

基于茶树芽叶转录组序列的 EST-SSR 分布特征研究

陈琪¹, 杨华^{1,2}, 韦朝领¹, 张正竹¹, 宛晓春^{1*}

(1. 安徽农业大学省部共建茶树生物学与资源利用国家重点实验室, 合肥 230036; 2. 安徽农业大学理学院, 合肥 230036)

摘要: 采用生物信息学方法对茶树 (*Camellia Sinensis*) 叶部不同发育阶段 (单芽/三叶) 转录组中的 EST-SSR 位点信息进行分析, 以期对茶树分子标记辅助育种提供有效的参考。利用 SSRFINDER 软件, 对茶树芽叶转录组的高通量测序获得的 97454 条 Unigenes 序列 (100.91Mb) 进行高通量 SSR 位点的筛选, 共获得 36249 个 SSR 位点, 分布于 27913 条 Unigenes 序列中, 其发生频率为 28.64%, 平均分布密度为 1/2.78 kb。在茶树芽叶的转录组序列中共发现 962 种碱基重复单元, 其中占主导的是以 (AG/CT)_n 为主 (占总 SSRs 的 23.78%) 的二核苷酸重复, 占到总 SSRs 的 59.19%, 其次是三核苷酸和单核苷酸重复, 其占比分别为 20.92% 和 13.13%。茶树芽叶转录组所含不同重复单元 SSRs 的平均长度, 相对全器官转录组较短, 其中占比最高的是重复长度为 18 bp 的中长重复序列, 占到总 SSRs 的 23.37%, 而大于 25 bp 的较长序列重复极少。同时对茶树不同器官转录组测序的 SSR 分布特性进行了分析比较, 以期为后续分子标记开发提供参考。

关键词: 茶树; 微卫星; 芽叶转录组

中图分类号: S571.1

文献标识码: A

文章编号: 1672-352X (2016)02-0170-06

Density and characterization of EST-SSR in tea (*Camellia sinensis*) bud and leaf transcriptome

CHEN Qi¹, YANG Hua^{1,2}, WEI Chaoling¹, ZHANG Zhengzhu¹, WAN Xiaochun¹

(1. State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei 230036;

2. School of Science, Anhui Agricultural University, Hefei 230036)

Abstract: Bioinformatics method is used to analysis the EST-SSR loci in different development stages of leaf (Single bud & 3rd leaf) transcriptome of tea plant (*Camellia Sinensis*), in order to provide effective reference for the tea plant molecular marker assisted breeding. In total 97454 unigenes (100.91 Mb) derived from deep sequencing of *camellia sinensis* bud and 3rd leaf transcriptome were used for the development of functional microsatellites or simple sequence repeats (EST-SSRs) molecular markers. 36249 SSRs were screened using SSRFINDER software in 27913 unigenes. The frequency of SSR was 28.64% and mean distribution density was 1/2.78 kb. Di-nucleotide was the major repeated type, accounting for 59.19%, and tri-nucleotide and mononucleotide accounted for 20.92% and 13.13%, respectively. Among all 962 SSR motifs, A/T, AG/CT and AAG/CTT were respectively the most frequent repeated motifs in mono-nucleotide, di-nucleotide and tri-nucleotide repeated type, and accounting for 12.98%, 45.95% and 5.59% in all the SSR repeat motifs. The microsatellites with length below 18 bp were in the maximum proportion, accounting for 23.37%, while the microsatellites over 25 bp were only 3.6%. The results indicated that the unigenes obtained from transcriptome sequence in *C. sinensis* can be used as an effective way for the development of SSR loci. More SSR primers will provide more markers for the research of genetic variation of *C. sinensis*.

Key words: *Camellia sinensis*, microsatellites; transcriptome of bud & leaf

茶叶 (*Camellia sinensis*) 作为世界三大饮料之一, 是一类重要的经济型栽培作物, 在我国栽培历史一悠久, 品种资源丰富, 目前在生产上以无性系地方品种推广程度较高。如何结合区域特性推进育种

收稿日期: 2015-11-24

基金项目: 安徽省科技专项 (13Z03012), 安徽省创新型省份建设专项 (15CZS08032) 和国家自然科学基金 (31500566) 共同资助。

作者简介: 陈琪, 博士, 讲师。

* 通信作者: 宛晓春, 博士, 教授, 博士生导师。E-mail: xcwan@ahau.edu.cn

进程, 提高品种选育效率和质量, 加快优良新品种的研发成为茶树育种工作者的重要任务。

DNA 分子标记是鉴定种质资源遗传多样性的重要手段, 也是一种重要的辅助育种方法, 其标记类型有多种, 例如限制性片段长度多态性 (restriction fragment length polymorphism, RFLP)、简单重复序列 (simple sequence repeat, SSR) 以及单核苷酸的多态性 (single nucleotide polymorphisms, SNP) 等^[1-2]。简单重复序列, 又称微卫星 DNA (microsatellite DNA), 是一类广泛存在于真核和原核生物基因组中由 1~6 个碱基组成的基元串联重复而成的 DNA 序列, 研究发现基因组中平均每 50 kb 就有 1 个 SSR^[3]。SSR 标记按来源分主要包括表达序列标签 SSR (EST-SSR) 和基因组 SSR (Genomic SSR)。EST-SSR 标记与 G-SSR 标记相比, 由于无需构建基因组文库、杂交、测序, 开发工作量与成本缩减, 而且其源于基因组的表达区域, 可直接反映相关基因的多样性^[4], 因此, 具有更高的通用性及转化率^[5]。随着新一代高通量 RNA 测序技术 (RNA-seq) 的推广, 完全改变了转录组学的研究模式, 使其迅速成为研究非模式生物转录组的先进技术^[6], 而 SSR 和 SNP 作为常用的分子标记, 也成为利用转录组数据开发最多的两类标记^[7], 特别是前者由于其检测方法简便, 一直为研究者所关注。在柑橘^[8]、茄子^[9]、洋葱^[10]、辣椒^[11]、杜仲^[12]、芝麻^[13]和橡胶^[14]等非模式植物中通过转录组信息开发 SSR 标记的方法均有报道。

在茶树资源和应用开发过程中, SSR 标记起到非常重要的作用, 金基强等^[15]率先对茶树 EST-SSR 进行了分析和分子标记的建立; 姚明哲等^[16]利用 25 对 EST-SSR 引物对江北茶区的 45 份茶树初级核心种质的遗传多样性、遗传结构和亲缘关系进行了分析, 认为江北茶区主要省份间茶树种质的遗传分化程度较低; 刘振等^[17-18]通过 EST-SSR 标记对福建茶区和西南茶区的茶树资源进行遗传多样性和亲缘关系分析, 发现上述茶区的茶树资源均具有相对较高的遗传多样性。但此阶段由于测序技术的限制, 所得序列和 SSR 数均较少。随后, 杨华等^[19]首次利用高通量测序技术 (RNA-Seq) 对茶树全器官转录组进行测序, 在获得的 127094 条 Unigenes 中发掘并获得了 12242 个 SSRs, 王丽鸾^[20]也通过对茶树花转录组测序获得的 75331 条 Unigenes 进行分析, 获得 12582 个 SSRs, 并对其在编码区和非编码区的出现频率及分布特性进行了研究。但是茶树作为经济作物其利用价值主要在叶部, 其遗传多样性和品质

特性差异也主要体现在叶部, 因而加大对其叶部特异性表达基因的 SSR 标记的开发, 从而进行功能成分的标记定位具有重要的意义。

本研究针对课题组前期茶树单芽/第三叶转录组测序获得的 97454 条序列, 利用软件进行 SSR 位点的筛选, 通过比较转录组序列中微卫星的发生频率和重复单元的组成特征, 确定茶树功能叶转录组序列中所含微卫星重复序列的分布特征和组成情况。此外, 本研究还对高通量测序获得的茶树全器官转录组和茶树花转录组中微卫星位点的分布和组成特性进行了比较, 以探讨不同组织器官含有微卫星差异, 以期为茶树功能基因表达调控的定位分析、微卫星标记开发、茶树种质资源多态性筛选标记的挖掘提供研究基础。

1 材料与方法

1.1 植物材料及数据源

茶树芽/叶转录组的数据来自于本课题组前期对茶树品种“舒茶早” [*Camellia sinensis* (L.) O. Kuntze cv. Shuchazao] 进行的全转录组 Illumina HiSeq™ 2000 双末端测序。在 2014 年 5 月中上旬, 于安徽农业大学农业产业示范园的大杨店茶园采集国家级良种“舒茶早”无性系茶树新梢的单芽和 1 芽 3 叶 2 个不同发育阶段的茶鲜叶, 经液氮瞬时冷冻后于 -86℃ 低温保存备用。将样品分别提取 mRNA, 利用 Illumina HiSeq2000 平台测序, 分别得到 10 Gp 的高质量 clean reads, 两样品的 clean reads 采用 Trinity 软件进行 *De novo* 混合组装后, 进一步进行去冗余拼接和同源转录本聚类, 获得最终 unigene 共 97454 条, 平均长度 1 036 nt, N50 值为 1748 nt。测序原始数据已经上传至美国国家生物技术信息中心 (National Center for Biotechnology Information, NCBI), 登录号为 SRR1928149。

1.2 基于 perl 语言简单重复序列分析 (SSR) 搜索

使用 SSRFINDER 软件 (<http://www.maizemap.org/bioinformatics/SSRFINDER/>) 进行转录组序列中 SSR 搜索。搜索 1~6 bp 的微卫星, 搜索标准为单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸最少重复次数分别为 15、6、5、4、4 和 3; 复合微卫星位点之间最大间隔碱基数为 100 bp。

2 结果与分析

2.1 茶树芽/叶转录组序列中 SSR 重复类型的分布特征

利用软件对茶树 (*Camellia sinensis*) 单芽/三叶

转录组中 97454 条 Unigene 的 cDNA 序列数据进行搜索, 序列拼接总长度为 100917 kb, 发现含不同重复基元 SSRs 的 Unigene 序列有 27913 条, 发生频率 (含有 SSR 的 Unigene 数目与总 Unigene 数目之比值) 为 28.64%。SSR 序列总长度约为 605.13 kb (0.59%), 说明在茶树单芽/三叶转录组中 SSR 序列约占整个转录组序列的 6%。对其分布特性进行分析发现, 含单个 SSR 位点的 Unigene 有 20415 条, 含 2~6 个 SSR 位点的 Unigene 有 7498 条, 共 36249 个 SSRs, 分布频率 (SSR 数目/总的 Unigene 数) 为 37.19%。从分布情况看, SSR 的平均距离较低 (茶树转录组 Unigene 长度与 SSR 数目之比值), 即是

茶树转录组中平均每 2.78 kb 就出现 1 个 SSR (见表 1)。

茶树芽/叶转录组 SSR 种类非常丰富, 对 1~6 bp 的 SSR 重复单元类型进行统计可知 (表 1), 在 36249 个 SSR 中共有 962 种重复基元 (motif), 与其他茶树组织器官转录组测序获得的 SSR 位点相一致, 重复单元类型以二核苷酸和三核苷酸为主, 分别占总数的 59.19% 和 20.92%, 其次是单核苷酸, 占 13.13%, 四、五和六核苷酸重复单元尽管数量较少, 但是类型非常丰富, 其中五核苷酸重复单元数量高于四、六核苷酸。

表 1 茶树芽/叶转录组中 SSR 重复类型的分布特征

Table 1 The characteristic of SSR repeat types in bud and leaf transcriptome of *C. sinensis*

重复类型 Repeat type	数目 Number	各类型比例/% Proportion	频率/% Frequency	平均距离/kb Average distance	总长度/bp Total length	平均长度/bp Average length
单核苷酸 Mononucleotide	4759	13.13	4.88	21.21	84439	18
二核苷酸 Dinucleotide	21456	59.19	22.02	4.70	341544	16
三核苷酸 Trinucleotide	7586	20.92	7.78	13.30	130032	17
四核苷酸 Tetranucleotide	558	1.53	0.57	180.85	11440	20
五核苷酸 Pentanucleotide	1018	2.81	1.04	99.13	20645	20
六核苷酸 Hexnucleotide	872	2.40	0.89	115.73	20934	24
总计 Total	36249	100.00	37.19	2.85	609034	16

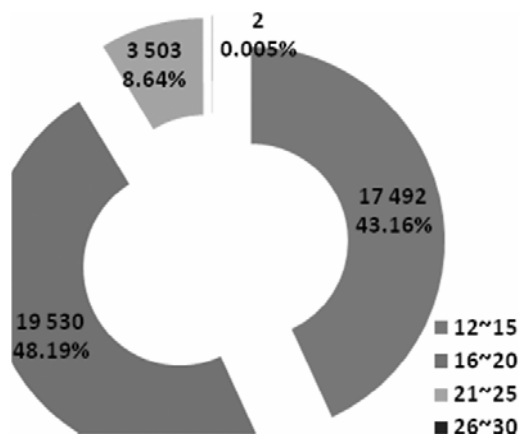


图 1 茶树芽/叶转录组中 SSR 长度分布

Figure 1 The distribution of SSR length in bud and leaf transcriptome of *C. sinensis*

2.2 茶树芽/叶转录组序列中 SSR 重复类型长度与基元重复次数

茶树转录组 SSR 位点的序列总长达到 609.03 kb, 其中单核苷酸重复基元、二核苷酸重复基元、三核苷酸重复基元、四核苷酸重复基元、五核苷酸重复基元和六核苷酸重复基元的 SSR 位点的碱基总长分别是 84439、341544、130032、11440、20645 和 20934 bp; 各类型 SSR 位点的平均长度分别是 18、

16、17、20、20 和 24 bp (表 1)。通过对茶树芽/叶转录组中所发现的 36249 个微卫星长度的变化进行分析, 结果发现茶树芽/叶转录组序列所含微卫星在长度上差异较小, 所发现的 SSR 重复片段的长度波动于 10~30 bp, 6 种类型的 SSR 平均长度为 16.7 个碱基 (单碱基重复 10 次以上的结果)。茶树芽/叶转录组所含微卫星序列长度的分布比例, 较茶树全器官和茶树花转录组中的结果略有差异^[19-20], 明显集中为中等偏短的序列。以 15~20 bp 的中等重复序列占多数 (48.19%), 其次是长度小于 15 bp 的序列重复, 占到 43.16%, 21~25 bp 的序列占比较少为 8.64%, 大于 26 个 bp 的较长序列基本可以忽略 (图 1)。

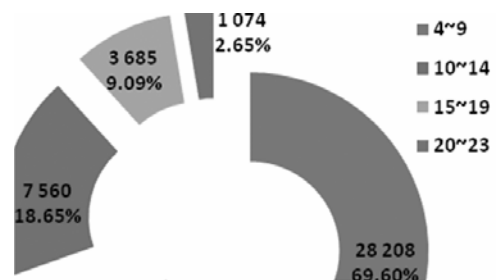


图 2 茶树芽/叶转录组中 SSR 重复次数分布

Figure 2 The distribution of the SSR repeats number in bud and leaf transcriptome of *C. sinensis*

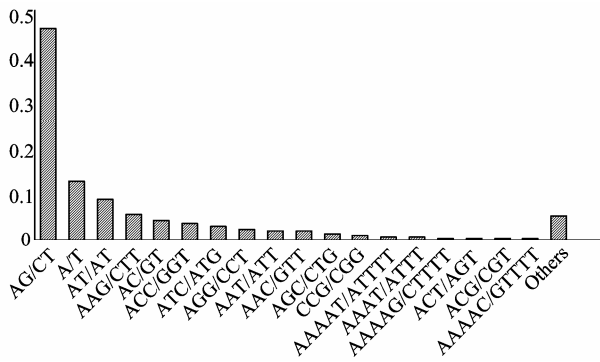


图 3 茶树芽叶转录组 SSR 中不同重复基元(motif)的分布 (考虑到碱基互补作用)

Figure 3 The distribution of different SSR repeat motifs in bud and leaf transcriptome of *C. sinensis*. (considering sequence complementary)

从重复次数看, 茶树芽叶转录组中各 SSR 重复类型的重复次数在 4~23 次波动, 除单核苷酸外, 重复基元以重复 6 次出现的频率最高, 有 7039 个,

占总 SSR 的 19.42%, 其次为 7 次、9 次、8 次和 5 次重复, 数量在 4320~5279 个之间波动。合计 4~9 次重复的 SSR 位点有 28208 个, 占 69.6%; 10~14 次重复的位点有 7560 个, 占 18.65%; 15~19 次重复的位点有 632 个, 占 5.16%; 最大重复数为单核苷酸的 23 次重复 (图 2)。从序列长度和重复次数所反映的多态性来看, 茶树芽叶转录组远少于茶树花转录组中的最大长度 90 bp 和最大 45 次重复以及茶树全器官转录组中的最大 60 次重复。

微卫星长度变化的情况反映了微卫星位点获得 (或失去) 重复单元的活跃程度, 因而是影响其多态性高低的重要因素, 有研究者提出^[21], 低级基元 SSR 多态性普遍比高级基元的高。对茶树芽叶转录组 SSR 分析发现, 长度在 16~20 bp 的 SSR 重复基元中, 多态性潜能高的低级基元 SSR 占到 81.2%, 共 6726 条。可以重点依据这部分多态性潜能高的 SSR 进行茶树分子标记的开发。

表 2 茶树芽叶转录组 SSR 不同重复类型的各基元数量

Table 2 Repeat motif number of each types of SSR in bud and leaf transcriptome of *C. sinensis*.

重复类型 Repeat type	重复基元 Repeat motif	数目 Number	比例/% Proportion	频率/% Frequency
单核苷酸	A/T	8877	12.86	4.78
Mononucleotide	C/G	160	0.27	0.10
二核苷酸*	AG/CT	16657	45.95	17.09
Dinucleotide	AT/AT	3216	8.87	3.30
	AC/GT	1547	4.26	1.59
	CG/CG	36	0.09	0.37
三核苷酸*	AAG/CTT	2026	5.59	2.08
Trinucleotide	ACC/GGT	1261	3.48	1.29
	ATC/ATG	1027	2.83	1.05
	AGG/CCT	773	2.13	0.79
	AAT/ATT	755	2.08	0.77
	AAC/GTT	705	1.94	0.72
	AGC/CTG	424	1.17	0.44
	CCG/CGG	339	0.94	0.35
	ACT/AGT	143	0.39	0.15
	ACG/CGT	133	0.37	0.14
四核苷酸*	AAAT/ATTT	187	0.52	0.19
Tetranucleotide	AAAC/GTTT	88	0.24	0.09
	AAAG/CTTT	78	0.22	0.08
	Others	205	0.57	0.21
五核苷酸*	AAAAT/ATTTT	189	0.52	0.19
Pentanucleotide	AAAAG/CTTTT	161	0.44	0.17
	AAAAC/GTTTT	100	0.28	0.10
	Others	568	1.57	0.58
六核苷酸*	AACCCT/AGGGTT	19	0.05	0.02
Hexnucleotide	AAAACC/GGTTTT	17	0.05	0.02
	AGCCTC/AGGCTG	17	0.05	0.02
	Others	819	2.26	0.84

表 3 不同茶树转录组 SSRs 位点比较

Table 3 Comparison of the characterization of SSRs in different *C. sinensis* transcriptomes

项目 Item	单芽/二叶转录组 Bud/Leaf transcriptome	全器官转录组 ^[19] Total tissue transcriptome	花转录组 ^[20] Floral transcriptome
总基因数 Total number of sequences	97454	127094	75531
含 SSR 基因数 Number of SSR containing sequences	27913	10325	10290
SSR 数目 Total number of SSR	36249	12242	12582
出现频率/% Frequency	37.19	9.63	16.66
拼接序列总长/kb Total size of sequences	100916	45069	30326
SSR 序列总长/kb Total size of SSR	605.13	197.99	278.55
平均长度/bp Average length of SSR	16	16	16
平均距离/kb Interval sequence length	1/2.78	1/3.68	1/2.41
SSR 重复模式 Type of repeat motifs	962	181	340

2.3 茶树芽叶转录组 SSR 碱基重复单元比较

茶树芽叶转录组共发现 962 种碱基重复单元, 为目前所有茶树转录组测序挖掘所得 SSR 之最, 其中占比最高的是 (AG/CT) n (45.95%), 其次是 (A/T) n (12.85%)、(AT/TA) n (8.87%)、(AAG/CTT) n (5.59%)、(AC/GT)n(4.27%)和(ACC/GGT) n (3.48%) 等。不同类型重复单元的 SSRs 分布见图 3。

此外, 从表 2 可以看出, 单、二、三、四、五和六碱基重复单元中出现频率最多的重复单元分别是(A/T)n、(AG/CT)n、(AAG/CTT)n、(AAAG/CTTT)n、(AAAAT/ATTTT)n 和(AAAAAC/GTTTTT)n。它们在各自重复单元类型中的比例分别是 92.24%、85.60%、27.35%、21.61%、20.55%和 4.71%。

3 讨论

本研究通过茶树芽叶转录组测序, 组装后得到 97454 条 Unigene, SSR 位点搜索, 获得 36249 个 SSR 位点, 平均分布距离为 2.78 kb, 分布频率为 37.19%。茶树转录组 SSRs 分布距离高于葡萄^[21]*Vitis vinifera* (1/14.5 kb)、柑橘^[22](1/ 5.7 kb)、欧洲大叶杨^[23]*Populus trichocarpa* (1/14.0 kb)、西洋参 *Panax quinquefolium* L.^[24] (1/7.5 kb)、云南松 *Pinus yunnanensis*^[25] (1/ 29 kb) 和杜仲^[12] *Eucommia ulmoides* (1/ 26.13 kb)等植物, 这表明茶树转录组中 SSR 数量很丰富。同时茶树芽叶组织中的 SSR 数量、重复单元类型与出现频率均远高于茶树其他同为高通量转录组测序所开发的 SSR (见表 3), 如茶树全器官转录组 (出现频率 9.63%) 和茶树花转录组 (出现频率 16.66%)。考虑到同为高通量转录组测序, 搜索的序列来源、搜索的标准、物种间 SSR 的分布等背景均较为一致, 因而这种差异的可能原因应出于器官特异性而非测序或样本方面^[24], 表明茶树不同

器官间 SSR 分布特征存在着较大的差异^[26]。

本研究测定出的茶树芽叶转录组 SSR 重复单元以二核苷酸为最多, 其次为三核苷酸与单核苷酸, 与其他茶树转录组测序结果一致。茶树芽叶转录组 SSR 中不同的重复单元出现频率最高的是 AG / CT, 占到总比的 45.95%, 高于茶树全器官转录组的 37.37%, 和花转录组的 44.99%接近。从茶树不同器官转录组测序结果来看, 茶树 SSR 中 2~3 bp 的类型占多数, 而其他核苷酸类型较少, 茶树芽叶转录组中二、三核苷酸占总数的 80%以上, 而全器官转录组中更是高达 86.42%, 重复单元碱基数少的 SSR 类型所占比例高, 可能进化的水平较高、进化的时间较长或突变频率较高等原因^[27], 因此推测茶树处于相对较高的进化水平, 进化时间较长。

针对不同重复单元的类型分析可知, A/T、AAG/CTT、AAAT/ATTT、AAAAT/ATTTT 分别为单核苷酸、三核苷酸、四核苷酸、五核苷酸的优势重复单元, 表现为 A/T 或 AT/TA 的出现的频率较高, 在其他植物中也有类似的现象发生, 在核苷酸重复类型中, 以 A/T 为最丰富的重复单元^[28], 这种趋势可能与碱基所含的能量有关, 打破 AT 键所需的能量低于 GC 键, 从而 AT 的波动较 GC 容易。SSR 中以 AT 出现较多, 表明富含 A/T-单元的 SSR 类型可能在植物中常见^[29-30]。但是, 本研究中二核苷酸中优势单元为 AG/CT (占总比 45.95%), 未表现为明显的 A/T-优势单元, 这可能是物种间的差异^[26]。

茶树芽叶组织中的 SSR 重复次数波动于 4~23, 以 4~9, 其次是 10~14 次重复次数较多, 相对其他植物, 以及茶树其他器官中的 SSR 重复次数较少, 但是重复单元类型非常丰富, 高达 962 种, 远较茶树全器官 (181 种) 和花转录组 (340 种) 丰富, 其中最为丰富的是六核苷酸, 达到 521 种, 表

明茶树芽叶转录组 SSR 具有非常高的多态性潜力^[31], 近些年来从转录组测序数据库中开发 SSR 是一个有效的途径^[6, 10-12], 获得的 SSR 标记具有通用性较好、开发方法简单及成本低廉等优点。迅速发展的 SSR 也用于种质资源的评价、多样性检测、图谱构建等方面^[16-18]。

总体来看, 茶树芽叶转录组中 SSR 出现的频率较高, 分布距离较近, 出现的重复类型及其重复单元在不同器官比较中最为丰富, 基于这些序列设计的引物具有高多态性的潜力。

参考文献:

- [1] POWELL W, MACHRAY G C, PROVAN J. Polymorphism revealed by simple sequence repeats [J]. *Trends Plant Science*, 1996, 1(7): 215-222.
- [2] 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.
- [3] 何平. 真核生物中的微卫星及其应用[J]. *遗传*, 1998, 20(4): 42-47.
- [4] MORGANTE M, OLIVIERI A M. PCR-amplified microsatellites as markers in plant genetics [J]. *The Plant Journal*, 1993, 3(1): 175-182.
- [5] EUJAYL I, SORRELLS M, BANM M, et al. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat [J]. *Theoretical and Applied Genetics*, 2002, 104(2): 399-407.
- [6] 李小白, 向林, 罗洁. 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, 35(5): 1-8.
- [7] 刘红亮, 郑丽明, 刘青青, 等. 非模式生物转录组研究[J]. *遗传*, 2013, 35(8): 955-970.
- [8] LURO FL, COSTANTINO G, TEROL J, et al. Transferability of the EST-SSRs developed on Nules clementine (*Citrus clementina* Hort ex Tan) to other Citrus species and their effectiveness for genetic mapping [J]. *BMC Genomics*, 2008, 9(1): 287.
- [9] 王利英, 乔军, 石瑶, 等. 茄子 SSR 多态性引物的筛选及品种纯度鉴定[J]. *华北农学报*, 2012, 27(4): 98-102.
- [10] 李满堂, 张仕林, 邓 鹏, 等. 洋葱转录组 SSR 信息分析及其多态性研究[J]. *园艺学报*, 2015, 42(6): 1103-1111.
- [11] 刘峰, 王运生, 田雪亮, 等. 辣椒转录组 SSR 挖掘及其多态性分析[J]. *园艺学报*, 2012, 39(1): 168-174.
- [12] 朱晓敏, 王弦云, 陈龙灿, 等. 杜仲 SSR-PCR 反应体系建立及引物筛选[J]. *安徽农业大学学报*, 2013, 40(3): 378-382.
- [13] WEI W, QI X, WANG L, et al. Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers [J]. *BMC Genomics*, 2011, 12(1): 451.
- [14] LI D, DENG Z, QIN B, et al. *De novo* assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.) [J]. *BMC Genomics*, 2012, 13(1): 192.
- [15] 金基强, 崔海瑞, 陈文岳, 等. 茶树 EST-SSR 的信息分析与标记建立[J]. *茶叶科学*, 2006, 26(1): 17-23.
- [16] 姚明哲, 刘振, 陈亮, 等. 利用 EST-SSR 分析江北茶区茶树资源的遗传多样性和遗传结构[J]. *茶叶科学*, 2009, 29(3): 243-250.
- [17] 刘振, 姚明哲, 王新超, 等. 基于 EST-SSR 的福建地区茶树资源遗传多样性和亲缘关系分析[J]. *中国农业科学*, 2009, 42(5): 1720-1727.
- [18] 刘振, 王新超, 赵丽萍, 等. 基于 EST-SSR 的西南地区茶树资源遗传多样性和亲缘关系分析[J]. *分子植物育种*, 2008, 6(1): 100-110.
- [19] 杨华, 陈琪, 韦朝领, 等. 茶树转录组中 SSR 位点的信息分析[J]. *安徽农业大学学报*, 2011, 38(6): 882-886.
- [20] 王丽鸳, 韦康, 张成才. 茶树花转录组微卫星分布特征[J]. *作物学报*, 2014, 40(1): 80-85.
- [21] DECROOCQ V, FAVE M G, HAGEN L, et al. Development and transferability of apricot and grape EST microsatellite markers across taxa[J]. *Theoretical and Applied Genetics*, 2003, 106(5): 912-922.
- [22] CHEN C, ZHOU P, CHOI Y A, et al. Mining and characterizing microsatellites from citrus ESTs [J]. *Theoretical and Applied Genetics*, 2006, 112(7): 1248-1257.
- [23] CARDLE L, RAMSAY L, MILBOURNE D, et al. Computational and experimental characterization of physically clustered simple sequence repeats in plants [J]. *Genetics*, 2000, 156(2): 847-854.
- [24] 杨维泽, 金航, 赵振玲, 等. 西洋参 EST 资源的 SSR 信息分析[J]. *西南农业学报*, 2011, 24(1): 275-278.
- [25] 蔡年辉, 许玉兰, 徐杨, 等. 云南松转录组 SSR 的分布及其序列特征[J]. *云南大学学报: 自然科学版*, 2015, 37(5): 770-778.
- [26] MASON A S. SSR genotyping [M]//*Plant Genotyping*. New York: Springer, 2015: 77-89.
- [27] TÓTH G, GÁSPÁRI Z, JURKA J. Microsatellites in different eukaryotic genomes: survey and analysis [J]. *Genome Research*, 2000, 10(7): 967-981.
- [28] THAO D V, YAMASHITA M, WATANABE A, et al. Development of tetranucleotide microsatellite markers in *Pinus kesiya* Royle ex Gordon [J]. *Conservation Genetics Resources*, 2013, 5(2): 405-407.
- [29] CAMASAMUDRAM V, FANG J K, AVADHANI N G. Transcription termination at the mouse mitochondrial H-strand promoter distal site requires an A/T rich sequence motif and sequence specific DNA binding proteins [J]. *European Journal of Biochemistry*, 2003, 270(6): 1128-1140.
- [30] POOTAKHAM W, CHANPRASERT J, JOMCHAI N, et al. Development of genomic-derived simple sequence repeat markers in *Hevea brasiliensis* from 454 genome shotgun sequences [J]. *Plant Breeding*, 2012, 131(4): 555-562.
- [31] ECKERT K A, MOWERY A, HILE S E. Misalignment-mediated DNA polymerase β mutations: comparison of microsatellite and frame-shift error rates using a forward mutation assay [J]. *Biochemistry*, 2002, 41(33): 10490-10498.