

DOI: 10.3724/SP.J.1006.2010.00744

浙江省茶树地方品种与选育品种遗传多样性和群体结构的 EST-SSR 分析

乔婷婷 马春雷** 周炎花 姚明哲 刘 饶 陈 亮*

中国农业科学院茶叶研究所茶树资源与改良研究中心 / 国家茶树改良中心, 浙江杭州 310008

摘 要: 以茶树品种龙井 43 幼根为材料, 构建 cDNA 文库并测序, 获得 4 833 条 EST 序列, 拼接后得到 3 482 条无冗余 EST, 总长 2 290 kb。对其进行 SSR 搜索, 共检测到 577 个 SSR 位点, 分布于 500 条茶树幼根 EST 中, 其中含有 EST-SSR 的序列占 14.36%, 平均每 3.97 kb 出现一个 SSR。利用 Primer premier 5.0, 对含有 SSR 的 EST 设计引物 416 对, 通过退火温度和多态性筛选, 确定可用的引物及其最佳退火温度, 并从筛到的引物中选取 63 对及 1 对已发表引物作为核心引物, 对浙江省茶树地方品种和选育品种进行遗传多样性和遗传结构分析。结果显示, 64 对引物均在供试材料中表现出多态性, 共检测到 232 个等位位点, 平均每对引物 3.6 个; 每对引物可鉴定的基因型为 2~13 个, 平均 4.3 个。供试材料多态性信息量(PIC)介于 0.02~0.84, 平均 0.44; 扩增位点的观测杂合度平均为 0.44, 期望杂合度平均为 0.48。地方品种的遗传多样性水平略高于选育品种(系)。不同地方资源群体多态性信息量为 0.24~0.36, 举岩群体的多样性最高, 惠明群体的多样性最低。浙江各地区以杭州资源的多态性最高, PIC 达 0.41; 丽水的多态性最低, PIC 为 0.24。Structure 2.2 群体结构分析和 UPGMA 聚类分析表明, 地方品种、选育品种(系)具有相对独立的群体结构, 选育品种(系)根据亲缘关系的不同形成不同的类群。

关键词: 茶树; 遗传多样性; 遗传结构; EST-SSR

EST-SSR Genetic Diversity and Population Structure of Tea Landraces and Developed Cultivars (Lines) in Zhejiang Province, China

QIAO Ting-Ting, MA Chun-Lei**, ZHOU Yan-Hua, YAO Ming-Zhe, LIU Rao, and CHEN Liang*

Research Center for Tea Germplasm and Improvement, Tea Research Institute, Chinese Academy of Agricultural Sciences / National Center for Tea Improvement, Hangzhou 310008, China

Abstract: Tea [*Camellia sinensis* (L.) O. Kuntze] has a long history in production and consumption in Zhejiang province, China. Improvement of tea variety, therefore, is of great importance and a good understanding of the genetic diversity and population structure of tea germplasm is a prerequisite to the improvement. In spite of great advances on the use of molecular markers in tea plant, achievement is still gotten very slowly compared with in other cereal crops and woody species. Expressed sequence tag derived simple sequence repeat (EST-SSR) is a less costly alternative way of developing new markers for genetic diversity analysis, functional markers development, and marker-assisted breeding of tea plant. A total of 4 833 ESTs generated from a cDNA library of tea young root were subjected to SSR mining using DNASTAR 5.0 software, 577 EST-SSRs were identified and 416 primer pairs were designed by Primer premier 5.0. After the determination of annealing temperatures and polymorphism of all the primers, 64 core primers were selected and used for genetic diversity and population structure analyses of tea landraces and improved cultivars in Zhejiang province. All selected primers were polymorphic and 232 alleles were amplified with 3.6 alleles per primer pair on an average. Each primer pair identified 2 to 13 genotypes, with an average of 4.3. The mean of polymorphism information content (PIC) was 0.44, ranging from 0.02 to 0.84. Observed heterozygosity (H_o) was 0.44, while expected heterozygosity (H_e) was 0.48. The level of genetic diversity among landraces was slightly higher than that among improved cultivars and breeding lines. There were 226 alleles amplified in 22 landraces with 14 of them that were special. In the thirty-seven improved cultivars, however, two hundred and eighteen alleles were amplified but only six were special. The PIC of the landrace groups

本研究由国家高技术研究发展计划(863 计划)项目(2006AA10Z171), 国家农业(茶叶)产业技术体系建设专项, 国家自然科学基金项目(30901159), 国家公益性行业(农业)科研专项(3-35)和浙江省科技计划项目(2006C12129)资助。

* 通讯作者(Corresponding author): 陈亮, E-mail: liangchen@mail.tricaas.com; Tel: 0571-86652835

第一作者联系方式: E-mail: qiao.tingting123@163.com ** 共同第一作者

Received(收稿日期): 2009-09-25; Accepted(接受日期): 2010-01-10.

varied from 0.24 to 0.36, in which Juyan Qunti was the highest and Huiming Qunti was the lowest. Deqing Qunti was closest to Juyan Qunti in genetic relationship, but farthest from Huiming Qunti. The genetic diversity of tea cultivars from Hangzhou was the highest with PIC of 0.41, while those from Lishui recorded the lowest PIC of 0.24. Population structure revealed by software Structure 2.2 and UPGMA cluster analysis showed that landraces and improved cultivars were relatively independent. The improved cultivars were further clustered into smaller groups according to their pedigree. Hybrid offspring from Fuding Dabaicha and Yunnan Dayezhong from different breeding organizations fell into similar groups.

Keywords: Tea; Genetic diversity; Genetic structure; EST-SSR

浙江省地处中国东南沿海, 位于北纬 $27^{\circ}01' \sim 31^{\circ}10'$, 东经 $118^{\circ}01' \sim 123^{\circ}08'$; 全省属亚热带季风气候, 年均日照时数在 1 710~2 100 h 之间, 年平均气温 $15 \sim 18^{\circ}\text{C}$, 属于茶树生态适宜性生长区, 尤其适宜绿茶生产。浙江茶产业历史悠久, 经过长期的栽培驯化, 形成了很多优良的地方品种, 如龙井群体种、紫笋群体、日铸群体等。目前浙江省通过国家审(认)定的茶树品种共有 14 个, 如龙井 43、浙农 12、迎霜等; 通过省级审(认)定品种有 20 个, 如嘉茗 1 号、黄叶早、霜峰、白叶 1 号等, 这些优良品种资源是茶树新品种选育的重要亲本材料, 对其遗传多样性和亲缘关系进行分析具有重要的理论价值及指导意义。

DNA 分子标记克服了形态和生化标记易受生长环境、发育阶段和组织器官影响等缺点, 是比较理想的遗传标记。近年来, RAPD、RFLP、AFLP 和 ISSR 等分子标记已广泛应用于茶树的遗传多样性分析等研究^[1]。简单序列重复又称微卫星标记, 包括基因组 SSR 和 EST-SSR 两种, 与其他分子标记相比, 因其具有多态性高、操作技术简单、重现性好等特点, 已成为目前最重要的分子标记之一^[2]。其中基因组 SSR 标记的开发需要构建基因组文库, 费用较高且效率较低, 且茶树也还没有大规模的基因组测序计划, 所以目前茶树可用的基因组 SSR 标记甚少^[1]。而利用测序获得的 EST 数据开发功能性的 EST-SSR 标记, 不仅可降低引物开发的成本, 而且能大大提高现有测序数据的利用效率, 使得标记开发更加经济、有效。茶树 EST-SSR 标记的发展将影响遗传多样性分析、功能标记开发、基因作图、关联分析、标记辅助育种等各方面, 目前已应用于茶树遗传多样性及群体结构分析。金基强等^[3]首先建立了茶树 EST-SSR 体系, 后又采用 16 对引物对 42 份茶树资源进行分析并确定其有效性。之后又有一系列报道表明 EST-SSR 是茶树遗传分析的有效标记^[4-6], 但到现在为止可用的茶树 EST-SSR 标记数量并不多, 这限制了它的进一步应用。本文利用自行开发的茶树幼根 EST-

SSR 分子标记, 对浙江省的选育茶树品种和地方品种进行比较分析, 以期揭示浙江省茶树种质资源的遗传多样性和遗传结构, 为茶树种质资源保存、核心种质构建和育种利用等提供依据。

1 材料与方法

1.1 试验材料

从国家种质杭州茶树圃保存的浙江茶树资源中选取 59 个浙江省茶树品种(表1), 其中包括选育品种(系) 37 个, 地方品种 22 个, 均属于山茶属茶组 [*Camellia* L. Sect. *Thea* (L.) Dyer] 中的茶 [*C. sinensis* (L.) O. Kuntze]。无性系品种按单株取样, 有性系品种随机选取 2~8 个单株作为群体代表, 采集一芽二叶经液氮迅速冷冻后, 保存于 -70°C 冰箱备用。

1.2 DNA 提取

采用改进的 SDS 法^[7]分别提取 59 个品种资源的基因组 DNA, 以 0.8% 琼脂糖凝胶电泳检测 DNA 质量, 用 ND-1000 UV-Vis 分光光度计 (Nanodrop Technologies, Inc., Wilmington, DE, USA) 检测 DNA 浓度后, 将其稀释到 $10 \text{ ng } \mu\text{L}^{-1}$ 的工作浓度, 用于 PCR 扩增。

1.3 引物设计及筛选

利用 DNASTar 5.0 软件 (DNASTar, Inc., Madison, WI, USA) 对茶树幼根 cDNA 文库测序获得的 4 833 条有效 EST 序列^[8]进行拼接, 对拼接得到的所有非冗余序列进行 SSR 搜索 (<http://www.gramene.org/db/markers/ssrtool>), 然后应用 Primer premier 5.0 软件 (Premier Biosoft International, Palo Alto, CA, USA) 对包含 SSR 的序列设计引物, 共设计引物 416 对。参照引物的理论 T_m 值, 采用温度梯度 PCR, 选择杂带较少、条带清晰的温度作为相应引物的最终退火温度; 再选用 12 份资源进行引物多态性筛选, 选择条带清晰、多态性高、重现性好的 63 对茶树幼根引物为核心引物对所有供试样品进行扩增, 另外选用 Freeman 等^[9]已发表的 1 对引物, 标记为 Camsin M5, 作为参照引物。

表 1 供试浙江茶树品种资源的名称及来源
Table 1 Name and origin of tea germplasm in Zhejiang province, China

No.	名称 Name	类型 Type	原产地 Origin
1	龙井 43 Longjing 43	选育品种 Developed cultivar	杭州 Hangzhou
2	龙井长叶 Longjing Changye	选育品种 Developed cultivar	杭州 Hangzhou
3	龙井圆叶 Longjing Yuanye	品系 Breeding line	杭州 Hangzhou
4	菊花春 Juhuachun	选育品种 Developed cultivar	杭州 Hangzhou
5	寒绿 Hanlü	选育品种 Developed cultivar	杭州 Hangzhou
6	中茶 102 Zhongcha 102	选育品种 Developed cultivar	杭州 Hangzhou
7	茂红 Maohong	品系 Breeding line	杭州 Hangzhou
8	霜峰 Shuangfeng	选育品种 Developed cultivar	杭州 Hangzhou
9	劲峰 Jinfeng	选育品种 Developed cultivar	杭州 Hangzhou
10	垂绿 Chuilü	品系 Breeding line	杭州 Hangzhou
11	迎霜 Yingshuang	选育品种 Developed cultivar	杭州 Hangzhou
12	翠峰 Cuifeng	选育品种 Developed cultivar	杭州 Hangzhou
13	青峰 Qingfeng	选育品种 Developed cultivar	杭州 Hangzhou
14	茂绿 Maolü	选育品种 Developed cultivar	杭州 Hangzhou
15	浙农 12 Zhenong 12	选育品种 Developed cultivar	杭州 Hangzhou
16	浙农 138 Zhenong 138	品系 Breeding line	杭州 Hangzhou
17	浙农 113 Zhenong 113	选育品种 Developed cultivar	杭州 Hangzhou
18	浙农 121 Zhenong 121	选育品种 Developed cultivar	杭州 Hangzhou
19	浙农 117 Zhenong 117	选育品种 Developed cultivar	杭州 Hangzhou
20	嘉茗 1 号 Jiaming 1	地方品种 Landrace	温州 Wenzhou
21	黄叶早 Huangyeyao	地方品种 Landrace	温州 Wenzhou
22	平阳特早 Pingyang Tezao	地方品种 Landrace	温州 Wenzhou
23	^a 德清群体 ^a Deqing group ⁸	地方品种 Landrace	湖州 Huzhou
24	^a 建德群体 ^a Jiande group ⁷	地方品种 Landrace	杭州 Hangzhou
25	^a 举岩群体 ^a Juyan group ⁸	地方品种 Landrace	金华 Jinhua
26	^a 紫笋群体 ^a Zisun group ⁶	地方品种 Landrace	湖州 Huzhou
27	^a 日铸群体 ^a Rizhu group ⁴	地方品种 Landrace	绍兴 Shaoxing
28	^a 天台群体 ^a Tiantai group ⁶	地方品种 Landrace	台州 Taizhou
29	^a 开化种 ^a Kaihuazhong	地方品种 Landrace	衢州 Quzhou
30	^a 余姚种 ^a Yuyaozhong	地方品种 Landrace	宁波 Ningbo
31	^a 诸暨种 ^a Zhujizhong ³	地方品种 Landrace	绍兴 Shaoxing
32	^a 嵊县种 ^a Shengxianzhong ²	地方品种 Landrace	绍兴 Shaoxing
33	藤茶 Tengcha	地方品种 Landrace	台州 Taizhou
34	早青茶 Zaoqingcha	地方品种 Landrace	台州 Taizhou
35	水古茶 Shuigucha	地方品种 Landrace	台州 Taizhou
36	香菇寮白毫 Xiangguliao Baihao	地方品种 Landrace	温州 Wenzhou
37	黑踎踎 Heitata	地方品种 Landrace	温州 Wenzhou
38	福云 10-27 Fuyun 10-27	品系 Breeding line	杭州 Hangzhou
39	福云 9-11 Fuyun 9-11	品系 Breeding line	杭州 Hangzhou
40	福云 9-6 Fuyun 9-6	品系 Breeding line	杭州 Hangzhou
41	福云 8-16 Fuyun 8-16	品系 Breeding line	杭州 Hangzhou
42	福云 9-22 Fuyun 9-22	品系 Breeding line	杭州 Hangzhou
43	福云 0-3 Fuyun 0-3	品系 Breeding line	杭州 Hangzhou
44	红魁 Hongkui	品系 Breeding line	杭州 Hangzhou

(续表 1)

No.	名称 Name	类型 Type	原产地 Origin
45	福云 11-35 Fuyun 11-35	品系 Breeding line	杭州 Hangzhou
46	红云 Hongyun	品系 Breeding line	杭州 Hangzhou
47	蓝海 Lanhai	株系 Breeding line	杭州 Hangzhou
48	杭州大叶 Hangzhou Daye	株系 Breeding line	杭州 Hangzhou
49	^a 平阳种 ² ^a Pingyangzhong ²	地方品种 Landrace	温州 Wenzhou
50	大叶乌皮 Daye Wupi	地方品种 Landrace	台州 Taizhou
51	^a 惠明群体 ⁴ ^a Huiming group ⁴	地方品种 Landrace	丽水 Lishui
52	中茶 108 Zhongcha 108	品系 Breeding line	杭州 Hangzhou
53	狗牯脑 F1-1 Gougunao F1-1	株系 Breeding line	杭州 Hangzhou
54	狗牯脑 F1-2 Gougunao F1-2	株系 Breeding line	杭州 Hangzhou
55	狗牯脑 F1-3 Gougunao F1-3	株系 Breeding line	杭州 Hangzhou
56	白叶 1 号 Baiye 1	地方品种 Landrace	湖州 Huzhou
57	横山早 Hengshanzao	选育品种 Developed cultivar	湖州 Huzhou
58	洛舍 Luoshe	选育品种 Developed cultivar	湖州 Huzhou
59	报福 1 号 Baofu 1	选育品种 Developed cultivar	湖州 Huzhou

带 ^a 的品种资源属于有性系, 其余为无性系, 上标为有性品种个体数。
The germplasm marked with ^a are the sexual lines, and the rest are clone. The superscripts are individual number of the sexual lines.

1.4 PCR 扩增与产物检测

扩增反应在 PTC-200 型 PCR 仪(MJ Research, Inc., Waltham, MA, USA)上进行, 反应体系与扩增程序参照刘振等^[4]的方法。扩增产物用 10%的聚丙烯酰胺凝胶, 在 CBS 垂直电泳系统(C.B.S. Scientific Company, Inc., San Diego, CA, USA)上进行电泳, 电压 150 V, 时间为 90 min, 电泳后参照 Charters 和 Wilkinson^[10]的方法银染显色。用 Bio-Rad Gel Doc 2000 (Bio-Rad Laboratories, Inc., Philadelphia, PA, USA)凝胶成像仪拍照记录。

1.5 数据处理

利用人工方法读带, 将电泳图上清晰的条带记为 1, 同一位置无带或不易分辨的弱带记为 0, 建立 0-1 原始数据库, 然后将相同带型的记为一种基因型。采用 Powermarker 3.25 软件^[11]计算每对引物扩增的等位位点数(alleles)、主要等位位点频率(major allele frequency)和基因型数(genotypes), 根据 Nei^[12]的方法计算等位位点的期望杂合度(expected heterozygosity, H_e)、观测杂合度(observed heterozygosity, H_o)和每对引物扩增位点的多态性信息量(polymorphism information content, PIC)^[13], 然后根据 Nei's 遗传距离^[14-15]绘制 UPGMA 聚类图。采用 Structure 2.2 软件^[16]模拟群体遗传结构, 分析的基本原理是, 假设样本存在 K 个(K 值未知)等位变异频率特征类型数, 将供试材料归到 K 个群体, 使得该群体位点

频率都遵循同一个 Hardy-Weinberg 平衡。本实验假定试材的群体数目(K)为 2~10, 并假定 64 个位点相互独立进行分析, 将 MCMC (Markov chain Monte Carlo)开始时的不作数迭代后(length of burn-in period)设为 100 000 次, 再将不作数迭代后的 MCMC 设为 10 000, 然后根据最大似然值原则选择合适的群体(K)值。

2 结果与分析

2.1 茶树幼根 EST-SSR 分布频率及特点

对幼根文库测序得到的 4 833 条序列进行拼接, 获得总长约 2 290 kb 的 3 482 条非冗余序列, 包括 901 个 contigs 和 2 581 个 singletons。对这些序列进行 SSR 搜索, 按照二、三、四、五、六和七以上核苷酸重复数为 8、5、4、3、3 和 3 进行分析, 最后获得 577 个 SSR, 包含于 500 条 EST 中, 含有 EST-SSR 的序列比例为 14.36%, 每 3.97 kb 出现一个 SSR, 其中二核苷酸重复和三核苷酸重复类型占全部 SSR 的比例高达 59.10% (表 2)。所用 64 对核心引物主要从三核苷酸重复(28.1%)、五核苷酸重复(25.0%)、六核苷酸重复(28.1%) 3 种 SSR 类型的侧翼设计引物, 目的片段在 80~330 bp 之间(表 3)。

2.2 茶树幼根 EST-SSR 引物特征及扩增结果

64 对 EST-SSR 引物在供试茶树品种资源中均表现出多态性, 多态性位点百分率为 100%, 共检测

表 2 茶树幼根 EST 中 SSR 的数量及出现的频率
Table 2 Number and frequency of SSR in tender root EST of tea plant

重复类型 Repeat type	SSR 数目 Number of SSRs	占全部 SSR 比例 Proportion in all SSRs (%)	出现频率 Frequency (%)	平均分布频率 Average distance (kb/SSR)
二核苷酸 Dinucleotide	213	36.92	6.12	10.75
三核苷酸 Trinucleotide	128	22.18	3.68	17.89
四核苷酸 Tetranucleotide	44	7.63	1.26	52.05
五核苷酸 Pentanucleotide	95	16.46	2.73	24.11
六核苷酸 Hexanucleotide	86	14.90	2.47	26.63
七核苷酸 Heptanucleotide	8	1.39	0.23	286.25
八核苷酸 Octanucleotide	2	0.35	0.06	1145.00
十核苷酸 Decanucleotide	1	0.17	0.03	2290.00
总计 Total	577	100.00	16.57	3.97

到 232 个等位位点, 平均每对引物可检测到 3.6 个, 其中最多为 6, 最少为 2。每个引物可鉴定的基因型数 2~13 之间, 平均 4.3 个。引物扩增位点的多态性信息量(PIC)变化范围在 0.02~0.84, 平均值为 0.44。等位位点的观测杂合度(H_o)平均值为 0.44, 期望杂合度(H_e)平均值为 0.48 (表 3)。

2.3 浙江省茶树品种资源的遗传多样性及其比较
地方品种的遗传多样性略高于选育品种, 所检测的基因型数平均分别为 4.0 和 3.7, 基因多样性指数平均分别为 0.47 和 0.45, 多态性信息量分别为 0.42 和 0.41 (图 1)。22 个地方品种共检测到 226 个等位基因, 占总数的 97.4%, 64 个位点上发现 14 个

表 3 茶树幼根 EST-SSR 引物特征及其扩增结果
Table 3 Characterization of EST-SSRs in tea young root and amplification results

引物编号 Primer code	重复位点 Repeat motif	目的片段长度 Fragment length (bp)	熔解温度 T_m (°C)	退火温度 T_a (°C)	等位 位点数 Allele	主要等位 位点频率 Main allele frequency	基因型数 Genotype	观测 杂合度 H_o	期望 杂合度 H_e	多态性 信息量 PIC
CamsinM5	(GT)15(GA)13	170–205	55.0	55	5	0.31	5	1.00	0.76	0.72
TM125	(GA)9	150–220	48.7	55	5	0.34	6	0.86	0.77	0.74
TM126	(TAT)7	260–295	45.7	55	4	0.31	5	0.64	0.76	0.72
TM127	(AATG)6	240–300	44.9	42	3	0.38	5	0.76	0.73	0.68
TM128	(AAATA)5	160–200	43.7	45	3	0.28	7	0.60	0.79	0.76
TM129	(GTG)5	175–215	50.0	51	4	0.78	4	0.87	0.36	0.34
TM130	(CCA)7	145–170	50.9	58	3	0.79	6	0.38	0.36	0.34
TM131	(GAGTGT)4	160–200	45.4	54	3	0.90	3	0.24	0.18	0.17
TM132	(TTGTG)3	95–125	43.1	51	5	0.25	13	0.68	0.85	0.84
TM133	(CCT)5	145–180	48.2	56	3	0.60	5	0.42	0.56	0.50
TM134	(CAT)8	230–260	47.8	56	3	0.46	4	0.55	0.58	0.49
TM135	(TAT)6	110–160	42.4	51	3	0.55	6	0.53	0.60	0.55
TM136	(TTTTG)5	220–270	45.7	54	4	0.50	6	0.5	0.66	0.61
TM137	(AGAGA)3	195–330	50.5	58	3	0.72	3	0.43	0.43	0.38
TM138	(AT)11	180–205	42.4	51	6	0.60	3	0.48	0.55	0.48
TM139	(ATTTTA)3	145–160	45.3	48	3	0.71	4	0.38	0.46	0.43
TM140	(GCACCA)3	250–290	54.5	58	3	0.52	3	0.50	0.55	0.45
TM141	(AAAAT)3	140–175	41.0	51	4	0.84	5	0.21	0.29	0.28
TM142	(TCCGAT)3	230–250	50.8	58	2	0.95	2	0.21	0.09	0.09
TM143	(TAC)6	200–250	44.8	54	5	0.33	8	0.95	0.74	0.70
TM144	(TAT)5	190–205	46.3	51	2	0.91	2	0.22	0.17	0.16

(续表 3)

引物编号 Primer code	重复位点 Repeat motif	目的片段长度 Fragment length (bp)	熔解温度 T_m (°C)	退火温度 T_a (°C)	等位 位点数 Allele	主要等位 位点频率 Main allele frequency	基因型数 Genotype	观测 杂合度 H_o	期望 杂合度 H_e	多态性 信息量 PIC
TM145	(TTC)7	175–200	42.5	48	6	0.56	7	0.98	0.63	0.60
TM146	(TTTCT)3	155–210	47.2	56	4	0.99	2	0.10	0.02	0.02
TM147	(GTG)5	150–170	49.1	48	3	0.95	2	0.19	0.09	0.09
TM148	(ACC)8	135–155	55.2	58	2	0.71	3	0.31	0.43	0.37
TM149	(TTTGC)3	120–130	44.5	51	2	0.49	3	0.35	0.62	0.55
TM150	(TTTTTA)3	220–250	46.2	51	3	0.57	4	0.64	0.52	0.43
TM151	(CATATG)4	140–205	45.2	51	4	0.65	4	1.00	0.52	0.46
TM152	(TTTTA)4	150–180	43.1	48	5	0.54	4	0.49	0.61	0.55
TM153	(ATTTT)4	220–260	45.9	56	4	0.92	3	0.13	0.14	0.13
TM154	(TCA)6	200–250	47.9	56	4	0.67	5	0.73	0.47	0.40
TM155	(TTTTTC)4	200–225	46.3	56	3	0.39	4	0.52	0.70	0.65
TM156	(ATTTT)5	220–255	45.5	54	4	0.36	8	0.59	0.77	0.73
TM157	(TTTTTA)5	80–120	43.3	51	6	0.27	8	0.43	0.80	0.77
TM158	(TCT)6	235–265	44.1	54	4	0.52	6	0.45	0.64	0.59
TM159	(AT)9	175–215	43.1	51	5	0.74	6	0.37	0.43	0.40
TM160	(GAT)6	240–295	41.3	51	3	0.67	3	0.34	0.46	0.37
TM161	(GCACCT)3	140–155	52.5	58	4	0.80	3	0.77	0.34	0.31
TM162	(CTCA)5	100–160	48.0	58	3	0.36	4	0.61	0.73	0.68
TM163	(ATCAGA)3	145–160	49.4	58	3	0.79	3	0.18	0.35	0.32
TM164	(GCTA)5	160–170	46.6	54	2	0.60	3	0.17	0.56	0.49
TM165	(ATAA)4	220–250	46.9	56	3	0.71	4	0.51	0.43	0.37
TM166	(AAGGTAC)3	180–250	44.4	51	3	0.91	3	0.07	0.16	0.15
TM167	(TGTATT)3	105–110	41.8	48	2	1.00	2	0.00	0.06	0.05
TM168	(ATATG)3	130–150	43.5	51	2	0.79	3	0.18	0.34	0.30
TM169	(TTTTTG)4	160–210	42.8	51	4	0.56	3	0.32	0.56	0.48
TM170	(CAG)5	175–195	49.6	56	3	0.84	4	0.11	0.27	0.25
TM171	(AAGGAG)3	275–305	50.8	51	3	0.55	3	0.40	0.51	0.40
TM172	(TTTTC)3	200–300	47.9	56	6	0.93	2	0.00	0.14	0.13
TM173	(TGAG)4	240–300	46.8	51	3	0.99	2	0.01	0.02	0.02
TM174	(TTTAT)5	180–200	47.9	56	3	0.66	3	0.06	0.48	0.41
TM175	(TCTCTG)3	90–145	46.4	54	5	0.38	6	0.56	0.76	0.72
TM176	(ACTCT)3	175–210	43.9	48	5	0.42	6	0.57	0.68	0.63
TM177	(TATTT)3	185–230	44.5	51	4	0.41	5	0.44	0.65	0.59
TM178	(ATC)5	178–220	51.6	58	4	0.66	3	1.00	0.48	0.40
TM179	(TGA)8	148–170	46.8	56	3	0.41	5	0.39	0.73	0.68
TM180	(GAAAG)3	160–175	47.6	56	2	0.46	3	0.45	0.63	0.56
TM181	(TCTTT)3	200–250	43.8	54	5	0.83	2	0.17	0.29	0.25
TM182	(TTCTT)3	240–255	46.9	54	4	0.89	3	0.12	0.20	0.19
TM183	(TA)10	110–145	41.4	48	5	0.42	7	0.27	0.73	0.70
TM184	(TTTC)4	120–140	44.7	54	3	0.94	2	0.60	0.12	0.11
TM185	(TTTATT)3	120–135	41.7	48	3	0.98	2	0.02	0.04	0.04
TM186	(CACCAA)3	280–300	50.6	51	4	0.43	7	0.54	0.64	0.58
TM187	(TTG)5	140–155	43.4	54	3	0.58	6	0.31	0.61	0.58
Mean	—	—	—	—	3.6	0.63	4.3	0.44	0.48	0.44

特异等位基因, 仅出现在地方品种中; 37 个选育品种共检测到 218 个等位基因, 较地方品种减少 3.5%, 检测到特异等位基因 6 个。

对地方资源群体的多样性分析结果表明, 基因多样性指数在 0.29~0.41 之间, 多态性信息量在 0.24~0.36 之间, 其中举岩群体的多样性最高, 惠明群体的多样性最低(图 2)。聚类分析表明, 惠明群体与其他群体的亲缘关系最远, 德清群体与举岩群体的亲缘关系最近; 同属湖州的德清群体和紫笋群体在亲缘关系上稍远(图 3)。

对浙江省不同地区茶树资源的遗传多样性分析结果显示杭州资源的多态性最高, 在 64 个位点上共检测到 197 个等位基因, 平均基因型数 4.0, PIC 达 0.41; 丽水的多态性最低, 共检测到 97 个等位基因, 基因型数 1.8, PIC 为 0.24 (图 4)。

2.4 浙江茶树品种资源的遗传结构

利用 Structure 2.2 软件对浙江茶树品种资源进行遗传结构分析, 当 $K=2$ 时, 能够取得最大的 $\ln P(D)$ 值和比较稳定的 α 值。即浙江省选育品种和地方品种的遗传结构可分为两大群体, 群体 I 主要包括大多地方品种和少数选育品种(系), 其中地方品种占 89.6%, 选育品种占 10.4%, 地方品种主要有德清

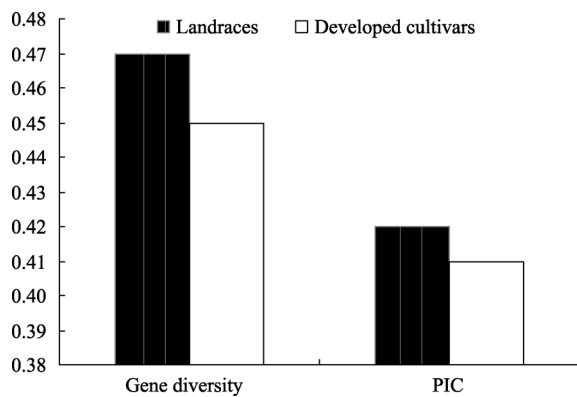


图 1 地方品种和选育品种的遗传多样性比较

Fig. 1 Comparison of genetic diversity between landraces and developed cultivars

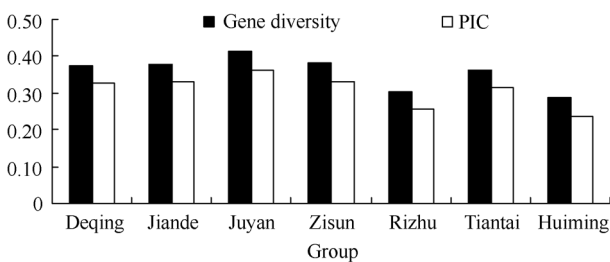


图 2 地方群体的遗传多样性水平比较

Fig. 2 Genetic diversity among land race groups

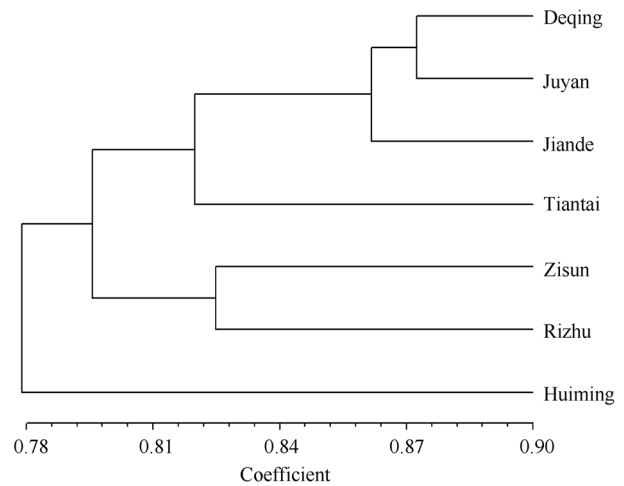


图 3 地方群体的亲缘关系图

Fig. 3 Dendrogram of landrace populations in tea germplasm

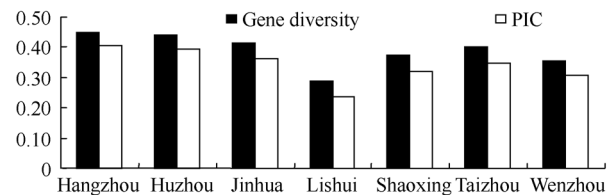


图 4 浙江省不同地区茶树种质资源的遗传多样性

Fig. 4 Genetic diversity of tea germplasms in different regions in Zhejiang province

群体、建德群体、举岩群体、紫笋群体和余姚种等; 群体 II 主要包括各地选育品种, 占 75.0%, 以及少量地方品种(占 25.0%), 其中选育品种中由福鼎大白茶与云南大叶茶后代单株育成的占 50%。从地理位置看, 浙江省不同地区茶树资源在两个群体里呈穿插分布, 未表现出明显的区域分化。

基于 Nei's 遗传距离构建 UPGMA 聚类图, 可以看出各个地方群体聚在一个小类群里, 如惠明群体、建德群体、举岩群体、德清群体、天台群体、紫笋群体等。大多选育品种根据遗传背景聚类, 不同选育单位的选育品种聚在不同的小类群中, 比如由浙江大学茶学系(原浙江农业大学茶学系)选育的浙农 12、浙农 113、浙农 138 聚在一个小类群里; 由杭州市农业科学研究院茶叶研究所(原浙江省余杭茶叶试验场)选育的劲峰、青峰、迎霜聚在一个小类群里。上述品种中由福鼎大白茶与云南大叶茶自然杂交后代单株育成又与福云系列同属同一类群。中国农业科学院茶叶研究所选育的品种龙井 43、龙井长叶、龙井圆叶、菊花春、寒绿在聚类群中呈分散分布(图 5)。基于 Structure 模型和遗传距离的分析结果可以看出, 浙江省的地方品种和选育品种具有相对独立的遗传结构。

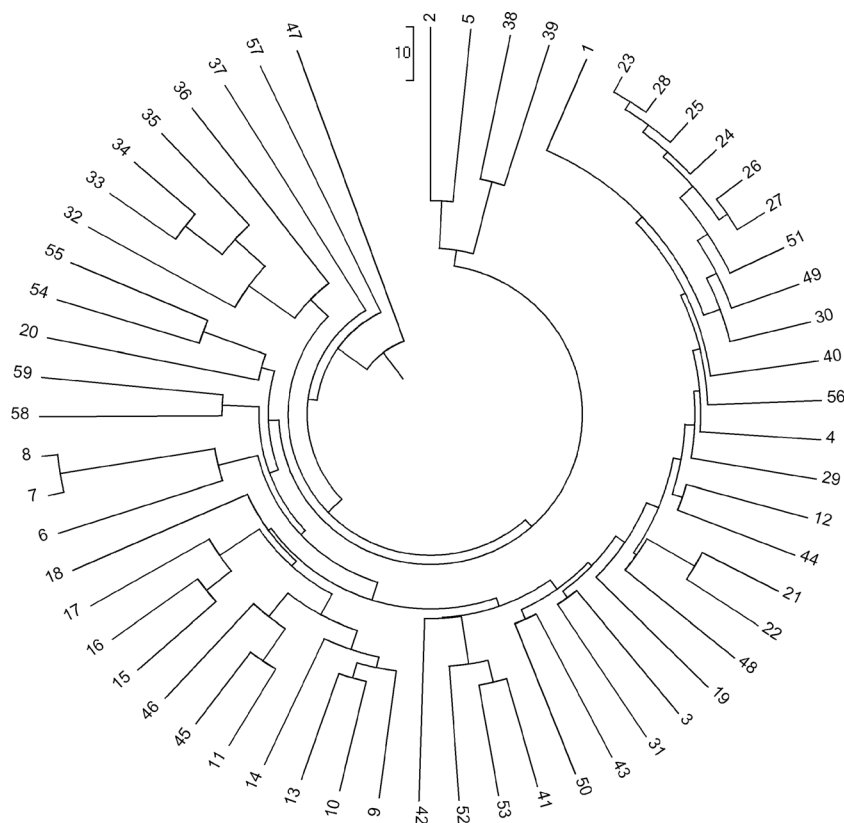


图 5 基于 Nei's 遗传距离的浙江省茶树种质资源聚类图

Fig. 5 Dendrogram of tea germplasm from Zhejiang province based on Nei's genetic distance

图中编号与表 1 同。The codes used in the dendrogram correspond with the number for germplasm given in Table 1.

3 讨论

3.1 茶树幼根 EST-SSR 分布频率及特征

一般 EST 中 SSR 序列统计结果表明, 三核苷酸重复含量最高^[17], 其次为二核苷酸重复与四核苷酸重复^[18]。但在本研究中, 茶树幼根 EST 中的 SSR 以二核苷酸重复和三核苷酸重复为主, 且以二核苷酸重复分布频率最高, 这与茶树新梢 EST 中 SSR 的搜索结果大致相同^[4,19], 这种差异产生的原因可能是不同学者所采用的搜索标准不同。另外, 在大麦、玉米、燕麦和小麦等禾谷类作物 EST 中, SSR 出现频率一般在 7%~10% 之间^[18]。Sharma 等^[20]对在 NCBI 上登录的茶树 EST 进行 SSR 搜索表明, 茶树中 EST-SSR 的出现频率为 8.9%, 而本研究中幼根文库的 SSR 频率却达到了 14.36%。不管是同其他作物相比, 还是与已有的茶树 EST 相比, 幼根构建的文库中 SSR 的出现频率都较高, 究其原因, 可能是前期所建的根文库质量较高, 测序获得的 EST 平均长度较长, 说明 cDNA 文库的质量可能直接影响 SSR 分布频率。

3.2 浙江省茶树品种资源的遗传多样性

从分子水平上看, PIC 越高, 物种遗传差异越大, 遗传多样性越高, 同时也就证明此物种的遗传背景越复杂。当 $PIC > 0.5$ 时, 该基因座为高度多态性; $0.25 < PIC < 0.5$ 时, 为中度多态性; $PIC < 0.25$ 时, 为低度多态性^[21]。本研究采用的 64 对引物扩增的 PIC 平均为 0.44, 扩增结果显示浙江省茶树品种资源多态性属于中度偏高, 低于福建种质 PIC 平均值 0.56^[22]和云南种质的 0.50^[23]。研究发现, 选育品种的多态信息含量和基因多样性均低于地方品种。这一现象在水稻中也有过报道, 在对中国普通野生稻与栽培稻 SSR 多样性的比较分析中发现, 地方品种等位基因数显著高于选育品种, 选育品种基因数仅占地方品种的 81.1%, 在品种改良过程中导致等位基因数下降, 多态性降低^[24]。这些结果都表明在选育品种及大力推广过程中, 长期的栽培驯化使得多样性有所降低, 部分等位基因损失, 因此应引起茶树育种者的重视, 增加亲本来源的多样性, 提高选育品种的遗传多样性。对各地方群体分析结果显示其多样性水平均低于地方品种的总体水平, 说明对

各个地方群体都要适当保护以保持浙江省茶树品种资源的多样性。对浙江省不同地区茶树资源的分析显示杭州的多态性最高,可能是因为杭州的茶树资源是比较特殊的群体,很多选育品种在品种改良过程中引入了外源基因或由不同群体选育、辐照而成,遗传背景和育种手段的不同导致多态性丰富,而其他地区大多是当地的地方群体,由于地理起源一致,来源于少量个体,经过繁衍形成了不同群体,导致遗传多样性较低。

3.3 浙江省茶树品种资源群体遗传结构的比较

对浙江省茶树资源遗传结构的分析发现,地方品种和选育品种具有相对独立的群体遗传结构,选育品种又形成几个小群体,由浙江大学茶学系和杭州市农业科学研究院茶叶研究所选育的品种聚在一个小类群里,而由中国农业科学院茶叶研究所选育的品种则分散在整个聚类群内,主要是因为亲本的关系导致遗传距离不同。前两者是由福鼎大白茶与云南大叶茶自然杂交后代单株选育的,亲本一致导致遗传距离较近而聚成一个小类群;而后者是由不同亲本选育的,如菊花春由平阳群体种和云南大叶茶自然杂交后代采用单株育种法育成,寒绿由格鲁吉亚8号有性后代中采用单株育种法育成,而龙井43等是从龙井群体采用单株育种法育成,导致遗传距离较远。因此在育种过程中应该适当考虑外源基因,这样有助于保留优异等位基因和品种的多样化,形成丰富多彩的茶树品种。

4 结论

用茶树幼根 EST-SSR 引物很好地揭示了浙江省茶树种质资源的遗传多样性和遗传结构,属于中度多态性。与地方品种相比,选育品种损失了部分等位基因,基因多样性下降,表明对地方品种的开发和利用相当重要。在选育过程中,要适当引入外源基因,丰富亲本以提高选育品种的多样性。

References

- [1] Ni S, Yao M Z, Chen L, Zhao L P, Wang X C. Germplasm and breeding research of tea plant based on DNA marker approaches. *Front Agric China*, 2008, 2: 200–207
- [2] Gupta P K, Rustgi S, Sharma S, Singh R, Kumar N, Balyan H S. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol Genet Genomics*, 2003, 270: 315–323
- [3] Jin J-Q(金基强), Cui H-R(崔海瑞), Chen W-Y(陈文岳), Lu M-Z(卢美贞), Yao Y-L(姚艳玲), Xin Y(忻雅), Gong X-C(龚晓春). Data mining for SSRs in ESTs and development of EST-SSR marker in tea plant (*Camellia sinensis*). *J Tea Sci* (茶叶科学), 2006, 26(1): 17–23 (in Chinese with English abstract)
- [4] Liu Z(刘振), Wang X-C(王新超), Zhao L-P(赵丽萍), Yao M-Z(姚明哲), Wang P-S(王平盛), Xu M(许玫), Tang Y-C(唐一春), Chen L(陈亮). Genetic diversity and relationship analysis of tea germplasms originated from south western China based on EST-SSR. *Mol Plant Breed* (分子植物育种), 2008, 6(1): 100–110 (in Chinese with English abstract)
- [5] Yang Y(杨阳), Liu Z(刘振), Zhao Y(赵洋), Liang G-Q(梁国强), Zhao X(赵熙). Genetic diversity and relationship of Huangjinchu cultivar based on EST-SSR markers. *J Tea Sci* (茶叶科学), 2009, 29(3): 236–242 (in Chinese with English abstract)
- [6] Yao M-Z(姚明哲), Liu Z(刘振), Chen L(陈亮), Wang X-C(王新超), Ma C-L(马春雷), Liang Y-R(梁月荣). Genetic diversity and structure of tea germplasm originated from region of north Yangtze River based on EST-SSR markers. *J Tea Sci* (茶叶科学), 2009, 29(3): 243–250 (in Chinese with English abstract)
- [7] Chen L(陈亮), Chen D-M(陈大明), Gao Q-K(高其康), Yang Y-J(杨亚军), Yu F-L(虞富莲). Isolation and appraisal of genomic DNA from tea plant [*Camellia sinensis* (L.) O. Kuntze]. *J Tea Sci* (茶叶科学), 1997, 17(2): 177–181 (in Chinese with English abstract)
- [8] Zhao L-P(赵丽萍), Ma C-L(马春雷), Chen L(陈亮). Construction and expressed sequence tags analysis of young roots cDNA library of tea plant. *Mol Plant Breed* (分子植物育种), 2008, 6(5): 893–898 (in Chinese with English abstract)
- [9] Freeman S, West J, James C, Lea V, Mayes S. Isolation and characterization of highly polymorphic microsatellites in tea (*Camellia sinensis*). *Mol Ecol Notes*, 2004, 4: 324–326
- [10] Charters Y M, Wilkinson M J. The use of self-pollinated progenies as ‘in-groups’ for the genetic characterization of cocoa germplasm. *Theor Appl Genet*, 2000, 100: 160–166
- [11] Liu K, Muse S V. Powermarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, 2005, 21: 2128–2129
- [12] Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 1978, 89: 583–590
- [13] Anderson J A, Churchill G A, Autrique J E, Tanksley S D, Sorrells M E. Optimizing parental selection for genetic linkage maps. *Genome*, 1993, 36: 181–186
- [14] Nei M, Tajima F A, Tateno Y. Accuracy of estimated phylogenetic trees from molecular data. *J Mol Evol*, 1982, 18: 387–404
- [15] Takezaki N, Nei M. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, 1996, 144: 389–399
- [16] Prichard J K, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*, 2000, 155: 945–959
- [17] Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R. Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics*, 2000, 156: 847–854

- [18] Kantety R V, La Rota M, Matthews D E, Sorrells M E. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol Biol*, 2002, 48: 501–510
- [19] Chen L, Zhao L P, Gao Q K. Generation and analysis of expressed sequence tags from the tender shoots cDNA library of tea plant (*Camellia sinensis*). *Plant Sci*, 2005, 168: 359–363
- [20] Sharma R K, Bhardwaj P, Negi R, Mohapatra T, Ahuja P S. Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.). *BMC Plant Biol*, 2009, 9: 53
- [21] Botstein D, White R L, Skolnick M, Davis R W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Human Genet*, 1980, 32: 314–331
- [22] Liu Z(刘振), Yao M-Z(姚明哲), Wang X-C(王新超), Chen L(陈亮). Analysis of genetic diversity and relationship of tea germplasms originated from Fujian province based on EST-SSR markers. *Sci Agric Sin (中国农业科学)*, 2009, 42(5): 1720–1727 (in Chinese with English abstract)
- [23] Liu B-Y(刘本英). Application Studies of EST-SSR and ISSR Markers in Tea Germplasm (*Camellia* spp.) of Yunnan. PhD Dissertation of Chinese Academy of Agricultural Sciences, 2009. pp 49–50 (in Chinese with English abstract)
- [24] Zhang X-L(张晓丽), Guo H(郭辉), Wang H-G(王海岗), Lü J-Z(吕建珍), Yuan X-P(袁筱萍), Peng S-T(彭锁堂), Wei X-H(魏兴华). Comparative assessment of SSR allelic diversity in wild and cultivated rice in China. *Acta Agron Sin (作物学报)*, 2008, 34(4): 591–597 (in Chinese with English abstract)