

楮叶齐茶树叶绿体基因组特征与进化分析

曾文娟^{1,2,3}, 张其湘^{1,2,3*}, 贺宁^{1,2,3}, 刘雅^{1,2,3}, 龚意辉^{1,2,3,4}, 陈致印^{1,2,3,4**}

1. 湖南人文科技学院农业与生物技术学院, 湖南娄底 417000; 2. 湖南人文科技学院湖南省园艺生产与加工类创新创业教育中心, 湖南娄底 417000; 3. 湖南人文科技学院湖南省创新创业示范基地, 湖南娄底 417000; 4. 湖南人文科技学院湖南省湘中特色农业资源开发利用与质量安全控制重点实验室, 湖南娄底 417000

摘要: 本研究对茶树品种楮叶齐叶绿体基因组进行测序, 并重点分析其结构特征、密码子偏好性及系统发育关系。结果显示: 其叶绿体基因组全长 157 072 bp, 共注释 135 个功能基因。研究发现其反向重复区 (inverted repeat, IR) 的 GC 含量最高 (42.96%), 并鉴定出 *clpP*、*ycf3* 等多个含内含子基因。密码子偏好性分析表明, 楮叶齐叶绿体基因组的密码子偏好性较弱, 且主要受自然选择主导 (贡献率 86.6%), 最优密码子多以 A/U 结尾, 这一结果与突变压力预期不符。系统发育分析进一步证实, 楮叶齐与普洱茶、白毛茶等栽培类型聚为一个高支持率的单系分支, 存在密切的母系遗传关联。本研究为深入解析山茶科物种进化及叶绿体功能优化提供重要依据。

关键词: 茶树; 楮叶齐; 叶绿体基因组; 密码子使用偏好性; 系统发育

中图分类号: S571.1 文献标志码: A

Chloroplast Genome Characteristics and Evolutionary Analysis of the Tea Plant Variety Zhuyeqi

ZENG Wenjuan^{1,2,3}, ZHANG Qixiang^{1,2,3*}, HE Ning^{1,2,3}, LIU Ya^{1,2,3}, GONG Yihui^{1,2,3,4}, CHEN Zhiyin^{1,2,3,4**}

1. College of Agriculture and Biotechnology, Hunan University of Humanities, Science and Technology, Loudi, Hunan 417000, China; 2. Innovation and Entrepreneurship Education Center for Horticultural Production and Processing in Hunan Province, Hunan University of Humanities, Science and Technology, Loudi, Hunan 417000, China; 3. Hunan Provincial Innovation and Entrepreneurship Demonstration Base, Hunan University of Humanities, Science and Technology, Loudi, Hunan 417000, China; 4. Key Laboratory of Characteristic Agricultural Resource Development and Quality Safety Control in Hunan Province, Hunan University of Humanities, Science and Technology, Loudi, Hunan 417000, China

Abstract: This study represents sequencing of the chloroplast genome of the tea plant variety Zhuyeqi, with a focus on analyzing its structural characteristics, codon usage bias, and phylogenetic relationships. The results revealed that its chloroplast genome had 157 072 bp, with 135 functional genes annotated. The GC content in the IR region was found to be the highest (42.96%), and several intron-containing genes, such as *clpP* and *ycf3*, were identified. Codon usage bias analysis indicated a relatively weak bias, primarily driven by natural selection (contribution rate of 86.6%). Optimal codons were mostly A/U-ending, deviating from expectations based on mutational pressure. Phylogenetic analysis further confirmed that Zhuyeqi formed a highly supported monophyletic clade with cultivated types such as Pu-erh tea and Baihao tea, demonstrating close maternal genetic relationships. This study would provide important insights into the evolution of *Theaceae* species and the functional optimization of chloroplast genomes.

Keywords: *Camellia sinensis*; Zhuyeqi; chloroplast genome; codon usage bias; phylogenetic analysis

DOI: 10.3969/j.issn.1000-2561.2026.02.004

收稿日期 2025-09-04; 接受日期 2025-11-07

基金项目 湖南省科技创新计划项目 (No. 2024RC8289); 国家级大学生创新创业训练项目 (No. 202510553020X); 省级大学生创新创业训练项目 (No. S202510553039)。

作者简介 曾文娟 (2004—), 女, 本科生, 研究方向: 茶学; *同等贡献作者: 张其湘 (2005—), 男, 本科生, 研究方向: 茶树遗传育种。**通信作者 (Corresponding author): 陈致印 (CHEN Zhiyin), E-mail: 772612626@qq.com。

茶树 [*Camellia sinensis* (L.) O. Kuntze] 是全球最重要的多年生木本经济作物之一, 其鲜叶加工而成的茶叶富含多酚、咖啡碱和芳香物质, 具有显著的保健功效与经济价值。2023 年全球茶叶产量已超 600 万 t, 中国作为最大的茶叶生产国和消费国, 在茶产业可持续发展中占据核心地位。茶树种质资源丰富, 包含多个变种 (如 var. *sinensis*、var. *assamica*、var. *pubilimba* 等) 及大量地方品种和选育品系, 构成了高度复杂的遗传多样性体系, 为茶树品种改良与适应性育种提供了宝贵基础。

叶绿体基因组因其母系遗传、结构保守、适度变异等特点, 已成为解析植物系统发育、物种分化及驯化历史的重要分子工具^[1]。在茶树研究中, 叶绿体基因组不仅有助于厘清近缘种间的亲缘关系, 还可为分子标记开发、种质鉴定及杂交亲本选择提供可靠依据, 对精准育种具有直接应用价值。近年来, 随着高通量测序技术的发展, 多个茶树及相关山茶属物种的叶绿体基因组相继被报道, 揭示其普遍具有典型的四分区环状结构 (约 157 kb), 编码约 130 个功能基因^[2]。尽管整体结构高度保守, 但在反向重复区 (inverted repeat, IR) 边界、简单序列重复 (simple sequence repeat, SSR) 分布及长重复序列等方面仍存在种/品种特异性变异, 这些特征已被成功用于构建系统发育框架和开发品种特异性分子标记。

在众多茶树品种中, 楮叶齐茶树是我国南方茶区广泛推广的国家级无性系良种, 由湖南省农业科学院茶叶研究所选育。该品种具有发芽整齐、持嫩性强、产量高、适制绿茶与红茶等突出农艺优势, 尤其在湖南、江西、广西等地大面积种植, 对提升区域茶叶品质与经济效益发挥了重要作用。然而, 目前尚缺乏关于楮叶齐茶树叶绿体基因组的系统研究, 限制了对其母系遗传背景、进化地位及种质特性的深入理解。

为此, 本研究基于 Illumina 高通量测序技术, 首次完整组装并注释楮叶齐茶树的叶绿体基因组, 系统分析其结构特征、重复序列组成及密码子使用偏好, 并结合山茶属近缘物种开展系统发育重建。研究旨在: (1) 揭示该重要栽培品种的叶绿体基因组独特性; (2) 评估自然选择与突变压力对其密码子偏好的影响; (3) 明确其在茶属系统演化中的分类位置。研究成果将为茶树遗传资源保护、分子辅助育种及叶绿体功能基因工程提供关键数据支撑。

1 材料与方法

1.1 材料

楮叶齐茶树 (*Camellia sinensis* cv. Zhuyeqi) 一芽一叶样本采自湖南省新化县川坳村标准化种植园 (27.73°N, 111.28°E)。样本采集后立即液氮速冻, 全程干冰保存并运输至南京集思慧远生物科技有限公司进行叶绿体基因组测序, 以确保 DNA 完整性和后续分析的可靠性^[3]。楮叶齐茶树叶绿体基因组序列已提交至美国国家生物技术信息中心 (NCBI) (登录号: OL450428.1), 相关数据可供公开检索。

1.2 方法

1.2.1 DNA 提取与测序 采用改良十六烷基三甲基溴化铵 (CTAB) 法提取叶片总 DNA, 通过琼脂糖凝胶电泳及 NanoDrop 分光光度计检测 DNA 纯度与浓度^[4-5]。合格样本经超声波机械打断后, 依次进行末端修复、3'端加 A、连接测序接头, 并通过片段选择构建测序文库。文库质检合格后, 使用 Illumina NovaSeq 6000 平台进行双端测序 (PE150), 原始数据经 fastp (v 0.20.0) 软件进行质量控制, 包括: ①截除测序接头及引物序列; ②剔除平均质量值低于 Q5 (碱基平均错误率高于 31.6%) 的短序列数据片段 (reads); ③排除 N 碱基数>5 的低质量片段 (reads)^[6]。

1.2.2 叶绿体基因组组装 为降低序列组装复杂度, 首先采用 bowtie2 (v2.2.4) 软件的极高灵敏度局部比对 (very-sensitive-local) 模式, 将测序序列比对至自建叶绿体基因组数据库, 筛选出目标样品的叶绿体 DNA^[7-8]。

组装核心采用 SPAdes v 3.10.1 软件进行非参考基因组引导的从头组装 (de novo) 组装, 设置 k-mer 参数为 55、87、121 以优化拼接效果^[8]。针对二代测序技术对重复序列和复杂结构的局限性, 建立 7 步迭代优化流程: (1) 基于 SPAdes 初步拼接获得初始输入 (SEED) 序列; (2) 通过 k-mer 迭代延伸种子序列, 若生成单一连续序列 (contig) 则进入步骤 6; (3) 利用 SSPACE v 2.0 软件连接 contig 构建支架 (scaffold); (4) 使用 GapFiller v 2.1.1 软件填补 scaffold 间隙^[8]; (5) 若仍存在缺口, 则设计引物进行聚合酶链式反应 (PCR) 扩增与测序, 直至获得完整伪基因组 (pseudo genome); (6) 通过测序数据回比校正基因组序列; (7) 依据叶绿体结构特征调整序列坐标, 最终输出完整环状或线状基因组^[9-10]。

1.2.3 叶绿体基因组注释与验证 为提高注释准确性, 采用多软件联合注释策略进行交叉验证。首先, 利用 prodigal v 2.6.3 软件对蛋白编码区 (CDS) 进行预测, 结合 hmmer v 3.1b2 和 aragorn v 1.2.38 软件, 分别注释核糖体核糖核酸 (rRNA) 和转移核糖核酸 (tRNA) 基因^[11]。其次, 通过 NCBI 数据库获取近缘物种叶绿体基因序列, 使用 BLAST v 2.6 软件对组装序列进行同源比对, 获取第二套注释结果。人工校验两套结果间的差异, 修正外显子边界并剔除冗余或错误注释, 最终整合形成高置信度注释集^[11-12]。

1.2.4 叶绿体基因组特征分析 重复序列分析采用 vmatch v 2.3.0 软件结合开源动态脚本语言 (Perl), 参数设置: 最小重复长度 30 bp、海明距离 3, 检测正向重复 (F)、回文重复 (P)、反向重复 (R)、互补重复 (C) 4 种重复类型^[11, 13]。针对简单序列重复 (SSR), 使用 MISA v 1.0 软件筛选叶绿体 SSR 位点, 参数为单碱基重复 ≥ 8 次, 二至六核苷酸重复 ≥ 3 次^[11, 14]。基于近缘物种叶绿体基因组数据, 利用 CGVIEW 软件进行基因组结构可视化及共线性比较, 参数为默认设置^[15]。

1.2.5 密码子偏好性分析 (1) 中性分析。采用密码子第 3 位碱基中 G+C 的平均含量 (GC_3) 与密码子第 1、2 位碱基中 G+C 的平均含量 (GC_{12}) 线性回归模型评估核苷酸突变与自然选择的相对贡献。若回归系数趋近于 1 且决定系数 (R^2) > 0.5 , 则表明密码子偏好性主要受突变压力驱动; 反之则提示自然选择占主导地位^[16-17]。

(2) ENC- GC_3 分析。基于 ENC 与 GC_3 的关联性, 以理论曲线 $ENC = 2 + GC_3 + 29/[GC_3^2 + (1 - GC_3)^2]$ 为参照, 若数据点显著低于理论曲线, 则表明翻译优化或自然选择对密码子使用具有显著调控作用^[16]。

(3) PR2 分析。通过密码子第 3 位碱基中 G 的含量占 G+C 含量的比例 [$G_3/(G_3+C_3)$] 和密码子第 3 位碱基中 A 含量占 A+T 含量比例 [$A_3/(A_3+T_3)$] 的分布偏离中心对称性 ($C_3=G_3$, $A_3=T_3$), 解析突变与选择压力的非对称性效应^[18-19]。

(4) 最优密码子筛选。依据 ENC 值将基因分为高表达组 (前 10%) 与低表达组 (后 10%), 通过 CodonW 软件计算相对同义密码子使用度的变化 $\Delta RSCU = RSCU_{高} - RSCU_{低}$ 。以 $\Delta RSCU \geq 0.08$ 且 $RSCU > 1$ 作为筛选最优密码子条件, 确保其编码

效率与翻译适应性。

1.2.6 叶绿体基因组系统发育分析 采用全基因组做进化树分析, 将环形序列设置相同起点, 物种间序列用 MAFFT 软件 (v7.427, --auto 模式) 进行多序列比对, 利用 RAxML v 8.2.10 (<https://cme.h-its.org/exelixis/software.html>) 在线软件^[20]的 GTRGAMMA 模型, 快速自举分析比对好的数据, bootstrap=1000, 构建最大似然进化树。

2 结果与分析

2.1 楮叶齐茶树的叶绿体基因组结构解析

基于 Illumina NovaSeq 6000 平台的高通量测序显示, 楮叶齐茶树叶绿体基因组测序数据质量较高, Q20 与 Q30 值分别达 96.76% 和 91.29%, 总碱基数为 6.36 Gb, 整体 GC 含量为 38.76%, 数据质量符合后续分析要求。楮叶齐茶树的叶绿体基因组呈典型四分体环状结构, 全长 157 072 bp, 包括大单拷贝区 (LSC, 86 628 bp)、小单拷贝区 (SSC, 18 282 bp) 以及一对反向重复区 (IRa/IRb, 各 26 081 bp)。全基因组 GC 含量为 37.29%, 其中 IR 区 GC 含量 (42.95%) 显著高于 LSC 区 (35.31%) 和 SSC 区 (30.54%), 与多数植物叶绿体基因组 GC 分布规律一致^[21-22]。该结构特征与栎属 (*Quercus*) 等物种的叶绿体基因组相似, 但 IR 区长度差异可能反映物种特异性进化事件^[4]。

2.2 楮叶齐茶树叶绿体基因组的基因组成

在楮叶齐茶树的叶绿体基因组共注释到 135 个功能基因, 其中 PCGs 90 个, tRNA 基因 37 个, rRNA 基因 8 个, 未检测到假基因 (表 1)。其基因组中的基因可分为 4 类: 光合作用基因 (45 个)、自我复制基因 (74 个)、其他基因 (6 个) 和功能未知的基因 (10 个)。共鉴定出 20 个重复基因, 涵盖多个功能类别: 1 个 NADH 脱氢酶亚基基因 (*ndhB*)、4 个与遗传信息传递相关的基因 (*rpl2*、*rpl23*、*rps12*、*rps7*)、4 个核糖体 RNA 基因 (*rrn16*、*rrn23*、*rrn4.5*、*rrn5*)、7 个转移 RNA 基因 (*trnA-UGC*、*trnI-CAU*、*trnI-GAU*、*trnL-CAA*、*trnN-GUU*、*trnR-ACG*、*trnV-GAC*)、4 个功能未知的开放阅读框基因 (*orf42*、*ycf1*、*ycf15*、*ycf2*)。这些重复基因以非编码 RNA 基因为主, 编码基因相对较少。结构特征分析表明, 多数基因不含内含子, 仅部分基因含单/双内含子。其中, 单内含子基因 16 个 (*ndhA*、*ndhB*、*petB*、*petD*、*atpF*、*rpl16*、*rpl2*、

表 1 楮叶齐茶树叶绿体基因组的基因组成
Tab. 1 Gene composition of chloroplast genome of *C. sinensis* cv. Zhuyeqi

基因编码产物/功能 Gene product/Function	基因名称 Gene name	数量 Number
光系统 I 亚基	<i>psaA</i> 、 <i>psaB</i> 、 <i>psaC</i> 、 <i>psaI</i> 、 <i>psaJ</i>	5
光系统 II 亚基	<i>psbA</i> 、 <i>psbB</i> 、 <i>psbC</i> 、 <i>psbD</i> 、 <i>psbE</i> 、 <i>psbF</i> 、 <i>psbH</i> 、 <i>psbI</i> 、 <i>psbJ</i> 、 <i>psbK</i> 、 <i>psbL</i> 、 <i>psbM</i> 、 <i>psbN</i> 、 <i>psbT</i> 、 <i>psbZ</i>	15
NADH 脱氢酶亚基	<i>ndhA</i> *, <i>ndhB</i> *(2)、 <i>ndhC</i> 、 <i>ndhD</i> 、 <i>ndhE</i> 、 <i>ndhF</i> 、 <i>ndhG</i> 、 <i>ndhH</i> 、 <i>ndhI</i> 、 <i>ndhJ</i> 、 <i>ndhK</i>	12
细胞色素 b/f 复合体亚基	<i>petA</i> 、 <i>petB</i> *, <i>petD</i> *, <i>petG</i> 、 <i>petL</i> 、 <i>petN</i>	6
ATP 合酶亚基	<i>atpA</i> 、 <i>atpB</i> 、 <i>atpE</i> 、 <i>atpF</i> *, <i>atpH</i> 、 <i>atpI</i>	6
核糖双加氧酶大亚基	<i>rbcL</i>	1
大核糖体亚基蛋白	<i>rpl14</i> 、 <i>rpl16</i> *, <i>rpl2</i> *(2)、 <i>rpl20</i> 、 <i>rpl22</i> 、 <i>rpl23</i> (2)、 <i>rpl32</i> 、 <i>rpl33</i> 、 <i>rpl36</i>	11
小核糖体亚基蛋白	<i>rps11</i> 、 <i>rps12</i> *(2)、 <i>rps14</i> 、 <i>rps15</i> 、 <i>rps16</i> *, <i>rps18</i> 、 <i>rps19</i> 、 <i>rps2</i> 、 <i>rps3</i> 、 <i>rps4</i> 、 <i>rps7</i> (2)、 <i>rps8</i>	14
RNA 聚合酶亚基	<i>rpoA</i> 、 <i>rpoB</i> 、 <i>rpoC1</i> *, <i>rpoC2</i>	4
核糖体 RNA	<i>rrn16</i> (2)、 <i>rrn23</i> (2)、 <i>rrn4.5</i> (2)、 <i>rrn5</i> (2)	8
转移 RNA	<i>trnA</i> -UGC*(2)、 <i>trnC</i> -GCA、 <i>trnD</i> -GUC、 <i>trnE</i> -UUC、 <i>trnF</i> -GAA、 <i>trnG</i> -UCC、 <i>trnH</i> -GUG、 <i>trnI</i> -CAU(2)、 <i>trnI</i> -GAU*(2)、 <i>trnK</i> -UUU*, <i>trnL</i> -CAA(2)、 <i>trnL</i> -UAA*, <i>trnL</i> -UAG、 <i>trnM</i> -CAU、 <i>trnN</i> -GUU(2)、 <i>trnP</i> -UGG、 <i>trnQ</i> -UUG、 <i>trnR</i> -ACG(2)、 <i>trnR</i> -UCU、 <i>trnS</i> -GCU、 <i>trnS</i> -GGA、 <i>trnS</i> -UGA、 <i>trnT</i> -CGU*, <i>trnT</i> -GGU、 <i>trnT</i> -UGU、 <i>trnV</i> -GAC(2)、 <i>trnV</i> -UAC*, <i>trnW</i> -CCA、 <i>trnY</i> -GUA、 <i>trnY</i> -CAU	37
成熟酶	<i>matK</i>	1
蛋白酶	<i>clpP</i> **	1
包膜蛋白	<i>cemA</i>	1
乙酰辅酶 A 羧化酶	<i>accD</i>	1
c 型细胞色素合成基因	<i>ccsA</i>	1
翻译起始因子	<i>infA</i>	1
保守的假设叶绿体开放阅读框	<i>orf42</i> (2)、 <i>ycf1</i> (2)、 <i>ycf15</i> (2)、 <i>ycf2</i> (2)、 <i>ycf3</i> *, <i>ycf4</i>	10

注：NADH 表示还原型烟酰胺腺嘌呤二核苷酸，ATP 表示三磷酸腺苷，RNA 表示核糖核酸，*表示为 1 个内含子，**表示为 2 个内含子，(2) 表示 2 个拷贝。

Notes: NADH refers to reduced nicotinamide adenine dinucleotide, ATP refers to adenosine triphosphate, and RNA refers to ribonucleic acid. A gene marked with * indicates that it contains one intron, ** indicates two introns, and (2) indicates two copies.

rps12、*rps16*、*rpoC1*、*trnA*-UGC、*trnI*-GAU、*trnK*-UUU、*trnL*-UAA、*trnT*-CGU 和 *trnV*-UAC)，双内含子基因 2 个（*clpP* 和 *ycf3*）。

重复序列分析结果显示，基因组含 40 个长重复序列，正向重复（F）和回文重复（P）各 20 个（表 2）。其中，正向重复序列长度介于 30~82 bp 之间，此单一重复类型特征较 *Medicago truncatula* 等物种更为保守^[23]。回文重复序列长度跨度大（30~26 081 bp），推测 IR 区可能存在大规模结构重组事件，该现象在 *Plantago ovata* 的叶绿体基因组中亦有报道^[24]。

2.3 SSR 与重复序列分析

在楮叶齐茶树叶绿体基因组中共鉴定出 247 个 SSR，涵盖单核苷酸至五核苷酸重复类型（表 3）。单核苷酸重复占据主导地位（63.56%，*n*=157），其中 A/T 重复基序占比高达 96.82%（*n*=152），显著高于 C/G 重复类型，与叶绿体基因组 AT 碱基

偏好性特征一致^[5, 25]。二核苷酸、三核苷酸、四核苷酸及五核苷酸重复数目依次为 4、72、12 和 2，显示出多核苷酸重复类型分布的异质性特征。值得注意的是，本研究中二核苷酸重复比例（1.62%）显著低于多数已报道物种（如 *Clusterbean* 中占 34.88%^[25]），推测楮叶齐茶树的叶绿体基因组可能具有独特的微卫星演化模式。此外，SSR 在非编码区的富集趋势与黄芩等物种相似^[26]，进一步证实叶绿体 SSR 分布与功能区域的选择压力相关。

2.4 楮叶齐茶树叶绿体基因组密码子使用偏好性分析

2.4.1 同义密码子的偏好性分析 本研究对楮叶齐茶树叶绿体基因组的 52 条蛋白质编码序列的同义密码子偏好性进行系统分析。结果显示，GC₃ 均值为 27.65%，显著低于 GC₁（46.77%）和 GC₂（39.51%），呈现 GC₁>GC₂>GC₃ 的递减模式

表 2 楮叶齐茶树叶绿体基因组长重复序列分析
Tab. 2 Analysis of long repeat sequences in chloroplast genome of *C. sinensis* cv. Zhuyeqi

编号 No.	第一部分的重复长度 Repeat length of the first part/bp	第一部分的起始位置 Starting position of the first part/bp	第二部分的重复 长度 Repeat length of the second part/bp	第二部分的起始 位置 Starting position of the second part/bp	匹配方向 Match direction	2 个重复序列间 距离 Distance of this repeat/bp	E 值 E-value
1	26 081	86 629	26 081	130 992	P	0	0
2	82	93 896	82	93 914	F	-3	7.10×10^{-34}
3	82	93 896	82	149 706	P	-3	7.10×10^{-34}
4	82	93 914	82	149 724	P	-3	7.10×10^{-34}
5	82	149 706	82	149 724	F	-3	7.10×10^{-34}
6	70	93 908	70	93 926	F	-2	1.08×10^{-28}
7	70	93 908	70	149 706	P	-2	1.08×10^{-28}
8	70	93 926	70	149 724	P	-2	1.08×10^{-28}
9	60	93 918	60	93 936	F	-1	9.40×10^{-25}
10	60	93 918	60	149 706	P	-1	9.40×10^{-25}
11	60	93 936	60	149 724	P	-1	9.40×10^{-25}
12	60	93 896	60	93 932	F	-3	4.82×10^{-21}
13	60	93 896	60	149 710	P	-3	4.82×10^{-21}
14	60	93 932	60	149 746	P	-3	4.82×10^{-21}
15	60	149 706	60	149 742	F	-3	4.82×10^{-21}
16	52	93 908	52	93 944	F	-2	4.08×10^{-18}
17	50	76 736	50	76 736	P	-2	6.03×10^{-17}
18	42	79 187	42	79 187	P	0	3.59×10^{-16}
19	42	100 979	42	122 910	F	0	3.59×10^{-16}
20	42	122 910	42	142 681	P	0	3.59×10^{-16}
21	42	93 918	42	93 954	F	-1	4.52×10^{-14}
22	42	45 571	42	122 909	F	-3	1.11×10^{-10}
23	42	93 896	42	93 950	F	-3	1.11×10^{-10}
24	42	149 706	42	149 760	F	-3	1.11×10^{-10}
25	39	45 574	39	100 981	F	-2	1.53×10^{-10}
26	39	45 574	39	142 682	P	-2	1.53×10^{-10}
27	38	60 749	38	60 765	F	-3	2.09×10^{-8}
28	35	40 527	35	42 751	F	-3	1.04×10^{-6}
29	34	93 908	34	93 962	F	-2	1.19×10^{-7}
30	32	9 032	32	37 370	F	-3	5.04×10^{-5}
31	31	38 692	31	38 692	P	-3	1.83×10^{-4}
32	30	9 034	30	47 287	P	0	6.02×10^{-9}
33	30	14 188	30	14 188	P	-2	2.36×10^{-5}
34	30	37 372	30	47 287	P	-3	6.60×10^{-4}
35	30	45 586	30	100 993	F	-3	6.60×10^{-4}
36	30	45 586	30	142 679	P	-3	6.60×10^{-4}
37	30	91 460	30	91 502	F	-3	6.60×10^{-4}
38	30	91 460	30	152 170	P	-3	6.60×10^{-4}
39	30	91 502	30	152 212	P	-3	6.60×10^{-4}
40	30	152 170	30	152 212	F	-3	6.60×10^{-4}

注：E 值为评估 2 个序列比对结果统计学显著性的指标，当 E 值小于 1.0×10^{-10} 时，认为比对结果具有显著性意义。
Notes: The E-value is a metric used to assess the statistical significance of sequence alignment results. When the E-value is less than 1.0×10^{-10} , the alignment result is considered statistically significant.

表 3 楮叶齐茶树叶绿体基因组中 SSR 的类型及分布
Tab.3 Types and distribution of SSRs in chloroplast genome of *C. sinensis* cv. Zhuyeqi

重复单元碱基类型 Repetitive unit base type	重复次数 Repeat number of repeat units																总 SSR 数量 Total number of SSRs
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
A/T	—	—	—	—	—	59	39	17	13	11	6	3	1	2	1	152	
C/G	—	—	—	—	—	5	—	—	—	—	—	—	—	—	—	5	
AT/AT	—	—	4	—	—	—	—	—	—	—	—	—	—	—	—	4	
AAT/AAG/AAC/ACC/AGA/AGC/ATA/ATC/ ATG/ATT	25	—	—	—	—	—	—	—	—	—	—	—	—	—	—	25	
TAA/TCA/TCT/TGA/TGC/TTA/TTC/TTG	28	—	—	—	—	—	—	—	—	—	—	—	—	—	—	28	
CAA/CAG/CTC/CTG/CTT	9	—	—	—	—	—	—	—	—	—	—	—	—	—	—	9	
GAA/GAG/GAT/GCA/GCT/GGA/GGT	10	—	—	—	—	—	—	—	—	—	—	—	—	—	—	10	
AAAT/AATA/AGAT/ATAG	5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	5	
CCCT	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	
GAAA/GAGG/GTCT	3	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3	
TCTA/TCTT/TTTC	3	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3	
AAAAAG/CTTTTT	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	

（图 1），该趋势与叶绿体基因组偏好 A/U 结尾密码子的典型特征相吻合^[17, 27]。有效密码子数（ENC）分布范围为 34.24~53.16（均值为 46.63），其中 51 条基因的 ENC>35，仅 *rps14* 基因（ENC=34.24）显示较弱偏好性，表明基因组整体密码子使用偏好性强度较弱，这一现象与蕤蕤苜蓿（ENC=37.11~51.91）及杧果（ENC 均值>45）等物种的叶绿体基因组特征相似^[18, 23]。GC₃ 与

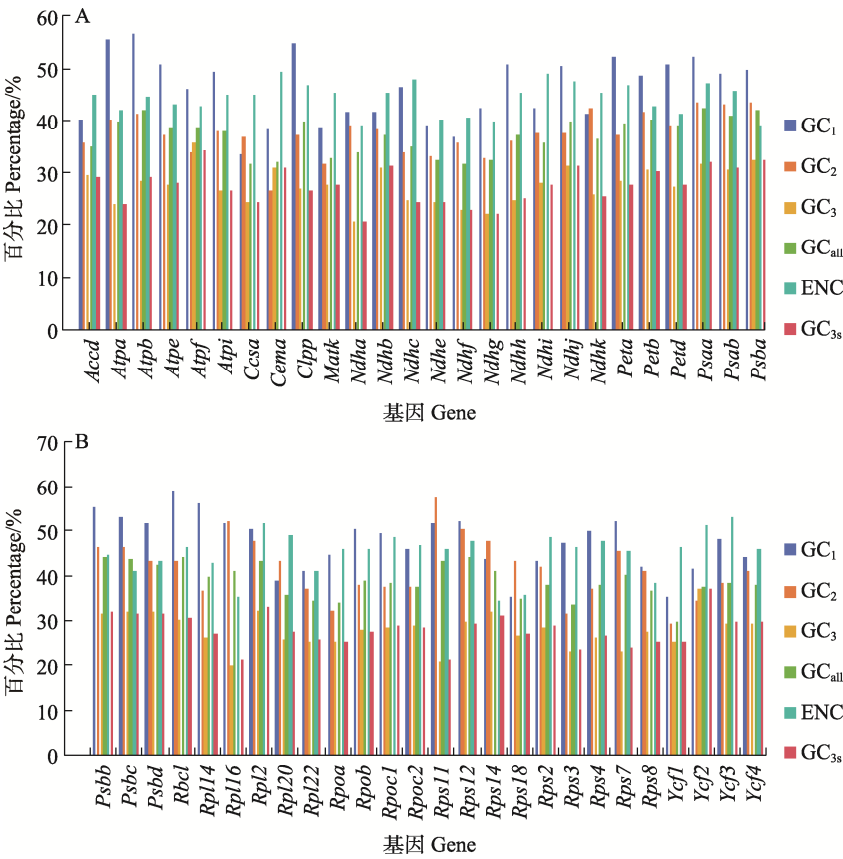


图 1 GC 含量与 ENC 值在楮叶齐茶树叶绿体基因组间的对比分析
Fig. 1 Comparative analysis of GC content and ENC values in chloroplast genome of *C. sinensis* cv. Zhuyeqi

ENC 值的关联性可能暗示突变压力与自然选择共同作用于密码子偏好性的形成机制^[17, 28]，而 *rps14* 基因的特殊性提示其可能受到额外的进化约束^[29]。

2.4.2 中性分析 (GC₃-GC₁₂ 分析) 楮叶齐茶树叶绿体基因组密码子的 GC₃ 分布区间为 19.85%~37.02% (表 4), 显著低于 GC₁₂ (32.13%~54.68%)。中性绘图分析显示, GC₁₂ 与 GC₃ 的回归斜率为

0.134 ($R^2_{adj}=0.01$), 表明二者无显著线性相关性^[17]。根据斜率推算, 自然选择对密码子偏好性的贡献率达 86.6%, 显著高于突变压力 (13.4%), 说明自然选择是该物种叶绿体基因组密码子偏好性形成的主要驱动力^[28]。该结果与茶科植物叶绿体基因组的进化特征一致, 即密码子偏好性主要由选择压力主导, 同时伴随微弱突变效应^[30]。

表 4 楮叶齐茶树的叶绿体基因组密码子分析
Tab. 4 Analysis in the codons of chloroplast genome of *C. sinensis* cv. Zhuyeqi

基因编码产物 Gene Product	基因 Gene	GC ₁₂ /%	GC ₃ /%	ENC	G ₃ /(G ₃ +C ₃)	A ₃ /(A ₃ +T ₃)	基因编码产物 Gene Product	基因 Gene	GC ₁₂ /%	GC ₃ /%	ENC	G ₃ /(G ₃ +C ₃)	A ₃ /(A ₃ +T ₃)
乙酰辅酶 A 羧化酶	<i>accD</i>	37.94	29.57	45.05	0.55	0.34	光系统 II 亚基	<i>psbB</i>	50.59	31.24	44.42	0.53	0.39
	<i>atpA</i>	47.84	24.02	41.86	0.55	0.48		<i>psbC</i>	49.79	31.65	41.28	0.46	0.44
	<i>atpB</i>	49.00	28.46	44.65	0.53	0.46		<i>psbD</i>	47.46	31.64	43.38	0.43	0.37
ATP 合酶亚基	<i>atpE</i>	44.03	27.61	43.11	0.62	0.44	核糖双加氧酶大亚基	<i>rbcL</i>	51.16	30.04	46.29	0.50	0.42
	<i>atpF</i>	40.00	35.68	42.79	0.62	0.50		<i>rpl14</i>	46.35	26.02	42.84	0.53	0.49
	<i>atpI</i>	43.55	26.61	44.95	0.44	0.43		<i>rpl16</i>	51.84	19.85	35.39	0.63	0.59
c 型细胞色素合成基因	<i>ccsA</i>	35.25	24.22	44.81	0.50	0.46	大核糖体亚基蛋白	<i>rpl2</i>	48.91	32.36	51.74	0.51	0.49
	<i>cemA</i>	32.54	31.03	49.42	0.47	0.43		<i>rpl20</i>	41.10	25.42	48.80	0.63	0.49
	<i>clpP</i>	46.05	26.98	46.69	0.53	0.48		<i>rpl22</i>	39.11	25.00	40.88	0.56	0.54
包膜膜蛋白	<i>matK</i>	35.20	27.60	45.45	0.54	0.43	RNA 聚合酶亚基	<i>rpoA</i>	38.39	25.00	46.13	0.56	0.48
	<i>ndhA</i>	40.39	20.60	39.22	0.49	0.49		<i>rpoB</i>	44.26	27.73	46.11	0.64	0.49
	<i>ndhB</i>	40.02	31.12	45.46	0.42	0.46		<i>rpoC1</i>	43.64	28.36	48.58	0.55	0.48
蛋白酶	<i>ndhC</i>	40.08	24.79	47.89	0.67	0.41		<i>rpoC2</i>	41.70	28.57	46.83	0.53	0.48
	<i>ndhE</i>	36.28	24.51	40.20	0.56	0.35		<i>rps11</i>	54.68	20.86	46.09	0.59	0.48
	<i>ndhF</i>	36.32	22.83	40.69	0.6	0.39		<i>rps12</i>	51.26	29.41	47.53	0.46	0.50
成熟酶	<i>ndhG</i>	37.57	22.03	39.83	0.54	0.41		<i>rps14</i>	45.54	31.68	34.24	0.72	0.58
	<i>ndhH</i>	43.53	24.62	45.37	0.61	0.49		<i>rps18</i>	39.22	26.47	35.74	0.67	0.47
	<i>ndhI</i>	39.88	27.98	48.87	0.49	0.44	小核糖体亚基蛋白	<i>rps2</i>	42.62	28.27	48.73	0.58	0.45
NADH 脱氢酶亚基	<i>ndhJ</i>	44.03	31.45	47.49	0.60	0.40		<i>rps3</i>	39.27	22.83	46.43	0.38	0.60
	<i>ndhK</i>	41.87	25.80	45.14	0.55	0.43		<i>rps4</i>	43.57	26.24	47.63	0.49	0.50
	<i>petA</i>	44.71	28.35	46.92	0.56	0.43		<i>rps7</i>	48.72	23.08	45.27	0.56	0.58
	<i>petB</i>	45.14	30.56	42.57	0.59	0.40		<i>rps8</i>	41.55	27.21	38.44	0.49	0.52
	<i>petD</i>	45.03	27.33	41.26	0.55	0.50	保守的假定叶绿体 开放阅读框	<i>ycf1</i>	32.13	25.03	46.23	0.54	0.52
细胞色素 b/f 复合体 亚基	<i>psaA</i>	47.87	31.82	46.96	0.49	0.42		<i>ycf2</i>	38.00	37.02	51.22	0.54	0.43
	<i>psaB</i>	45.92	30.61	45.67	0.56	0.38		<i>ycf3</i>	43.20	28.99	53.16	0.55	0.47
	<i>psbA</i>	46.61	32.49	39.09	0.33	0.33		<i>ycf4</i>	42.70	29.19	46.09	0.54	0.41

2.4.3 GC₃-ENC 分析 根据 GC₃-ENC 分析结果显示, 本研究筛选的 52 个基因的 ENC 值主要分布于 40~50 区间 (表 4), 其中 83.7% 的基因实测 ENC 值低于中性突变理论预期值, 表明密码子偏好性形成机制中自然选择压力占据主导作用^[18, 28]。GC₃ 呈显著窄幅分布特征 (19.85%~37.02%), 该数值域显著低于中性突变模型预测的 GC₃ 波动范围, 进一步证实了非随机进化力量对密码子使

用模式的塑造作用^[31]。结合 GC₃-ENC 双参数分析, 基因组的密码子偏好性主要受翻译选择等定向进化压力的影响, 而非单纯由碱基突变偏倚驱动^[18-19]。

2.4.4 PR2 分析 PR2 分析表明, 楮叶齐茶树叶绿体基因组的密码子第 3 位碱基使用偏好性分布呈现显著异质性。G₃/(G₃+C₃) 和 A₃/(A₃+T₃) 比值显示, 系统偏向 G₃> C₃ 及 T₃> A₃ 的非对称分布 (表

4), 其中光合系统核心基因 *psbA* 的 $G_3/(G_3+C_3)$ (0.33) 和 $A_3/(A_3+T_3)$ (0.33) 均处于最小值, 核糖体蛋白编码基因 *rps14* 的 $G_3/(G_3+C_3)$ 达到峰值 (0.72), *rps3* 的 $A_3/(A_3+T_3)$ 则为 0.6。该分布模式显著偏离突变压力作用下的理论平衡值 ($C_3=G_3$, $A_3=T_3$), 推测自然选择对碱基使用频率的调控占据主导地位^[17-18, 32]。值得注意的是, *psbA* 基因的极端偏倚特征与其他研究中对光合核心基因受强选择压力作用的结果具有一致性^[33, 34], 而核糖体蛋白基因的变异幅度则可能反映其功能分化

的选择强度差异^[32]。

2.4.5 最优密码子确定 本研究在楮叶齐茶叶叶绿体基因组中共鉴定出 20 个最优密码子 (表 5)。其中, 以 U/A 结尾的密码子占比达 90% (如 GCA、GCU、CAA 等), 表明其密码子第 3 位碱基存在显著的 A/U 偏好性。该特征与多数植物叶绿体基因组中密码子偏好性受自然选择主导的进化模式一致^[17, 28]。值得注意的是, 最优密码子的碱基组成特征可能反映出基因组在翻译效率优化与环境适应性间的平衡机制^[23]。

表 5 楮叶齐茶叶叶绿体基因最优密码子
Tab. 5 Optimal codons in chloroplast genome of *Camellia sinensis* cv. Zhuyeqi

氨基酸 Amino acid	密码子 Codon	基因组 Genome		高表达基因 High expressed genes		低表达基因 Low expressed genes		Δ RSCU
		数量 Number	RSCU	数量 Number	RSCU	数量 Number	RSCU	
丙氨酸 (Ala)	GCU***	526	1.83	37	2.51	41	1.67	0.84
	GCA*	332	1.16	17	1.15	25	1.02	0.13
半胱氨酸 (Cys)	UGU***	173	1.52	6	2.00	30	1.43	0.57
谷氨酸 (Glu)	GAA*	859	1.53	33	1.53	120	1.34	0.19
甘氨酸 (Gly)	GGU***	465	1.29	31	1.91	39	1.04	0.87
赖氨酸 (Lys)	AAA**	850	1.53	31	1.72	127	1.30	0.42
亮氨酸 (Leu)	UUA**	707	1.98	18	1.59	65	1.22	0.37
	CUU*	445	1.25	18	1.59	76	1.42	0.17
	CUA*	278	0.78	13	1.15	49	0.92	0.23
	UUG*	440	1.23	18	1.59	75	1.40	0.19
脯氨酸 (Pro)	CCU**	361	1.66	15	1.76	43	1.39	0.37
谷酰胺酸 (Gln)	CAA**	581	1.53	19	1.73	82	1.43	0.30
精氨酸 (Arg)	CGA**	296	1.42	26	1.84	46	1.37	0.47
	CGU**	290	1.39	20	1.41	33	0.99	0.42
丝氨酸 (Ser)	UCU**	471	1.79	22	1.97	79	1.60	0.37
	AGU**	344	1.32	17	1.52	52	1.05	0.47
苏氨酸 (Thr)	ACC**	192	0.74	11	1.07	24	0.71	0.36
	ACU***	430	1.66	20	1.95	44	1.30	0.65
缬氨酸 (Val)	GUU***	422	1.48	19	1.95	40	1.36	0.59
	GUA***	437	1.53	18	1.85	37	1.25	0.62

注: *表示 $0.08<\Delta$ RSCU <0.30 ; **表示 $0.30\leq\Delta$ RSCU <0.50 ; ***表示 Δ RSCU >0.50 。
Notes: * indicates $0.08<\Delta$ RSCU <0.30 ; ** indicates $0.30\leq\Delta$ RSCU <0.50 ; and *** indicates Δ RSCU >0.50 .

2.5 系统发育分析

基于叶绿体基因组序列的系统发育分析表明 (图 2), 楮叶齐茶树与多个茶树栽培类型[包括普洱茶 (*C. sinensis* var. *assamica*) 和白毛茶 (*C. sinensis* var. *pubilimba*)]共同聚类于一个支持率极高 (Bootstrap 值=100%) 的单系分支, 证实其与

这些栽培类群存在密切的母系遗传关联^[35-36]。该结果印证了楮叶齐属于茶树 (*C. sinensis*) 种内广泛栽培的类群, 其叶绿体基因组在进化上高度保守, 与主要茶用类型共享近期共同祖先。

在更广泛的山茶属系统发育框架中, 茶树及其变种 (含楮叶齐) 与大理茶 (*C. taliensis*)、秃房茶

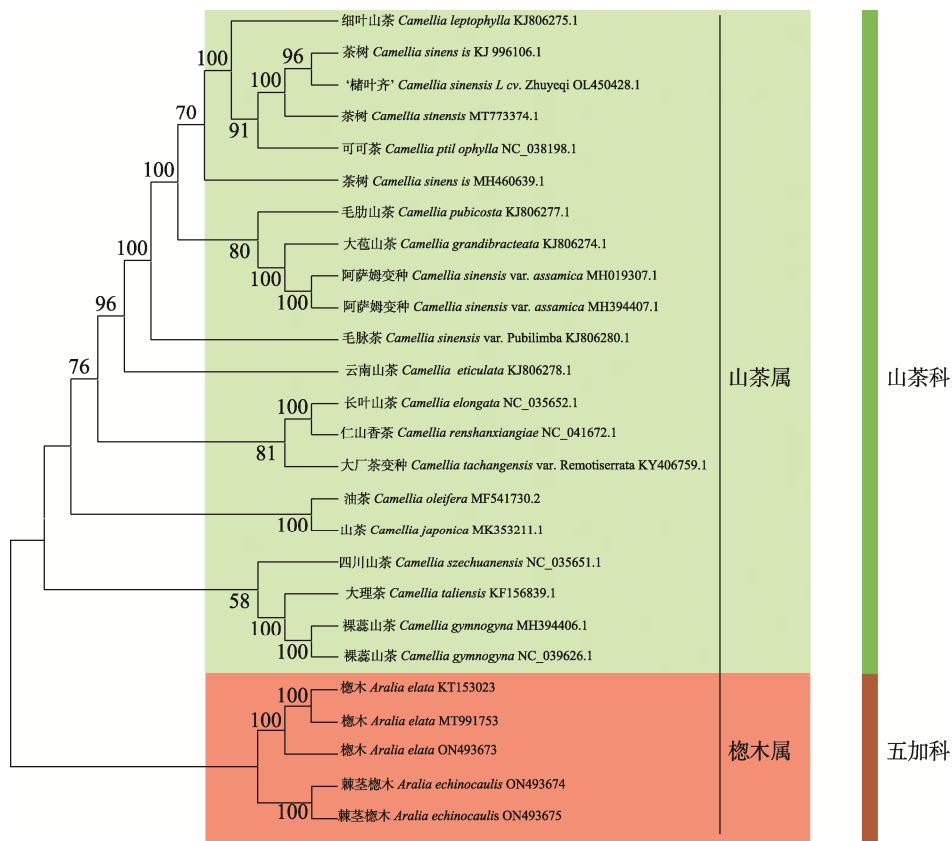


图 2 楮叶齐茶树的叶绿体系统进化分析

Fig. 2 Phylogenetic analysis of chloroplast genome in *C. sinensis* cv. Zhuyeqi

(*C. gymnogyna*) 等野生近缘种构成核心茶组分支。其中，大理茶与栽培茶树类群亲缘关系尤为接近，暗示其可能作为栽培茶树的野生祖先或存在基因交流^[35-36]，而疏齿大厂茶 (*C. tachangensis* var. *remotiserrata*) 等类群则代表了茶组内较早分化的谱系^[35]。非茶用山茶物种[如油茶 (*C. oleifera*)、山茶花 (*C. japonica*)]位于系统发育树远端，反映了种间显著的遗传距离^[36]。此外，以五加科 (*Araliaceae*) 楸木属 (*Aralia*) 为外群，其与山茶属形成显著分离 (Bootstrap 值= 100%)，验证了跨科系统发育距离的可靠性。

3 讨论

本研究完成了茶树品种楮叶齐的叶绿体基因组的全面解析。通过与近缘物种的比较，发现其基因组在保持核心结构保守性的同时，在重复序列和密码子使用模式上展现出独特的进化特征。

3.1 基因组结构保守性与 IR 区稳定性

楮叶齐茶树叶绿体基因组呈现典型的四分体环状结构，其大小、GC 含量及基因数量与已报道

的茶树 (如龙井 43) 及山茶科其他物种高度一致，印证了该科物种叶绿体基因组的整体保守性^[37-38]。值得注意的是，IR 区的 GC 含量显著高于 LSC 和 SSC 区，这一现象在槭属、木姜子属等植物中普遍存在^[39]。这主要归因于 IR 区富含 GC 含量较高的 rRNA 基因，同时，其特殊的结构可能通过抑制突变或进行偏向性修复来维持其序列的稳定性，这对于保障核糖体功能的稳定至关重要。

3.2 重复序列特征及其进化意义

在重复序列方面，楮叶齐茶树叶绿体基因组中检测到的大量 A/T 富集的 SSR 位点，与信阳 10 号、大面白等茶树品种的报道高度吻合^[40-42]。这种跨物种的一致性表明，A/T 碱基的富集并非随机，而可能是叶绿体基因组在进化过程中形成的一种适应性策略。A/T 碱基对所需的能量较少，可能有助于基因组的快速复制或特定环境压力下的适应性演化。此外，本研究观察到的 IR 区长度与已报道茶树的高度相似，与山茶科 IR 区边界相对稳定^[43]的观点一致，但其中回文重复序列在长度上的巨大变异，表明该区域在历史上可能发生过小规模的重组或滑动事件。

3.3 密码子使用偏性受自然选择主导

本研究最关键的发现之一是揭示了自然选择在塑造楮叶齐茶树叶绿体基因组密码子使用偏好性中的压倒性作用。中性绘图显示,其自然选择的贡献率高达 86.6%,显著高于突变压力,这一模式在森林茶园古茶树、米楮及籽粒苋等物种中也得到验证^[16, 42, 44]。这表明,叶绿体基因的密码子优化并非由随机的碱基突变主导,而是为了适应翻译效率的最大化或维持特定蛋白质结构的稳定性,从而受到强烈的纯化选择。最优密码子绝大多数以 A/U 结尾的特征,与山茶科叶绿体基因组的整体趋势相符。但茶树核基因组中的 *CsNRT1.1* 基因却偏好以 G/C 结尾^[45],这一鲜明对比揭示了不同细胞器基因组(叶绿体与细胞核)可能处于截然不同的进化压力之下,叶绿体基因更倾向于通过优化 A/U 结尾的密码子来协调其基因表达系统。

4 结论

本研究成功组装并注释了茶树品种楮叶齐完整的叶绿体基因组(157 072 bp),其具备典型的四分体结构和 135 个功能基因,IR 区表现出最高的 GC 含量(42.96%)。基因组中共鉴定出 40 个长重复序列和 247 个 SSR 位点,后者以 A/T 富集为主,这为开发高多态性的叶绿体分子标记提供了丰富的资源。密码子偏好性分析表明,其偏好性较弱,且主要受自然选择(86.6%)而非突变压力(13.4%)驱动,最优密码子普遍以 A/U 结尾,反映了叶绿体基因在进化过程中对翻译效率的优化。系统发育分析将楮叶齐茶树与普洱茶、白毛茶等栽培类型紧密聚类于一个高支持率的单系分支,明确了其系统发育地位,为其在茶树育种和进化研究中的应用提供了坚实的母系遗传学依据。

参考文献

- [1] 杨雨青, 谭娟, 汪芳, 彭顺利, 陈婕, 谭明燕, 吕美艳, 周富裕, 刘声传. 茶树叶体基因组的研究与应用进展[J]. 生物技术通报, 2024, 40(2): 20-30.
YANG Y Q, TAN J, WANG F, PENG S L, CHEN J, TAN M Y, LYU M Y, ZHOU F Y, LIU S C. Research and application progress in chloroplast genome of tea plant (*Camellia sinensis*)[J]. Biotechnology Bulletin, 2024, 40(2): 20-30. (in Chinese)
- [2] 曾文娟, 周伟, 何诗恬, 贺宁, 龚意辉, 陈致印. 茶树品种白毫早叶绿体基因组结构特征及其密码子偏好性分析[J]. 江苏农业学报, 2025, 41(7): 1398-1411.
ZENG W J, ZHOU W, HE S T, HE N, GONG Y H, CHEN Z Y. Structural characteristics and codon usage bias analysis of the chloroplast genome in the tea cultivar *Camellia sinensis* cv. Baihaozao[J]. Jiangsu Journal of Agricultural Sciences, 2025, 41(7): 1398-1411. (in Chinese)
- [3] LI L, HU Y F, HE M, ZHANG B, WU W, CAI P M, HUO D, HONG Y C. Comparative chloroplast genomes: insights into the evolution of the chloroplast genome of *Camellia sinensis* and the phylogeny of *Camellia*[J]. BMC Genomics, 2021, 22(1): 138-160.
- [4] LI X, LI Y F, ZANG M Y, LI M Z, FANG Y M. Complete chloroplast genome sequence and phylogenetic analysis of *Quercus acutissima*[J]. International Journal of Molecular Sciences, 2018, 19: 2443.
- [5] SOMARATNE Y, GUAN D L, WANG W Q, ZHAO L, XU S Q. The complete chloroplast genomes of two *Lespedeza* species: insights into codon usage bias, RNA editing sites, and phylogenetic relationships in Desmodieae (Fabaceae Papilionoideae)[J]. Plants, 2020, 9: 51.
- [6] YU Y F, OUYANG Z, GUO J, ZENG W, ZHAO Y J, HUANG L Q. Complete chloroplast genome sequence of *Erigeron breviscapus* and characterization of chloroplast regulatory elements[J]. Frontiers in Plant Science, 2021, 12: 758290.
- [7] LI L, WANG W Y, ZHANG G Q, WU K L, FANG L, LI Z M, LIU Z J, ZENG S J. Comparative analyses and phylogenetic relationships of thirteen *Pholidota* species (Orchidaceae) inferred from complete chloroplast genomes[J]. BMC Plant Biology, 2023, 23: 269.
- [8] GRUENSTAEUDL M, GERSCHLER N, BORSCH T. Bioinformatic workflows for generating complete plastid genome sequences: an example from *Cabomba* (Cabombaceae) in the context of the phylogenomic analysis of the water-lily clade[J]. Life, 2018, 8: 25.
- [9] TWYFORD A D, NESS R W. Strategies for complete plastid genome sequencing[J]. Molecular Ecology Resources, 2017, 17: 858-868.
- [10] TOMASELLO S, MANZO E, KARBSTEIN K. Comparative plastome assembly of the yellow ironweed (*Verbesina alternifolia*) using Nanopore and Illumina reads[J]. Frontiers in Plant Science, 2024, 15: 1429494.
- [11] 陆奇丰, 黄至欢, 骆文华. 极小种群濒危植物广西火桐、丹霞梧桐的叶绿体基因组特征[J]. 生物多样性, 2021, 29(5): 586-595.
LU Q F, HUANG Z H, LUO W H. Characterization of complete chloroplast genome in *Firmiana kwangsiensis* and *F. danxiaensis* with extremely small populations[J]. Biodiver-

- sity Science, 2021, 29(5): 586-595. (in Chinese)
- [12] QU X J, ZOU D, ZHANG R Y, STULL G W, YI T S. Progress, challenge and prospect of plant plastome annotation[J]. *Frontiers in Plant Science*, 2023, 14: 1166140.
- [13] GAO R, WANG W Z, HUANG Q Y, FAN R F, WANG X, FENG P, ZHAO G M, BIAN S, REN H L, CHANG Y. Complete chloroplast genome sequence of *Dryopteris fragrans* (L.) Schott and the repeat structures against the thermal environment[J]. *Scientific Reports*, 2018, 8: 16635.
- [14] 阳秀玫, 陈亮明, 禹霖, 柏文富, 李建挥, 吴思政, 李柏海, 熊颖, 聂东伶, 严佳文. 华中樱叶绿体基因组特征与分子标记[J]. 中南林业科技大学学报, 2024, 44(1): 175-185.
YANG X M, CHEN L M, YU L, BAI W F, LI J H, WU S Z, LI B H, XIONG Y, NIE D L, YAN J W. Chloroplast genomic characteristics and molecular markers of *Prunus conradinae*[J]. *Journal of Central South University of Forestry & Technology*, 2024, 44(1): 175-185. (in Chinese)
- [15] GICHIRA A W, LI Z Z, SAINA J K, LONG Z C, HU G W, GITURU R W, WANG Q F, CHEN J M. The complete chloroplast genome sequence of an endemic monotypic genus *Hagenia* (Rosaceae): structural comparative analysis, gene content and microsatellite detection[J]. *Peer J*, 2017, 5: e2846.
- [16] 冯瑞云, 梅超, 王慧杰, 雷梦林, 田翔, 侯雅静. 籽粒苋叶绿体基因组密码子偏好性分析[J]. 中国草地学报, 2019, 41(4): 8-16.
FENG R Y, MEI C, WANG H J, LEI M L, TIAN X, HOU Y J. Analysis of codon usage in the chloroplast genome of grain amaranth (*Amaranthus hypochondriacus* L.)[J]. *Chinese Journal of Grassland*, 2019, 41(4): 8-16. (in Chinese)
- [17] 赵月梅, 杨贵清, 徐其碧, 尹鑫, 徐林, 丁波. 建始槭叶绿体基因组密码子使用偏性分析[J]. 福建农林大学学报(自然科学版), 2022, 51(6): 792-800.
ZHAO Y M, YANG G Q, XU Q B, YIN X, XU L, DING B. Codon usage bias of chloroplast genome in *Acer henryi*[J]. *Journal of Fujian Agriculture and Forestry University (Natural Science Edition)*, 2022, 51(6): 792-800. (in Chinese)
- [18] 辛雅萱, 黎若竹, 李鑫, 陈丽琼, 唐军荣, 屈亚亚, 杨琳懿, 辛培尧, 李云芳. 杧果叶绿体基因组密码子使用偏好性分析[J]. 中南林业科技大学学报, 2021, 41(9): 148-156, 165.
XIN Y X, LI R Z, LI X, CHEN L Q, TANG J R, QU Y Y, YANG L Y, XIN P Y, LI Y F. Analysis on codon usage bias of chloroplast genome in *Mangifera indica*[J]. *Journal of Central South University of Forestry & Technology*, 2021, 41(9): 148-156, 165. (in Chinese)
- [19] 王茹, 赵仁菲, 卢新雄, 尹广鹏, 邓志军, 罗永坚. 珙桐叶绿体基因组同义密码子使用偏好性分析[J]. 北京林业大学学报, 2024, 46(3): 8-17.
WANG R, ZHAO R F, LU X X, YIN G K, DENG Z J, LUO Y J. Analysis of synonymous codon usage bias in the chloroplast genome of *Davidia involucrata*[J]. *Journal of Beijing Forestry University*, 2024, 46(3): 8-17. (in Chinese)
- [20] STAMATAKIS A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies[J]. *Bioinformatics*, 2014, 30(9): 1312-1313.
- [21] SONG W C, CHEN Z M, SHI W B, HAN W Q, FENG Q, SHI C, ENGEL M S, WANG S. Comparative analysis of complete chloroplast genomes of nine species of *Litsea* (Lauraceae): hypervariable regions, positive selection, and phylogenetic relationships[J]. *Genes*, 2022, 13: 1550.
- [22] 钟浩天, 蒋莉萍, 江宇慧, 胡志刚, 余坤, 刘义飞, 森林. 菹荇叶绿体基因组的组装分析与分子标记研究[J]. 世界科学技术-中医药现代化, 2022, 42(4): 1410-1421.
ZHONG H T, JIANG L P, JIANG Y H, HU Z G, YU K, LIU Y F, SEN L. Assembly and molecular markers of *Smilax china* chloroplast genome[J]. *Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology*, 2022, 42(4): 1410-1421. (in Chinese)
- [23] 杨国锋, 苏昆龙, 赵怡然, 宋智斌, 孙娟. 蒺藜苜蓿叶绿体密码子偏好性分析[J]. 草业学报, 2015, 24(12): 171-179.
YANG G F, SU K L, ZHAO Y R, SONG Z B, SUN J. Analysis of codon usage in the chloroplast genome of *Medicago truncatula*[J]. *Acta Prataculturae Sinica*, 2015, 24(12): 171-179. (in Chinese)
- [24] ASAF S, KHAN A L, LUBNA, KHAN A, KHAN A, KHAN G, LEE I J, AL-HARRASI A. Expanded inverted repeat region with large scale inversion in the first complete plastid genome sequence of *Plantago ovata*[J]. *Scientific Reports*, 2020, 10: 3881.
- [25] KAILA T, CHADUVLA P K, RAWAL H C, SAXENA S, TYAGI A, MITHRA S V A, SOLANKE A U, KALIA P, SHARMA T R, SINGH N K, GAIKWAD K. Chloroplast genome sequence of clusterbean (*Cyamopsis tetragonoloba* L.): genome structure and comparative analysis[J]. *Genes*, 2017, 8: 212.
- [26] 王文斌, 于欢, 邱相坡. 黄芩叶绿体基因组重复序列及密码子偏好性分析[J]. 分子植物育种, 2018, 16(8): 2445-2452.
WANG W B, YU H, QIU X P. Analysis of repeat sequence and codon bias of chloroplast genome in *Scutellaria baicalensis*[J]. *Molecular Plant Breeding*, 2018, 16(8): 2445-2452. (in Chinese)
- [27] CHEN S, ZHANG H, WANG X, ZHANG Y H, RUAN G H, MA J. Analysis of codon usage bias in the chloroplast genome of *Helianthus annuus* J-01[J]. *IOP Conference Series: Earth and Environmental Science*, 2021, 792: 012009.
- [28] 徐荣, 宋文政, 乔振升, 张晓琳, 于国栋, 何承忠, 纵丹. 富民枳叶绿体基因组密码子偏好性分析[J]. 中草药, 2024,

- 55(20): 7011-7093.
- XU R, SONG W Z, QIAO Z S, ZHANG X L, YU G D, HE C Z, ZONG D. Codon preference analysis of chloroplast genome in medicinal plant *Poncirus polyandra*[J]. Chinese Traditional and Herbal Drugs, 2024, 55(20): 7011-7093. (in Chinese)
- [29] YANG X, LUO X N, CAI X P. Analysis of codon usage pattern in *Taenia saginata* based on a transcriptome dataset[J]. Parasites & Vectors, 2014, 7(1): 527.
- [30] WANG Z J, CAI Q W, WANG Y, LI M H, WANG C C, WANG Z X, JIAO C Y, XU C C, WANG H Y, ZHANG Z L. Comparative analysis of codon bias in the chloroplast genomes of Theaceae species[J]. Frontiers in Genetics, 2022, 13: 824610.
- [31] 柳燕杰, 田旭平, 李倩. 美国红栲叶绿体基因组密码子偏好性分析[J]. 江苏农业科学, 2020, 48(15): 83-88.
- LIU Y J, TIAN X P, LI Q. Analysis of codon usage bias in the chloroplast genome of *Fraxinus americana*[J]. Jiangsu Agricultural Sciences, 2020, 48(15): 83-88. (in Chinese)
- [32] YANG Y, WANG X L, SHI Z J. Comparative study on codon usage patterns across chloroplast genomes of eighteen *Taraxacum* species[J]. Horticulturae, 2024, 10(5): 492.
- [33] 晁岳恩, 吴政卿, 杨会民, 何宁, 杨攀. 11 种植物 *psbA* 基因的密码子偏好性及聚类分析[J]. 核农学报, 2011, 25(5): 927-932.
- CHAO Y E, WU Z Q, YANG H M, HE N, YANG P. Cluster analysis and codon usage bias studies on *psbA* genes from 11 plant species[J]. Journal of Nuclear Agricultural Sciences, 2011, 25(5): 927-932. (in Chinese)
- [34] MORTON B R. Context-dependent mutation dynamics, not selection, explains the codon usage bias of most angiosperm chloroplast genes[J]. Journal of Molecular Evolution, 2022, 90: 17-29.
- [35] HAO W J, MA J Q, MA C L, JIN J Q, CHEN L. The complete chloroplast genome sequence of *Camellia tachangensis* F. C. Zhang (Theaceae)[J]. Mitochondrial DNA Part B, 2019, 4(2): 3344-3345.
- [36] PENG J, ZHAO Y L, DONG M, LIU S Q, HU Z Y, ZHONG X F, XU Z G. Exploring the evolutionary characteristics between cultivated tea and its wild relatives using complete chloroplast genomes[J]. BMC Ecology and Evolution, 2021, 21(1): 71-88.
- [37] WU Z H, LIAO R, YANG T G, DONG X, LAN D Q, QIN R, LIU H. Analysis of six chloroplast genomes provides insight into the evolution of *Chrysosplenium* (Saxifragaceae)[J]. BMC Genomics, 2020, 21: 621.
- [38] 叶晓倩, 赵忠辉, 朱全武, 王营营, 林张翔, 叶楚玉, 樊龙江. 茶树“龙井 43”叶绿体基因组测序及其系统进化[J]. 浙江大学学报(农业与生命科学版), 2014, 40(4): 404-412.
- YE X Q, ZHAO Z H, ZHU Q W, WANG Y Y, LIN Z X, YE C Y, FAN L J. Entire chloroplast genome sequence of tea (*Camellia sinensis* cv. Longjing 43): a molecular phylogenetic analysis[J]. Journal of Zhejiang University (Agriculture and Life Sciences), 2014, 40(4): 404-412. (in Chinese)
- [39] AFZAL N. Comparative chloroplast genomics of Aceraceae species: insights into sequence variations and phylogenetic evolution[D]. Xi'an: Northwest University, 2020. (in Chinese)
- [40] 闫明慧, 刘柯, 王满, 吕颖, 张倩. 信阳 10 号叶绿体基因组及其系统进化[J]. 茶叶科学, 2021, 41(6): 777-788.
- YAN M H, LIU K, WANG M, LYU Y, ZHANG Q. Complete chloroplast genome of *Camellia sinensis* cv. Xinyang 10 and its phylogenetic evolution[J]. Journal of Tea Science, 2021, 41(6): 777-788. (in Chinese)
- [41] 尹明华, 张嘉, 乐芸, 何凡凡, 黄添慧, 张牧彤. 茶树大面白叶绿体基因组特征、密码子偏好性及其系统发育分析[J]. 茶叶科学, 2024, 44(3): 411-430.
- YIN M H, ZHANG J X, LE Y, HE F F, HUANG T H, ZHANG M T. Genomic characteristics, codon preference, and phylogenetic analysis of chloroplasts of *Camellia sinensis* cv. 'Damianbai'[J]. Journal of Tea Science, 2024, 44(3): 411-430. (in Chinese)
- [42] 佟岩, 黄芸, 王雨华. 森林茶园古茶树大理茶叶叶绿体基因组密码子偏好性及系统发育研究[J]. 茶叶科学, 2023, 43(3): 297-309.
- TONG Y, HUANG H, WANG Y H. Analysis of codon usage bias and phylogenesis in the chloroplast genome of ancient tea tree *Camellia taliensis* in forest-tea garden[J]. Journal of Tea Science, 2023, 43(3): 297-309. (in Chinese)
- [43] SUN R X, YE X M, WANG Z L, LIN X F. The complete chloroplast genome of *Castanopsis carlesii* (Hemsl.) Hay[J]. Mitochondrial DNA Part B, 2019, 4(2): 2591-2592.
- [44] 江淑珍, 连辉, 熊远芳, 张赛, 陈世品. 米槠叶绿体基因组密码子偏好性分析[J]. 分子植物育种, 2024, 22(16): 5307-5316.
- JIANG S Z, LIAN H, XIONG Y F, ZHANG S, CHEN S P. Analysis of codon bias in chloroplast genome of *Castanopsis carlesii*[J]. Molecular Plant Breeding, 2024, 22(16): 5307-5316. (in Chinese)
- [45] 胡振民, 万青, 李欢, 李荣林, 王枫, 杨亦扬. 茶树 *CsNRT1.1* 基因密码子使用特性分析[J]. 江苏农业学报, 2019, 35(4): 896-903.
- HU Z M, WAN Q, LI H, LI R L, WANG F, YANG Y Y. Analysis of codon usage features of *CsNRT1.1* gene in *Camellia sinensis*[J]. Jiangsu Journal of Agricultural Sciences, 2019, 35(4): 896-903. (in Chinese)