中 SSR 基序重复类型和频率特征

从山茶叶片转录组 SSR 核苷酸基序类型来看,以单核苷酸 A/T 出现频率最高,达 6 301 次,占 SSR 位点总数的 47.746%,占单核苷酸基序总数的 98.61%,C/G 则仅占 1.39%(图 2A)。其次是 AG/CT 和 AT/TA 基序类型,出现了 3 146 和 1 059 次,分别占 SSR 位点总数的 23.839% 和 8.025%,占二核苷酸

重复总数的 68.27% 和 22.98%(图 2B),上述 3 种基序类型是山茶转录组中 SSR 的主要类型,占了总 SSR 位点的 79.61%,而 CG/GC 基序仅出现 5 次,表现出明显的偏倚性。三核苷酸重复基序中以 AAG/CTT、ACC/GGT 和 AAT/ATT 居多,分别占三核苷酸基序总数的 33.40%、24.00% 和 17.88%,占 SSR 总数的 11.647%(图 2C)。而在四、五、六核苷酸重复中,相同基序类型出现较少,出现频率极低。

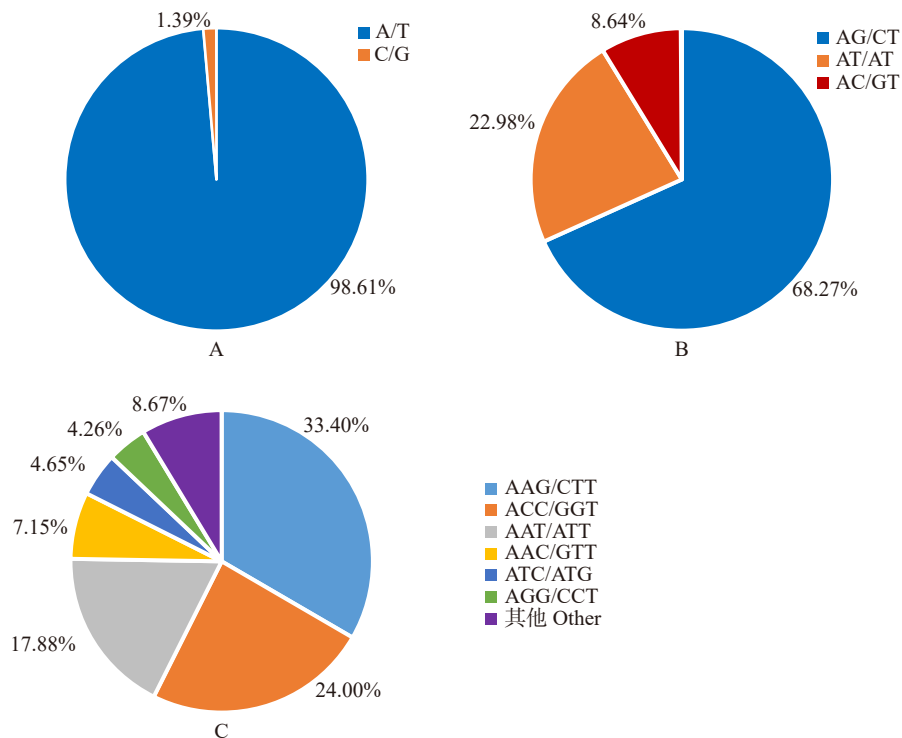
2.3 SSR 引物设计与多态性分析

利用 Primer3.0 对含有 SSR 位点的 9 859 条 Unigenes 序列进行引物设计,共设计出 10 974 对引物。随机挑选 90 对不同重复单元(以二、三、四核苷酸为主)的引物进行 PCR 扩增,其中有 73 对引物能够扩增出与预期产物大小相符的产物,有效扩增率为 81.11%。其中 29 对引物存在扩增多态性,占有效扩增引物的 39.73%,扩增信号如图 3。

利用 29 对多态性引物对 8 份山茶种质进行遗传多样性分析,共扩增得到 72 个等位基因。各 SSR 位点之间的有效等位基因数平均值为 3.264,其变化范围为 1.556~5.818 之间。观察杂合度和期望杂合度的平均值分别为 0.208 和 0.638,引物的多态信息含量平均为 0.496(表 3)。

2.4 山茶种质的遗传多样性分析

遗传相似系数反映了山茶种质间的亲缘关系



A、B、C 分别表示单核苷酸重复、二核苷酸重复、三核苷酸重复的基元比例。A, B, C, represent the ratio of mono-nucleotide repeats, di-nucleotide repeats and tri-nucleotide repeats in *Camellia*.

图 2 山茶转录组 SSR 的基序类型分布
Fig. 2 Distribution of SSR types in *Camellia*

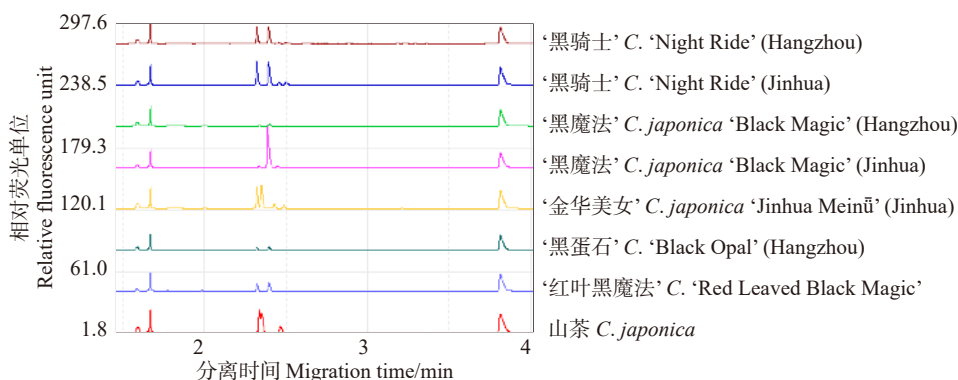


图 3 引物 32 对 8 份山茶种质的 PCR 扩增信号图

Fig. 3 PCR amplification signal figure of primer 32 in 8 *Camellia* germplasm

表 3 山茶转录组 29 对 SSR 引物信息及多态性信息

Tab. 3 29 pairs of SSR primers and polymorphic information of *C. japonica* transcriptome

引物编号 Primer No.	来源 Gene ID	引物序列 Primer sequence	SSR 基元 SSR motif	有效等位基因 Effective alleles	观察杂合度 Observed heterozygosity	期望杂合度 Expected heterozygosity	多态信息含量 Polymorphism information content
P1	c105034_graph_c0	F:CTTCTCTCGATCCACAGCC R:CGATCTCCTCCGTAACAAGC	(GAG)5	4.414	0.250	0.773	0.556
P6	c123499_graph_c0	F:GGACAATCTTTTTGGGAGCA R:CCCTACACAACCAGGAAACC	(ATT)5	1.556	0.143	0.357	0.375
P10	c126182_graph_c0	F:CCTTAACAATCAGCAATGCC R:TGCCATGTACCACATACCCA	(AAC)5	3.122	0.375	0.680	0.455
P12	c129722_graph_c0	F:TCAAAAAGAGACCTTGGGCTG R:GGGGAATTCGATAACACAA	(GTA)7	4.741	0.375	0.789	0.545
P16	c133766_graph_c0	F:GTCCCGAAAAATCCCAAAT R:AATTTGTCTGCAATGGCTCC	(CT)6	4.235	0.500	0.764	0.778
P17	c133791_graph_c0	F:TGAATCACAATCTTGGCTGG R:GGTGGCCTAATACAAGCTGC	(GA)8	2.909	0.500	0.656	0.417
P18	c133791_graph_c0	F:TGAATCACAATCTTGGCTGG R:GGTGGCCTAATACAAGCTGC	(A)12	2.667	0.500	0.625	0.417
P21	c134194_graph_c1	F:ATGGTGCAAGGAATCAAAG R:TGTAAGCTCCTGTGCTGTGG	(TGG)5	4.000	0.750	0.750	0.500
P22	c135100_graph_c0	F:TTCCTCTTTCAAATGCCAATG R:TTAACGGGGAGCCATATCAA	(TA)6	2.844	0.125	0.648	0.333
P24	c135126_graph_c1	F:GTGAAACAAAGCCGGAGAGT R:ACCTGGTTCAATCTATGGCG	(GA)7	5.565	0.625	0.820	0.615
P25	c136711_graph_c1	F:CGTTTCAAGGCAATATCGT R:GCTACCATGAAGCTCCAACC	(TTG)5	2.390	0.429	0.582	0.400
P26	c137110_graph_c0	F:AGATTTGCAAGGTTGGGTTG R:TCTACCACACTCCCCTCCC	(GCA)6	2.579	0.286	0.612	0.375
P28	c137365_graph_c0	F:TGGTTGCTGTTGTTGAGGAG R:CCGCCTAATCAGAACCCTTT	(TGG)6	3.200	0.250	0.688	0.444
P32	c138380_graph_c1	F:TCGTTAAGGCAGCTCTCGAT R:CCACCCTGTGTACGGTCAA	(CGT)7	2.723	0.750	0.633	0.286
P37	c141066_graph_c0	F:ACACACGAACCACTCCATCA R:TTTGGTTGTTGGCATTTCAT	(TAA)6	2.844	0.375	0.648	0.273

表 3(续)
Tab.3(Continued)

引物编号 Primer No.	来源 Gene ID	引物序列 Primer sequence	SSR基元 SSR motif	有效等位基因 Effective alleles	观察杂合度 Observed heterozygosity	期望杂合度 Expected heterozygosity	多态信息含量 Polymorphism information content
P40	c142086.graph_c0	F:AACAATACCCGACTCCTCCC R:CCTATGGCGAGACGTTCAT	(CAT)6	5.333	0.500	0.813	0.583
P41	c142086.graph_c1	F:AATAGCACGGTAATCACGGC R:GAATTTCTGGGCCATCTGA	(CAT)7	4.129	0.125	0.758	0.556
P42	c143100.graph_c0	F:TCCACAATTTCCACCTTA R:CCAGTCAAGCCCTGTAGCTC	(ACC)5	3.200	0.250	0.688	0.500
P44	c143532.graph_c1	F:CACCATCACCAAAGAAGGCT R:TGCAAGAATTTTAACCAAACG	(CCA)6	4.455	0.429	0.776	0.600
P45	c144082.graph_c1	F:AATACCTTGGCAAATGACGC R:CGCCAACCTATCTCAAAGC	(AGA)5	3.161	0.571	0.684	0.364
P47	c144342.graph_c0	F:CGTCCAGCATTCTCCATTT R:CAAGAAGGCTCTGGAGGATG	(TTC)6	3.122	0.125	0.680	0.556
P61	c147427.graph_c0	F:GGGGAGTAGAAGAAGGGACG R:TGATCTCTCACTCCGACACG	(GTG)5	3.657	0.125	0.727	0.556
P70	c148755.graph_c0	F:TGGTGGAGCTCAGAACAAGA R:TCCATTGAAGTATCCACCGC	(GGT)6	3.657	0.375	0.727	0.455
P71	c148777.graph_c3	F:CTCGGTCTCAAACTTTCC R:ATCACTGGCTCATCTCCGTC	(AAG)5	2.977	0.125	0.664	0.556
P75	c151957.graph_c0	F:GGGGGCAGGTAACTTTGTT R:CCCGTCCTGATCTACCTCCT	(TA)7	2.977	0.250	0.664	0.500
P78	c151961.graph_c0	F:TCCCATGTAGACTCTTCCG R:AAGACATGTTCCGGTCCGTC	(CCG)5	5.818	0.750	0.828	0.500
P84	c153951.graph_c1	F:CCCACATGTTTCTCCACTT R:TAGGGCAGAATTTGGGTTTG	(TTA)5	3.556	0.625	0.719	0.385
P85	c154034.graph_c0	F:GCGTTGATCATGGTTTATCG R:CCGTTGATCCCTTCGACTTA	(ATG)5	4.900	0.571	0.796	0.636
P89	c76886.graph_c0	F:AGATCTATTGGCCACGGATG R:GTTGCGAAAAGACGAAGAGG	(TCT)5	4.414	0.375	0.773	0.456
平均值 Average	-	-	-	3.264 ± 0.158	0.208 ± 0.029	0.638 ± 0.020	0.496 ± 0.038

(表 4)。8 份山茶种质间的相似系数为 0.217 4 ~ 0.814 8。来自于不同苗圃地的 2 个‘黑骑士’(C. ‘Night Rider’)品种之间的遗传相似系数最大,为 0.814 8,表明其确实为同名同物品种;其次是‘黑蛋石’(C. ‘Black Opal’)和‘红叶黑魔法’,为 0.752 2,表明这两份种质间具有较近的亲缘关系。山茶原种与其他 7 份种质之间的相似系数均较低,其中与‘红叶黑魔法’‘黑蛋石’之间的值最低,分别仅为 0.217 4 和 0.260 9。

对 8 份山茶种质进行了 UPGMA 聚类(图 4),可以看出,在相似系数为 0.58 时,8 份种质可以聚为 4 类。山茶原种为 I 类,‘黑蛋石’‘红叶黑魔法’和

‘黑骑士’3 份为 II 类,‘黑魔法’(C. japonica ‘Black Magic’)和‘金华美女’(C. japonica ‘Jinhua Meinū’)分别为 III 类和 IV 类。来自不同苗圃地的 2 份‘黑骑士’品种之间的相似系数最大,聚在一小类;‘红叶黑魔法’与‘黑蛋石’也被聚在一小类中。

3 讨 论

3.1 山茶转录组 SSR 位点特征

本研究中,利用 MISA 软件搜索到 13 197 个 SSR 位点,SSR 的发生频率为 19.52%,低于同属茶树(C. sinensis)的 28.64%^[15],油茶(C. oleifera)的 26.75%^[16]; EST-SSR 分布频率为 26.12%,平均每 4.33 kb 出现

表 4 8 份山茶种质的遗传相似系数

Tab. 4 Genetic similarity coefficients of 8 *Camellia* germplasms

编号 No.	1	2	3	4	5	6	7	8
1	1.000 0							
2	0.260 9	1.000 0						
3	0.217 4	0.752 2	1.000 0					
4	0.449 0	0.693 9	0.489 8	1.000 0				
5	0.367 3	0.612 2	0.693 9	0.576 9	1.000 0			
6	0.347 8	0.521 7	0.565 2	0.445 0	0.671 4	1.000 0		
7	0.461 5	0.692 3	0.692 3	0.545 5	0.618 2	0.538 5	1.000 0	
8	0.375 0	0.666 7	0.625 0	0.549 0	0.627 5	0.500 0	0.814 8	1.000 0

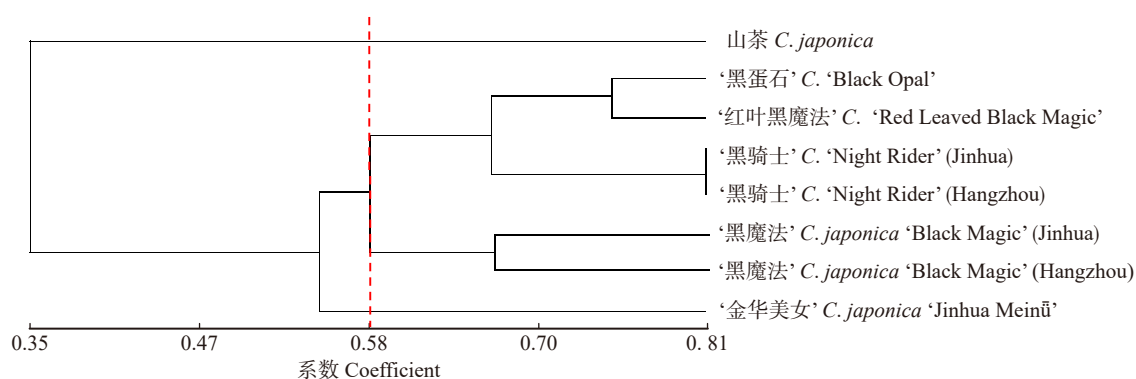


图 4 8 份山茶种质的 UPGMA 聚类图

Fig. 4 UPGMA dendrogram for 8 *Camellia* samples

一个 SSR 位点,也低于同属油茶 的 39.67% 和 2.33 kb,但高于崇左金花茶(*C. chuongtsoensis*)的 21.90%^[17]。不同植物转录组中 SSR 位点的出现频率各不相同,产生这种差异的原因是多方面的,一是统计参数设置引起的偏差,由于 SSR 的搜索标准、数据库大小以及序列的冗余级别的差异,导致统计发生频率及分布频率等参数时有所差异^[18-19],更重要的是植物种间特异性可引起 SSR 分布频率和发生频率较大差异,如在松属(*Pinus*)植物研究中发现,湿地松(*Pinus elliottii*) SSR 分布频率及发生频率分别为 6.25%、5.37%^[20],红松(*P. koraiensis*) SSR 发生频率为 3.74%^[21],而火炬松(*P. taeda*) SSR 分布频率仅为 1.2%^[22]。

核苷酸重复类型是 SSR 位点重要特征之一。本研究中,山茶 SSR 类型以单、二、三核苷酸重复类型为主,其中单核苷酸重复类型(A/T)数量最大,占重复总数的 47.746%,其次是二核苷酸(AG/CT、AT/AT)和三核苷酸(AAG/CTT)重复类型,而四、五、六核苷酸重复类型很少。这与崇左金花茶^[17]、柿树(*Diospyros kaki*)^[23]、马尾松(*Pinus massoniana*)^[24]、蒙古黄芪(*Astragalus mongholicus*)^[25]的研究结果较为一致,均以单核苷酸重复类型为主。而在较多的

植物中,SSR 重复类型是以二核苷酸、三核苷酸为主的,如桤木(*Alnus cremastogyne*)^[26]、长梗变光杜鹃(*Rhododendron calvescens* var. *duseimatum*)^[27]、蜡梅(*Chimonanthus praecox*)^[28]、紫鹃茶(*C. sinensis* var. *assamica* 'Zijuan')^[29]等。导致这种差异的原因除了物种本身的基因型差异外,还可能与参数设置有关,在设置 SSR 搜索参数时,如未将单核苷酸重复类型设置为搜索对象,则其重复类型以二、三核苷酸为主。同一物种不同器官的转录组 SSR 位点特征也表现出较大的差异,由于 EST-SSR 以功能基因为切入点,不同器官中基因的差异表达也可导致核苷酸重复类型的不同。如茶树花的转录组序列开发出的 SSR 以三核苷酸重复为最多,其次是六核苷酸^[30],这与茶树芽叶的转录组^[16]开发的 SSR 以二核苷酸类型为主的试验结果是不同的。

3.2 山茶转录组 SSR 标记的多态性

PIC 是用来评估基因变异程度的主要指标之一。Botstein D 等^[31]指出:当 $PIC > 0.5$ 时位点具有高度多态性, $0.5 > PIC > 0.25$ 时位点具有中度多态性。本研究中,利用 29 对多态性引物对 8 份山茶材料进行了遗传多样性分析,共扩增得到 72 个等位基因,多态信息含量(PIC)平均值为 0.496。Jia 等^[32]利

用油茶种子转录组 SSR 检测了 18 份油茶和 15 份山茶种质的遗传多样性,结果显示,其 PIC 值为 0.498;李海波等^[16]利用油茶叶转录组开发 SSR 标记,并对 57 份油茶样品进行了遗传关系分析,其多态信息含量为 0.439 9,这些结果与本研究结果相差不大,表明基于山茶叶片转录组开发的 SSR 标记具有较高的可靠性。另一方面,与基因组 SSR 相比,本研究的 PIC 值要低于同属植物基因组 SSR 位点的多态信息含量,如低于茶树的 0.51^[33]、大叶茶(*C. sinensis* var. *assamica*)的 0.64^[34]、黔南野生茶树的 0.572^[35]。张亚利等^[8]采用基因组 SSR 标记研究了茶花 28 份品种和金花茶(*C. nitidissima*)、浙江红山茶(*C. chekiangoleosa*)、毛花连蕊茶(*C. fraterna*)、微花连蕊茶(*C. minutiflora*)、岳麓连蕊茶(*C. handelii*)等种之间的遗传多样性,其 PIC 平均值高达 0.74,远高于本研究以及他人研究中的 PIC。本研究 SSR 位点多态性较低的原因是多方面的。首先是 SSR 长度,较多的研究表明,微卫星位点多态性与 SSR 长度存在一定的正相关^[36-38]。Temnykh 等^[39]认为,当 SSR 长度在 20 bp 以上时,不同品种间显示出较高的多态性,长度低于 20 bp,则多态性会降低。本研究中随机选用的 SSR 长度大多在 18~21 bp 之间,可能在一定程度上低估了 SSR 位点的多态性。在后续的 SSR 标记开发中,需要选择具有较长碱基序列的 SSR 位点来设计引物,以提高多态性检测效率。其次,基因组 SSR 标记可能比来自于转录组的 EST-SSR 标记含有更全面的基因组信息,由于 EST-SSR 来自于转录区域 DNA,仅反映了相关基因外显子的信息,因而比基因组 DNA 的多态性低^[40]。

3.3 山茶种质间的亲缘关系

利用 NTSYS-pc2.10e 软件计算遗传相似系数,并对 8 份种质进行了 UPGMA 聚类。在相似系数为 0.58 处,可将 8 份样品聚为 4 类,聚类结果与这些种质的遗传背景是高度一致的。山茶原种以及山茶种内的品种‘黑魔法’和‘金华美女’,属种内实生苗变异,均各自聚成一类。‘黑蛋石’‘黑骑士’和‘红叶黑魔法’三者聚为一大类,表明三者之间遗传关系较为密切,‘黑蛋石’与‘黑骑士’均为怒江红山茶杂交种‘宝石钟’(*C. hybrid* ‘Ruby Bell’)与山茶‘黑椿’(*C. japonica* ‘Kuro-tsubaki’)的全同胞子代^[11-12],聚在一小类。

‘红叶黑魔法’是从‘黑魔法’自由授粉子代中选出育出的红叶类品种,其母本为‘黑魔法’,但其父本究竟是‘黑蛋石’,还是‘黑骑士’文献记载不清楚。本研究中可以看出,‘红叶黑魔法’与‘黑蛋石’的相似系数最大,为 0.752 2,表明两者间的亲缘关系较为接近。因此,根据 SSR 的亲缘关系及聚类分析的结果,推

测‘黑蛋石’极有可能是‘红叶黑魔法’的父本。而‘红叶黑魔法’与其母本‘黑魔法’的遗传相似系数相对较小,这一聚类结果与这 3 个品种的表型较为一致。从表型上看,子代‘红叶黑魔法’嫩叶叶色红艳,红叶持续时间长,其父本‘黑蛋石’的新叶为暗红色,而母本‘黑魔法’新叶仅叶尖及叶缘呈浅红色,且在叶展开后迅速转绿。另外,潘丽芹等^[41]检测了这 3 个品种叶片的花青苷含量及比例,‘红叶黑魔法’和父本‘黑蛋石’红叶的花青苷比值分别达到 87.88% 和 79.08%,而母本仅为 42.62%。可见,子代‘红叶黑魔法’与其父本‘黑蛋石’的表型更为接近,因此,相对于母本而言,父本‘黑蛋石’对子代具有较高的遗传贡献。

4 结 论

本研究中,以山茶转录组数据为背景材料,从 50 518 条 Unigene 序列中搜索得到 13 197 个 SSR 位点,发生频率和分布频率分别为 19.52%、26.12%,平均 4.33 kb 存在一个 SSR 位点;SSR 类型以单、二、三核苷酸重复类型为主;根据 SSR 分析结果,随机选取了 90 对引物进行扩增筛选,有效扩增率为 81.11%,其中 29 对具有较好的扩增多态性,其平均多态信息含量为 0.496。利用 29 对引物对 8 个山茶种质进行 PCR 扩增和 UPGMA 聚类,聚类结果与材料的遗传背景基本一致,表明利用山茶转录组数据开发的 SSR 标记具有较好的实用性,开发的引物可为山茶的遗传多样性研究、亲缘关系鉴定以及分子标记辅助育种等提供可靠的分子标记来源。

参 考 文 献

- [1] Noormohammadi Z, Rahnama A, Sheidai M. EST-SSR and SSR analyses of genetic diversity in diploid cotton genotypes from Iran[J]. *The Nucleus*, 2013, 56(3): 171-178.
- [2] Kirungu J N, Deng Y, Cai X, et al. Simple sequence repeat (SSR) genetic linkage map of D genome diploid cotton derived from an interspecific cross between *Gossypium davidsonii* and *Gossypium klotzschianum*[J]. *International Journal of Molecular Sciences*, 2018, 19(1): 204-225.
- [3] Taheri S, Lee A T, Yusop M R, et al. Mining and development of novel SSR markers using next generation sequencing (NGS) data in plants[J]. *Molecules*, 2018, 23(2): 399-419.
- [4] Han Z, Ma X, Wei M, et al. SSR marker development and intraspecific genetic divergence exploration of *Chrysanthemum indicum* based on transcriptome analysis[J]. *BMC Genomics*, 2018, 19(1): 291-301.
- [5] Oliya B K, Kim M Y, Lee S H. Development of genic-SSR markers and genetic diversity of Indian lettuce (*Lactuca indica* L.) in South Korea[J]. *Genes & Genomics*, 2018, 40(6): 615-623.
- [6] Zhai S H, Yin G S, Yang X H. Population genetics of the endangered and wild edible plant *Ottelia acuminata* in

- Southwestern China using novel SSR markers[J]. *Biochemical Genetics*, 2018, 56(8): 1–20.
- [7] 魏秀清, 许玲, 章希娟, 等. 莲雾转录组 SSR 信息分析及其分子标记开发 [J]. *园艺学报*, 2018, 45(3): 541–551.
- Wei X Q, Xu L, Zhang X J, et al. Analysis on SSR information in transcriptome and development of molecular markers in wax apple[J]. *Acta Horticulturae Sinica*, 2018, 45(3): 541–551.
- [8] 张亚利, 宋焘, 奉树成. SSR 分子标记在山茶属观赏资源遗传多样性研究中的应用 [J]. *植物科学学报*, 2016, 34(5): 755–764.
- Zhang Y L, Song Y, Feng S C. Application of SSR for the analysis of genetic diversity in *Camellia*[J]. *Plant Science Journal*, 2016, 34(5): 755–764.
- [9] 胡兴华, 王燕, 邹玲俐, 等. 茶花品种 SSR 指纹图谱分型技术反应体系优化 [J]. *中国农学通报*, 2013, 29(1): 127–131.
- Hu X H, Wang Y, Zou L L, et al. The study on SSR-PCR fingerprints map classification for *Camellia japonica* cultivars[J]. *Chinese Agricultural Science Bulletin*, 2013, 29(1): 127–131.
- [10] 李琳琳, 黄万坚, 刘信凯, 等. 应用 SSR 分子标记技术鉴定张氏红山茶杂交 F1 代真实性的研究 [J]. *广东园林*, 2014, 36(2): 44–47.
- Li L L, Huang W J, Liu X K, et al. Verity identification of hybrid F1 derived from *Camellia changii* using SSR markers[J]. *Guangdong Landscape Architecture*, 2014, 36(2): 44–47.
- [11] 高继银, 陈绍云, 徐碧玉. 世界名贵茶花 [M]. 杭州: 浙江科学技术出版社, 1998.
- Gao J Y, Chen S Y, Xu B Y. The world's best *Camellia* cultivars[M]. Hangzhou: Zhejiang Science and Technology Publishing House, 1998.
- [12] 高继银, 苏玉华, 胡羨聪. 国内外茶花名种识别与欣赏 [M]. 杭州: 浙江科学技术出版社, 2007.
- Gao J Y, Su Y H, Hu X C. The identification and appreciation of the world's outstanding *Camellias*[M]. Hangzhou: Zhejiang Science and Technology Publishing House, 2007.
- [13] Savige J T. The international *Camellia* Register [M]. Sydney: Fine Arts Press Pty. Limited, 1993.
- [14] 管开云, 李纪元, 王仲朗. 中国茶花图鉴 [M]. 杭州: 浙江科学技术出版社, 2014.
- Guan K Y, Li J Y, Wang Z L. *Camellias of China*[M]. Hangzhou: Zhejiang Science and Technology Publishing House, 2014.
- [15] 陈琪, 杨华, 韦朝领, 等. 基于茶树芽叶转录组序列的 EST-SSR 分布特征研究 [J]. *安徽农业大学学报*, 2016, 43(2): 170–175.
- Chen Q, Yang H, Wei C L, et al. Density and characterization of EST-SSR in tea (*Camellia sinensis*) bud and leaf transcriptome[J]. *Journal of Anhui Agricultural University*, 2016, 43(2): 170–175.
- [16] 李海波, 王珊, 丁红梅, 等. 普通油茶转录组 EST-SSR 分子标记开发 [J]. *植物生理学报*, 2017, 53(7): 1267–1278.
- Li H B, Wang S, Ding H M, et al. Development of EST-SSR molecular markers based on transcriptome sequencing of *Camellia oleifera*[J]. *Plant Physiology Journal*, 2017, 53(7): 1267–1278.
- [17] 邵阳, 范文, 黄连冬, 等. 基于 RNA-seq 的崇左金花茶 EST-SSR 标记开发 [J]. *复旦学报(自然科学版)*, 2015, 54(6): 761–767.
- Shao Y, Fan W, Huang L D, et al. Development of EST-SSR markers in *Camellia chuongtsoensis* based on RNA-seq[J]. *Journal of Fudan University (Natural Science)*, 2015, 54(6): 761–767.
- [18] Biswas M K, Chai L, Mayer C, et al. Exploiting BAC-end sequences for the mining, characterization and utility of new short sequences repeat (SSR) markers in *Citrus*[J]. *Molecular Biology Reports*, 2012, 39(5): 5373–5386.
- [19] Sharma R K, Bhardwaj P, Negi R, et al. Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.)[J]. *BMC Plant Biology*, 2009, 9(1): 53–77.
- [20] 赵衡. 湿地松 EST-SSR 引物开发及其亲缘关系的研究 [D]. 南昌: 江西农业大学, 2016.
- Zhao H. EST-SSR primer design and analysis the genetic relationship in *Pinus elliottii*[D]. Nanchang: Jiangxi Agricultural University, 2016.
- [21] 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, 51(8): 114–120.
- Zhang Z, Zhang H G, Mo C, et al. Transcriptome sequencing analysis and development of EST-SSR markers for *Pinus koraiensis*[J]. *Scientia Silvae Sinicae*, 2015, 51(8): 114–120.
- [22] Chagné D, Chaumeil P, Ramboer A, et al. Cross-species transferability and mapping of genomic and cDNA SSRs in pines[J]. *Theoretical & Applied Genetics*, 2004, 109(6): 1204–1214.
- [23] 杜改改, 孙鹏, 索玉静, 等. 基于柿雌雄花芽转录组测序的 SSR 和 SNP 多态性分析 [J]. *中国农业大学学报*, 2017, 22(10): 45–55.
- Du G G, Sun P, Suo Y J, et al. SSR and SNP polymorphism analysis based on persimmon (*Diospyros kaki*) transcriptome[J]. *Journal of China Agricultural University*, 2017, 22(10): 45–55.
- [24] 梅利那, 范付华, 崔博文, 等. 基于马尾松转录组的 SSR 分子标记开发及种质鉴定 [J]. *农业生物技术学报*, 2017, 22(6): 991–1002.
- Mei L N, Fan F H, Cui B W, et al. Development of SSR molecular markers based on transcriptome sequences and germplasm identification in masson pine (*Pinus massoniana*) [J]. *Journal of Agricultural Biotechnology*, 2017, 22(6): 991–1002.
- [25] 贺润丽, 韩毅丽, 王芳, 等. 蒙古黄芪转录组 SSR 标记开发及遗传多样性研究 [J]. *中国中药杂志*, 2018, 43(9): 1838–1843.
- He R L, Han Y L, Wang F, et al. Development of SSR markers and genetic diversity of *Astragalus mongolicus* transcriptome[J]. *China Journal of Chinese Materia Medica*, 2018, 43(9): 1838–1843.
- [26] 饶龙兵, 杨汉波, 郭洪英, 等. 基于柃木属转录组测序的 SSR 分子标记的开发 [J]. *林业科学研究*, 2016, 29(6): 875–882.
- Rao L B, Yang H B, Guo H Y, et al. Development of SSR molecular markers based on transcriptome sequences of *Alnus*[J].

- Forest Research, 2016, 29(6): 875–882.
- [27] 李太强, 刘雄芳, 万友名, 等. 滇东南濒危植物长梗杜鹃转录组微卫星特征分析 [J]. 林业科学研究, 2017, 30(4): 533–541.
Li T Q, Liu X F, Wan Y M, et al. Characteristic analysis of microsatellites in the t of *Rhododendron longipedicellatum*, an endangered species endemic to Southeastern Yunnan, China [J]. Forest Research, 2017, 30(4): 533–541.
- [28] 李响, 杨楠, 赵凯歌, 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, 35(增刊 1): 25–32.
Li X, Yang N, Zhao K G, et al. Development and primer selection of EST-SSR molecular markers based on transcriptome sequencing of *Chimonanthus praecox* [J]. Journal of Beijing Forestry University, 2013, 35(Suppl. 1): 25–32.
- [29] 周萌, 宋维希, 鄢文光, 等. 基于紫娟茶叶片转录组的 SSR 引物开发 [J]. 华北农学报, 2016, 31(增刊 1): 146–151.
Zhou M, Song W X, Yan W G, et al. SSR Primers development based on Zijuancha leaves' transcription group [J]. Acta Agriculturae Boreali-Sinica, 2016, 31(Suppl. 1): 146–151.
- [30] 王丽鹭, 韦康, 张成才, 等. 茶树花转录组微卫星分布特征 [J]. 作物学报, 2014, 40(1): 80–85.
Wang L Y, Wei K, Zhang C C, et al. Characterization of microsatellites in tea (*Camellia sinensis*) floral transcriptome [J]. Acta Agronomica Sinica, 2014, 40(1): 80–85.
- [31] Botstein D, Cherry J M, Ashburner M, et al. Gene Ontology: tool for the unification of biology [J]. Nature Genetics, 2000, 25(1): 25–29.
- [32] Jia B, Lin Q, Lin Z, et al. Development of 15 genic-ssr markers in oil-tea tree (*Camellia oleifera*) based on transcriptome sequencing [J]. Genetika, 2014, 46(3): 789–797.
- [33] Wu H, Chen D, Li J, et al. De novo characterization of leaf transcriptome using 454 sequencing and development of EST-SSR markers in tea (*Camellia sinensis*) [J]. Plant Molecular Biology Reporter, 2013, 31(3): 524–538.
- [34] 尚卫琼, 段志芬, 杨毅坚, 等. 基于 EST-SSR 标记的云南大叶茶资源遗传多样性分析 [J]. 山东农业科学, 2018, 50(1): 16–22.
Shang W Q, Duan Z F, Yang Y J, et al. Genetic diversity analysis of Daye tea germplasm in Yunnan based on EST-SSR markers [J]. Shandong Agricultural Sciences, 2018, 50(1): 16–22.
- [35] 陈世军, 张明泽, 姚玉仙, 等. 基于 SSR 标记的黔南茶树种质资源 DNA 指纹图谱构建 [J]. 植物遗传资源学报, 2017, 18(1): 106–111.
Chen S J, Zhang M Z, Yao Y X, et al. Establishment of DNA fingerprinting for tea germplasm from Qiannan Prefecture by SSR markers [J]. Journal of Plant Genetic Resources, 2017, 18(1): 106–111.
- [36] He C, Poysa V, Yu K. Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars [J]. Tag. theoretical & Applied Genetics. theoretische Und Angewandte Genetik, 2003, 106(2): 363–373.
- [37] Smulders M J M, Bredemeijer G, Rus-Kortekaas W, et al. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species [J]. Theoretical & Applied Genetics, 1997, 94(2): 264–272.
- [38] 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(1): 136–152.
- [39] Temnykh S, Park W D, Ayres N, et al. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.) [J]. Theoretical & Applied Genetics, 2000, 100(5): 697–712.
- [40] Morgante M, Hanafey M, Powell W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes [J]. Nature Genetics, 2002, 30(2): 194–200.
- [41] 潘丽芹, 李纪元, 韦海忠, 等. 红叶山茶 (*Camellia japonica*) 叶片色素含量与叶色参数的相关性分析 [J/OL]. 分子植物育种 (2019–03–13) [2019–03–19]. <http://kns.cnki.net/kcms/detail/46.1068.s.20190313.0941.004.html>.
Pan L Q, Li J Y, Wei H Z, et al. Correlation analysis on pigment content changes and leaf color parameters of *Camellia japonica* cultivars [J/OL]. Molecular Plant Breeding (2019–03–13) [2019–03–19]. <http://kns.cnki.net/kcms/detail/46.1068.s.20190313.0941.004.html>.

(责任编辑 赵勃
责任编辑 康向阳)