

DOI: 10.13718/j.cnki.xdzk.2019.08.006

# 贵阳花溪古茶树遗传进化的 SNP 分析<sup>①</sup>

陈立杰<sup>1</sup>, 张素勤<sup>2</sup>, 尹杰<sup>3</sup>, 宋勤飞<sup>3</sup>, 牛素贞<sup>3</sup>,  
赵基英<sup>4</sup>, 陈丹萍<sup>4</sup>, 王胜威<sup>4</sup>, 耿广东<sup>2</sup>

1. 贵州大学 资产经营办公室, 贵阳 550025; 2. 贵州大学 农学院, 贵阳 550025;  
3. 贵州大学 茶学院, 贵阳 550025; 4. 贵阳市花溪区农业局, 贵阳 550025

**摘要:** 试验采用 SNP(单核苷酸多态性)技术对贵阳花溪古茶树资源的遗传多样性、群体结构和遗传进化进行了分析. 结果表明, 利用生物信息学分析古茶树样品的重测序数据, 获得了 1 656 258 个 SLAF 标签, 古茶树样品平均测序深度为 24.21x, 其中多态性的 SLAF 标签有 462 897 个, 共得到 283 376 个高一一致性的群体 SNP 标记. 从基于 SNP 标记获得的进化树可以看出, 来自相近地方的古茶树在进化树上基本上处于相近位置, 但是花溪古茶树存在着广泛的遗传变异, 主成分分析(PCA)获得了与进化树一致的结果. 群体结构分析可清晰地把古茶树资源分为乔木型和灌木型 2 个类群. 乔木型古茶树位于进化树下部, 而灌木型古茶树处于进化树上部, 表明茶树是由乔木型向灌木型进化的.

**关键词:** 古茶树; 特异位点扩增片段测序(SLAF-seq); SNP; 进化树; PCA; 遗传结构

**中图分类号:** S571.1

**文献标志码:** A

**文章编号:** 1673-9868(2019)08-0033-08

茶树(*Camellia sinensis*)是世界上三大饮料作物之一, 具有重要的经济、保健和生态价值. 目前茶树的分布最北界限达北纬 42°, 最南达南纬 33°, 遍及四大洲的 50 多个国家<sup>[1]</sup>. 茶树的原产地位于我国云贵高原以大娄山脉为中心的地域, 其中云南普洱茶和贵州绿茶均以独特的风味品质而被大家熟知<sup>[2-3]</sup>; 茶树原产地得天独厚的自然环境和悠久的种茶历史造就了丰富的古茶树资源, 由于其具有广泛的遗传变异和大量原始的优良基因, 近年来愈来愈受到茶叶界的关注<sup>[4]</sup>. 贵阳花溪古茶树栽培历史悠久, 分布广泛, 资源丰富, 仅久安乡境内就具有保护价值的古茶树 54 000 多株, 占地 266.67 余 hm<sup>2</sup>, 其中 2000 年以上的古茶树有 19 株, 1 000~1 500 年之间的古茶树有 1 450 株. 因此该区域的古茶树资源是开展茶树起源、演化、种质创新和新品种选育等研究的天然宝库. 随着分子生物学发展, 以 DNA 分子标记为标志的基因组研究进展迅速, RAPD<sup>[5-7]</sup>, RFLP<sup>[8]</sup>, AFLP<sup>[9-10]</sup>, SSR<sup>[11-12]</sup>, ISSR<sup>[13-14]</sup> 和 EST-SSR<sup>[15]</sup> 等标记已用于茶树种质资源和品种鉴别、遗传多样性、亲缘关系和遗传演化等研究. 单核苷酸多态性(Single nucleotide polymorphism, SNP)技术具有自动化程度高、通量大、速度快、易于标准化操作、适合大规模研究及基因分型等优点. 李志远等<sup>[16]</sup>搜集生产上主要甘蓝品种, 利用重测序数据与参考基因组进行比对开发了 2.54×10<sup>6</sup> 个 SNP 标记. 马宇等<sup>[17]</sup>利用 SLAF-seq 测序技术获得 1 105 347 个向日葵 SLAF 标签, 其中多态性 SLAF 标签共有 86 985 个, 获得了 414 692 个群体 SNP 标记. 李敏等<sup>[18]</sup>采用 SLAF-seq 技术获得了 277 333 个乔木柳 SLAF 标签, 其中多态性标签 99 526 个, 开发了 9 488 个 SNP 位点. 刘丽华等<sup>[19]</sup>仅需 14 个 SNP 位点就可

① 收稿日期: 2018-08-04

基金项目: 中央财政农业生产发展基金(黔财农[2015]240号); 贵州省农业攻关项目(黔科合支撑[2016]2603号, 黔科合支撑[2018]2330号).

作者简介: 陈立杰(1964-), 男, 高级实验师, 主要从事茶树育种与技术推广的研究.

通信作者: 耿广东, 博士, 教授, 硕士研究生导师.

将 378 份测试小麦中的 99.5% 的材料区分开. 本试验拟采用 SNP 技术, 对贵阳花溪古茶树进行分子标记开发, 获得全基因组范围内的分子标记, 了解贵阳花溪古茶树的遗传多样性、群体结构和遗传进化, 为茶树特异 SNP 标记开发、资源鉴定分析、高密度遗传连锁图谱构建和重要农艺性状的关联分析等奠定基础.

## 1 材料与方法

### 1.1 材料

供试材料来自贵州省不同地区的 48 份典型古茶树, 其中 30 份古茶树材料来自贵阳花溪, 其余 18 份古茶树材料来自贵州其它不同地区, 古茶树的来源及树型的详细信息见表 1.

表 1 供试古茶树的来源和树型

序号	古茶树资源	来源	树型	序号	古茶树资源	来源	树型
1	KNZ-1	贵阳开阳	灌木	25	HJ08	贵阳花溪久安乡	灌木
2	KNZ-2	贵阳开阳	灌木	26	HJ09	贵阳花溪久安乡	灌木
3	KNZ-3	贵阳开阳	灌木	27	HJ10	贵阳花溪久安乡	灌木
4	HGL-1	贵阳花溪高坡乡	乔木	28	N1	贵阳花溪久安乡	灌木
5	HGL-2	贵阳花溪高坡乡	灌木	29	N12	贵阳花溪久安乡	灌木
6	HGL-3	贵阳花溪高坡乡	乔木	30	N14	贵阳花溪久安乡	灌木
7	HGK-1	贵阳花溪高坡乡	乔木	31	N22	贵阳花溪久安乡	灌木
8	HGK-2	贵阳花溪高坡乡	乔木	32	N30	贵阳花溪久安乡	灌木
9	HGK-3	贵阳花溪高坡乡	乔木	33	N31	贵阳花溪久安乡	灌木
10	HGK-4	贵阳花溪高坡乡	乔木	34	N252	贵州普定	灌木
11	HQZ-1	贵阳花溪黔陶乡	灌木	35	N254	贵州普定	乔木
12	HQZ-2	贵阳花溪黔陶乡	灌木	36	N282	贵州务川	灌木
13	HQZ-3	贵阳花溪黔陶乡	灌木	37	N283	贵州务川	乔木
14	XSHTLZ-1	贵州习水	乔木	38	N291	贵州务川	灌木
15	XSHTLZ-2	贵州习水	乔木	39	N315	贵州安顺	灌木
16	XSHTLZ-3	贵州习水	乔木	40	N318	贵州安顺	灌木
17	XSHTLZ-4	贵州习水	灌木	41	N319	贵州安顺	灌木
18	HJ01	贵阳花溪久安乡	灌木	42	N323	贵阳花溪久安乡	灌木
19	HJ02	贵阳花溪久安乡	灌木	43	N629	贵阳花溪久安乡	灌木
20	HJ03	贵阳花溪久安乡	灌木	44	N636	贵州安顺	灌木
21	HJ04	贵阳花溪久安乡	灌木	45	N638	贵州安顺	灌木
22	HJ05	贵阳花溪久安乡	灌木	46	N645	贵州安顺	灌木
23	HJ06	贵阳花溪久安乡	灌木	47	J9	贵阳花溪久安乡	灌木
24	HJ07	贵阳花溪久安乡	灌木	48	SH8	贵阳花溪高坡乡	灌木

### 1.2 方法

#### 1.2.1 古茶树基因组 DNA 提取

古茶树叶片基因组的提取采用 CTAB 法<sup>[20]</sup>, 其具体方法为: 称取新鲜幼嫩叶片 0.2 g, 液氮研磨至粉末, 移入盛有 600  $\mu$ L 2% CTAB 提取液的离心管中, 混匀; 65  $^{\circ}$ C 水浴 1 h; 冷却至室温后, 加入 600  $\mu$ L 的氯仿: 异戊醇 (V: V=24: 1), 上下混匀 10 min; 4  $^{\circ}$ C 13 000 r/min 离心 10 min; 吸取上清液, 放入 600  $\mu$ L 异丙醇的离心管中, 轻轻混匀; 4  $^{\circ}$ C 13 000 r/min 离心 10 min; 弃上清, 加入 1 mL 70% 乙醇清洗 DNA 沉淀, 弃上清, 用吸水纸吸干管口水分, 气干; 加入 50  $\mu$ L 已灭菌的超纯水, 溶解及混匀 DNA. 所提取 DNA 经超微量分光光度计检测质量和浓度后, -20  $^{\circ}$ C 保存备用.

### 1.2.2 古茶树基因组 SNP 标记分析

利用 SLAF-seq(Specific-locus amplified fragment sequencing)技术对供试材料进行分子标记开发, SNP 标记的开发是以每个 SLAF 标签中深度最高的序列类型作为参考序列, 利用 bwa<sup>[21]</sup>将测序 reads 比对到参考序列上, 并使用 GATK<sup>[22]</sup>和 samtools<sup>[23]</sup>两种方法开发 SNP, 以两种方法得到的 SNP 标记交集作为最终可靠的 SNP 标记数集, 最终获得全基因组范围的具有代表性的高质量 SNP 标记。

### 1.2.3 数据分析

通过 MEGA5 软件<sup>[24]</sup>和 Neighbor-joining<sup>[25]</sup>算法进行古茶树的进化分析; 通过 Admixture 软件<sup>[26]</sup>分析古茶树样品的群体结构; 通过 Cluster<sup>[27]</sup>软件进行主成分分析。

## 2 结果与分析

### 2.1 古茶树 SLAF 标签与 SNP 标记开发

本试验利用生物信息学分析古茶树样品重测序数据, 共开发 1 656 258 个古茶树 SLAF 标签(特异长度的 DNA 片段), 古茶树样品平均测序深度为 24.21x, 其中多态性 SLAF 标签有 462 897 个, 共得到 2 690 638 个群体 SNP 标记, 根据完整度 > 0.8, MAF > 0.05 过滤, 共获得 283 376 个高一一致性的群体 SNP 标记(表 2)。

表 2 古茶树 SNP 信息统计表

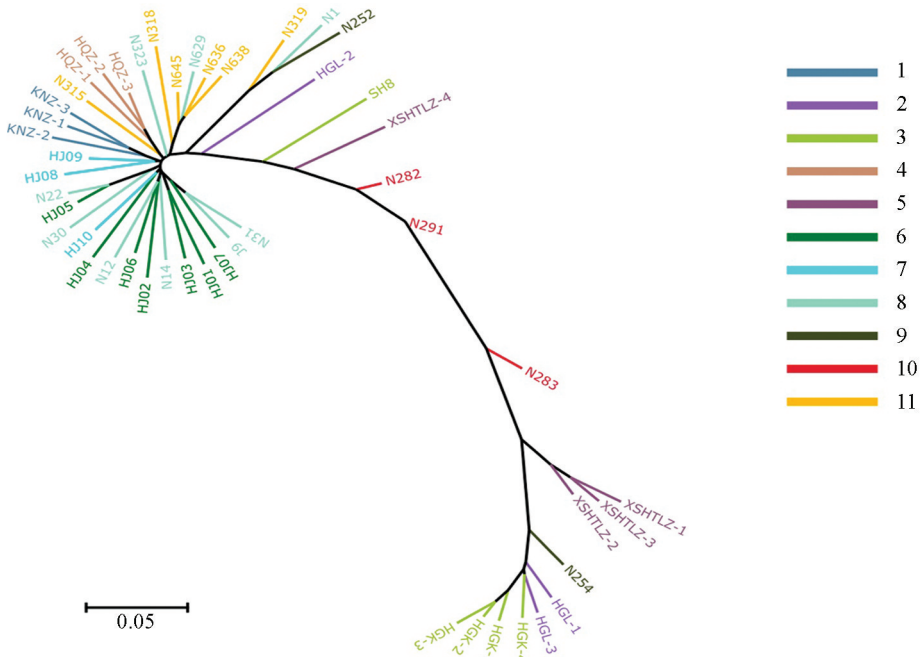
样品编号	SNP 数	完整度/%	杂合率/%
KNZ-1	267 182	94.28	17.20
KNZ-2	261 148	92.15	15.99
KNZ-3	265 078	93.54	16.39
HGL-1	235 165	82.98	11.39
HGL-2	264 280	93.26	18.26
HGL-3	239 824	84.63	11.53
HGK-1	237 268	83.72	10.83
HGK-2	238 075	84.01	11.37
HGK-3	232 226	81.94	9.66
HGK-4	234 489	82.74	10.63
HQZ-1	262 409	92.60	16.74
HQZ-2	262 712	92.70	16.77
HQZ-3	262 814	92.74	21.37
XSHTLZ-1	234 679	82.81	10.60
XSHTLZ-2	228 576	80.66	11.37
XSHTLZ-3	233 921	82.54	11.20
XSHTLZ-4	261 791	92.38	22.38
HJ01	262 087	92.48	17.55
HJ02	258 688	91.28	15.88
HJ03	269 149	94.97	18.33
HJ04	268 100	94.60	17.56
HJ05	272 163	96.04	18.76
HJ06	271 077	95.65	18.68
HJ07	268 399	94.71	18.30
HJ08	269 246	95.01	17.30
HJ09	272 653	96.21	21.52
HJ10	272 658	96.21	19.21

续表 2

样品编号	SNP 数	完整度/%	杂合率/%
N1	257 112	90.73	14.73
N12	268 056	94.59	17.20
N14	268 668	94.80	18.05
N22	270 118	95.32	18.55
N30	268 319	94.68	16.35
N31	270 924	95.60	19.21
N252	252 640	89.15	13.85
N254	235 434	83.08	10.30
N282	269 566	95.12	22.64
N283	253 691	89.52	16.28
N291	266 381	94.00	22.24
N315	267 690	94.46	16.69
N318	265 994	93.86	14.68
N319	262 401	92.59	16.21
N323	269 870	95.23	17.98
N629	268 818	94.86	17.43
N636	271 418	95.78	18.19
N638	269 035	94.93	17.90
N645	273 777	96.61	21.85
J9	268 917	94.89	18.71
SH8	270 537	95.46	25.21

## 2.2 古茶树的系统进化树分析

进化树用来表示物种之间的进化关系, 根据各类生物间亲缘关系的远近, 把各类生物安置在有分枝的树状图上, 简明地表示生物的亲缘关系和进化历程. 本试验基于 SNP 标记, 对 48 份贵州古茶树资源进行了进化分析, 计算得到古茶树的进化树(图 1).



图中每个分枝为一个样品. 0.05 表示在固定长度的样品间的遗传距离. 1 组材料来自开阳, 2~3 组材料来自花溪高坡乡, 4 组和 5 组材料分别来自花溪黔陶乡和习水, 6 组、7 组和 8 组材料来自花溪久安乡, 9 组、10 组和 11 组材料分别来自普定、务川和安顺; 下同.

图 1 贵州古茶树的进化树

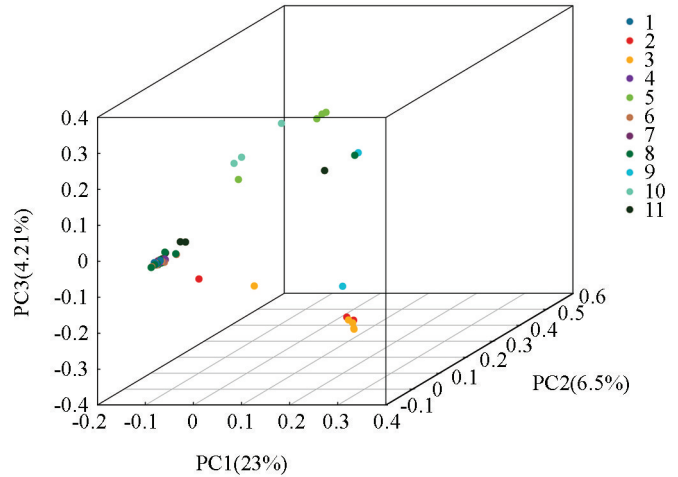
从进化树可以看出, 古茶树资源间存在丰富的遗传变异, 来自相近地方的古茶树在进化树上基本处于相近位置, 但是贵阳花溪古茶树存在着广泛的遗传差异. 位于进化树最下端的古茶树资源分别为 HGK-1, HGK-2, HGK-3, HGK-4, HGL-1 和 HGL-3(来自贵阳花溪高坡乡). 顺着进化树向上为来自贵州普定的 N254、来自贵州习水的 XSHTLZ-1, XSHTLZ-2 和 XSHTLZ-3、来自贵州务川的 N283, 以上古茶树资源的树型均为乔木. 顺着进化树继续向上的古茶树资源树型开始变为灌木, 最上端的古茶树资源分别为 N31, J9, HJ07, HJ01, HJ03, N14 和 HJ02 等, 均来自贵阳花溪, 由此可见, 花溪古茶树分布非常广泛, 从进化树下部到上部均有分布.

### 2.3 古茶树主成分分析

本试验基于 SNP 标记, 通过 Cluster 软件, 对供试古茶树资源进行了主成分分析(Principal component analysis, PCA), 得到 48 份古茶树的主成分聚类图(图 2). 从图 2 可以看出, 样品间存在广泛的差异, 来自同一地方的茶树基本上聚在一起, 说明它们之间的亲缘关系比较近, 但是花溪古茶树的遗传差异最大, 分布范围广泛, 该结果和前面进化树的分析结果是一致的.

### 2.4 古茶树群体结构分析

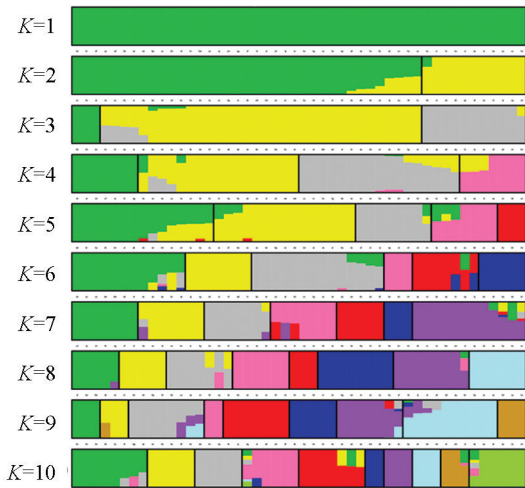
群体遗传结构分析能够提供个体的系统来源及其组成信息, 是一种重要的遗传关系分析工具. 本试验分别假设 48 份古茶树资源的分群数( $K$  值)为 1~10 进行聚类, 当  $K$  为 2 时,  $CV$  值(变异系数)最小, 根据  $CV$  值来确定分群数为 2(图 3), 可将 48 份古茶树资源分为 2 个类群——灌木型古茶树(类群 1)和乔木型古茶树(类群 2); 类群 2 中古茶树包括 11 份材料(HGK-1, HGK-2, HGK-3, HGK-4, HGL-1, HGL-3, XSHTLZ-1, XSHTLZ-2, XSHTLZ-3, N254 和 N283), 属于高大的乔木型古茶树, 其余 37 份古茶树材料皆属于类群 1, 它们是较低矮的灌木型古茶树, 这与进化树的分析结果一致.



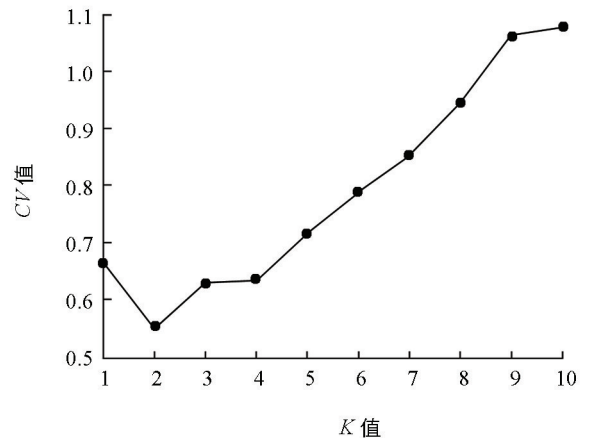
一个点代表一个样品, 一种颜色代表一个分组.

图 2 古茶树的 PCA 聚类图

群体结构



(a) 群体结构



(b)  $CV$  值

每种颜色代表一个群, 每行代表一个分群值的情况.

图 3 古茶树的群体结构图

### 3 讨 论

茶树世代周期长, 自交不亲和, 遗传组成高度异质杂合, 利用常规育种方法培育新品种费时耗力<sup>[28]</sup>. 近年来, 随着分子生物学的发展, 以 DNA 分子标记为标志的基因组研究进展迅速, 已广泛用于茶树种质资源和品种鉴别、遗传多样性、遗传演化及稳定性等研究<sup>[4]</sup>. 金惠淑等<sup>[29]</sup>利用 RAPD 技术研究了中国、韩国和日本的 46 个茶树样品的遗传多态性, 19 个随机引物扩增出 200 条带, 发现中国茶树品种的遗传多样性高于韩国茶树品种. Lai 等<sup>[30]</sup>利用 53 个 RAPD 和 56 个 ISSR 标记对台湾野生茶和栽培品种进行了研究, 结果表明, 台湾本土野生茶与阿萨姆类型亲缘关系更为接近, 在所研究的 3 个居群中, 台湾野生茶的遗传多样性最为丰富. 牛素贞等<sup>[15]</sup>利用 29 个 EST-SSR 标记对贵州 25 份地方茶树群体种和 3 份育成品种的遗传多样性、亲缘关系和遗传分化进行了分析, 发现贵州地方茶树品种资源间的遗传差异较大, 黔南地区茶树资源遗传多样性程度高于黔北地区, 两地区遗传分化较高, 基因交流频率较低. 本试验开发了 1 656 258 个古茶树 SLAF 标签, 其中多态性 SLAF 标签有 462 897 个, 共得到 283 376 个高一一致性的群体 SNP 标记. 从基于 SNP 标记获得的进化树可以看出, 贵州茶树样品间存在丰富的遗传变异, 进化树最下端是来自花溪高坡的乔木型古茶树(HGK-1~HGK-4 等), 顺着进化树向上依次是来自普定、习水、务川、花溪久安、安顺、开阳和花溪的古茶树, 由此可得出花溪古茶树分布非常广泛, 从进化树下部到上部均有分布. 通过 PCA 分析发现, 来自同一地方的古茶树在 PCA 聚类图上基本处于相近位置, 但是花溪古茶树存在着广泛的遗传变异, 获得了与进化树一致的结果.

Matsumoto 等<sup>[8]</sup>利用 PAL(丙氨酸解氨酶)cDNA 作为 RFLP 探针, 对 29 个日本绿茶栽培品种和地方品种的遗传变异进行了研究, 将日本绿茶品种分为 5 类. Kaundun 等<sup>[6]</sup>利用 SSR 技术将 24 份茶树资源分为 Assamica 和 Sinensis 两个类群. Mondal<sup>[11]</sup>利用 SSR 引物对 25 份茶树资源进行分析, 将所有资源分为禅叶型、阿萨姆型和中国型 3 个簇. Wachira 等<sup>[31]</sup>采用 RAPD 和 AFLP 两种标记, 将来源于印度、斯里兰卡、中国、日本、越南和肯尼亚等 48 个品种划分为阿萨姆变种、中国变种和栽培型与野生型茶树品种 3 个类型. Balasatavanan 等<sup>[32]</sup>应用 AFLP 标记将 49 个茶树品种资源划分为阿萨姆、中国和阿富汗过渡型 3 个类型. 本试验通过古茶树的群体结构分析, 清晰地把古茶树资源分为 2 个类群——乔木型古茶树与灌木型古茶树, 其中 11 份材料属于高大的乔木型古茶树, 其余 37 份材料为较低矮的灌木型古茶树. 陈亮<sup>[5]</sup>认为茶组的系统演化途径由乔木向灌木演化, 本试验通过基于 283 376 个 SNP 标记的古茶树进化树分析, 发现乔木型古茶树分布于进化树下部, 而灌木型古茶树位于进化树上部, 证实了茶树是由乔木型向灌木型进化的.

#### 参考文献:

- [1] 闵天禄. 世界山茶属的研究 [M]. 昆明: 云南科技出版社, 2000.
- [2] 罗美玲, 田洪敏, 杨雪梅, 等. 电子鼻技术对普洱熟茶香气判别的研究 [J]. 西南大学学报(自然科学版), 2018, 40(8): 16-24.
- [3] 李俊, 蔡滔, 周雪丽, 等. 贵州绿茶品质分析研究 [J]. 中国茶叶, 2017, 39(7): 22-26.
- [4] GUNASEKARE M T K. Applications of Molecular Markers to the Genetic Improvement of *Camellia Sinensis* L. (Tea) [J]. Journal of Horticultural Science and Biotechnology, 2007, 82(2): 161-169.
- [5] 陈亮. 茶组植物的分子系统学研究——遗传多样性与分类 [D]. 杭州: 浙江大学, 2002: 63-64.
- [6] KAUNDUN S S, MATSUMOTO S. Heterologous Nuclear and Chloroplast Microsatellite Amplification and Variation in Tea, *Camellia Sinensis* [J]. Genome, 2002, 45(6): 1041-1048.
- [7] KAUNDUN S, PARK Y G. Genetic Structure of Six Korean Tea Populations as Revealed by RAPD-PCR Markers [J]. Crop Science, 2002, 42(2): 594-601.
- [8] MATSUMOTO S, KIRIWA Y, TAKEDA Y. Differentiation of Japanese Green Tea Cultivars as Revealed by RFLP A-

- analysis of Phenylalanine Ammonia-lyase DNA [J]. *Theoretical and Applied Genetics*, 2002, 104(6): 998-1002.
- [9] PAUL S, WACHIRA F N, POWELL W, et al. Diversity and Genetic Differentiation among Populations of Indian and Kenyan Tea (*Camellia Sinensis* (L.) O. Kuntze) Revealed by AFLP Markers [J]. *Theoretical and Applied Genetics*, 1997, 94(2): 255-263.
- [10] MISHRA R K, CHAUDHARY S, AHMAD A, et al. Molecular Analysis of Tea Clones (*Camellia Sinensis*) using AFLP Markers [J]. *International Journal of Integrative Biology*, 2009, 5(2): 130-135.
- [11] MONDAL T K. Assessment of Genetic Diversity of Tea (*Camellia sinensis* (L.) O. Kuntze) by Inter-simple Sequence Repeat Polymerase Chain Reaction [J]. *Euphytica*, 2002, 128(3): 307-315.
- [12] FANG W, CHENG H, DUAN Y, et al. Genetic Diversity and Relationship of Clonal Tea (*Camellia sinensis*) Cultivars in China as Revealed by SSR Markers [J]. *Plant Systematics and Evolution*, 2012, 298(2): 469-483.
- [13] 王亨洪, 索化夷, 杨 坚, 等. 川渝地区重要野生大茶树遗传多样性的 ISSR 分析 [J]. *茶叶科学*, 2009, 29(2): 168-172.
- [14] 侯渝嘉, 何 桥, 梁国鲁, 等. 茶树杂交后代的 ISSR 分析 [J]. *西南农业大学学报(自然科学版)*, 2006, 28(2): 267-270.
- [15] 牛素贞, 刘玉倩, 杨锦标, 等. 贵州茶树地方品种的 EST-SSR 遗传多样性分析 [J]. *浙江农业学报*, 2012, 24(5): 836-841.
- [16] 李志远, 于海龙, 方智远, 等. 甘蓝 SNP 标记开发及主要品种的 DNA 指纹图谱构建 [J]. *中国农业科学*, 2018, 51(14): 2771-2787.
- [17] 马 宇, 于海峰, 侯建华, 等. 基于 SLAF-seq 技术的向日葵 SNP 标记开发与利用 [J]. *分子植物育种*, 2018, 16(18): 5994-6000.
- [18] 李 敏, 郭 聪, 王 莹, 等. 基于 SLAF-seq 技术的乔木柳 SNP 位点开发 [J]. *西南农业学报*, 2018, 31(5): 891-895.
- [19] 刘丽华, 庞斌双, 刘阳娜, 等. 基于 SNP 标记的小麦高通量身份鉴定模式 [J]. *麦类作物学报*, 2018, 38(5): 529-534.
- [20] 盛中飞. 提莫菲维与野燕麦杂交后代外源遗传物质的检测 [D]. 贵阳: 贵州大学, 2010: 16.
- [21] LI H, DURBIN R. Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform [J]. *Bioinformatics*, 2009, 25(14): 1754-1760.
- [22] MCKENNA A, HANNA M, BANKS E, et al. The Genome Analysis Toolkit: a MapReduce Framework for Analyzing Next-generation DNA Sequencing Data [J]. *Genome Research*, 2010, 20(9): 1297-1303.
- [23] LI H, HANDSAKER B, WYSOKER A, et al. The Sequence Alignment/map Format and SAMtools [J]. *Bioinformatics*, 2009, 25(16): 2078-2079.
- [24] TAMURA K, PETERSON D, PETERSON N, et al. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods [J]. *Molecular Biology and Evolution*, 2011, 28(10): 2731-2739.
- [25] SAITOU N, NEI M. The Neighbor-joining Method: a New Method for Reconstructing Phylogenetic Trees [J]. *Molecular Biology and Evolution*, 1987, 4(4): 406-425.
- [26] ALEXANDER D H, NOVEMBRE J, LANGE K. Fast Model-based Estimation of Ancestry in Unrelated Individuals [J]. *Genome Research*, 2009, 19(9): 1655-1664.
- [27] DE HOON M J L, IMOTO S, NOLAN J, et al. Open Source Clustering Software [J]. *Bioinformatics*, 2004, 20(9): 1453-1454.
- [28] 马建强, 姚明哲, 陈 亮. 茶树遗传图谱研究进展 [J]. *茶叶科学*, 2010, 30(5): 329-335.
- [29] 金惠淑, 梁月荣, 陆建良. 中、韩两国主要茶树品种基因组 DNA 多态性比较研究 [J]. *茶叶科学*, 2001, 21(2): 103-107.
- [30] LAI J A, YANG W C, HSIAO J Y. An Assessment of Genetic Relationships in Cultivated Tea Clones and Native Wild Tea in Taiwan using RAPD and ISSR Markers [J]. *Botanical Bulletin of Academia Sinica*, 2001, 42(2): 93-100.

- [31] WACHIRA F, TANAKA J, TAKEDA Y. Genetic Variation and Differentiation in Tea (*Camellia sinensis*) Germplasm Revealed by RAPD and AFLP Variation [J]. The Journal of Horticultural Science and Biotechnology, 2001, 76(5): 557-563.
- [32] BALASARAVANAN T, PIUS P K, RAJ KUMAR R, et al. Genetic Diversity among South Indian Tea Germplasm (*Camellia Sinensis*, *C. Assamica* and *C. Assamica* Spp. *Lasiocalyx*) using AFLP Markers [J]. Plant Science, 2003, 165(2): 365-372.

## SNP Analysis of the Genetic Evolution of Ancient *Camellia sinensis* Trees from Huaxi, Guiyang

CHEN Li-jie<sup>1</sup>, ZHANG Su-qin<sup>2</sup>, YIN Jie<sup>3</sup>,  
SONG Qin-fei<sup>3</sup>, NIU Su-zhen<sup>3</sup>, ZHAO Ji-ying<sup>4</sup>,  
CHEN Dan-ping<sup>4</sup>, WANG Sheng-wei<sup>4</sup>, GENG Guang-dong<sup>2</sup>

1. Asset Management Office, Guizhou University, Guiyang 550025, China;

2. College of Agriculture, Guizhou University, Guiyang 550025, China;

3. College of Tea Science, Guizhou University, Guiyang 550025, China;

4. Huaxi District Agricultural Bureau, Guiyang 550025, China

**Abstract:** In the present study, the SNP (single nucleotide polymorphism) technique was used to analyze the genetic diversity, population structure and genetic evolution of ancient *Camellia sinensis* trees from Huaxi, Guiyang. In bioinformatics analysis of the resequencing data, with an average sequencing depth of 24.21x, a total of 1, 656, 258 SLAF tags were obtained, of which 462, 897 were polymorphic SLAF tags and 283, 376 were highly consistent population SNP markers. The phylogenetic tree analysis based on these SNP markers indicated that there existed abundant genetic variation among Huaxi tea accessions, and the accessions from the same place were generally clustered together, which was further verified by principal component analysis (PCA). Population structure analysis divided these resources into two distinct groups: an arbor-type group and a shrub-type group. The arbor group and the shrub group were located at the bottom and the top of the phylogenetic tree, respectively, indicating that the teas evolved from the arbor type to the shrub type.

**Key words:** ancient *Camellia sinensis*; specific-locus amplified fragment sequencing (SLAF-seq); SNP (single nucleotide polymorphism); phylogenetic tree; PCA; genetic structure

责任编辑 汤振金



