

心叶毛蕊茶叶叶绿体基因组特征及系统发育分析

胡悦, 刘兵兵^①

(山西大学黄土高原研究所, 山西太原 030006)

摘要: 为了探明心叶毛蕊茶 [*Camellia cordifolia* (Metc.) Nakai] 叶绿体基因组结构特征及系统发育关系, 基于 Illumina NovaSeq 测序平台的高通量测序技术首次对心叶毛蕊茶叶叶绿体基因组进行测序, 并利用生物信息学相关软件对该基因组进行了组装、注释、图谱绘制以及序列特征分析, 并基于叶绿体基因组构建了山茶属 (*Camellia* Linn.) 16 个物种的系统发育树。结果显示: 心叶毛蕊茶的叶绿体基因组呈现典型的四分体结构, 全长 156 745 bp, 总 GC 含量为 37.3%; 基因组有 135 个基因, 其中包括 90 个蛋白质编码基因、37 个 tRNA 基因以及 8 个 rRNA 基因。密码子偏好性分析发现心叶毛蕊茶叶叶绿体基因密码子偏好以 A 或 U 结尾, 但整体上密码子使用偏好性较弱。重复序列分析共检测到 57 个长重复序列和 73 个简单重复序列 (SSR) 位点。反向重复区 (IR) 边界结构分析发现不同物种 *rps19*、*ndhF* 和 *yef1* (假基因) 基因长度以及与边界的距离存在一定差别。系统发育分析表明心叶毛蕊茶与连蕊茶组 (Sect. *Theopsis* Coh. St.) 的尖连蕊茶 [*Camellia cuspidata* (Kochs) Wright ex Gard. Chron.] 亲缘关系较近 (自展支持率为 100%)。综合研究结果显示: 心叶毛蕊茶叶叶绿体基因组的长度、结构及排列顺序具有高度保守性, IR 区边界在进化过程中发生了一定程度的收缩, 最常用到的密码子是编码异亮氨酸 (Ile) 的 AUU, 以单核苷酸 A/T 重复序列为主的 SSR 位点较丰富, 叶绿体基因组的碱基组成偏好使用 A 或 T; 从分子系统发育的角度支持毛蕊茶组 (Sect. *Eriandria* Coh. St.) 与连蕊茶组合并的观点。

关键词: 心叶毛蕊茶; 叶绿体基因组; 序列比对; 系统发育分析

中图分类号: Q943.2; Q949.9 文献标志码: A 文章编号: 1674-7895(2024)03-0001-13

DOI: 10.3969/j.issn.1674-7895.2024.03.01

Analyses on chloroplast genome characteristics and phylogeny of *Camellia cordifolia* HU Yue, LIU Bingbing^① (Institute of Loess Plateau, Shanxi University, Taiyuan 030006, China), *J. Plant Resour. & Environ.*, 2024, 33(3): 1-13

Abstract: To clarify the structural characteristics of chloroplast genome and phylogenetic relationships of *Camellia cordifolia* (Metc.) Nakai, the chloroplast genome of *C. cordifolia* was sequenced for the first time by using the high-throughput sequencing technology based on the Illumina NovaSeq sequencing platform, assembly, annotation, mapping, and sequence characteristic analysis of this genome were performed by using bioinformatics-related softwares, and a phylogenetic tree of 16 species in *Camellia* Linn. was constructed based on their chloroplast genomes. The results show that the chloroplast genome of *C. cordifolia* exhibits a typical quadripartite structure, its total length is 156 745 bp, and its total GC content is 37.3%; the genome has 135 genes, including 90 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. The codon usage bias analysis shows that the codons of chloroplast genes of *C. cordifolia* prefer to end with A or U, but the codon usage bias is relatively weak in general. A total of 57 long repeated sequences and 73 simple sequence repeat (SSR) loci are detected in the repeated sequence analysis. The boundary structure analysis of the inverted repeat (IR) region reveals that the gene length and distance from the boundary of *rps19*, *ndhF*, and *yef1* (pseudogene) are different in different species.

收稿日期: 2023-07-26

基金项目: 国家自然科学基金项目 (41601055)

作者简介: 胡悦 (1998—), 女, 山西晋中人, 硕士研究生, 主要从事资源利用与植物保护研究。

^①通信作者 E-mail: lbb2015@sxu.edu.cn

引用格式: 胡悦, 刘兵兵. 心叶毛蕊茶叶叶绿体基因组特征及系统发育分析[J]. 植物资源与环境学报, 2024, 33(3): 1-13.

The phylogenetic analysis suggests that *C. cordifolia* has a relatively close genetic relationship with *Camellia cuspidata* (Kochs) Wright ex Gard. Chron. in the Sect. *Theopsis* Coh. St. (the bootstrap support rate is 100%). The comprehensive study results show that the length, structure, and arrangement order of chloroplast genome of *C. cordifolia* are highly conserved, the boundary of IR region has contracted to a certain extent during evolution process, the most frequently used codon is AUU, which encodes isoleucine (Ile), SSR loci dominated by mononucleotide A/T repeats are relatively rich, and the base composition of the chloroplast genome prefers to use A or T; the viewpoint of merging Sect. *Eriandria* Coh. St. and Sect. *Theopsis* is supported in terms of molecular phylogeny.

Key words: *Camellia cordifolia* (Metc.) Nakai; chloroplast genome; sequence alignment; phylogenetic analysis

叶绿体是绿色植物进行光合作用的特有细胞器,对植株生长发育具有重要的意义^[1]。叶绿体基因组独立于核基因组之外,并且在绝大多数被子植物中通过母系遗传,其结构简单且较为保守,进化速率适中^[2],发生重组的几率很小^[3],测序成本低,可为属及以下水平的分类和系统发育研究提供高效的信息位点,因此叶绿体基因组被广泛应用于物种鉴定、物种进化和系统发育关系研究^[4-5]。

山茶属 (*Camellia* Linn.) 为山茶科 (Theaceae) 中最大的属,该属的传统分类研究主要基于张宏达分类系统^{[6]59-61}和闵天禄分类系统^[7]展开,其中依据张宏达分类系统,山茶属共包含 4 个亚属 18 个组,包括红山茶组 [Sect. *Camellia* (Linn.) Dyer]、糙果茶组 (Sect. *Furfuracea* H. T. Chang)、小黄花茶组 (Sect. *Luteoflora* H. T. Chang)、油茶组 (Sect. *Oleifera* H. T. Chang)、短柱茶组 (Sect. *Paracamellia* Sealy)、半宿萼茶组 (Sect. *Pseudocamellia* Sealy)、瘤果茶组 (Sect. *Tuberculata* Chang)、古茶组 (Sect. *Archecamellia* Sealy)、实果茶组 [Sect. *Stereocarpus* (Pierre) Sealy]、短蕊茶组 (Sect. *Brachyandra* Chang)、金花茶组 (Sect. *Chrysantha* Chang)、离生雄蕊组 (Sect. *Corallina* Sealy)、秃茶组 (Sect. *Glaberrima* Chang)、长柄山茶组 (Sect. *Longipedicellata* Chang)、超长柄茶组 (Sect. *Longissima* Chang)、茶组 [Sect. *Thea* (Linn.) Dyer]、毛蕊茶组 (Sect. *Eriandria* Coh. St.) 和连蕊茶组 (Sect. *Theopsis* Coh. St.)。但对于山茶属不同组间以及组内种间的关系一直争议不断。张宏达将连蕊茶组和毛蕊茶组分为 2 组,共属于后生山茶亚属 (Subgen. *Metacamellia* Chang)^{[6]141}。叶创兴等^[8]认为这 2 个组在形态学上的差异只有子房被毛与否,不支持将其分为 2 组。江正栋^[9]和 Vijayan 等^[10]对山茶属植物进行系统发育分析,结果显示连蕊茶组和毛蕊茶组植物

聚为一支,建议将 2 组合并。陈春梅等^[11]利用叶绿体 *trnL-trnF* 片段构建了 42 个山茶属植物的系统发育树,发现虽然该系统中的茶组物种聚成一个独立分支,但糙果茶组的阔柄糙果茶 [*Camellia furfuracea* var. *latipetiolata* (C. W. Chi) T. L. Ming] 也聚到该组中。方伟等^[12]基于 *rpl16*、*psbA-trnH*、*trnL-F* 和 *rpl32-trnL* 4 个叶绿体基因组片段构建山茶属 28 个物种的系统发育树,结果显示长柄山茶组和金花茶组间亲缘关系较近,但基于少数几个叶绿体基因片段构建的系统发育树存在许多多歧分支和支持率较低的情况。而利用叶绿体全基因组数据构建的系统发育树支持率更高、更准确^[13]。随着 DNA 测序技术的快速发展和测序成本的下降,对叶绿体全基因组的分析被视为山茶属分类学和系统发育研究中的重要工具。Yang 等^[14]对山茶属 6 个物种的叶绿体全基因组进行了测序分析,发现山茶属叶绿体基因组遗传多样性高,利用其进行物种鉴定和系统发育研究来解决山茶属内争议是一条可行途径。

心叶毛蕊茶 [*Camellia cordifolia* (Metc.) Nakai] 隶属于山茶科山茶属的毛蕊茶组,主要分布于广西、广东和江西等地海拔 100~850 m 的山谷、水旁或疏林中^{[6]187},具有较高的观赏价值^[15]。此外,毛蕊茶组植物具有一定的抗逆性,其生态价值也逐渐被学者们关注。李辛雷等^[16]对山茶属 130 个物种的相对电导率和丙二醛含量等生理指标进行测定和分析,结果显示毛蕊茶组植物耐热性较强。但由于心叶毛蕊茶植株数量较少,分布范围也不大^{[6]181},目前尚未有其叶绿体基因组特征、系统发育关系以及遗传进化等方面的相关研究。

近年来,许多山茶属植物完成了叶绿体基因组组装,如油茶 (*Camellia oleifera* Abel)^[17]、多齿红山茶 (*C. polyodonta* How ex Hu)^[18]、杜鹃叶山茶 (*C. azalea*

C. F. Wei)^[19]、贵州金花茶(*C. huana* T. L. Ming et W. J. Zhang)和小黄花茶(*C. luteoflora* Li ex Chang)^[20]等,而心叶毛蕊茶缺少叶绿体基因组完整信息,对于毛蕊茶组在山茶属中系统发育位置的研究结论缺乏完整性和可靠性,影响了该物种的进一步开发和利用以及山茶属分类和进化的研究。

鉴于此,本研究利用 Illumina NovaSeq 测序平台的高通量测序技术首次对心叶毛蕊茶叶绿体基因组进行测序及组装,并和山茶属 15 个组的代表物种进行序列对比,从基因组成、序列特征及系统发育关系等方面展开研究,以期为中心叶毛蕊茶的分类鉴定、种质资源开发和利用、毛蕊茶组的系统发育关系研究以及后续起源进化和遗传多样性研究提供参考依据和理论基础。

1 材料和方法

1.1 材料

供试心叶毛蕊茶采自广东省佛山植物园(东经 113°00′00″、北纬 23°06′12″),凭证标本(标本号 20160709002)保存于广西壮族自治区中国科学院广西植物研究所。选取 1 株健康植株,采集其新梢嫩叶 5~7 枚,放入茶叶袋后装入有变色硅胶的密封袋中干燥,随后放入 -4 °C 冰箱中保存、备用。

1.2 方法

1.2.1 DNA 的提取和测序 用液氮将干燥后的叶片快速冷冻,研磨成粉末,采用改良的溴化十六烷基三甲基铵(CTAB)法^[21]提取心叶毛蕊茶基因组总 DNA,用质量体积分数 2% 的琼脂糖凝胶电泳检测 DNA 的完整性,使用美国 Thermo Scientific 公司的 NanoDrop 2000 分光光度计检测 DNA 浓度和纯度,检测完毕后放入 -20 °C 冰箱备用。将检测合格的 DNA 样品送到生工生物工程(上海)股份有限公司进行末端修复、片段纯化、PCR 扩增等文库的构建,合格的文库通过美国 Illumina 公司的 Illumina NovaSeq 测序平台进行双末端测序,原始数据过滤掉低质量序列后得到 1 Gb 的干净数据(clean reads)。

1.2.2 叶绿体基因组的组装和注释 使用 NOVOPlasty(v4.3.1)软件对心叶毛蕊茶叶绿体基因组进行从头组装,保持默认参数^[22],得到 1 条完整环状叶绿体基因组序列。利用 NCBI 数据库(<https://www.ncbi.nlm.nih.gov>)的 BLAST 功能检索出近缘物

种长尾毛蕊茶(*Camellia caudata* Wall.)(GenBank 登录号 NC_072329)并以其基因信息作为参考序列,分别使用 Plann(v1.1)软件和 Sequin(v16.0)软件进行叶绿体基因组的注释和手动校正。将最终得到的心叶毛蕊茶叶绿体基因组注释信息和基因组序列提交至 NCBI 的 GenBank 数据库(登录号 OQ630970)。使用 OGDRAW 在线软件(<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>)绘制心叶毛蕊茶的叶绿体基因组图谱^[23]。

1.2.3 密码子偏好性分析 使用 CodonW(v1.4.2)在线软件(<http://codonw.sourceforge.net>)对心叶毛蕊茶叶绿体基因组密码子的使用情况、有效密码子数(ENC)和相对同义密码子使用度(RSCU)进行分析。ENC 值范围为 20~61,20 表示每个氨基酸只使用 1 个密码子,61 表示各个密码子被平均使用,ENC 值越大表示密码子使用偏好性越弱;通常高表达基因的密码子种类较少,偏好性强,ENC 值较小,而低表达基因则含有较多种类的稀有密码子,偏好性弱,ENC 值较大,因此也可以通过比较 ENC 值来确定内源基因表达量的相对水平^[24]。RSCU 值以 1 为界限,RSCU 值为 1 时,表示编码蛋白质时该密码子没有偏好性;RSCU 值大于 1 时,表示编码蛋白质时该密码子使用频率较高;RSCU 值小于 1 时,则表示编码蛋白质时该密码子使用频率较低^[25]。

1.2.4 重复序列分析 使用 REPuter 在线软件(<http://bibiserv.cebitec.uni-bielefeld.de/reputer>)对心叶毛蕊茶叶绿体基因组的长重复序列(长度大于 30 bp)进行检索分析^[26],其中最大计算重复次数、最小重复序列长度和汉明距离(Hamming distance)分别设置为 1 000、30 和 3,其余参数保持默认。重复序列包含正向重复(F)、反向重复(R)、互补重复(C)和回文重复(P) 4 种类型。使用 MISA 在线软件(<https://webblast.ipk-gatersleben.de/misa/index.php>)搜索简单重复序列(SSR)位点^[27],单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸、六核苷酸的最小重复数分别设置为 10、5、4、3、3、3。

1.2.5 反向重复区边界结构分析 以张宏达分类系统^{[6]59-61}为依据,从 NCBI 数据库下载山茶属 15 个组中代表物种的完整叶绿体基因组序列,并选取圆籽荷属(*Apterosperma* Hung T. Chang)的圆籽荷(*A. oblata* Chang)(GenBank 登录号 KY406751)作为外类群^[28]。使用 IRscope 可视化在线软件(<https://irscope>。

基因组注释结果见表1。结果显示:心叶毛蕊茶叶叶绿体基因组由1个大单拷贝区、1个小单拷贝区和2个反向重复区组成,为典型的四分体结构。叶绿体基因组序列全长156 745 bp,总GC含量为37.3%。大单拷贝区序列长度为86 390 bp,小单拷贝区序列长度为18 267 bp,反向重复区序列长度为26 044 bp,2个反向重复区的起始位置分别为86 391和130 702 bp,大单拷贝区、小单拷贝区和反向重复区的GC含量分别为35.3%、30.6%和43.0%。心叶毛蕊茶共有135个基因,包括90个蛋白质编码基因、37个tRNA基因和8个rRNA基因。根据功能的不同又可将其分为光合作用基因、自我复制基因、其他功能基因和未知功能基因4类。大单拷贝区包含61个蛋白质编码基因和22个tRNA基因;小单拷贝区包含11个蛋白质编码

基因和1个tRNA基因;反向重复区包含20个双拷贝基因,分别为4个rRNA基因(*rrn4.5*、*rrn5*、*rrn16*和*rrn23*),7个tRNA基因(*trnA-UGC*、*trnI-CAU*、*trnI-GAU*、*trnL-CAA*、*trnN-GUU*、*trnR-ACG*和*trnV-GAC*)和9个蛋白质编码基因(*ndhB*、*rps7*、*rps12*、*rpl2*、*rpl23*、*ycf1*、*ycf2*、*ycf15*和*orf42*),其中*ycf1*基因位于反向重复区和小单拷贝区交界处。

对具有内含子的叶绿体基因进行分析,结果见表2。结果显示:存在内含子的叶绿体基因共17个,其中15个基因(9个蛋白质编码基因和6个tRNA基因)有1个内含子;而*ycf3*和*clpP*基因有2个内含子;*rps12*为反式剪切基因^[30],有3个外显子,但外显子II、III位于反向重复区,因此仅统计2个外显子,不统计其内含子信息。

表1 心叶毛蕊茶叶叶绿体基因组注释信息

Table 1 Annotation information of chloroplast genome of *Camellia cordifolia* (Metc.) Nakai

功能 Function	基因 ¹⁾ Gene ¹⁾	数量 Number
光合作用基因 Photosynthesis gene		
光系统 I Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>	5
光系统 II Photosystem II	<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
核酮糖-1,5-二磷酸羧化酶大亚基 Large subunit of rubisco	<i>rbcL</i>	1
细胞色素 b/f 复合体 Cytochrome b/f complex	<i>petA</i> , <i>petB</i> *, <i>petD</i> *, <i>petG</i> , <i>petL</i> , <i>petN</i>	6
ATP 合成酶 ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> *, <i>atpH</i> , <i>atpI</i>	6
NADH 脱氢酶 NADH dehydrogenase	<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
自我复制基因 Self-replication gene		
rRNA	<i>rrn4.5</i> (×2), <i>rrn5</i> (×2), <i>rrn16</i> (×2), <i>rrn23</i> (×2)	8
tRNA	<i>trnA-UGC</i> * (×2), <i>trnC-GCA</i> , <i>trnD-GUC</i> , <i>trnE-UUC</i> , <i>trnF-GAA</i> , <i>trnI-M-CAU</i> , <i>trnG-GCC</i> *, <i>trnG-UCC</i> , <i>trnH-GUG</i> , <i>trnI-CAU</i> (×2), <i>trnI-GAU</i> * (×2), <i>trnK-UUU</i> *, <i>trnL-CAA</i> (×2), <i>trnL-UAA</i> *, <i>trnL-UAG</i> , <i>trnM-CAU</i> , <i>trnN-GUU</i> (×2), <i>trnP-UGG</i> , <i>trnQ-UUG</i> , <i>trnR-ACG</i> (×2), <i>trnR-UCU</i> , <i>trnS-GCU</i> , <i>trnS-GGA</i> , <i>trnS-UGA</i> , <i>trnT-GGU</i> , <i>trnT-UGU</i> , <i>trnV-GAC</i> (×2), <i>trnV-UAC</i> *, <i>trnW-CCA</i> , <i>trnY-GUA</i>	37
RNA 聚合酶 RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> *, <i>rpoC2</i>	4
核糖体小亚基 Small subunit of ribosome	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (×2), <i>rps8</i> , <i>rps11</i> , <i>rps12</i> (×2), <i>rps14</i> , <i>rps15</i> , <i>rps16</i> *, <i>rps18</i> , <i>rps19</i>	14
核糖体大亚基 Large subunit of ribosome	<i>rpl2</i> * (×2), <i>rpl14</i> , <i>rpl16</i> *, <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (×2), <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>	11
其他功能基因 Other functional gene		
成熟酶 Maturase	<i>matK</i>	1
蛋白酶 Protease	<i>clpP</i> **	1
被膜蛋白 Envelop membrane protein	<i>cemA</i>	1
乙酰辅酶 A 羧化酶亚基 Subunit of acetyl-CoA-carboxylase	<i>accD</i>	1
翻译起始因子 Translation initiation factor	<i>infA</i>	1
细胞色素 C 合成酶 Cytochrom C synthetase gene	<i>ccsA</i>	1
未知功能基因 Functionally unknown gene		
叶绿体开放阅读框 Chloroplast open reading frame	<i>ycf1</i> (×2), <i>ycf2</i> (×2), <i>ycf3</i> **, <i>ycf4</i> , <i>ycf15</i> (×2), <i>orf42</i> (×2)	10

¹⁾ * : 具有1个内含子 Containing one intron; ** : 具有2个内含子 Containing two introns; (×2) : 2份拷贝 Two copies.

2.2 密码子偏好性分析

通过分析心叶毛蕊茶叶叶绿体基因组中 90 个蛋白质编码基因序列密码子偏好性的常用参数,发现心叶毛蕊茶叶叶绿体基因组中有效密码子数(ENC)为 55.3,表明叶绿体基因的密码子使用偏好性较弱,基因中含有的稀有密码子种类较多,基因的表达量相对较低。密码子的 GC 含量为 38.3%,其中密码子第 3 位点的 GC 含量为 35.8%。

从心叶毛蕊茶叶叶绿体基因组的相对同义密码子使用度(表 3)看,心叶毛蕊茶叶叶绿体基因组中,编码异亮氨酸(Ile)的 AUU、编码赖氨酸(Lys)的 AAA、编码谷氨酸(Glu)的 GAA、编码天冬酰胺(Asn)的 AAU 以及编码苯丙氨酸(Phe)的 UUU 数量较多,分别为 1 121、1 083、1 053、1 017 和 979。此外,在不统计终止密码子的情况下,检测到 29 个密码子的相对同义密码子使用度大于 1,其中,编码亮氨酸(Leu)的 UUA

表 2 心叶毛蕊茶叶叶绿体基因组中具有内含子的基因信息

Table 2 Information of genes with intron in chloroplast genome of *Camellia cordifolia* (Metc.) Nakai

基因 Gene	位置 ¹⁾ Strand ¹⁾	序列长度/bp Length of sequence			基因 Gene	位置 ¹⁾ Strand ¹⁾	序列长度/bp Length of sequence		
		外显子 I Exon I	内含子 I Intron I	外显子 II Exon II			内含子 II Intron II	外显子 III Exon III	
<i>rpl2</i>	IR	435	670	390	<i>rpoC1</i>	LSC	1 617	741	435
<i>ndhB</i>	IR	756	684	777	<i>ycf3</i>	LSC	153	737	228
<i>trnI-GAU</i>	IR	42	941	35	<i>trnL-UAA</i>	LSC	37	523	50
<i>trnA-UGC</i>	IR	38	812	35	<i>trnV-UAC</i>	LSC	37	585	39
<i>rps12</i> ²⁾	LSC-IR	114	—	243	—				
<i>trnK-UUU</i>	LSC	35	2 489	37	<i>clpP</i>	LSC	228	596	289
<i>rps16</i>	LSC	227	870	40	<i>petB</i>	LSC	6	794	642
<i>trnG-GCC</i>	LSC	23	705	37	<i>petD</i>	LSC	8	747	475
<i>atpF</i>	LSC	411	716	144	<i>rpl16</i>	LSC	399	1 021	9
					<i>ndhA</i>	SSC	540	1 084	552

¹⁾ LSC: 大单拷贝区 Large single copy region; SSC: 小单拷贝区 Small single copy region; IR: 反向重复区 Inverted repeat region.

²⁾ —: *rps12* 为反式剪切基因,不统计其内含子信息 *rps12* is a trans-splicing gene, and its intron information is not counted.

表 3 心叶毛蕊茶叶叶绿体基因组的相对同义密码子使用度(RSCU)

Table 3 Relative synonymous codon usage (RSCU) of chloroplast genome of *Camellia cordifolia* (Metc.) Nakai

氨基酸 Amino acid	密码子 Codon	数量 Number	RSCU	氨基酸 Amino acid	密码子 Codon	数量 Number	RSCU	氨基酸 Amino acid	密码子 Codon	数量 Number	RSCU
Phe	UUU	979	1.28	His	CAU	509	1.57	Ala	ACC	254	0.75
	UUC	554	0.72		CAC	141	0.43		ACA	416	1.22
Leu	UUA	876	1.88	Arg	CGU	360	1.33	ACG	142	0.42	
	UUG	564	1.21		CGC	88	0.33	GCU	638	1.83	
	CUU	587	1.26		CGA	388	1.44	GCC	225	0.64	
	CUC	210	0.45		CGG	115	0.43	GCA	399	1.14	
	CUA	377	0.81		AGA	506	1.87	GCG	135	0.39	
Ser	CUG	186	0.40	AGG	164	0.61	Tyr	UAU	808	1.60	
	UCU	631	1.80	Gln	CAA	716		1.52	UAC	202	0.40
	UCC	331	0.94		CAG	227	0.48	Asp	GAU	914	1.63
	UCA	415	1.18	Ile	AUU	1 121	1.44		GAC	207	0.37
	UCG	182	0.52		AUC	462	0.60	Glu	GAA	1 053	1.51
	AGU	431	1.23	AUA	746	0.96	GAG		342	0.49	
Pro	AGC	118	0.34	Met	AUG	654	1.00	Lys	AAA	1 083	1.48
	CCU	450	1.59		Val	GUU	531		1.47	AAG	379
	CCC	196	0.69	GUC		172	0.48	Gly	GGU	574	1.27
	CCA	339	1.20	GUA	540	1.50	GGC		186	0.41	
Cys	CCG	146	0.52	Asn	GUG	200	0.55	GGA	743	1.65	
	UGU	221	1.50		AAU	1 017	1.54	GGG	300	0.67	
Trp	UGC	74	0.50	Thr	AAC	303	0.46				
	UGG	482	1.00		ACU	551	1.62				

具有最高的相对同义密码子使用度(1.88);29个密码子中有28个以A或U结尾,仅有1个以G结尾,表明心叶毛蕊茶叶绿体基因组密码子对A或U结尾具有偏好性。

2.3 长重复序列和简单重复序列(SSR)分析

长重复序列分析结果(表4)显示:在心叶毛蕊茶叶绿体基因组中共检测到57个长度大于或等于30bp的长重复序列,包括23个正向重复和34个回文重复,位于大单拷贝区、小单拷贝区和反向重复区的重复序列数量占比分别为22.8%、1.8%和75.4%,其中,5个正向重复和8个回文重复位于大单拷贝区,1个回文重复位于小单拷贝区,18个正向重复和25个回文重复位于反向重复区。大部分重复序列集中分布于ycf2基因上(42个),1个位于psaB和psaA基因

上,14个位于基因间隔区或内含子区域。最长的重复序列长度为26044bp,起始位置分别在86391和130702bp处。在整个心叶毛蕊茶叶绿体基因组中未检测到反向重复序列和互补重复序列。

SSR分析结果(表5)显示:心叶毛蕊茶叶绿体基因组中共有73个SSR位点,其中单核苷酸重复有53个,占72.6%,均为A/T重复序列;二核苷酸重复有4个,占5.5%,均为AT/AT重复序列;三核苷酸重复有1个,占1.4%,为AAG/CTT重复序列;四核苷酸重复有13个,占17.8%,包括AAAG/CTTT、AAAT/ATTT、ACAG/CTGT、AGAT/ATCT、AGGG/CCCT 5种类型;六核苷酸重复有2个,占2.7%,均为AAAAAG/CTTTTT重复序列,全序列未检测到五核苷酸重复序列。据统计,SSR位点最丰富的类型为A/T(53),其

表4 心叶毛蕊茶叶绿体基因组的长重复序列
Table 4 Long repeated sequence in chloroplast genome of *Camellia cordifolia* (Metc.) Nakai

序列长度/bp Sequence length	起始位置 1/bp Start position 1	起始位置 2/bp Start position 2	重复类型 ¹⁾ Repetition type ¹⁾	序列长度/bp Sequence length	起始位置 1/bp Start position 1	起始位置 2/bp Start position 2	重复类型 ¹⁾ Repetition type ¹⁾
26 044	86 390	130 701	P	50	76 451	76 451	P
96	93 613	93 631	F	42	100 714	122 637	F
96	93 613	149 408	P	42	122 637	142 379	P
96	93 613	149 426	P	38	93 635	93 689	F
96	149 408	149 426	F	38	93 635	149 408	P
84	93 625	93 643	F	38	93 689	149 462	P
84	93 625	149 408	P	42	93 617	93 689	F
84	93 643	149 426	P	42	93 617	149 404	P
74	93 635	93 653	F	42	93 689	149 476	P
74	93 635	149 408	P	42	149 408	149 480	F
74	93 653	149 426	P	39	45 264	100 716	F
78	93 613	93 649	F	39	45 264	122 639	F
78	93 613	149 408	P	39	45 264	142 380	P
78	93 649	149 444	P	30	9 059	47 007	P
78	149 404	149 440	F	34	93 625	93 697	F
66	93 625	93 661	F	34	93 625	149 404	P
66	93 625	149 408	P	34	93 697	149 476	P
66	93 661	149 444	P	35	40 219	42 443	F
56	93 635	93 671	F	30	14 212	14 212	P
56	93 635	149 408	P	32	9 057	37 061	F
56	93 671	149 444	P	31	38 384	38 384	P
60	93 617	93 671	F	30	37 063	47 007	P
60	93 617	149 404	P	30	82 785	101 066	F
60	93 671	149 458	P	30	82 785	142 039	P
60	149 408	149 462	F	30	91 177	91 219	F
46	78 917	78 917	P	30	91 177	151 886	P
48	93 625	93 679	F	30	91 219	151 928	P
48	93 625	149 408	P	30	151 886	151 928	F
48	93 679	149 462	P				

¹⁾ P: 回文重复 Palindromic repeat; F: 正向重复 Forward repeat.

表5 心叶毛蕊茶叶叶绿体基因组简单重复序列信息

Table 5 Simple repeat sequence information of chloroplast genome of *Camellia cordifolia* (Metc.) Nakai

重复类型 Repeat type	重复单元 Repeat unit	重复次数 Number of repeats														总计 Total	占比/% Proportion
		3	4	5	6	7	8	9	10	11	12	13	14	15			
单核苷酸 Mononucleotide	A/T								18	11	13	5	2	4	53	72.6	
二核苷酸 Dinucleotide	AT/AT			4											4	5.5	
三核苷酸 Trinucleotide	AAG/CTT		1												1	1.4	
四核苷酸 Tetranucleotide	AAAG/CTTT	4													13	17.8	
	AAAT/ATTT	3															
	ACAG/CTGT	1															
	AGAT/ATCT	3															
	AGGG/CCCT	2															
六核苷酸 Hexanucleotide	AAAAAG/CTTTTT	2													2	2.7	

中 T 重复数量最多(33);重复单元为 A 时,重复次数为 11 的 SSR 位点数量最多(7),重复单元为 T 时,重复次数为 10 的 SSR 位点数量最多(14)。

2.4 反向重复区边界结构分析

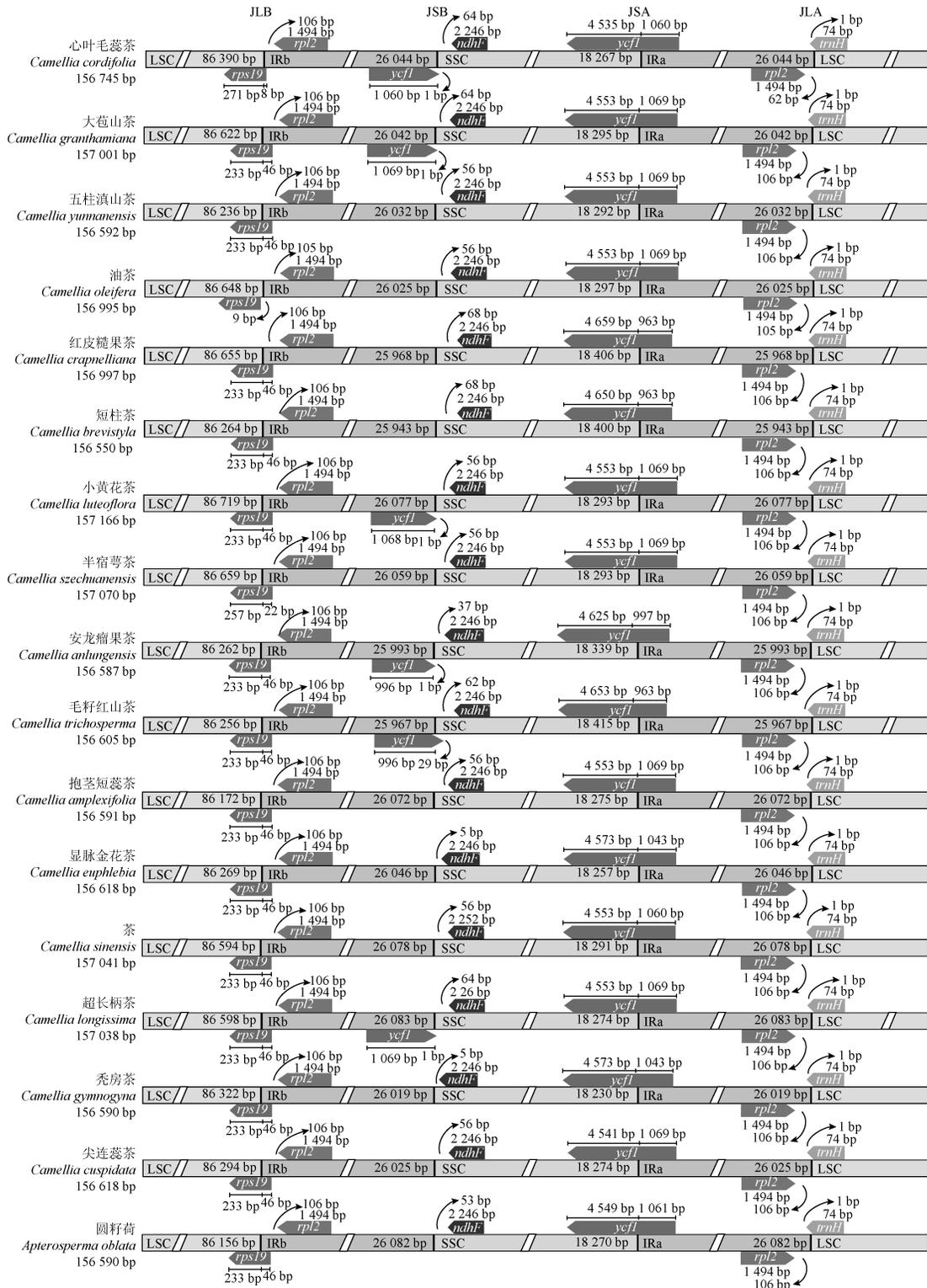
选取的山茶属各组代表物种分别为古茶组的大苞山茶(*Camellia granthamiana* Sealy)(GenBank 登录号 NC_038181)、实果茶组的五柱滇山茶[*C. yunnanensis* (Pitard) Coh. St.](GenBank 登录号 NC_022463)、油茶组的油茶(GenBank 登录号 MF541730)、糙果茶组的红皮糙果茶(*C. crapnelliana* Tutch.)(GenBank 登录号 NC_024541)、短柱茶组的短柱茶[*C. brevistyla* (Hayata) Coh. St.](GenBank 登录号 MW256435)、小黄花茶组的小黄花茶(登录号 KY626042)、半宿萼茶组的半宿萼茶(*C. szechuanensis* C. W. Chi)(GenBank 登录号 NC_035651)、瘤果茶组的安龙瘤果茶(*C. anlungensis* Chi)(GenBank 登录号 NC_050354)、红山茶组的毛籽红山茶(*C. trichosperma* Chang)(GenBank 登录号 OK181904)、短蕊茶组的抱茎短蕊茶(*C. amplexifolia* Merr. et Chun)(GenBank 登录号 NC_061610)、金花茶组的显脉金花茶(*C. euphlebica* Merr. ex Sealy)(GenBank 登录号 OL405564)、茶组的茶[*C. sinensis* (Linn.) O. Ktze.](GenBank 登录号 MZ153237)、超长柄茶组的超长柄茶(*C. longissima* Chang et Liang)(GenBank 登录号 NC_067609)、秃茶组的秃房茶(*C. gymnogyna* Chang)(GenBank 登录号 NC_039626)、连蕊茶组的尖连蕊茶[*C. cuspidata* (Kochs) Wright ex Gard. Chron.](GenBank 登录号 NC_022459),对这 15 个物种及心叶毛蕊茶和外类群圆籽荷共 17 个物种的叶绿体基因组反向重复区(IR)边界区域和相邻基因进行比较,

并分析连接区扩张和收缩的变化。结果(图 2)显示:在 17 个物种叶绿体基因组 4 个区域中,反向重复区较为保守,长度在 25 943~26 083 bp 之间,心叶毛蕊茶反向重复区长度为 26 044 bp。这 17 个物种各个区域间对应的边界基因基本相同,在大单拷贝区和 IRb 区交界处(JLB)的基因为 *rps19* 和 *rpl2*;在 IRb 区和小单拷贝区交界处(JSB)的基因为 *ndhF*;在小单拷贝区和 IRa 区交界处(JSA)的基因为 *ycf1*,且不为假基因;在 IRa 区和大单拷贝区交界处(JLA)的基因为 *rpl2* 和 *trnH*。在心叶毛蕊茶 JSB 处注释到的 *ycf1* 为假基因,大苞山茶、小黄花茶、安龙瘤果茶、毛籽红山茶和超长柄茶也出现了相同的情况。

对反向重复区边界扩张和收缩的分析结果(图 2)显示:不同物种 *rpl2* 和 *trnH* 基因的位置以及与 JLA 处的距离基本一致,相对保守,仅心叶毛蕊茶 *rpl2* 基因与 JLA 处的距离为 62 bp,其余物种 *rpl2* 基因与 JLA 处的距离均为 106 bp。但 *rps19*、*ndhF* 和 *ycf1* 基因(假基因)则在位置以及与边界的距离具有一定特异性,不同物种间差异明显,在 JLB 处,心叶毛蕊茶 *rps19* 基因有 271 bp 位于大单拷贝区,8 bp 位于 IRb 区;油茶 *rps19* 基因均位于大单拷贝区;半宿萼茶 *rps19* 基因有 257 bp 位于大单拷贝区,22 bp 位于 IRb 区;其余物种 *rps19* 基因均有 233 bp 位于大单拷贝区,46 bp 位于 IRb 区;17 个物种的 *ndhF* 和 *ycf1* 基因(假基因)的位置以及与边界的距离总体存在细微差别。表明心叶毛蕊茶叶叶绿体基因组反向重复区边界在进化过程中发生了一定程度的收缩。

2.5 系统发育分析

16 个山茶属植物和外类群圆籽荷的叶绿体基因组序列经 MAFFT 软件多重比对后,使用 MEGA-X 软

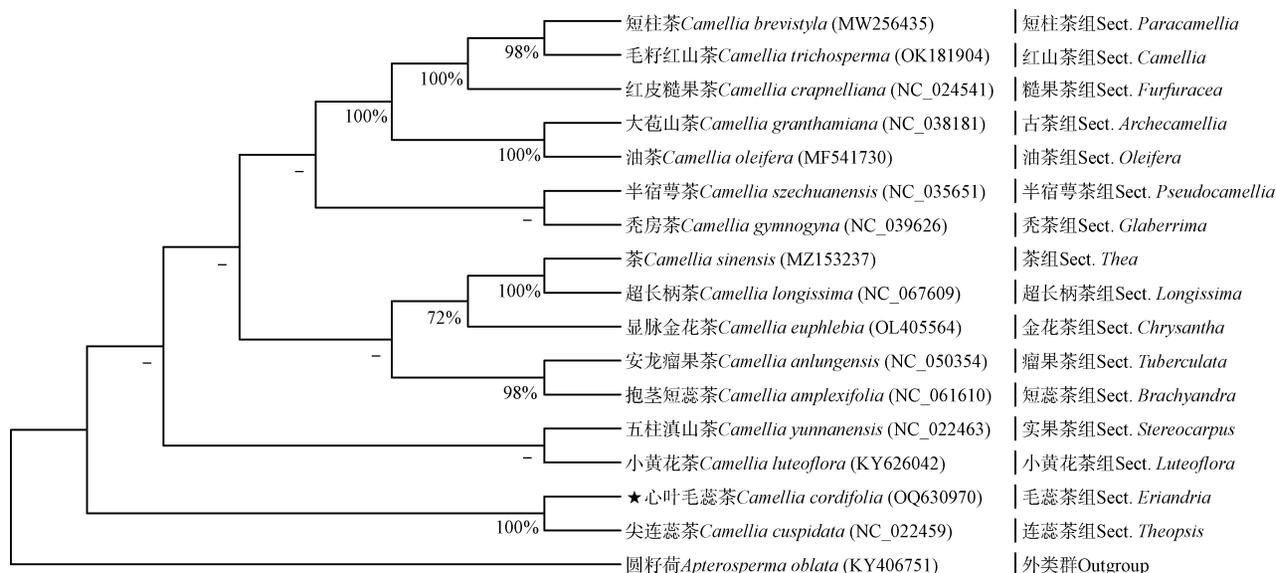


JLB: 大单拷贝区 (LSC) 和 IRb 区交界处 The junction of large single copy (LSC) and IRb regions; JSB: IRb 区和小单拷贝区 (SSC) 交界处 The junction of IRb and small single copy (SSC) regions; JSA: SSC 区和 IRa 区交界处 The junction of SSC and IRa regions; JLA: IRa 区和 LSC 区交界处 The junction of IRa and LSC regions.

图 2 山茶科 17 种植物叶绿体基因组反向重复区 (IR) 边界结构比较
 Fig. 2 Comparison on the boundary structures of inverted repeat region (IR) in chloroplast genomes of 17 species of Theaceae

件计算出的最佳 ML 建树模型为 GTR+G+I, 构建的系统发育树见图 3。结果显示: 山茶属植物总体上聚为 2 个大分支, 毛蕊茶组的心叶毛蕊茶与连蕊茶组的尖连蕊茶聚为一支, 自展支持率为 100%, 结果可信度高, 其余物种聚为一支。瘤果茶组的安龙瘤果茶与短蕊茶组的抱茎短蕊茶聚为一小支, 自展支持率为

98%; 茶组的茶和超长柄茶组的超长柄茶聚为一小支, 自展支持率为 100%; 糙果茶组的红皮糙果茶单独聚为一小支, 短柱茶组的短柱茶与红山茶组的毛籽红山茶聚为一小支(自展支持率为 98%), 古茶组的大苞山茶和油茶组的油茶聚为一小支(自展支持率为 100%), 这 3 小支间的亲缘关系较近。



进化树分支上的数据为自展支持率, 括号内编号为 GenBank 登录号, “-”表示自展支持率小于 70% The data on the branches of phylogenetic tree are bootstrap values and the Nos. in brackets are the GenBank accession numbers, and “-” indicates that the bootstrap value is less than 70%.

图 3 基于最大似然法构建的山茶科 17 种植物系统发育树
Fig. 3 Phylogenetic tree of 17 species of Theaceae constructed by maximum likelihood method

3 讨 论

光合植物的叶绿体基因组传递遗传物质的方式为单性遗传, 相对保守, 大小范围通常在 120 000 ~ 170 000 bp^[31]。本研究中, 心叶毛蕊茶叶绿体基因组序列全长 156 745 bp, 叶绿体基因组呈现典型的四分体结构, 由 1 个大单拷贝区 (LSC)、1 个小单拷贝区 (SSC) 和 2 个反向重复区 (IR) 共同组成环状结构, 这与已报道的球花石斛 (*Dendrobium thyrsiflorum* Rehb. f.)、插田泡 (*Rubus coreanus* Miq.) 和光滑高粱泡 (*Rubus lambertianus* var. *glaber* Hemsl.) 等物种^[4, 32]一致。心叶毛蕊茶共注释到 135 个基因, 其中, 蛋白质编码基因有 90 个, tRNA 基因有 37 个, rRNA 基因有 8 个, 与山茶属其他植物的注释结果相似。例如: 薄叶金花茶 (*Camellia chrysanthoides* Hung T. Chang)、簇

蕊金花茶 (*C. fascicularis* Hung T. Chang) 和毛籽红山茶^[33-34]等, 说明植物叶绿体基因组的长度、结构及排列顺序具有高度保守性。高等植物叶绿体基因组的 GC 含量比较稳定, 一般在 34% ~ 40% 之间^[35], 且被子植物 4 个区域 GC 含量通常分布不均匀^[36]。心叶毛蕊茶叶绿体基因组的总 GC 含量为 37.3%, 其中 IR 区的 GC 含量为 43.0%, 明显高于 LSC 区 (35.3%) 和 SSC 区 (30.6%), 这种 GC 含量的分布差异是由 IR 区存在的 4 个 rRNA 基因引起的, 其 GC 含量比其他基因高, 且 SSC 区中 NADH 脱氢酶基因的 GC 含量比其他基因低^[37]。

在植物叶绿体基因组中密码子偏好性是普遍存在的, 这可能与基因突变、遗传漂变、环境以及自然选择等因子有关, 且对物种进化具有重要意义^[38]。密码子使用偏好性可以通过相对同义密码子使用度 (RSCU) 来反映, 本文中心叶毛蕊茶 RSCU 值大于

1 的密码子共有 29 个, 其中大部分以 A 或 U 结尾, 表明心叶毛蕊茶偏好使用 A 或 U 结尾的密码子, 可能与叶绿体基因组中高 AT 含量有关, 这与大多数被子植物研究结果相似^[39], 该研究结果可为山茶属物种性状改良和进化研究等提供参考依据。重复序列可能导致叶绿体基因组出现重排或使其结构发生一定变化, 进而推动叶绿体基因组的进化。心叶毛蕊茶叶绿体基因组中共含有 57 个长重复序列, 均为正向重复(23 个)和回文重复(34 个), 重复序列类型与多数被子植物研究结果一致^[40]。不同物种的简单重复序列(SSR)数量和类型不同, 有时甚至在同一物种不同个体间也存在差异, 可用于物种鉴定和遗传进化分析等^[41]。心叶毛蕊茶中共检测到 73 个 SSR 位点, 其中单核苷酸 A/T 类型最多, 占 72.6%, 没有出现五核苷酸重复类型, 单核苷酸类型中也未检测到 C/G 重复单元, 由此可见简单重复序列的碱基组成的偏好性, 即高 A/T 占比。本研究检出的 SSR 位点数量和类型与抱茎金花茶(*Camellia tienii* Ninh)^[42]、多齿红山茶^[18]和毛籽红山茶^[34]等山茶属其他物种略有差异, 原因可能是检索 SSR 位点时设定的参数不同, 导致输出的结果不同。设置重复单元的间距较大、范围较宽且长度最小值较低时, 如本研究设置的最小重复数从单核苷酸至六核苷酸分别为 10、5、4、3、3、3, 可以提高检出率, 挖掘出更多的重复类型, 但也能挖掘出无效位点^[43], 检测 SSR 位点设置重复单元范围较窄且长度最小值较高时, 虽然在一定程度上减少了无效检出, 但 SSR 位点检出率较低, 因此对重复单元范围和长度进行设定时应根据不同研究的需求进行调整。

在植物进化过程中, 叶绿体基因组 IR 区边界的收缩与扩张现象时有发生。而 IR 区边界处的基因种类相对保守, 在维持叶绿体基因组稳定方面发挥重要作用^[44]。在 LSC 区和 IRb 区交界处基因一般为 *rps19* 和 *rpl2*, IRb 区和 SSC 区交界处基因为 *ycf1* (假基因) 和 *ndhF*, 此处的假基因通常是因为拷贝不完整形成的^[45], SSC 区和 IRa 区交界处基因为 *ycf1*, 且不为假基因, IRa 区和 LSC 区交界处基因通常为 *rpl2* 与 *trnH*。本研究中山茶科 17 个物种 IR 区边界基因种类与上述一致, 但不同物种 *rps19*、*ndhF* 和 *ycf1* (假基因) 基因长度以及与边界的距离存在一定差别, 这些差别可能是 IR 区发生收缩与扩张导致的, 进而影响叶绿体基因组的整体长度^[46]。此外, 在山茶属与圆

籽荷属进行 IR 区边界结构对比分析时发现, 2 个属物种 IR 区边界附近的基因片段所在区间大致相同, 说明山茶属与圆籽荷属物种的叶绿体基因组均高度保守且进化历程较为相似。评估 IR 区边界基因间的差异可以为物种鉴定和物种进化过程研究提供理论依据。

基于植物叶绿体基因组构建的系统发育树支持度更高, 比基于 DNA 片段构建的系统发育树具有更高的可信度^[39]。基于最大似然法构建的系统发育树中自展支持率大于 70% 表示支持该支系关系, 在 50% ~ 70% 之间为弱支持, 小于 70% 则视为不支持^[31]。本文中, 心叶毛蕊茶在分类学上属于毛蕊茶组, 在系统发育树中与连蕊茶组的尖连蕊茶聚为一支, 自展支持率为 100%, 说明在山茶科供试 17 个物种中心叶毛蕊茶与尖连蕊茶亲缘关系更近, 与形态学分类结果一致, 从分子系统发育的角度支持闵天禄^[7]和 Vijayan 等^[10]将毛蕊茶组与连蕊茶组合并的观点。从山茶属各组物种来看, 毛蕊茶组和连蕊茶组系统位置处于基部, 与其他 14 个组的物种亲缘关系较远。此外, 茶和超长柄茶聚为一支也支持了将超长柄茶组与茶组的归并处理^[7]; 大苞山茶没有与显脉金花茶聚为一支, 不支持闵天禄^[7]对古茶组和金花茶组的归并。本研究仅涉及山茶属 16 个组的 16 个物种, 还有很大一部分山茶属物种没有涉及, 对山茶属植物系统发育关系的研究有待进一步深入。

4 结 论

心叶毛蕊茶叶绿体基因组全长 156 745 bp, 包含 135 个基因, 呈现典型的四分体结构; 最常用密码子是编码异亮氨酸(Ile)的 AUU; 心叶毛蕊茶的简单重复序列(SSR)较丰富, 主要是单核苷酸 A/T 重复序列; 密码子偏好性和重复序列分析均表明心叶毛蕊茶叶绿体基因组的碱基组成偏好使用 A 或 T。山茶属植物反向重复区(IR)存在差异的主要原因可能是 IR 区边界发生了一定程度的收缩, 导致 *rps19*、*ndhF* 和 *ycf1* (假基因) 基因发生了改变。系统发育分析结果显示心叶毛蕊茶与尖连蕊茶亲缘关系更近。

参考文献:

- [1] LIU H Z, YE H, ZHANG N Y, et al. Comparative analyses of chloroplast genomes provide comprehensive insights into the adaptive evolution of *Paphiopedilum* (Orchidaceae) [J]. Horticulturae,

- 2022, 8: 391.
- [2] SMITH D R, ARRIGO K R, ALDERKAMP A C, et al. Massive difference in synonymous substitution rates among mitochondrial, plastid, and nuclear genes of *Phaeocystis* algae [J]. *Molecular Phylogenetics and Evolution*, 2014, 71: 36-40.
- [3] HUANG H, SHI C, LIU Y, et al. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships [J]. *BMC Evolutionary Biology*, 2014, 14: 151.
- [4] 朱 斌, 甘晨晨, 王洪程. 球花石斛(*Dendrobium thyrsiflorum*)叶绿体基因组特征及亲缘关系解析[J]. *生物技术通报*, 2021, 37(5): 38-47.
- [5] 姜汶君, 郭梦月, 庞晓慧. 叶绿体基因组在药用植物鉴定及系统进化研究中的应用[J]. *世界中医药*, 2020, 15(5): 702-708, 716.
- [6] 中国科学院中国植物志编辑委员会. 中国植物志: 第四十九卷第三分册[M]. 北京: 科学出版社, 1998.
- [7] 闵天禄. 山茶属的系统大纲[J]. *云南植物研究*, 1999, 21(2): 149-159.
- [8] 叶创兴, 张宏达. 山茶科系统发育诠释IX. 山茶属的原始特征及其演化趋向[J]. *中山大学学报(自然科学版)*, 1997, 36(3): 76-81.
- [9] 江正栋. 基于叶绿体 DNA 的山茶属植物的分子系统学和生物地理学初探[D]. 杭州: 浙江理工大学, 2017: 20-31.
- [10] VIJAYAN K, ZHANG W J, TSOU C H. Molecular taxonomy of *Camellia* (Theaceae) inferred from nrITS sequences[J]. *American Journal Botany*, 2009, 96: 1348-1360.
- [11] 陈春梅, 马春雷, 马建强, 等. 茶树 cpDNA 测序及基于 cpDNA 序列的山茶属植物亲缘关系研究[J]. *茶叶科学*, 2014, 34(4): 371-380.
- [12] 方 伟, 杨俊波, 杨世雄, 等. 基于叶绿体四个 DNA 片段联合分析探讨山茶属长柄山茶组、金花茶组和超长柄茶组的系统位置与亲缘关系[J]. *云南植物研究*, 2010, 32(1): 1-13.
- [13] TWYFORD A D, NESS R W. Strategies for complete plastid genome sequencing[J]. *Molecular Ecology Resources*, 2017, 17: 858-868.
- [14] YANG J B, YANG S X, LI H T, et al. Comparative chloroplast genomes of *Camellia* species [J]. *PLoS ONE*, 2013, 8(8): e73053.
- [15] 葛丽俊, 郭晓娜, 陈永鸿, 等. 浅谈山茶属植物在园林绿化中的应用[J]. *现代农村科技*, 2013(21): 57.
- [16] 李辛雷, 李纪元, 范妙华, 等. 山茶属主要物种耐热性研究[J]. *西北植物学报*, 2006, 26(9): 1803-1810.
- [17] 周安佩, 陈曦颜, 梁坚坤. 油茶叶绿体基因组特征及其系统分类[J]. *九江学院学报(自然科学版)*, 2021(1): 5-8, 14.
- [18] 童一涵, 郑 倩, 杜新明, 等. 多齿红山茶叶叶绿体基因组序列特征分析[J]. *植物资源与环境学报*, 2022, 31(5): 27-36.
- [19] XU X D, ZHENG W, WEN J. The complete chloroplast genome of the long blooming and critically endangered *Camellia azalea* [J]. *Conservation Genetics Resources*, 2018, 10: 5-7.
- [20] WANG G, LUO Y, HOU N, et al. The complete chloroplast genomes of three rare and endangered camellias (*Camellia huana*, *C. liberofilamenta* and *C. luteoflora*) endemic to Southwest China [J]. *Conservation Genetics Resources*, 2017, 9: 583-585.
- [21] 陈昆松, 李 方, 徐昌杰, 等. 改良 CTAB 法用于多年生植物组织基因组 DNA 的大量提取[J]. *遗传*, 2004, 26(4): 529-531.
- [22] DIERCKXSENS N, MARDULYN P, SMITS G. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data [J]. *Nucleic Acids Research*, 2017, 45(4): e18.
- [23] LOHSE M, DRECHSEL O, KAHLAU S, et al. OrganellarGenomeDRAW: a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets [J]. *Nucleic Acids Research*, 2013, 41: W575-W581.
- [24] WRIGHT F. The 'effective number of codons' used in a gene [J]. *Gene*, 1990, 87: 23-29.
- [25] 马录花, 宁佳奇, 王永杰, 等. 桃儿七叶绿体比较基因组学分析[J]. *生物工程学报*, 2022, 38(10): 3695-3712.
- [26] KURTZ S, CHOUDHURI J V, OHLEBUSCH E, et al. REPuter: the manifold applications of repeat analysis on a genomic scale [J]. *Nucleic Acids Research*, 2001, 29(22): 4633-4642.
- [27] BEIER S, THIEL T, MVNCH T, et al. MISA-web: a web server for microsatellite prediction [J]. *Bioinformatics*, 2017, 33(16): 2583-2585.
- [28] 张 立, 程永琴, 姜在民, 等. 中国山茶科植物区系及叶绿体基因组结构进化分析[J]. *西北林学院学报*, 2020, 35(5): 47-53.
- [29] LETUNIC I, BORK P. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation [J]. *Nucleic Acids Research*, 2021, 49: W293-W296.
- [30] 平晶耀, 冯佩沛, 郝 静, 等. *rps12* 基因在裸子植物中的分子进化式样[J]. *科学通报*, 2021, 66(24): 3182-3193.
- [31] SMITH D R. Does cell size impact chloroplast genome size? [J]. *Frontiers in Plant Science*, 2017, 8: 2116.
- [32] 杨 芮. 三个树莓类资源叶绿体基因组比较分析[D]. 雅安: 四川农业大学, 2019: 18-19.
- [33] 丁祥青, 李文芳, 吴丽君, 等. 4 种金花茶叶叶绿体基因组比较分析[J]. *福建农林大学学报(自然科学版)*, 2023, 52(3): 300-308.
- [34] 郑 倩, 童一涵, 孔庆博, 等. 毛籽红山茶叶叶绿体基因组特征及其系统发育分析[J]. *四川农业大学学报*, 2022, 40(4): 574-582.
- [35] 邢 钰, 冯慧喆. 叶绿体基因组的比较分析及系统发育研究[J]. *乡村科技*, 2021(24): 82-83.
- [36] KIM K J, LEE H L. Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants [J]. *DNA Research*, 2004, 11: 247-261.
- [37] RAMAN G, PARK K T, KIM J H, et al. Characteristics of the

- completed chloroplast genome sequence of *Xanthium spinosum*: comparative analyses, identification of mutational hotspots and phylogenetic implications[J]. *BMC Genomics*, 2020, 21: 855.
- [38] XU C, CAI X N, CHEN Q Z, et al. Factors affecting synonymous codon usage bias in chloroplast genome of *Oncidium Gower Ramsey* [J]. *Evolutionary Bioinformatics*, 2011, 7: 271-278.
- [39] 王子豪, 郭佳乐, 范琪, 等. 红边龙血树叶叶绿体基因组特征及其系统发育分析[J]. *生物工程学报*, 2023, 39(7): 2926-2938.
- [40] YAN C, DU J C, GAO L, et al. The complete chloroplast genome sequence of watercress (*Nasturtium officinale* R. Br.): genome organization, adaptive evolution and phylogenetic relationships in *Cardamineae*[J]. *Gene*, 2019, 699: 24-36.
- [41] HE S L, YANG Y, LI Z W, et al. Comparative analysis of four *Zantedeschia* chloroplast genomes: expansion and contraction of the IR region, phylogenetic analyses and SSR genetic diversity assessment[J]. *PeerJ*, 2020, 8: e9132.
- [42] 丁祥青, 毕远洋, 陈佳婷, 等. 抱茎金花茶(*Camellia tienii*)的叶绿体基因组特征分析[J]. *江苏农业科学*, 2022, 50(7): 1-8.
- [43] 王希, 陈丽, 赵春雷. 利用 MISA 工具对不同类型序列进行 SSR 标记位点挖掘的探讨[J]. *中国农学通报*, 2016, 32(10): 150-156.
- [44] FU J M, LIU H M, HU J J, et al. Five complete chloroplast genome sequences from *Diospyros*: genome organization and comparative analysis[J]. *PLoS ONE*, 2016, 11(7): e0159566.
- [45] 杨家鑫. 掌叶木和伞花木的叶绿体全基因组研究及其系统位置分析[D]. 贵阳: 贵州大学, 2019: 46-47.
- [46] WANG W C, CHEN S Y, ZHANG X Z. Whole-genome comparison reveals divergent IR borders and mutation hotspots in chloroplast genomes of herbaceous bamboos (*Bambusoideae: Olyreae*) [J]. *Molecules*, 2018, 23: 1537.

(责任编辑: 郭严冬)

公益宣传: 世界环境日

2024年世界环境日(World Environment Day)主题活动聚焦土地修复、荒漠化和干旱抗灾能力。该活动强调了全球对恢复退化土地、防止荒漠化并增强对于旱抵抗力的紧迫需求。土地是地球生态系统的基础,对维护生物多样性、保护水资源、调节气候等具有重要作用。土地修复是恢复和改善土地生态功能、促进可持续发展的关键措施。荒漠化和干旱是全球面临的严重环境问题,对生态环境、经济发展和社会稳定造成了巨大威胁。应对土地荒漠化和干旱挑战,需要全球共同努力,共同推动绿色发展。该活动呼吁个人、政府和组织共同参与土地恢复行动,通过改善生态系统、提高生物多样性等措施应对土地退化,增强生态系统的韧性。

中国在推进美丽中国建设、实现人与自然和谐共生方面取得了突出成效,通过倡导社会公众践行绿色低碳生活方式,减少资源消耗和环境污染,弘扬生态文化,培育全社会的生态文明意识,促进人与自然和谐共生。