

六堡茶群体种叶绿体基因组捕获历史与遗传多样性研究

姚淑婷^{1,2}, 张强¹, 苏敏³, 吴雨婷³, 庞月兰³, 梁燕妮⁴, 蔡爱华¹, 覃信梅^{1*}

(1. 广西喀斯特植物保育与恢复生态学重点实验室, 广西壮族自治区中国科学院广西植物研究所, 广西 桂林 541006; 2. 广西师范大学 生命科学学院, 广西 桂林 541004; 3. 广西壮族自治区茶叶科学研究所, 广西 桂林 541004; 4. 梧州学院, 广西 梧州 543002)

摘要: 六堡茶群体种 (*Camellia sinensis* var. *sinensis* cv. 'Liupao') 是制作六堡茶的主要原料植物。为探索六堡茶群体种的遗传背景, 特别是其系统发育和演化历史, 该文对 27 份六堡茶群体种和 5 份突肋茶开展了浅层基因组和转录组测序, 并进行了六堡茶群体种的系统发育关系、分歧进化时间及遗传多样性分析。结果表明: (1) 在叶绿体树上, 六堡茶群体种样品被分为距离较远的两组, 一组与茶亲缘关系最近, 嵌入了主要由茶组成的分支内部, 分别与不同栽培茶聚在一起, 另一组则形成单系分支, 与突肋茶亲缘关系最近且嵌入其内部; 然而, 在核基因树上, 六堡茶群体种样品与栽培茶及山茶属其他几个物种聚为一支, 但与突肋茶关系较远; 核质基因组系统发育矛盾表明, 部分六堡茶群体种的祖先曾与突肋茶发生过杂交事件, 捕获了突肋茶的叶绿体基因组。(2) 分子钟分析进一步表明, 该杂交事件发生在 1.55 百万年前的第四纪时期, 远早于人类栽培和制作茶的历史。(3) 基于叶绿体基因组和核基因的遗传多样性分析均发现, 六堡茶群体种具有较高的单倍型多样性, 有较大的演化潜力。该研究将为六堡茶群体种种质资源保护、品种选育及开发利用提供科学依据和指导。

关键词: 六堡茶群体种, 叶绿体基因组, 转录组, 叶绿体基因组捕获, 杂交, 遗传多样性
中图分类号: Q943 **文献标识码:** A

Chloroplast genome capture history and genetic diversity of *Camellia sinensis* var. *sinensis* cv. 'Liupao'

YAO Shuting^{1,2}, ZHANG Qiang¹, SU Min³, WU Yuting³, PANG Yuelan³,
LIANG Yanni⁴, CAI Aihua¹, QIN Xinmei^{1*}

(1. *Guangxi Key Laboratory of Plant Conservation and Restoration Ecology in Karst Terrain, Guangxi Institute of Botany, Guangxi Zhuang Autonomous Region and Chinese Academy of Sciences, Guilin 541006, Guangxi, China;* 2. *College of Life Sciences, Guangxi Normal University, Guilin 541004, Guangxi, China;* 3. *Tea Science and Research Institute, Guangxi Zhuang Autonomous Region, Guilin 541004, Guangxi, China;* 4. *Wuzhou University, Wuzhou 543002, Guangxi, China*)

Abstract: *Camellia sinensis* var. *sinensis* cv. 'Liupao' is the main raw material for making Liupao tea. To explore the genetic background of *C. sinensis* var. *sinensis* cv. 'Liupao', especially its phylogenetic position and evolutionary history, 27 and five individuals of *C. sinensis* var. *sinensis* cv. 'Liupao' and *C. costata* were sampled for genome skimming and transcriptomic sequencing. The

基金项目: 广西创新驱动发展专项资金项目 (桂科 AA20302018-1); 广西自然科学基金项目 (2023GXNSFAA026221); 广西喀斯特植物保育与恢复生态学重点实验室项目 (22-035-26)。

第一作者: 姚淑婷 (1998-), 硕士研究生, 主要从事植物系统发育研究, (E-mail) 1150487855@qq.com。

***通信作者:** 覃信梅, 硕士, 助理研究员, 主要从事植物系统演化和生物信息学研究, (E-mail) xinmeiqin2018@163.com。

phylogenetic relationship between *C. sinensis* var. *sinensis* cv. ‘Liupao’ and other *Camellia* was reconstructed, the divergence time of the chloroplast genome sequences of *C. sinensis* var. *sinensis* cv. ‘Liupao’ was estimated, and the genetic diversity was investigated. The results were as follows: (1) On the chloroplast tree, 27 samples of *C. sinensis* var. *sinensis* cv. ‘Liupao’ were divided into two distantly-related groups, one was nested within a clade mainly consisting of *C. sinensis*, being interspersed among other cultivated *C. sinensis*, while the other group formed as a well supported lineage that was most closely-related to and nested within *C. costata*; On the nuclear gene tree, however, all the 27 samples formed a clade with other *C. sinensis* as well as some other *Camellia* species with generally unresolved relationships among them. Despite the lack of resolution in this clade, it was definitely far separated from *C. costata*; The cytonuclear phylogenetic conflict suggested once ancient introgression hybridization of *C. costata* with the ancestor of some *C. sinensis* var. *sinensis* cv. ‘Liupao’ individuals so that the latter captured the chloroplast genome of the former. (2) The time estimate indicated that the introgression hybridization event occurred in the Quaternary period, ca. 1.55 million years ago, long before the history of tea cultivation and production by humans. (3) In addition, both the chloroplast genomes and nuclear genes revealed that the *C. sinensis* var. *sinensis* cv. ‘Liupao’ had high haplotype diversity, possessing high evolutionary potential. This study could provide important references for the germplasm protection, breeding, and utilization of high-quality *C. sinensis* var. *sinensis* cv. ‘Liupao’ strains.

Key words: *Camellia sinensis* var. *sinensis* cv. ‘Liupao’, chloroplast genomes, transcriptomes, chloroplast genome capture, hybridization, genetic diversity

广义的茶树指山茶科 (Theaceae) 山茶属 (*Camellia*) 茶组 (Sect. *Thea*) 植物, 而狭义的茶树则指茶 (*Camellia sinensis*) 这一种 (含变种、变型等种下等级) 植物 (杨世雄, 2021a)。茶包含了众多品种或地方种, 具有不同的风味品质。茶的不同品种以及山茶属其他种类被认为存在频繁的杂交和多倍化, 具有复杂的演化历史 (Huang et al., 2013; Wang et al., 2020; Zan et al., 2023)。例如, 山茶 (*C. japonica*) 和短柄山茶 (*C. rusticana*) 之间能发生自然杂交 (Ueno & Tsumura, 2009)。Meegahakumbura 等 (2016) 对 392 份茶样品的遗传分析发现 30% 的样品具有不同类型茶树的遗传混合, 表明它们为杂交起源, 其中, 柬埔寨茶 (Cambod tea) 就被认为是阿萨姆茶 (Assam tea) 和中国茶 (China tea) 的杂交品种。李苗苗等利用核基因微卫星标记对普洱茶 (*C. sinensis* var. *assamica*) 和大理茶 (*C. taliensis*) 进行 STRUCTURE 聚类分析, 发现在混栽的大理茶和普洱茶居群间, 存在由大理茶向普洱茶的基因渐渗 (李苗苗等, 2015)。Li 等研究发现武夷菜茶水仙茶树品种 (三倍体) 和山茶属其他 4 个物种在 *trnE-trnT* 间隔区具有同样 335 bp 的序列缺失, 但 292 份栽培茶都没有这段缺失, 由此推测该茶树在三倍化过程中发生了叶绿体转移, 即通过渐渗杂交捕获了其他物种的叶绿体 (Li et al., 2021)。虽然叶绿体基因已广泛应用于山茶属植物系统发育和遗传多样性研究 (Ye et al., 2014; Li et al., 2021; 闫明慧等, 2021), 但是, 叶绿体基因组通常为母系遗传, 单独使用无法准确解析杂交或基因渐渗等网状演化, 而利用核低拷贝直系同源基因序列则已成功解决了许多疑难的系统发育尤其复杂的网状演化问题 (Yang et al., 2015; Wu et al., 2022; Zan et al., 2023)。

广西茶树种质资源极为丰富, 根据闵天禄的茶组分类系统, 茶组共包含 12 种 6 变种, 而广西有 8 种, 是中国茶组植物资源最丰富的地区 (Ming & Bartholomew, 2007; 杨世雄, 2021b)。广西也拥有很多历史悠久的地方品种及名茶 (陈宗懋, 1992), 其中, 广西梧州的六堡茶, 为中国二十四大名茶之一, 以其“红、浓、陈、醇”及独特的槟榔香特点闻名于世 (陈佳等, 2020)。六堡茶群体种作为制作六堡茶的主要原材料, 它属于茶组 (Sect. *Thea*) 茶 (*C. sinensis*)

下众多地方种系中的一员。根据《中国茶树品种志》中的相关记载，六堡茶群体种为广西地方有性系茶树品种，为中叶种，主要分布在广西苍梧、贺县、蒙山和昭平等地（中国茶树品种志编写委员会，2001）。对六堡茶群体种开展的研究主要为形态特征、资源调查和开发利用（陈佳等，2010；邱瑞瑾等，2020；滕翠琴等，2020）以及基于 EST-SSR 和 SSR 分子标记进行的亲缘关系与遗传多样性分析（周炎花等，2011；黄厚宸等，2021；王留彬等，2022），而基于序列分析的遗传背景及遗传多样性研究还未见报道，其起源和演化历史仍不清楚。值得注意的是，茶组植物中突肋茶（*C. costata*）的模式产地为广西昭平县，与六堡茶群体种分布区重叠。目前对突肋茶的研究只限于对广西昭平县突肋茶的资源调查和开发利用（李朝昌等，2019；龚明寿等，2021），以及贵州地方茶组植物（包含突肋茶）的亲缘关系研究（郭灿等，2021）。六堡茶群体种与突肋茶地理分布重叠，它们之间及它们与同组或同属其他种类间是否存在杂交或基因渐渗尚无研究。

本研究以 27 份六堡茶群体种和 5 份突肋茶为实验材料，通过浅层基因组和转录组测序，拼接组装叶绿体基因组并筛选单拷贝核基因，结合公共数据库中已公布的山茶科相应的核苣酸序列，重建六堡茶群体种与其他山茶属物种间的系统发育关系，估算六堡茶群体种样品的叶绿体基因组分歧进化时间，并计算遗传多样性指数，以期解决或探讨如下问题：（1）六堡茶群体种的系统发育位置；（2）六堡茶群体种与其他山茶属物种是否存在杂交或基因渐渗；（3）六堡茶群体种的遗传多样性。六堡茶群体种种质资源对促进当地经济发展及乡村振兴具有重要意义，本研究可为六堡茶群体种资源保护、品种选育及开发利用提供科学依据和指导。

1 材料与方法

1.1 实验材料

本研究使用的 27 份六堡茶群体种均来源于梧州市，其中 12 份采集于梧州市苍梧县六堡镇石牛村，编号为 H1~H12，10 份采集于梧州市苍梧县六堡镇四柳村，编号为 Q13~Q22；5 份采集于梧州市藤县大益村马欧岭，编号为 D1~D5。27 份六堡茶群体种叶片多样性较高，小、中、大叶类均有，绿芽、紫芽、黄白芽均有。另外还从广西贺州昭平县农业科学研究所资源圃采集了 5 份突肋茶，编号为 T1~T5。选择无病虫害、生长状况良好的新鲜幼嫩叶片用于后续实验。

1.2 DNA、RNA 提取和测序

基因组浅层测序和转录组测序主要委托青岛百迈克生物科技有限公司进行。32 份样品的总 DNA 和 RNA 分别采用 CTAB 法和天根 DP411 试剂盒进行提取，并检测其浓度和完整性。使用检验合格的 DNA 和 RNA 构建测序文库，并用 Illumina NovaSeq 6000 测序仪进行双端测序（PE150 bp）。最终获得浅层基因组原始数据每个样约 3 Gb，转录组原始数据每个样约 6 Gb。

1.3 数据处理

1.3.1 叶绿体基因组的组装和注释

对获得的浅层基因组原始数据用 fastp 0.20.1 软件（Chen et al., 2018）进行质量控制，去掉接头和低质量的片段（reads）后得到净化数据（clean data），使用 GetOrganelle 1.7.5 软件（Jin et al., 2020）在默认参数设置下进行拼接组装。以茶（*C. sinensis* var. *Sinensis*）为参考序列（GenBank 登录号为 NC_020019），利用 PGA 软件（Qu et al., 2019）进行叶绿体基因组注释，并使用 OGDRAW（<https://chlorobox.mpimp-golm.mpg.de/OGDraw>）绘制叶绿体基因组图谱。

1.3.2 转录组拼接和筛选单拷贝核基因

除了本研究新获得的转录组测序数据，我们还从 NCBI 数据库下载了山茶科 26 个物种（38 份样品）的转录组测序数据（表 1），包括 1 个大头茶属（*Polyspora*）物种，25 个山茶属物

种。用上述 1.3.1 中的质控方式获得净化数据，使用 Trinity 2.11.0 软件 (Haas et al., 2013) 在默认参数设置下把净化数据从头组装成转录本。使用 TransDecoder 5.5.0 软件 (Haas et al., 2013) 识别转录本序列中的候选编码区。使用 CD-HIT 4.8.1 软件 (Fu et al., 2012) 根据序列的相似度对序列进行聚类以去除冗余的序列，序列相似性阈值设置为 0.98。因为一些样品测序质量不高会影响筛选直系同源基因数量，所以选择 11 个测序质量较高且具有系统发育代表性的样品，使用 OrthoFinder 2.5.2 软件 (Emms & Kelly, 2019) 搜索出 11 个样品之间的 798 个单拷贝核基因 (一对一核直系同源基因)。在其余样品中，使用 blast++ 软件 (Wang et al., 2003) 搜索并提取 798 个单拷贝核基因对应的得分最高的同源序列。

表 1 38 份样品转录组数据的 GenBank 登录号

Table 1 GenBank accession numbers of the transcriptome data of the 38 samples

物种名 Species name	SSR 号 SSR number	物种名 Species name	SSR 号 SSR number
狭叶茶 <i>Camellia angustifolia</i>	SRR19266728	汝城毛叶茶 <i>C. pubescens</i>	SRR19266726
大树茶 <i>C. arborescens</i>	SRR19266723	茶梅 <i>C. sasanqua</i>	SRR19266679
老黑茶 <i>C. atrothea</i>	SRR19266735	安吉白茶 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'AJBC'	SRR1849929
厚轴茶 <i>C. crassicolumna</i>	SRR19266722	英红 9 号 <i>C. sinensis</i> var. <i>assamica</i> cv. 'YH9'	SRR19266668
丹寨秃茶 <i>C. danzaiensis</i>	SRR19266665	云抗 10 号 <i>C. sinensis</i> var. <i>assamica</i> cv. 'YK10'	SRR19266725
糙果茶 <i>C. furfuracea</i>	SRR19266754	紫鹃 <i>C. sinensis</i> var. <i>assamica</i> cv. 'ZJ'	SRR19266669
弄岗金花茶 <i>C. grandis</i>	SRR19266743	白鸡冠 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'BJG'	SRR2781847
秃房茶 <i>C. gymnogyna</i>	SRR19266731	黄金芽 1 号 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'HJ1'	SRR5075641
河口超长柄茶 <i>C. hekouensis</i>	SRR19266667	金萱 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'JX'	SRR2191769
缙云山茶 <i>C. jingyunshanica</i>	SRR19266730	宁州 2 号 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'NZ2'	SRR2730141
金沙江红山茶 <i>C. jinshajiangica</i>	SRR19266700	舒茶早 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'SCZ'	SRR1928149
广西茶 <i>C. kwangsiensis</i>	SRR19266738	福鼎大白 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'FDDB'	SRR19266672
广东秃茶 <i>C. kwangtungensis</i>	SRR19266666	龙井 43 号 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'LJ'	SRR19266721
膜叶茶 <i>C. leptophylla</i>	SRR19266732	铁观音 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'TGY'	SRR19266671
马关茶 <i>C. makuanica</i>	SRR19266734	大厂茶 <i>C. tachangensis</i>	SRR19266737
小长尾连蕊茶	SRR19266761	大理茶	SRR19266733

<i>C. parvicaudata</i>		<i>C. taliensis</i>	
细萼茶	SRR19266724	四球茶	SRR19266736
<i>C. parvisepala</i>		<i>C. tetracocca</i>	
西南红山茶	SRR19266689	瘤果茶	SRR19266711
<i>C. pitardii</i>		<i>C. tuberculata</i>	
毛叶茶	SRR19266727	四川大头茶	SRR19266757
<i>C. ptilophylla</i>		<i>Polyspora speciosa</i>	

1.3.3 序列比对与构建系统树

选取一条组装好的六堡茶群体种叶绿体基因组序列作为参考序列在 NCBI 上进行相似性搜索，下载山茶科 77 种（含变种）植物的 103 条叶绿体基因组序列（表 2），包含山茶属 76 种（含变种），圆籽荷属（*Apterosperma*）1 种。除茶这一物种共下载 29 条序列外，其余每种下载一条序列。使用 MAFFT 7.490 软件（Katoh & Standley, 2013）对所有序列进行比对，再用 BioEdit 7.2.5 软件（Hall, 1999）手工调整。使用 MEGA 7.0 软件（Kumar et al., 2016）计算叶绿体基因组矩阵的长度、变异位点以及碱基组成频率。以圆籽荷（*A. oblata*）为外类群（闫明慧等, 2021），基于叶绿体基因组序列，用 RaxML 7.2.6 软件（Stamatakis, 2006）中的最大似然法（maximum likelihood, ML）构建系统发育树，碱基替换模型设定为 GTRGAMMA，用随机起始树，矩阵重复抽样 100 次计算支持率（bootstrap support）。

用 muscle 3.8.1 软件（Edgar, 2004）对 798 个单拷贝核基因序列分别进行比对，然后用自创 R 包 alignmentFilter（<https://github.com/qiangzhang04/alignmentFilter>）过滤比对不可靠区段，删除缺失（gap）超过 90% 的位点，并去掉矩阵中序列长度低于 100 bp 的序列及矩阵长度低于 500 bp 的基因，最终获得 689 个基因用于后续分析。通过 PhyloSuite 1.2.2 软件（Zhang et al., 2020）将 689 个核基因进行串联，计算串联矩阵的长度、变异位点以及碱基组成频率，并以山茶科四川大头茶（*Polyspora speciosa*）为外类群（Wu et al., 2022）构建系统发育树，方法与参数设置与上述相同。

表 2 103 份样品叶绿体基因组数据的 GenBank 登录号

Table 2 GenBank accession numbers of the chloroplast genome of the 103 samples

物种名 Species name	登录号 Accession number	物种名 Species name	登录号 Accession number	物种名 Species name	登录号 Accession number
越南抱茎茶 <i>Camellia amplexicaulis</i>	MT317095	龙州金花茶 <i>C. lungzhouensis</i>	MN579509	<i>C. sinensis</i> cv. 'Chamnok'	LC625850
抱茎短蕊茶 <i>C. amplexifolia</i>	OL435568	云南管蕊茶 <i>C. luteocalpandria</i>	OK546696	大红袍 <i>C. sinensis</i> cv. 'Dahongpao'	MT773374
安龙瘤果茶 <i>C. anlungensis</i>	MN756594	小黄花茶 <i>C. luteoflora</i>	KY626042	龙井 43 <i>C. sinensis</i> cv. 'Longjing43'	KF562708
老黑茶 <i>C. atrothea</i>	OK382089	毛蕊红山茶 <i>C. mairei</i>	KY406767	肉桂 <i>C. sinensis</i> cv. 'Rougui'	MT773375
杜鹃叶山茶 <i>C. azalea</i>	KY856741	小果油茶 <i>C. meiocarpa</i>	MT956593	<i>C. sinensis</i> cv. 'Sangmok'	LC488797
短柱茶 <i>C. brevistyla</i>	NC_052752	小花金花茶 <i>C. micrantha</i>	NC_058646	水金龟 <i>C. sinensis</i> cv. 'Shuijingui'	MT773376
浙江红山茶 <i>C. chekiangoleosa</i>	MG431968	富宁金花茶 <i>C. mingii</i>	NC_046699	铁罗汉 <i>C. sinensis</i> cv. 'Tieluohan'	MT773377

薄叶金花茶 <i>C. chrysanthoides</i>	MZ618349	金花茶 <i>C. nitidissima</i> var. <i>nitidissima</i>	MT157617	武夷岩茶水仙 <i>C. sinensis</i> cv. 'Wuyi Narcissus'	MT612435
崇左金花茶 <i>C. chuongtsoensis</i>	MT663341	八瓣糙果茶 <i>C. octopetala</i>	MN095258	柚花香 <i>C. sinensis</i> cv. 'Youhuaxiang'	MZ379786
小果短柱茶 <i>C. confusa</i>	NC_058854	油茶 <i>C. oleifera</i>	KY406750	普洱茶 <i>C. sinensis</i> var. <i>assamica</i>	JQ975030
突肋茶 <i>C. costata</i>	NC_061904	香花油茶 <i>C. osmantha</i>	MZ128138	普洱茶 <i>C. sinensis</i> var. <i>assamica</i>	MH394408
红皮糙果茶 <i>C. crapnelliana</i>	KF753632	细萼茶 <i>C. parvisepala</i>	NC_058924	普洱茶 <i>C. sinensis</i> var. <i>assamica</i>	MH394407
尖连蕊茶 <i>C. cuspidata</i>	KF156833	小果金花茶 <i>C. petelotii</i> var. <i>microcarpa</i>	MT157619	城步峒茶 <i>C. sinensis</i> var. <i>assamica</i> cv. 'Duntsa'	OL450397
丹寨秃茶 <i>C. danzaiensis</i>	KF156834	顶生金花茶 <i>C. pingguoensis</i> var. <i>terminalis</i>	OK149109	云抗 10 号 <i>C. sinensis</i> var. <i>assamica</i> cv. 'Yunkang10'	MH019307
德保金花茶 <i>C. debaoensis</i>	NC_057958	西南红山茶 <i>C. pitardii</i>	KF156837	德宏茶 <i>C. sinensis</i> var. <i>dehungensis</i>	KJ806279
长管连蕊茶 <i>C. elongata</i>	KY406791	多齿红山茶 <i>C. polyodonta</i>	NC_060777	白毛茶 <i>C. sinensis</i> var. <i>pubilimba</i>	NC_061691
显脉金花茶 <i>C. euphlebica</i>	OL405564	毛叶茶 <i>C. ptilophylla</i>	MT721159	白毛茶 <i>C. sinensis</i> var. <i>pubilimba</i>	OM262114
簇蕊金花茶 <i>C. fascicularis</i>	NC_053896	毛肋茶 <i>C. pubicosta</i>	KJ806277	白毛茶 <i>C. sinensis</i> var. <i>pubilimba</i>	KJ806280
毛柄连蕊茶 <i>C. fraterna</i>	MT663342	毛瓣金花茶 <i>C. pubipetala</i>	NC_054365	茶 <i>C. sinensis</i> var. <i>sinensis</i>	KJ806281
高州油茶 <i>C. gauchowensis</i>	NC_053541	红花三江瘤果茶 <i>C. pyxidiacea</i> var. <i>rubituberculata</i>	MZ424202	黔茶 1 号 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'Qiancha1'	MZ043860
博白大果油茶 <i>C. gigantocarpa</i>	NC_058879	毛药山茶 <i>C. renshanxiangiae</i>	MH253889	铁观音 <i>C. sinensis</i> var. <i>sinensis</i> cv. 'Tieguanyin'	MW148820
大苞茶 <i>C. grandibracteata</i>	KJ806274	滇山茶 <i>C. reticulata</i>	KY406793	香花糙果茶 <i>C. suaveolens</i>	MT123282
大苞山茶 <i>C. granthamiana</i>	MG782842	皱叶瘤果茶 <i>C. rhytidophylla</i>	MT663343	半宿萼茶 <i>C. szechuanensis</i>	KY406778
珍珠茶（长瓣短柱茶） <i>C. grijsii</i> cv. 'Zhenzhucha'	MT916932	喙果金花茶 <i>C. rostrata</i>	NC_058254	兴义 6 号（大厂茶） <i>C. tachangensis</i> cv. 'Xingyi6'	NC_056149
秃房茶	MH394403	茶梅	MH782189	大理茶	JQ975032

<i>C. gymnogyna</i>		<i>C. sasanqua</i>		<i>C. taliensis</i>	
贵州金花茶		南山茶		四球茶	
<i>C. huana</i>	KY626040	<i>C. semiserrata</i>	MT317096	<i>C. tetracocca</i>	OK405020
		楮叶齐		毛籽红山茶	
夏日七心		<i>C. sinensis</i> cv.	OL450428	<i>C. trichosperma</i>	OK181904
<i>C. 'Xiari Qixin'</i>	MZ004951	'Chuyeqi'			
凹脉金花茶		茶		越南油茶	
<i>C. impressinervis</i>	KF156835	<i>C. sinensis</i>	KC143082	<i>C. vietnamensis</i>	MN078085
东兴金花茶		茶		威宁短柱油茶	
<i>C. indochinensis</i> var.	OK135162	<i>C. sinensis</i>	MZ153237	<i>C. weiningensis</i>	MK820035
<i>tunghinensis</i>		安华		夏石金花茶	
山茶		<i>C. sinensis</i> cv.	MH042531	<i>C. xiashiensis</i>	NC_061602
<i>C. japonica</i>	MW602996	'Anhua'			
落瓣油茶		茶		攸县油茶	
<i>C. kissii</i>	NC_053915	<i>C. sinensis</i>	MH460639	<i>C. yuhsienensis</i>	NC_053622
膜叶茶		白鸡冠		五柱滇山茶	
<i>C. leptophylla</i>	KJ806275	<i>C. sinensis</i> cv.	MT773373	<i>C. yunnanensis</i>	KF156838
		'Baijiguan'			
乐业瘤果茶		白叶1号		红花长柄茶	
<i>C. leyeensis</i>	OK046127	<i>C. sinensis</i> cv.	MN086819	<i>C. zhaiana</i>	NC_058253
		'Baiyel'			
离蕊金花茶		半天腰		圆籽荷	
<i>C. liberofilamenta</i>	NC_061603	<i>C. sinensis</i> cv.	MW046255	<i>Apterosperma oblata</i>	KY406751
		'Bantianyao'			
柠檬金花茶					
<i>C. limonia</i>	OM238071				

1.3.4 基于核基因的基因流分析

为检验六堡茶群体种与突肋茶核基因组之间是否存在基因流，分别开展了 STRUCTURE 分析和 D 检验（ABBA-BABA 统计检验）。STRUCTURE 分析：利用 PGDSpider 2.1.1.5 软件（Lischer & Excoffier, 2012）将 32 份六堡茶群体种和突肋茶样品的核基因串联矩阵从 FASTA 格式转换为 VCF 格式。通过 Admixture 1.3.0 软件（Alexander et al., 2009）分析 32 份样品的种群结构和遗传混合程度，分别假设样品的分群数（ K 值）为 1~10 进行聚类，根据 CV error（Cross validation error）得到最佳 K 值（ CV error 最小）。D 检验：使用 Dsuite 0.4 r42 软件进行了 ABBA-BABA 统计检验（Malinsky et al., 2021）。该检验利用四个分类单元 P1、P2、P3 和 O，以 O 为外群，将 P1 和 P2 作为一对姐妹种，将 P3 作为与 P1 或 P2 具有潜在基因流的物种。统计具有 ABBA 和 BABA 等位基因模式的位点数量。D 统计量来自于计算 $D = (nABBA - nBABA) / (nABBA + nBABA)$ ，其中 $nABBA$ 和 $nBABA$ 分别为具有 ABBA 和 BABA 模式的位点总数。如果 $D \neq 0$ ， Z -score > 3 和 P 值 < 0.05 ，则存在基因流。根据叶绿体系统树结果，与突肋茶近缘的六堡茶群体种样品可能与突肋茶有基因流，为此，我们进一步根据核基因系统发育关系选择样品分别作为 P1、P2、P3、O 进行 D 检验：P1 分别用与六堡茶群体种近缘的 12 份栽培茶样品（简称：ZPC）、与六堡茶群体种近缘的其余 4 个茶组植物（简称：NO-ZPC）和与茶近

缘的六堡茶群体种 13 份样品（简称：LP）；P2 则分别用与突肋茶近缘的六堡茶群体种 14 份样品（简称：LP-TLC）和与茶近缘的六堡茶群体种 13 份样品（简称：LP）；P3 分别用突肋茶 5 份样品（简称：TLC）和与突肋茶近缘的 2 个物种（简称：NO-TLC）；O 为外类群，选择系统发育关系在茶、突肋茶和茶组物种外且相对近缘的包含 3 个物种的一分支为外类群。

1.3.5 六堡茶群体种样品的叶绿体基因组分歧进化时间估算

根据叶绿体基因组构建的六堡茶群体种系统发育关系，选择具有系统发育代表性的 6 份样品，即分散在不同支系，代表不同来源的两大类叶绿体基因组序列（见结果部分），将 6 份样品的叶绿体基因组序列整合到已发表的山茶科分化时间估算所用的叶绿体基因组数据中（Yu et al., 2017），共 77 份样品，比对后矩阵长度为 156 954 bp。参照 Yu 等（2017）的方法，用 8 个化石标定点，使用 BEAST 2.6.6 软件（Bouckaert et al., 2014）估算六堡茶群体种叶绿体基因组序列分歧进化暨关键演化事件发生的时间，具体设置包括：先验树采用 birth-death 模型，采用松弛分子钟模型 lognormal (UCLN) relaxed clock，运行 5 亿代，每 5 万代取样一次。使用 Tracer 1.7.2 软件（Rambaut et al., 2018）检查收敛情况，以确保有效采样值（ESS）均大于 100。使用 TreeAnnotator 2.6.6 软件（Bouckaert et al., 2014）舍弃掉初始的 20% 样本，利用剩余 80% 生成最大可信度时间树（maximum credibility chronogram; MCC）。最后用 Figtree 1.3.1 软件（Rambaut, 2010）对时间树进行展示和美化。

1.3.6 遗传多样性分析

利用 Dnasp 5.10 软件（Librado & Rozas, 2009）对六堡茶群体种的叶绿体基因组序列矩阵和串联核基因序列矩阵进行遗传多样性分析：分别计算单倍型数目（ Hh ）、单倍型多样性（ Hd ）、核苷酸多样性（ Pi ）、平均核苷酸差异（ K ）、单倍型多样性标准差（ Sh ）以及单倍型多样性方差（ Vh ）。

2 结果与分析

2.1 叶绿体基因组的基本结构

27 份六堡茶群体种的叶绿体基因组总长度范围为 156 430~157 131 bp，序列平均长度为 156 928 bp，GC 含量均为 37.3%。六堡茶群体种的叶绿体基因组结构为典型的环状四分体，由一个大单拷贝区（large single copy, LSC），2 个反向的重复区（inverted repeat, IR）以及一个小单拷贝区（small single copy, SSC）组成（图 1）。LSC 的长度范围为 85 531~86 669 bp，平均长度为 86 513 bp；IR 的长度范围为 26 018~26 090 bp，平均长度为 26 051 bp；SSC 的长度范围为 18 258~18 294 bp，平均长度为 18 276 bp。27 份六堡茶群体种样品均包含 132 个基因，其中 87 个蛋白质编码基因，37 个 tRNA 基因，8 个 rRNA 基因（表 3）。

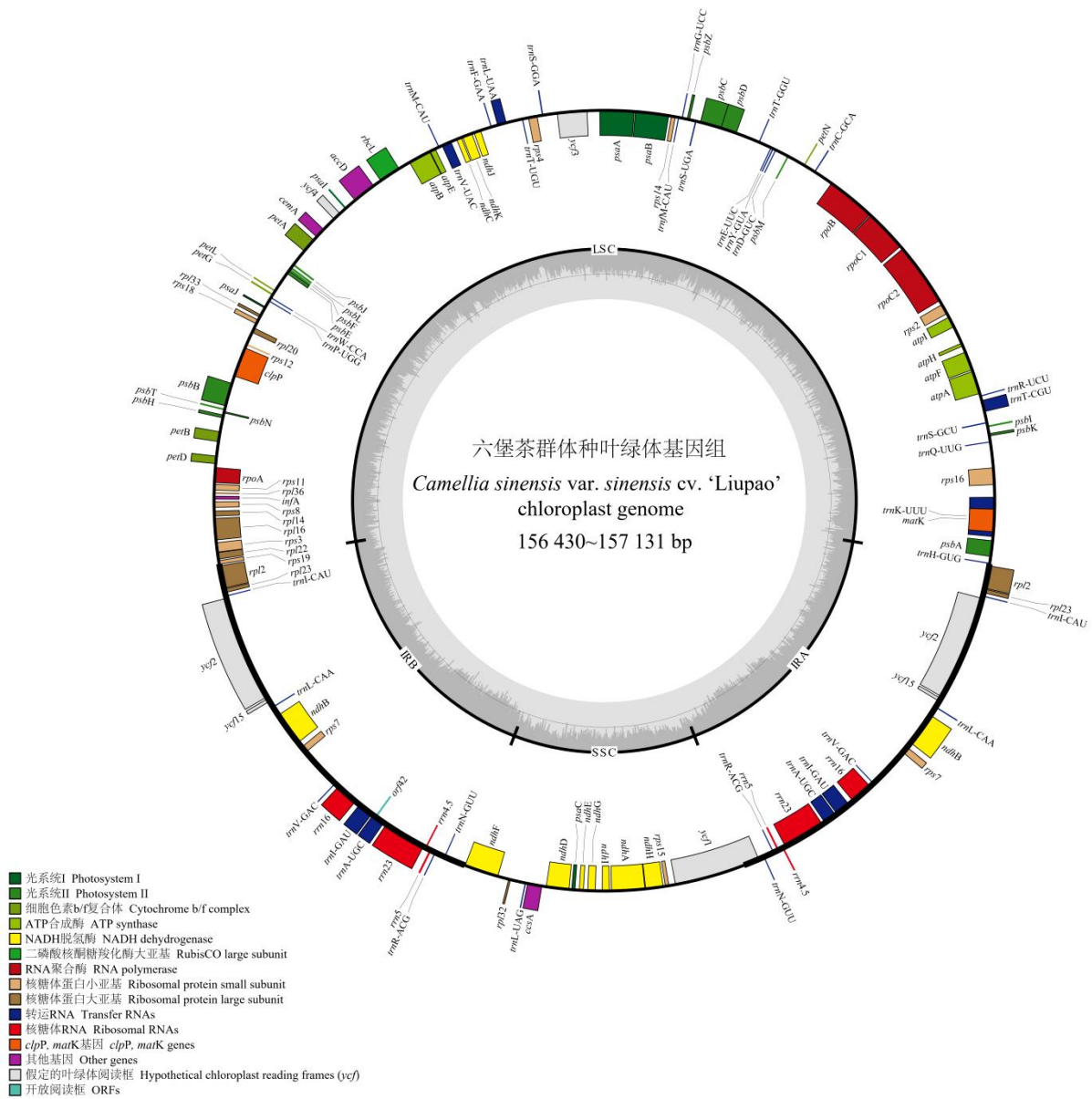


图 1 六堡茶群体种的叶绿体基因组图谱

Fig. 1 The chloroplast genome map of *Camellia sinensis* var. *sinensis* cv. 'Liupao'

表 3 27 份六堡茶群体种叶绿体基因组的基本信息

Table 3 Basic information of the 27 chloroplast genomes of the *Camellia sinensis* var. *sinensis* cv. 'Liupao'

样品	基因组大小	LSC 区长度	IR 区长度	SSC 区长度	基因数量	蛋白质编码基因数量	rRNA 数量	tRNA 数量	CG 含量
Samples	Genome size (bp)	LSC length (bp)	IR length (bp)	SSC length (bp)	Number of genes	Number of protein-coding genes	Number of rRNA	Number of tRNA	CG content (%)
H1	156 876	86 562	26 025	18 264	132	87	8	37	37.3

H2	156 850	86 531	26 025	18 269	132	87	8	37	37.3
H3	157 062	86 626	26 078	18 280	132	87	8	37	37.3
H4	156 884	86 560	26 025	18 274	132	87	8	37	37.3
H5	156 884	86 560	26 025	18 274	132	87	8	37	37.3
H6	156 829	86 535	26 018	18 258	132	87	8	37	37.3
H7	156 875	86 531	26 025	18 294	132	87	8	37	37.3
H8	157 098	86 656	26 080	18 282	132	87	8	37	37.3
H9	156 850	86 531	26 025	18 269	132	87	8	37	37.3
H10	156 875	85 531	26 025	18 294	132	87	8	37	37.3
H11	157 131	86 669	26 090	18 282	132	87	8	37	37.3
H12	157 063	86 627	26 078	18 280	132	87	8	37	37.3
Q13	157 026	86 587	26 081	18 277	132	87	8	37	37.3
Q14	156 850	86 531	26 025	18 269	132	87	8	37	37.3
Q15	156 430	86 123	26 019	18 269	132	87	8	37	37.3
Q16	156 829	86 535	26 018	18 258	132	87	8	37	37.3
Q17	157 131	86 669	26 090	18 282	132	87	8	37	37.3
Q18	156 880	86 561	26 026	18 267	132	87	8	37	37.3
Q19	156 430	86 123	26 019	18 269	132	87	8	37	37.3
Q20	156 866	86 562	26 020	18 264	132	87	8	37	37.3
Q21	157 066	86 642	26 072	18 280	132	87	8	37	37.3
Q22	157 025	86 586	26 081	18 277	132	87	8	37	37.3
D1	157 088	86 625	26 085	18 293	132	87	8	37	37.3
D2	157 088	86 625	26 085	18 293	132	87	8	37	37.3
D3	157 025	86 586	26 081	18 277	132	87	8	37	37.3
D4	157 025	86 586	26 081	18 277	132	87	8	37	37.3
D5	157 025	86 586	26 081	18 277	132	87	8	37	37.3
平均值									
Average value	156 928	86 513	26 051	18 276	132	87	8	37	37.3

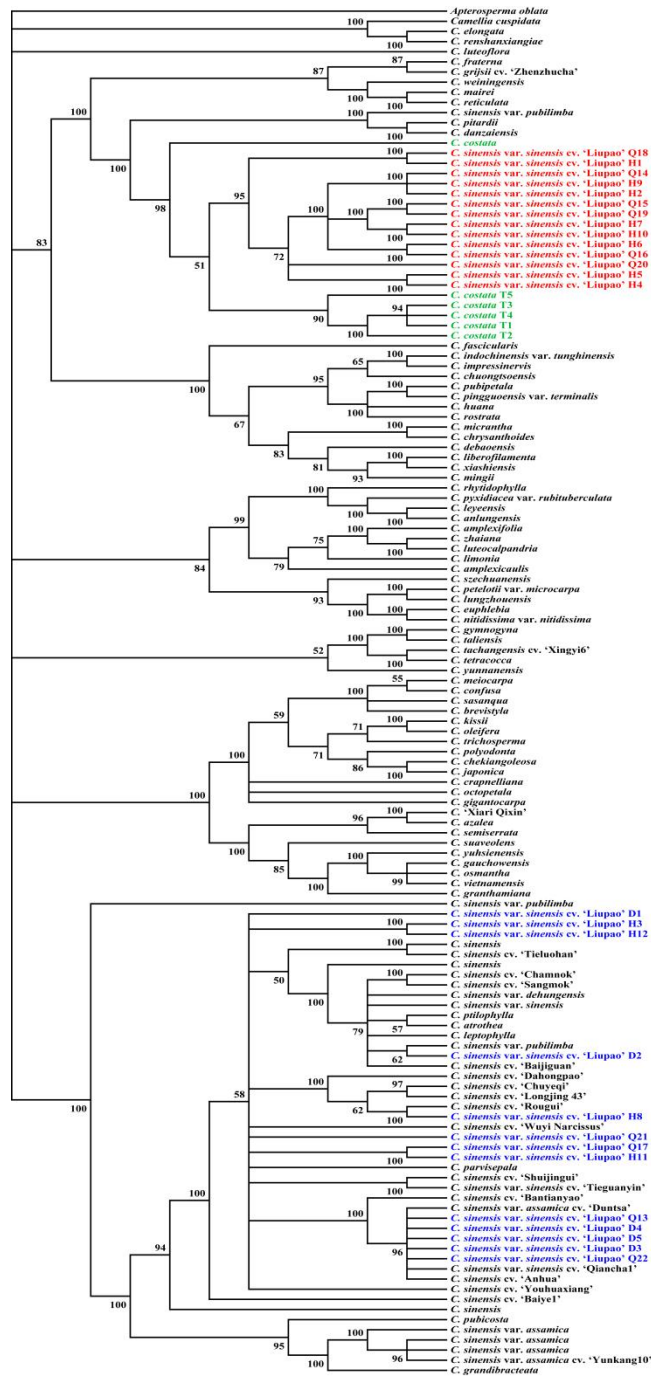
2.2 矩阵基本信息

叶绿体基因组构成的矩阵比对后序列全长为 162 774 bp, 包含 152 625 个保守位点和 5 457 个变异位点; 变异位点中包含 1 578 个简约信息位点。135 条序列的碱基含量平均值为 A = 31.1%, T=31.6%, C=19.0%, G=18.3%, 其中 A+T (62.7%) 的含量较高。

689 个单拷贝核基因串联矩阵全长为 843 771 bp, 包含 772 399 个保守位点和 71 345 个变异位点; 变异位点中包含 31 251 个简约信息位点; 70 条序列的碱基含量平均值为 A = 27.1%, T = 26.5%, C = 21.5%, G = 24.9%, 其中 A + T (53.6%) 的含量略高。

2.3 系统发育关系和种间杂交

叶绿体基因组最大似然树显示 (图 2), 27 份六堡茶群体种可分为系统发育关系较远的 2 大组。其中一组包含 13 份六堡茶群体种样品, 它们分散嵌入到主要由茶 (*C. sinensis*) 组成的一支内部 (BS=100%), 该分支还嵌入几个山茶属其他物种。另一组包括了剩余的 14 份六堡茶群体种样品, 它们自成单系 (BS=95%), 与突肋茶 (*C. costata*) 最近缘 (BS=98%), 且嵌入突肋茶内部与其中一支突肋茶互为姐妹 (BS=51%)。叶绿体基因组系统树表明了 27 份六堡茶群体种的母本至少包括两种来源, 其中一组来自于茶, 另一组来自于突肋茶。



两组具有不同来源叶绿体基因组序列的样品分别用红色和蓝色标注，突肋茶用绿色标注。下同。

The two groups of samples possessing plastid genomes of different sources are highlighted in red and blue, respectively, and the *Camellia costata* accessions are marked in green. The same below.

图 2 基于叶绿体基因组构建的 135 份样品最大似然树

Fig. 2 The maximum likelihood (ML) phylogenetic tree of the 135 samples based on the chloroplast genome sequences

基于核基因串联矩阵构建的最大似然树显示（图 3），27 份六堡茶群体种与栽培茶及其他几个山茶属物种构成一个多歧（polytomy）分支（BS=68%），该分支与突肋茶系统发育关系较远。由此表明，部分六堡茶群体种样品在叶绿体基因组树上与突肋茶最近缘最可能是因

为杂交，部分六堡茶群体种的祖先捕获了突肋茶的叶绿体基因组。

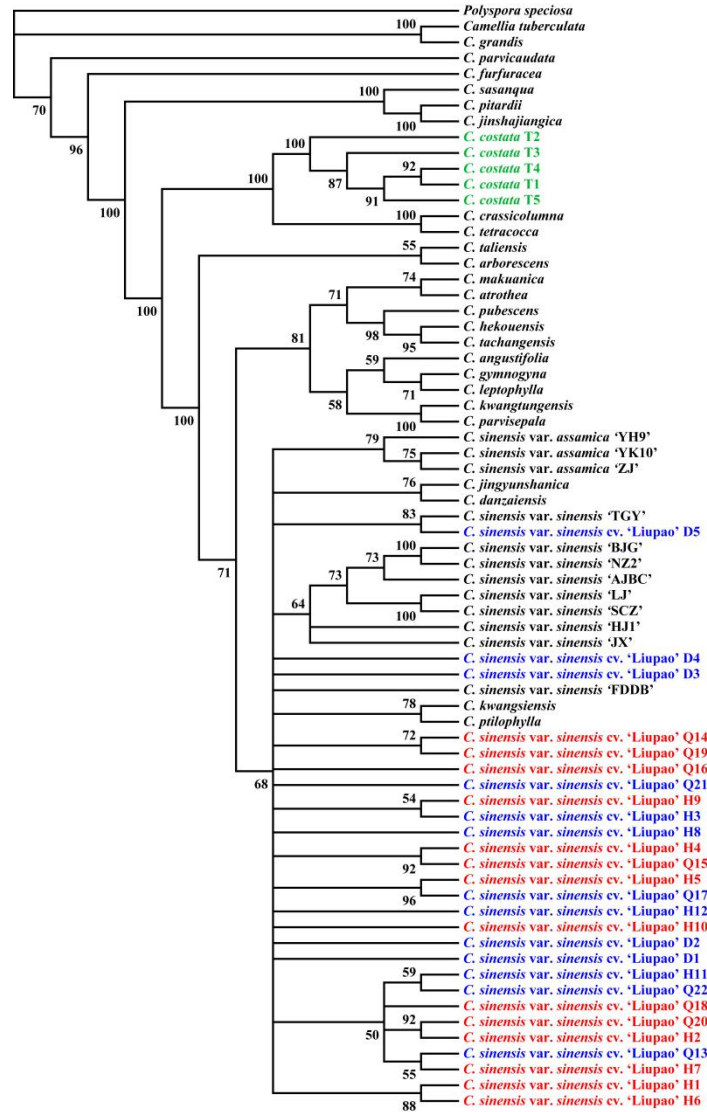


图 3 基于 689 个单拷贝核基因串联矩阵构建的 70 份样品的最大似然树

Fig. 3 The maximum likelihood (ML) phylogenetic tree of the 70 samples based on the concatenated data of 688 single-copy nuclear genes

2.4 基于核基因的六堡茶群体种和突肋茶基因流分析

基于核基因，对 32 份六堡茶群体种和突肋茶进行遗传混合程度分析结果表明，当 $K=2$ 时， CV error 最小，即为最佳 K 值（图 4: A）。32 份样品可分为 2 个类群即六堡茶群体种和突肋茶，两个类群之间没有遗传物质混合，即未检测到基因流（图 4: B）。D 检验同样检测不到六堡茶群体种和突肋茶之间有基因流（表 4）。上述核质基因组系统发育矛盾已表明，部分六堡茶群体种的祖先曾与突肋茶发生过杂交事件，但可能因为杂交后代作为母本又反复与父本六堡茶群体种发生了回交，经过多代回交后，这些捕获了突肋茶叶绿体基因组的杂交后代核基因组中的突肋茶基因已被充分稀释，所以根据核基因序列已检测不到基因流。

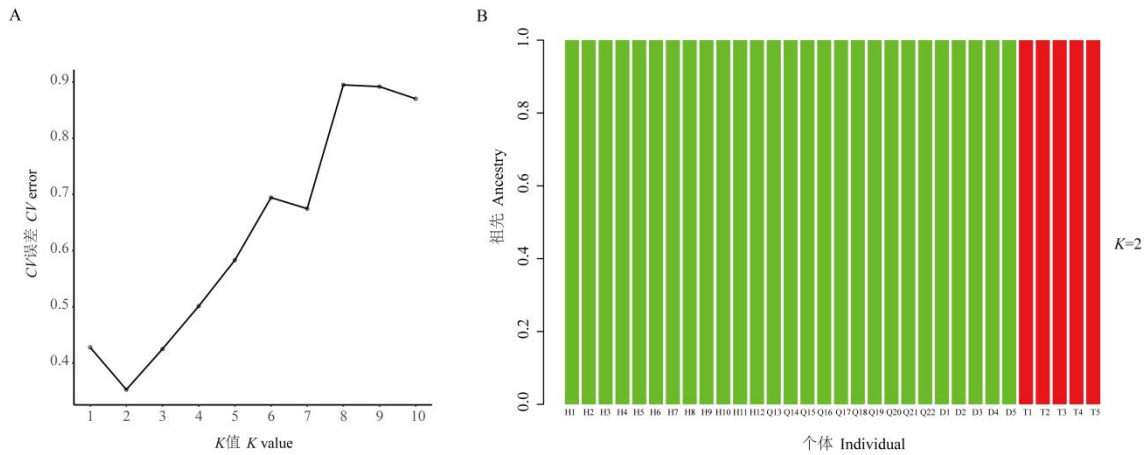


图 4 K 值最佳路线图 (A) 和群体结构图 (B)

Fig. 4 The optimal K value (A) and the population structure diagrams (B)

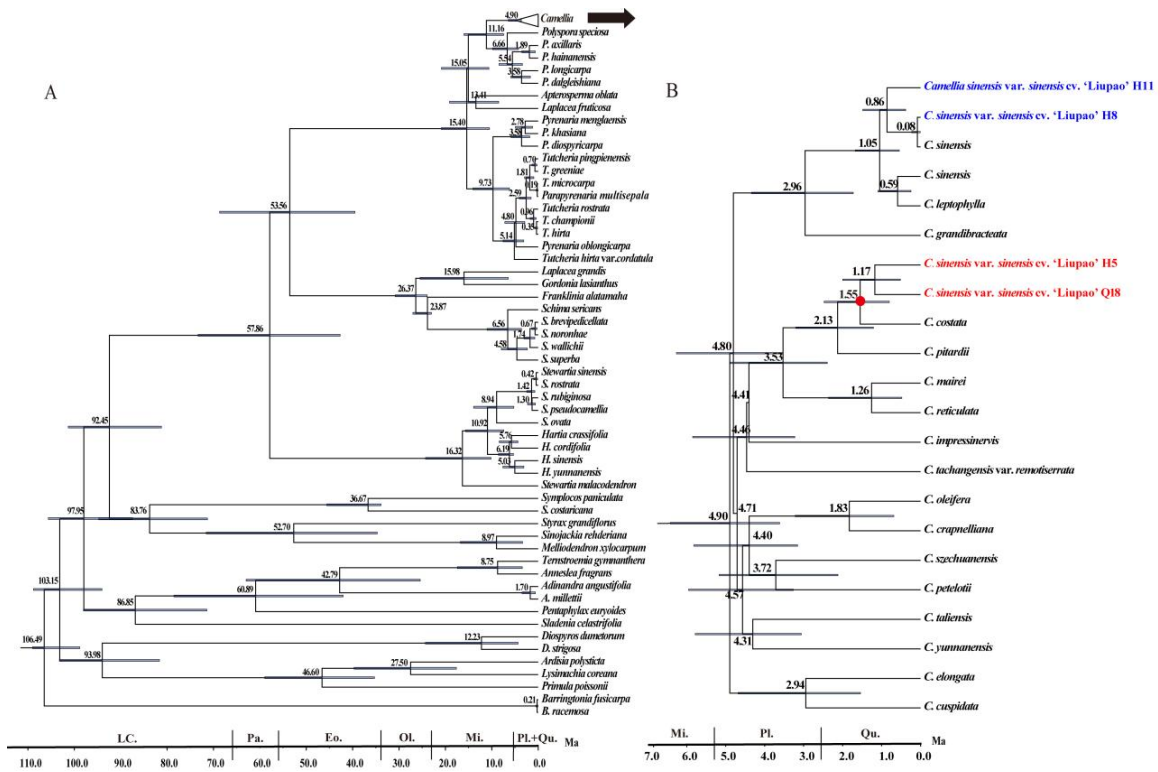
表 4 基于 D 统计检验六堡茶群体种和突肋茶之间的基因流结果

Table 4 Gene flow between *Camellia sinensis* var. *sinensis* cv. 'Liupao' and *C. costata* revealed by D-statistic test

P1	P2	P3	D 统计量 Dstatistic	Z 分数 Z score	P 值 P value	f4 比率 f4 ratio	BBAA	ABBA	BABA
ZPC	LP-TLC	TLC	0.04	1.51	0.13	0.03	415.55	134.14	124.77
ZPC	LP	TLC	0.02	1.02	0.31	0.02	444.29	133.50	128.37
ZPC	LP-TLC	NO-TLC	0.01	0.66	0.51	0.04	358.28	141.23	137.39
ZPC	LP	NO-TLC	0.01	0.56	0.57	0.02	375.56	136.79	134.10
NO-ZPC	LP	TLC	0.02	0.61	0.54	0.02	396.32	122.65	118.21
NO-ZPC	LP-TLC	TLC	0.01	0.19	0.85	0.01	399.27	127.66	126.21
NO-ZPC	LP	NO-TLC	0.01	0.36	0.72	0.02	351.60	137.59	134.71
NO-ZPC	LP-TLC	NO-TLC	0.01	0.28	0.78	0.02	346.56	133.35	130.75
LP	LP-TLC	TLC	0.00	0.21	0.83	0.00	382.56	138.59	138.06
LP	LP-TLC	TLC	0.00	0.16	0.88	0.00	444.13	128.01	127.60

2.5 六堡茶群体种与突肋茶杂交的时间

基于叶绿体基因组数据估算的分子钟结果表明 (图 5), 部分六堡茶群体种的祖先大约于 1.55 百万年前 (95% HPD: 0.80~2.47 Ma) 的第四纪时期与部分突肋茶的祖先发生了杂交, 捕获了突肋茶的叶绿体基因组。杂交事件远早于人类栽培和制作茶的历史。



A 展示属间分化时间，包含六堡茶群体种的山茶属分支用三角形代替；B 展示包含六堡茶群体种在内的山茶属分支的分化时间，红色圆点表示六堡茶群体种发生杂交事件的时间节点。地层边界的年代来自国际年代地层图（Cohen et al., 2013）（LC. 晚白垩世；Pa. 古新世；Eo. 始新世；Ol. 渐新世；Mi. 中新世；Pl. 上新世；Qu. 第四纪），单位为百万年前（Ma）。

A shows the intergeneric divergence times and the *Camellia* clade is compressed into a triangle; B shows the divergence times in *Camellia* including *C. sinensis* var. *sinensis* cv. 'Liupao', the red dot in the figure indicates the time of the introgression hybridization event in *C. sinensis* var. *sinensis* cv. 'Liupao'. The ages of stratigraphic boundaries were obtained from the International Chronostratigraphic Chart (Cohen et al., 2013) (LC. Late Cretaceous; Pa. Paleocene; Eo. Eocene; Ol. Oligocene; Mi. Miocene; Pl. Pliocene; Qu. Quaternary), with a scale of millions of years ago (Ma).

图 5 基于 77 份叶绿体基因组序列估算的时间树

Fig. 5 The chronogram based on the 77 chloroplast genome sequences

2.6 遗传多样性

六堡茶群体种的遗传多样性分析结果显示（表 5），27 份叶绿体基因组序列共定义了 16 个单倍型，单倍型多样性 (Hd) 为 0.949，核苷酸多样性 (Pi) 为 0.000 9，平均核苷酸差异 (K) 为 140.057，单倍型多样性方差 (Vh) 为 0.000 62，单倍型多样性标准差 (Sh) 为 0.025。而 27 份六堡茶群体种单拷贝核基因串联序列共定义了 27 个单倍型，单倍型多样性 (Hd) 为 1，核苷酸多样性 (Pi) 为 0.004 72，平均核苷酸差异 (K) 为 2 598.37，单倍型多样性方差 (Vh) 为 0.000 1，单倍型多样性标准差 (Sh) 为 0.01。

表 5 六堡茶群体种不同基因组序列的遗传多样性参数

Table 5 Genetic diversity parameters of sequences from different genomes of *Camellia sinensis* var. *sinensis* cv. 'Liupao'

数据类型 Data type	单倍型数量 Number of haplotypes	单倍型多样性 Haplotype diversity	核苷酸多样性 Nucleotide diversity	平均核苷酸差异 Average number of nucleotide differences	单倍型多样性方差 Variance of haplotype diversity	单倍型多样性标准差 Standard deviation of haplotype diversity
叶绿体基因组 Chloroplast genome	16	0.949	0.000 9	140.057	0.000 62	0.025
核基因 Nuclear gene	27	1	0.004 72	2 598.37	0.000 1	0.01

3 讨论与结论

3.1 六堡茶群体种杂交捕获突肋茶叶绿体基因组的历史过程

部分六堡茶群体种样品在基于叶绿体基因组和核基因重建的系统树上的位置显著不同（核质结果不一致）。在叶绿体树上，部分六堡茶群体种样品与突肋茶最近且嵌入其内部，而在核基因树上，这些样品与突肋茶亲缘关系较远。除了上述部分六堡茶群体种样品外，其余茶样品在叶绿体和核基因树上都与突肋茶亲缘关系较远。已有研究表明，叶绿体捕获事件是核质系统发育关系不一致的重要原因之一（Fehrer et al., 2007; Yi et al., 2015; Liu et al., 2017），叶绿体捕获在同域分布（或分布区重叠）和生殖相容的物种间可能频繁发生（Acosta & Premoli, 2010）。根据记载，野生突肋茶在广东、广西和贵州均有分布，尤其是广西贺州昭平县有丰富的野生突肋茶资源（李朝昌等，2019）。这表明，六堡茶群体种与突肋茶的分布区有重叠，为杂交提供了可能性。所以，造成六堡茶群体种部分样品在叶绿体树上与突肋茶最近缘的最可能原因是它们的祖先与部分突肋茶的祖先发生了杂交，捕获了突肋茶的叶绿体基因组。在串联核基因树上，含有突肋茶类型叶绿体基因组的六堡茶群体种样品与其他六堡茶群体种样品均不与突肋茶最近，表明这些样品虽然为杂交后代，但可能因为杂交后代作为母本又反复与父本六堡茶群体种发生了回交，经过多代回交后，这些杂交后代核基因组中的突肋茶基因已被充分稀释，所以没有杂交后代个体在核基因树上与突肋茶最近缘。基于核基因串联矩阵的基因流分析也未检测到六堡茶群体种样品与突肋茶之间的基因流，这进一步支持了杂交后代个体核基因组中突肋茶基因已被充分稀释的推测。这一推测也与先前的研究相符：当不同物种发生杂交后，杂交后代与亲本反复回交会导导致叶绿体基因组从一个物种渗入到另一个物种，最终导致回交后代的细胞核与细胞质遗传信息来自不同的物种（Rieseberg & Soltis, 1991; Acosta & Premoli, 2010; Kleinkopf et al., 2019）。此外，杂交后代个体中，叶绿体基因组序列和核基因序列都有相当程度的分化，这有两种可能的解释：一种是一次古老杂交产生的杂交后代发生了分化；另一种是含有不同单倍型的六堡茶群体种与突肋茶发生多次近期独立杂交。鉴于杂交后代核基因与纯系六堡茶群体种及其他茶品种最近而与突肋茶较远，且杂交后代样品在叶绿体树上为单系并嵌入突肋茶内部，以及杂交后代样品与突肋茶在核基因组中未检测到基因流，所以我们推测所取样的六堡茶群体种杂交样品更可能是其祖先与部分突肋茶祖先发生一次古老杂交的后代。分子钟结果进一步表明，六堡茶群体种父本祖先与突肋茶母本祖先在 1.55 百万年前的第四纪时期发生了杂交和叶绿体基因组捕获，该事件的发生远远早于人类栽培和制作茶的历史，因此，现有六堡茶群体种多样性应是选育了具有不同母系来源的野生六堡茶，而非在栽培过程中人为介导了杂交。

3.2 六堡茶群体种遗传多样性

无论在叶绿体基因组上还是核基因上，六堡茶群体种的单倍型多样性与核苷酸多样性相比，单倍型多样性都较高，这可能是由于少量碱基突变能迅速增加单倍型多样性，但核苷酸多样性却难以在短期内迅速增加。六堡茶群体种叶绿体基因组的单倍型多样性 ($Hd = 0.949$) 远高于 Petit 等计算的 170 种植物的叶绿体 DNA 多样性的平均值 0.67 (Petit et al., 2005)，这一定程度上表明六堡茶群体种的遗传多样性较高，与前人基于 SSR 对六堡茶群体种进行的遗传多样性分析得出六堡茶群体种遗传多样性高的结论一致 (王留彬等, 2022)。已有研究发现杂交或基因渐渗会提高茶树的遗传多样性 (Zhang et al., 2021)，六堡茶群体种也通过杂交捕获了突肋茶的叶绿体基因组，从而使得六堡茶群体种遗传多样性较高。六堡茶群体种具有丰富的遗传多样性和较高的演化潜力，应充分保护具有不同遗传背景的个体，这有利于六堡茶产业的可持续发展，并为进一步的遗传改良和品质提升提供更多的选择和可能性。

3.3 六堡茶品种选育和资源开发利用

优质的茶树品种是茶叶优异品质形成的重要因素 (刘佳等, 2013)。茶树品种影响茶叶化学成分组成与含量，从而影响茶叶的品质 (郭桂义等, 2010)。随着六堡茶市场不断扩大，为了满足日益增长的需求，六堡茶的制作除了使用传统的六堡茶群体种外，还引进了其他外地的适制品种，比如凌云白毫茶、临桂白毫茶、资源大桂茶等，加之六堡茶茶园里茶树品种来源不一，导致市场上部分六堡茶产品失去了其原来的品质特点 (刘晓东等, 2013)。本研究显示六堡茶群体种具有不同的遗传背景和演化历史，具有较高遗传多样性。这为进一步探索不同遗传背景与农艺性状、茶品质之间的关联提供了契机。未来的研究可以通过深入的遗传分析和农艺性状评估，探索不同品种间的遗传差异与茶叶品质之间的关系，为高品质的六堡茶株系或品种选育、鉴定和开发利用提供新的思路和方案。此外，借鉴六堡茶与突肋茶间的自然杂交，可以开展种间或品种间人工杂交。通过定向杂交引入其他品种或种类的优质基因，可以进一步改良六堡茶的品质特点，获得高品质的新品种。

参考文献:

- 陈佳, 廖传盛, 刘春梅, 等, 2020. 适制六堡茶的广西无性系茶树良种介绍[J]. 南方农业, 14(19): 53-56.
- 陈佳, 覃秀菊, 赖兆荣, 等, 2010. 六堡茶原生种的改良技术探讨[J]. 中国园艺文摘, 26(9): 183-184.[CHEN J, QIN XJ, LAI ZR, et al., 2010. Study on improved technologies of 'Liubao' tea native species[J]. Chin Horticult Abstr, 26(9): 183-184.]
- 陈宗懋, 1992. 中国茶经[M]. 上海: 上海文化出版社: 130.
- 中国茶树品种志编写委员会, 2001. 中国茶树品种志[M]. 上海: 上海科学技术出版社: 187.
- 龚明寿, 陆超丽, 卢一科, 等, 2021. 昭平县野生茶树种质资源开发利用初探[J]. 农业研究与应用, 34(4): 68-72. [GONG MS, LU CL, LU YK, et al., 2021. Exploitation and utilization of wild tea germplasm resources in Zhaoping County[J]. Agric Res Appl, 34(4): 68-72.]
- 郭灿, 皮发娟, 吴昌敏, 等, 2021. 基于 GBS 测序的全基因组 SNP 揭示贵州地方茶组植物资源的亲缘关系[J]. 南方农业学报, 52(3): 660-670. [GUO C, PI FJ, WU CM, et al., 2021. Genome-wide SNP developed by genotyping-by-sequencing revealed the phylogenetic relationship of Sect. *Thea* (L.) Dyer resources in Guizhou[J]. J South Agric, 52(3): 660-670.]
- 郭桂义, 严佩峰, 文宏, 等, 2010. 安吉白茶与信阳群体种信阳毛尖茶化学成分和品质的比较[J]. 食品科技, 35(6): 118-121.[GUO GY, YAN PF, WEN H, et al., 2010. Comparison of

- chemical components and organoleptic quality of Xinyang maojian tea made by Anjibaicha and the Xinyang community variety[J]. *Food Sci Technol*, 35(6): 118-121.]
- 黄厚宸, 蒋维昕, 梁彩霞, 等, 2021. 广西六堡茶古茶树及其子代遗传多样性的 EST-SSR 标记分析[J]. *分子植物育种*, 19(7): 2410-2418.[HUANG HC, JIANG WX, LIANG CX, et al., 2021. EST-SSR-based analysis on genetic diversity of ancient Liupao tea trees and their progeny in Guangxi[J]. *Mol Plant Breed*, 19(7): 2410-2418.]
- 李朝昌, 邓慧群, 区胜基, 等, 2019. 广西昭平县野生突肋茶调查初报[J]. *广西农学报*, 34(5): 34-36.[LI CC, DENG HQ, OU SJ, et al., 2019. A preliminary report on investigating wild *Camellia costata* in Zhaoping County, Guangxi[J]. *J Guangxi Agric*, 34(5): 34-36.]
- 李苗苗, MEEGAHAKUMBURA MK, 严丽君, 等, 2015. 核基因组微卫星标记揭示大理茶参与了普洱茶的驯化过程[J]. *植物分类与资源学报*, 37(1): 29-37.[LI MM, MEEGAHAKUMBURA MK, YAN LJ, et al., 2015. Genetic involvement of *Camellia taliensis* in the domestication of *C. sinensis* var. *assamica* (Assamica tea) revealed by nuclear microsatellite markers[J]. *Plant Diversity Resour*, 37(1): 29-37.]
- 刘佳, 梁丽云, 冯建灿, 等, 2013. 信阳毛尖品质影响因素研究进展[J]. *河南农业科学*, 42(7): 1-5.[LIU J, LIANG LY, FENG JC, et al., 2013. Research progress of factors influencing the quality of Xinyang maojian[J]. *J Henan Agric Sci*, 42(7): 1-5.]
- 刘晓东, 杨春, 汤周斌, 等, 2013. 适制六堡茶茶树品种部分遗传标记特征及其在六堡茶选育种中的作用[J]. *中国农学通报*, 29(4): 136-140.[LIU XD, YANG C, TANG ZB, et al., 2013. Research on the genetic marks of tea plants that suitable to make Liubao tea and their roles in the breeding of Liubao tea[J]. *Chin Agric Sci Bull*, 29(4): 136-140.]
- 邱瑞瑾, 龙志荣, 于翠平, 等, 2020. 六堡茶群体种资源的调查、开发和保护[J]. *中国茶叶加工*, (3): 15-19.[QIU RJ, LONG ZR, YU CP, et al., 2020. Investigation, development and protection of the population resources of the Liupao tea[J]. *China Tea Proc*, (3): 15-19.]
- 滕翠琴, 邱瑞瑾, 于翠平, 等, 2020. 岑溪市野生古茶树资源特征与开发利用现状调查[J]. *南方农业*, 14(4): 45-48.
- 王留彬, 黄丽蕴, 滕翠琴, 等, 2022. 梧州茶树种质资源的遗传多样性及亲缘关系分析[J]. *茶叶科学*, 42(5): 601-609.[WANG LB, HUANG LY, TENG CQ, et al., 2022. Genetic and phylogenetic analysis for germplasm resources of *Camellia sinensis* from Wuzhou city[J]. *J Tea Sci*, 42(5): 601-609.]
- 闫明慧, 刘柯, 王满, 等, 2021. 信阳 10 号叶绿体基因组及其系统进化[J]. *茶叶科学*, 41(6): 777-788.[YAN MH, LIU K, WANG M, et al., 2021. Complete chloroplast genome of *Camellia sinensis* cv. Xinyang 10 and its phylogenetic evolution[J]. *J Tea Sci*, 41(6): 777-788.]
- 杨世雄, 2021a. 茶组植物的分类历史与思考[J]. *茶叶科学*, 41(4): 439-453.[YANG SX, 2021a. Thinking on the taxonomy of *Camellia* sect. *Thea*[J]. *J Tea Sci*, 41(4): 439-453.]
- 杨世雄, 2021b. 广西的茶树资源[J]. *广西林业科学*, 50(4): 414-416.[YANG SX, 2021b. Tea-plant (*Camellia* sect. *Thea*) in Guangxi, China[J]. *Guangxi For Sci*, 50(4): 414-416.]
- 周炎花, 乔小燕, 马春雷, 等, 2011. 广西茶树地方品种遗传多样性和遗传结构的 EST-SSR 分析[J]. *林业科学*, 47(3): 59-67.[ZHOU YH, QIAO XY, MA CL, et al., 2011. Genetic diversity and structure of tea landraces from Guangxi, based on EST-SSR analysis[J]. *Sci Silv Sin*, 47(3): 59-67.]
- ACOSTA MC, PREMOLI AC, 2010. Evidence of chloroplast capture in South American *Nothofagus* (subgenus *Nothofagus*, Nothofagaceae)[J]. *Mol Phylogenet Evol*, 54: 235-242.

- ALEXANDER DH, NOVEMBRE J, LANGE K, 2009. Fast model-based estimation of ancestry in unrelated individuals[J]. *Genome Res*, 19(9): 1655-1664.
- BOUCKAERT R, HELED J, KÜHNERT D, et al., 2014. BEAST 2: a software platform for bayesian evolutionary analysis[J]. *Plos Comput Biol*, 10(4): e1003537.
- CHEN SF, ZHOU YQ, CHEN YR, et al., 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor[J]. *Bioinformatics*, 34(17): i884-i890.
- COHEN KM, FINNRY SC, GIBBARD PL, et al., 2013. The ICS international chronostratigraphic chart[J]. *Episodes*, 36: 199-204.
- EDGAR RC, 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput[J]. *Nucl Acid Res*, 32(5): 1792-1797.
- EMMS DM, KELLY S, 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics[J]. *Genome Biol*, 20(1): 238.
- FEHRER J, GEMEINHOLZER B, CHRTEK J, et al., 2007. Incongruent plastid and nuclear DNA phylogenies reveal ancient intergeneric hybridization in *Pilosella* hawkweeds (*Hieracium*, Cichorieae, Asteraceae)[J]. *Mol Phylogenet Evol*, 42(2): 347-361.
- FU LM, NIU BF, ZHU ZW, et al., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data[J]. *Bioinformatics*, 28(23): 3150-3152.
- HAAS BJ, PAPANICOLAOU A, YASSOUR M, et al., 2013. De novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis[J]. *Nat Protoc*, 8(8): 1494-1512.
- HALL TA, 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Window 95/98/NT[J]. *Nucl Acid Symp Ser*, 41: 95-98.
- HUANG H, TONG Y, ZHANG QJ, et al., 2013. Genome size variation among and within *Camellia* species by using flow cytometric analysis[J]. *PLoS ONE*, 8(5): e64981.
- JIN JJ, YU WB, YANG JB, et al., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes[J]. *Genome Biol*, 21(1): 241.
- KATO H, STANDLEY DM, 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability[J]. *Mol Biol Evol*, 30(4): 772-780.
- KLEINKOPF JA, ROBERTS WR, WAGNER WL, et al., 2019. Diversification of Hawaiian *Cyrtandra* (Gesneriaceae) under the influence of incomplete lineage sorting and hybridization[J]. *J Syst Evol*, 57(6): 561-578.
- KUMAR S, STECHER G, TAMURA K, 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets[J]. *Mol Biol Evol*, 33(7): 1870-1874.
- LI L, HU YF, HE M, et al., 2021. Comparative chloroplast genomes: insights into the evolution of the chloroplast genome of *Camellia sinensis* and the phylogeny of *Camellia*[J]. *BMC Genomics*, 22(1): 138.
- LIBRADO P, ROZAS J, 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data[J]. *Bioinformatics*, 25(11): 1451-1452.
- LISCHER HE, EXCOFFIER L, 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs[J]. *Bioinformatics*, 28(2): 298-299.
- LIU X, WANG Z, SHAO W, et al., 2017. Phylogenetic and taxonomic status analyses of the abaso section from multiple nuclear genes and plastid fragments reveal new insights into the North America origin of *Populus* (Salicaceae)[J]. *Front Plant Sci*, 7: 2022.

- MALINSKY M, MATSCHINER M, SVARDAL H, 2021. Dsuite-Fast D-statistics and related admixture evidence from VCF files[J]. *Mol Ecol Resour*, 21(2): 584-595.
- MEEGAHAKUMBURA MK, WAMBULWA MC, THAPA KK, et al., 2016. Indications for three independent domestication events for the tea plant (*Camellia sinensis* (L.) O. Kuntze) and new insights into the origin of tea germplasm in China and India revealed by nuclear microsatellites[J]. *Plos One*, 11(5): e0155369.
- MING TL, BARTHOLOMEW B. Theaceae[M]//WU ZY, RAVEN PH, HONG DY, 2007. *Flora of China*: Vol. 12. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press.
- PETIT RJ, DUMINIL J, FINESCHI S, et al., 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations[J]. *Mol Ecol*, 14(3): 689-701.
- QU XJ, MOORE MJ, LI DZ, et al., 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes[J]. *Plant Methods*, 15: 50.
- RAMBAUT A, 2010. FigTree v1.3.1[CP]. Edinburgh, UK: Institute of Evolutionary Biology, University of Edinburgh. <http://tree.bio.ed.ac.uk/software/figtree/>.
- RAMBAUT A, DRUMMOND AJ, XIE D, et al., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7[J]. *Syst Biol*, 67(5): 901-904.
- RIESEBERG LH, SOLTIS DE, 1991. Phylogenetic consequences of cytoplasmic gene flow in plants[J]. *Am J Bot*, 5(1): 65-84.
- STAMATAKIS, A, 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models[J]. *Bioinformatics*, 22(21): 2688-2690.
- UENO S, TSUMURA Y, 2009. Development of microsatellite and amplicon length polymorphism markers for *Camellia japonica* L. from tea plant (*Camellia sinensis*) expressed sequence tags[J]. *Mol Ecol Resour*, 9(3): 814-816.
- WANG H, OOI BC, TAN KL, et al., 2003. BLAST++: BLASTing queries in batches[J]. *Bioinformatics*, 19(17): 2323-2324.
- WANG XC, FENG H, CHANG YX, et al., 2020. Population sequencing enhances understanding of tea plant evolution[J]. *Nat Commun*, 11(1): 4447.
- WU Q, TONG W, ZHAO HJ, et al., 2022. Comparative transcriptomic analysis unveils the deep phylogeny and secondary metabolite evolution of 116 *Camellia* plants[J]. *Plant J*, 111(2): 406-421.
- YANG Y, MOORE MJ, BROCKINGTON SF, et al., 2015. Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing[J]. *Mol Biol Evol*, 32(8): 2001-2014.
- YE XQ, ZHAO ZH, ZHU QW, et al., 2014. Entire chloroplast genome sequence of tea (*Camellia sinensis* cv. Longjing 43): a molecular phylogenetic analysis[J]. *J Zhejiang Univ*, 40(4): 404-412.
- YI TS, JIN GH, WEN J, 2015. Chloroplast capture and intra- and inter-continental biogeographic diversification in the Asian-New World disjunct plant genus *Osmorhiza* (Apiaceae)[J]. *Mol Phylogenet Evol*, 85: 10-21.
- YU XQ, GAO LM, SOLTIS DE, et al., 2017. Insights into the historical assembly of East Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family[J]. *New Phytol*, 215(3): 1235-1248.
- ZAN T, HE YT, ZHANG M, et al., 2023. Phylogenomic analyses of *Camellia* support reticulate

evolution among major clades[J]. *Mol Phylogenet Evol*, 182: 107744.

ZHANG D, GAO FL, JAKOVLIĆ I, et al., 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies[J]. *Mol Ecol Resour*, 20(1): 348-355.

ZHANG XT, CHEN S, SHI LQ, et al., 2021. Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant *Camellia sinensis*[J]. *Nat Genet*, 53(8): 1250-1259.