

• 药材与资源 •

不同产地小构树与藤构叶绿体基因组特征与系统进化研究

刘佳升¹, 周丽娜¹, 林诗童¹, 王媚臻³, 李攀², 金鑫杰¹, 张永华^{1*}

1. 温州大学生命与环境科学学院, 浙江 温州 325035

2. 浙江大学生命科学学院, 浙江 杭州 310058

3. 浙江树人学院生物与环境工程学院, 浙江 杭州 310015

摘要: 目的 解析各 5 个产地的小构树 *Broussonetia monoica* 和藤构 *Broussonetia kaempferi* 的叶绿体基因组序列特征与系统进化关系。方法 基于深圳华大 BGISEQ-500 测序平台, 对各 5 个产地的小构树和藤构个体进行高通量基因组重测序。采用 GetOrganelle、Geneious Prime、CpGAVAS2、MISA、REPuter、IRscope、DnaSP、CodonW、IQ-tree 等生物信息学工具, 进行叶绿体基因组组装与注释、比较基因组分析、密码子偏好分析和系统发育分析。结果 小构树和藤构的各 5 个体完整叶绿体基因组均呈现典型的环状四分体结构, 序列全长分别为 160 799~160 877 bp 和 160 414~160 622 bp, 均注释得到 135 个基因, 包括 87 个蛋白质编码基因、37 个 tRNA 基因、8 个 rRNA 基因和 3 个假基因。10 个叶绿体基因组序列边界基因较为保守, 扩张收缩现象不明显。从它们的叶绿体基因组中分别鉴定出 98~101 个和 108~115 个 SSR 位点, 以及 42~48 个和 43~54 个长重复序列。小构树中的 *psbC* (8 bp) + *psbC_trnS-UGA* 区域以及 *atpF*、*ndhF*、*ndhI* 基因和藤构中的 *ccsA* (35 bp) + *ccsA_ndhD* 区域以及 *rps16*、*psbI*、*petG* 基因变异度最大。两者叶绿体基因组中均发现 30 个高频密码子, 并筛选出 7 个相同的最优密码子。系统发育分析显示, 小构树与藤构的亲缘关系较构树更近, 种内小构树和藤构个体之间的亲缘关系与地理位置有关。结论 小构树和藤构的叶绿体基因组结构、基因数量和密码子使用偏好保守性高, 但在重复序列和基因间隔区上存在较大变异; 以叶绿体基因数据构建的系统发育树指示了 2 个物种种内的亲缘关系与样本取样地地理位置有关。

关键词: 小构树; 藤构; 叶绿体基因组; 密码子偏好性; 系统发育基因组学

中图分类号: R286.2

文献标志码: A

文章编号: 0253-2670(2025)10-3637-14

DOI: 10.7501/j.issn.0253-2670.2025.10.021

Characteristics and phylogenetic evolution of complete chloroplast genomes of *Broussonetia kazinoki* and *Broussonetia kaempferi* from multiple provenances

LIU Jiasheng¹, ZHOU Lina¹, LIN Shitong¹, WANG Meizhen³, LI Pan², JIN Xinjie¹, ZHANG Yonghua¹

1. College of Life and Environmental Science, Wenzhou University, Wenzhou 325035, China

2. College of Life Sciences, Zhejiang University, Hangzhou 310058, China

3. College of Biology and Environmental Engineering, Zhejiang Shuren University, Hangzhou 310015, China

Abstract: Objective To analysis the characteristics of chloroplast genome sequences and the phylogenetic relationships of *Broussonetia kazinoki* and *Broussonetia kaempferi*, utilizing samples from five distinct provenances of each species. **Methods** High-throughput genome resequencing was conducted on individuals of *B. kazinoki* and *B. kaempferi* from five each different provenances using BGISEQ-500 sequencing platform. The chloroplast genomes were assembled and annotated, followed by comparative genomic analysis, codon usage preference analysis, and phylogenetic analysis using bioinformatics tools such as GetOrganelle, Geneious Prime, CpGAVAS2, MISA, REPuter, IRscope, DnaSP, CodonW, and IQ-tree. **Results** The complete chloroplast genomes of each five individuals from *B. kazinoki* and *B. kaempferi* exhibited the typical quadripartite circular structure, with total lengths of 160 799—160 877 bp for *B. kazinoki* and 160 414—160 622 bp for *B. kaempferi*. A total of 135 genes were annotated for both species, including 87 protein-coding genes, 37 tRNA genes, eight rRNA genes, and three pseudogenes. The boundary

收稿日期: 2024-10-23

基金项目: 国家自然科学基金项目 (31800309); 温州市基础性公益科研项目 (N20240014); 浙江省自然科学基金项目 (LY21C030002)

作者简介: 刘佳升, 本科生, 研究方向为植物细胞器基因组学。Tel: 13023742535 E-mail: liujiashengxx@outlook.com

*通信作者: 张永华, 副教授, 研究方向为植物分子系统学、中药植物资源鉴定。Tel: 13567782963 E-mail: zhangyonghua@wzu.edu.cn

genes of 10 chloroplast genome sequences were relatively conserved, with no distinct expansion or contraction detected. In *B. kazinoki*, a total of 98—101 SSRs and 42—48 long sequence repeats were identified, while *B. kaempferi* showed 108—115 SSRs and 43—54 long sequence repeats. The highest levels of sequence variation were observed in the *psbC* (8 bp) + *psbC_trnS*-UGA sequence, as well as in the *atpF*, *ndhF*, and *ndhI* genes for *B. kazinoki*. In contrast, *B. kaempferi* exhibited variation in the *ccsA* (35 bp) + *ccsA_ndhD* sequence, as well as in the *rps16*, *psbI*, and *petG* genes. Additionally, 30 high-frequency codons were shared between the chloroplast genomes of both species, and seven identical optimal codons were identified. Phylogenetic analysis indicated that *B. kazinoki* was more closely related to *B. kaempferi* than to *B. papyrifera*. Furthermore, intraspecific relationships within *B. kazinoki* and *B. kaempferi* were found to correlate with geographical origin. **Conclusion** The chloroplast genome structure, gene number, and codon usage preference of *B. kazinoki* and *B. kaempferi* were highly conserved; however, there were substantial variations in repeated sequences and intergenic regions. The phylogenetic tree constructed using chloroplast genome database indicated that the genetic relationships within these two species correlated with the geographical locations of the samples.

Key words: *Broussonetia kazinoki* Sieb.; *Broussonetia kaempferi* Siebold; chloroplast genome; codon usage bias; phylogenomics

叶绿体作为植物光合作用和能量转化的核心场所, 承载着独立的遗传物质, 并在大多数被子植物中以母系遗传的方式传递^[1-2]。高等植物的叶绿体基因组通常介于 120~180 kbp, 呈现典型的双链闭合环状四分体结构, 由 1 个大单拷贝区 (large single copy, LSC)、1 个小单拷贝区 (small single copy, SSC) 以及 1 对反向重复区 (inverted repeats, IRa/b) 构成^[3-4]。叶绿体基因组功能多样, 编码 110~130 个基因, 涵盖了参与光合作用及叶绿体基因表达的关键基因^[5]。由于其序列组成和结构高度保守、碱基变异速率适中, 且基因组相对紧凑的特性^[6-7], 叶绿体基因组已成为植物系统发育和物种鉴定研究中不可或缺的重要分子标记^[8-9]。这种独特的基因组特性为揭示植物进化机制和生态适应提供了宝贵的分子线索。

构属 *Broussonetia* L'Hér. ex Vent. 隶属于桑科 (Moraceae), 属下有构树 *B. papyrifera* (L.) L'Hér. ex Vent.、小构树 *B. kazinoki* Sieb. 和藤构 *B. kaempferi* Siebold 等 3 个物种以及 1 个自然杂交种楮构 *Broussonetia* × *kazinoki*, 这些物种在我国均有分布, 主要集中在黄河流域以南地区^[10]。构树为高大落叶乔木, 叶片呈广卵形至长椭圆状卵形, 背面密被细绒毛, 边缘不裂或 3~5 裂, 雌花序头状, 雄花序柔荑状; 小构树则为直立灌木, 其叶片形态为斜卵形或卵形, 基部呈圆至截形, 边缘锯齿粗, 雌、雄花序均呈球形头状; 藤构为蔓生藤状灌木, 叶片近对称, 呈卵状椭圆形, 基部心形或心状截形, 边缘锯齿细, 雌花序头状, 雄花序短圆柱状^[11]。构属植物具有广泛的经济价值和药用用途, 其树皮韧性强、拉力大, 是造纸和造人造棉的优质原材料^[11]。构树的成熟种子作为中药楮实子入药, 为常见的传统中药, 历版《中国药典》均有收载, 具有补肾清肝、明目利尿的功效^[12];

小构树始载于《名医别录》, 被列为中药上品, 其树枝、根茎、树皮等在东方医学中常作为改善视力以及治疗炎症和传染病的药物^[13]。构属植物提取物被发现具有抗氧化^[14-15]、抗菌^[16]、抗血小板凝聚^[17]和抗炎^[18]等作用, 而其中起药效作用的主要为黄酮类和生物碱类 2 大类活性物质。尽管构属植物种类较少, 但其种内形态多样性较大, 同时, 长期的栽培、选择和杂交育种历史导致了分类上的许多混淆和争议^[10]。近年来, Kuo 等^[19]和 Zhou 等^[20]通过研究桑科疏桑族 (Dorstenieae) 的构属、牛筋藤属 *Malaisia* Blanco 和落叶花桑属 *Allaeanthus* Thwaites 的叶绿体基因组, 验证了二者在桑科中的分类地位, 并支持将落叶花桑属与构属分开的观点。Yang 等^[21]组装了构树、楮构和藤构的叶绿体完整基因组, 研究其进化过程, 为构属植物在桑科植物中的分类提供了依据。另外, Peng 等^[22]对构树基因组学的研究揭示了其造纸、入药和饲用遗传基础, 随后, Hu 等^[23-24]又进一步研究了构树自然种群局部适应低温以及异叶性的遗传机制。目前, 虽然在构属植物分类、系统发育以及构树的基因组和种群适应基因组学等方面的研究取得了一定的进展, 但构属内部小构树和藤构种内进化关系尚未深入研究, 不同种源地的小构树和藤构叶绿体基因组资源也亟待进一步挖掘和分析。

本研究采用高通量测序技术, 获取了小构树和藤构不同产地个体的基因组重测序数据, 并成功组装了完整的叶绿体全基因组序列。通过基因注释与详尽的生物信息学分析, 对其基因组成、序列特征、SSR 位点、种间和种内变异、密码子使用偏好等方面进行了系统性的比较与解析, 并构建了系统发育树, 进一步开展了种间和种内比较基因组学和系统发育分析。本研究旨在全面解析小构树和藤构叶绿

体全基因组的序列与结构特征,深入探讨其序列变异情况,为构属植物资源的开发利用、分子鉴定及系统进化研究奠定了坚实的基础,并提供了重要的理论参考。

1 材料

小构树 5 个个体分别采自广西桂林市龙胜花坪、贵州铜仁市梵净山、浙江宁波市宁海、安徽池

州市区、江苏南京宝华山,藤构 5 个个体分别采自广东连山县笔架山、广西桂林市龙胜花坪、贵州雷山县方祥乡、贵州铜仁市梵净山、浙江宁波市宁海,经张永华副教授鉴定,确认为桑科构属植物小构树 *B. kazinoki* 和藤构 *B. kaempferi*。将采集的叶片装入取样袋,用硅胶干燥处理,带回实验室,用于后续

表 1 植物样本来源

Table 1 Source of plant samples

物种	取样的标本号	编号	产地	GenBank 登录号
小构树	WMZ737-2020-08	<i>B. kazinoki</i> _1	广西桂林市龙胜花坪	PQ427290
小构树	WMZ556-2020-08	<i>B. kazinoki</i> _2	贵州铜仁市梵净山	PQ427291
小构树	ZYH03-2022-07	<i>B. kazinoki</i> _3	浙江宁波市宁海	PQ427292
小构树	ZYH01-2020-10	<i>B. kazinoki</i> _4	安徽池州市区	PQ427293
小构树	HY01-2022-07	<i>B. kazinoki</i> _5	江苏南京宝华山	PQ427294
藤构	CJL13-2020-10	<i>B. kaempferi</i> _1	广东连山县笔架山	PQ427285
藤构	WMZ755-2020-08	<i>B. kaempferi</i> _2	广西桂林市龙胜花坪	PQ427286
藤构	WMZ454-2020-08	<i>B. kaempferi</i> _3	贵州雷山县方祥乡	PQ427287
藤构	WMZ552-2020-08	<i>B. kaempferi</i> _4	贵州铜仁市梵净山	PQ427288
藤构	ZYH01-2022-07	<i>B. kaempferi</i> _5	浙江宁波市宁海	PQ427289

2 方法

2.1 DNA 的提取和高通量测序

采用植物 DNA 提取试剂盒 (Tiagen Biotech 有限公司, 中国) 提取样品的总 DNA。用 NanoDrop2000 微量分光光度计 (Thermo Scientific, 美国) 检测总 DNA 的纯度和浓度,以 1 μ g DNA 起始量建库,随机打断检测合格的基因组总 DNA,构建长度约 350 bp 的插入片段文库。构建好的文库经过质检后,利用深圳华大自研的 BGISEQ-500 平台进行双末端测序,序列读长为 150 bp,测序数据量为 20 G。测序完成后,利用华大自主开发的过滤软件 SOAPnuke 过滤原始测序数据,最后获得 Clean reads。得到的数据以 FASTQ 格式储存,用于后续的拼接和注释

2.2 叶绿体基因组组装与注释

用 GetOrganelle v 1.6.3 程序将过滤后得到的数据组装成叶绿体基因组^[25]。组装完成后的叶绿体基因组数据用 Geneious Prime v 2021.2.2 中的 MAFFT 工具进行比对后,以构树 (GenBank 登录号: NC_035569) 叶绿体基因组为参考序列,使用在线工具 CpGAVAS 2 (http://47.96.249.172:16019/analyzer/home)^[26]对小构树和藤构叶绿体基因组进行注释,并在 Geneious Prime v 2021.2.2 软件中进行

手动校正。最后通过 OGDRAW (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>)^[27]在线网站绘制小构树和藤构叶绿体全基因组图谱。使用 Geneious Prime v 2021.2.2 软件分析 5 个小构树和 5 个藤构叶绿体基因组的基本结构,统计其 LSC、SSC 和 IR 区域的长度、GC 含量以及基因数目。

2.3 叶绿体基因组 SSR 与重复序列分析

通过 MISA (<https://webblast.ipk-gatersleben.de/misa>)在线网站^[28]检测 5 个小构树和 5 个藤构叶绿体基因组中简单重复序列 (simple sequence repeats, SSR) 的分布,参数设置为 10 (单核苷酸重复 mono-nucleotide)、5 (二核苷酸重复 di-nucleotide)、4 (三核苷酸重复 tri-nucleotide)、3 (四核苷酸重复 tetra-nucleotide)、3 (五核苷酸重复 penta-nucleotide)、3 (六核苷酸重复 hexa-nucleotide)。

通过 REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer>)在线网站检测 10 个个体测算叶绿体基因组中正向重复 (forward match)、反向重复 (reverse match)、互补重复 (complement match) 和回文重复 (palindromic match) 的分布,参数设置为最小重复序列 30 bp,汉明距离 (Hamming distance) 设置为 3^[29]。

2.4 叶绿体基因组的比较分析

使用 R v4.0.2 软件运行 IRscope^[30], 对 10 个个体的叶绿体基因组序列的相邻边界区域进行绘制, 比较分析边界区域的扩张和收缩情况。

通过 mVISTA (<https://genome.lbl.gov/vista>) 在线网站^[31]以 Shuffle-LAGAN 模型并以 *B.monoica*_1 的叶绿体基因组作为参考, 对 10 个个体进行叶绿体基因组序列比较分析, 评估各基因组序列之间的相似性和差异。分别提取 5 个小构树和 5 个藤构叶绿体基因组中的共有基因, 使用 Geneious Prime v 2021.2.2 中的 MAFFT 工具对齐后, 使用 DnaSP 6.0 软件计算核酸变异值 (P_i)^[32]。将 5 个小构树和 5 个藤构的叶绿体基因组分别使用 Geneious Prime v 2021.2.2 中的 MAFFT 工具进行比对, 通过 DnaSP v 5.10.0.1 软件, 计算 P_i , 参数设置为滑动窗口 200 bp, 步长为 100 bp^[33]。

2.5 密码子偏好性分析

使用 Geneious Prime v 2021.2.2 软件从 5 个小构树样本和 5 个藤构样本中提取叶绿体基因组原始编码序列 (coding sequence, CDS), 剔除重复序列, 保留长度 ≥ 300 bp 目标序列。使用 CodonW 1.4.4 (<https://codonw.sourceforge.net/>) 软件对筛选出的序列进行密码子使用情况分析, 包括相对同义密码子使用度 (relative synonymous codon usage, RSCU)、有效密码字数 (effective number of codons, ENC) 和密码子的第 3 个碱基出现 G/C 的概率

(guanine-cytosine content at the third position of codons, GC3s) 等。对小构树和藤构分别进行 ENC-plot 分析和 PR2-plot 分析和最优密码子分析。

2.6 系统发育分析

基于本研究新测序获得的 10 个个体与 NCBI 数据库中下载的桑科 6 个族 9 个个体的叶绿体基因组序列, 以大麻科 (Cannabaceae) 大麻属 *Cannabis* L. 的大麻 *Cannabis sativa* L. 作为外类群进行系统发育分析。使用 Geneious Prime v 2021.2.2 中的 MAFFT 工具进行叶绿体基因组多序列比对, 所使用的数据为叶绿体全基因组去除其中一个反向重复区 (IR 区)。使用 IQ-tree 软件^[34]的 ModelFinder 筛选最优模型 (K3Pu+F+I+G4), 以最大似然法 (Maximum likelihood, ML) 构建系统发育树, 自展值为 1 000。

3 结果与分析

3.1 叶绿体基因组基本特征

获得不同产地的 5 个小构树个体和 5 个藤构个体的叶绿体基因组呈现典型的环状四分体结构, 包括 LSC 区, SSC 区和 2 个 IR 区 (图 1)。小构树叶绿体基因组序列全长为 160 799~160 877 bp, 其中 LSC 的长度为 89 035~89 102 bp, SSC 的长度为 20 077~20 093 bp, IR 的长度为 25 840~25 841 bp。藤构叶绿体基因组序列全长为 160 414~160 622 bp, 其中 LSC 的长度为 88 996~89 094 bp, SSC 的长度为 19 810~19 998 bp, IR 的长度为 25 794~25 798 bp。

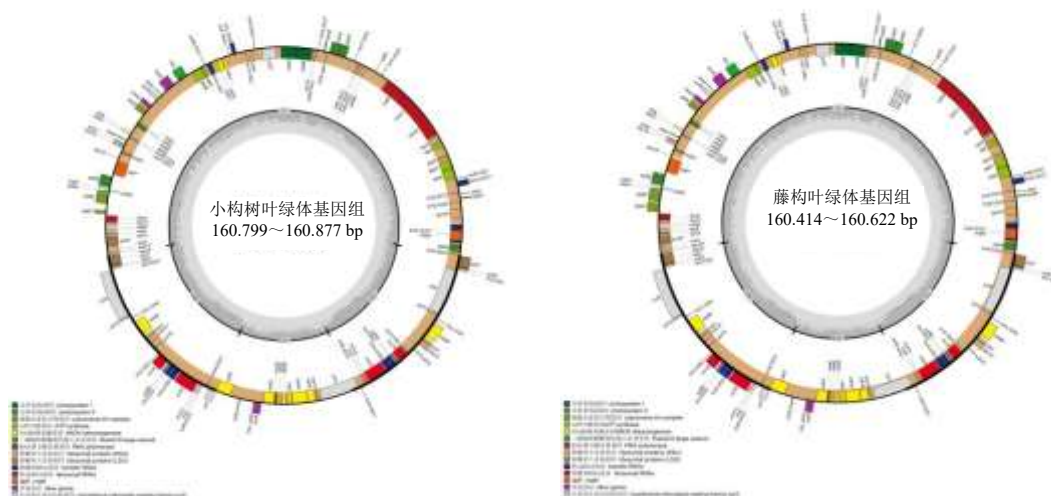


图 1 小构树和藤构叶绿体基因组图谱

Fig. 1 Chloroplast genome maps for both *B. kazinoki* and *B. kaempferi*

小构树叶绿体基因组序列总 GC 含量为 35.7%~35.8%，IR 区的 GC 含量最高（42.7%），其次是 LSC 区（33.4%）和 SSC 区（28.5%）；藤构叶

叶绿体基因组序列总 GC 含量为 35.6%~35.7%，IR 区的 GC 含量同样最高（42.7%），其次是 LSC 区（33.2%）和 SSC 区（28.3%~28.4%）（表 2）。

表 2 小构树和藤构叶绿体基因组基本特征

Table 2 Basics characteristics of chloroplast genomes of *B. kazinoki* and *B. kaempferi*

基因组特征	<i>B. monoica</i>					<i>B. kaempferi</i>				
	1	2	3	4	5	1	2	3	4	5
序列全长/bp	160 877	160 803	160 822	160 799	160 805	160 622	160 598	160 414	160 576	160 593
LSC区长度/bp	89 102	89 035	89 065	89 038	89 044	89 094	89 004	89 008	88 996	89 029
SSC区长度/bp	20 093	20 086	20 077	20 079	20 079	19 934	19 998	19 810	19 986	19 976
IR区长度/bp	25 841	25 841	25 840	25 841	25 841	25 797	25 798	25 798	25 797	25 794
GC含量/%	35.7	35.8	35.8	35.8	35.8	35.6	35.6	35.7	35.6	35.6
LSC区GC含量/%	33.4	33.4	33.4	33.4	33.4	33.2	33.2	33.2	33.2	33.2
SSC区GC含量/%	28.5	28.5	28.5	28.5	28.5	28.4	28.3	28.4	28.3	28.4
IR区GC含量/%	42.7	42.7	42.7	42.7	42.7	42.7	42.7	42.7	42.7	42.7
CDS区GC含量/%	37.2	37.2	37.3	37.3	37.3	37.2	37.2	37.2	37.2	37.2
tRNA区GC含量/%	53.4	53.4	53.3	53.3	53.3	53.3	53.3	53.3	53.3	53.3
rRNA区GC含量/%	55.3	55.3	55.3	55.3	55.3	55.3	55.3	55.3	55.3	55.3
基因总数	135	135	135	135	135	135	135	135	135	135
假基因数	3	3	3	3	3	3	3	3	3	3
编码蛋白基因数	87	87	87	87	87	87	87	87	87	87
tRNA基因数	37	37	37	37	37	37	37	37	37	37
rRNA基因数	8	8	8	8	8	8	8	8	8	8

在 5 个小构树个体和 5 个藤构个体叶绿体基因组中注释得到 135 个基因，包括 87 个蛋白质编码基因、37 个 tRNA 基因、8 个 rRNA 基因以及 3 个假基因（*ycf1*、2 个 *ycf15*）。根据其不同的功能可以把它们分为 4 大类：自我复制相关基因、光合作用相关基因、其他基因和未知功能基因。其中 18 个基因在 IR 区重复且有 2 个拷贝，包括 7 个蛋白质编码基因（*rpl2*、*rpl23*、*rps7*、*rps12*、*ndhB*、*ycf2*、*ycf15*）、7 个 tRNA 基因（*trnA*-UGC、*trnI*-CAU、*trnI*-GAU、*trnL*-CAA、*trnN*-GUU、*trnR*-ACG、*trnV*-GAC）和 4 个 rRNA 基因（*rrn4.5*、*rrn5*、*rrn16*、*rrn23*）。16 个基因有 1 个内含子，包括 10 个蛋白编码基因（*rpl2*、*rpl16*、*rps16*、*rpoC1*、*atpF*、*ndhA*、*ndhB*、*petB*、*petD*、*ycf68*）和 6 个 tRNA 基因（*trnA*-UGC、*trnG*-UCC、*trnI*-GAU、*trnK*-UUU、*trnL*-UAA、*trnV*-UAC）；有 3 个基因含有 2 个内含子（*rps12*、*clpP*、*ycf3*），见表 3。

3.2 边界分析

边界分析（图 2）结果显示，5 个小构树个体和 5 个藤构个体叶绿体基因组序列相似度较高，*rps19*、*rpl2*、*ndhF*、*ycf1* 和 *trnH* 基因位于 10 条序列的 LSC/IRb、IRb/SSC、SSC/IRa、IRa/LSC 边界或边界

附近。*rps19* 基因位于 LSC 区内，离 LSC/IRb 边界 57 bp；重复基因 *rpl2* 位于 IRb 区，离 LSC/IRb 边界 44~56 bp；重复基因 *ycf1* 一个位于 IRb/SSC 边界，有 1 040 bp 在 IRb 区，仅有 1~3 bp 在 SSC 区，另一个位于 SSC/IRa 边界，有 4 607~4 616 bp 在 SSC 区，有 1 037~1 039 bp 在 IRa 区；*ndhF* 基因位于 IRb/SSC 边界，有 16~18 bp 在 IRb 区，有 2 258~2 261 bp 在 SSC 区；*trnH* 基因位于 LSC 区，离 IRa/LSC 边界 10~12 bp。

3.3 重复序列分析

利用 MISA 软件在 5 个小构树个体叶绿体基因组中检测到 98~101 个 SSR 位点，包括 67~71 个单核苷酸重复（以 A/T 碱基重复为主）、12~13 个二核苷酸重复、4 个三核苷酸重复、6 个四核苷酸重复、6~7 个五核苷酸重复以及 0~1 个六核苷酸重复类型；在 5 个藤构个体叶绿体基因组中检测到 108~115 个 SSR 位点，包括 68~75 个单核苷酸重复（以 A/T 碱基重复为主）、18~19 个二核苷酸重复、4 个三核苷酸重复、9 个四核苷酸重复、7~9 个五核苷酸重复以及 0~2 个六核苷酸重复类型。其中小构树在 SSR 总数、二核苷酸重复以及四核苷酸重复上均少于藤构，小构树中只有 *B. kazinoki*_1 中

表 3 小构树和藤构叶绿体基因组注释基因信息

Table 3 Annotate genetic information in chloroplast genomes of *B. kazinoki* and *B. kaempferi*

基因类别	基因组	基因名称	数量
自我复制相关基因	转RNA基因	<i>trnA</i> -UGC ^{a*,b} (×2)、 <i>trnC</i> -GCA、 <i>trnD</i> -GUC、 <i>trnE</i> -UUC、 <i>trnF</i> -GAA、 <i>trnG</i> -UCC、 <i>trnG</i> -UCC ^{a*} 、 <i>trnH</i> -GUG、 <i>trnI</i> -CAU ^b (×2)、 <i>trnI</i> -GAU ^{a*,b} (×2)、 <i>trnK</i> -UUU ^{a*} 、 <i>trnL</i> -CAA ^b (×2)、 <i>trnL</i> -UAA ^{a*} 、 <i>trnL</i> -UAG、 <i>trnM</i> -CAU(×2)、 <i>trnN</i> -GUU ^b (×2)、 <i>trnP</i> -UGG、 <i>trnQ</i> -UUG、 <i>trnR</i> -ACG ^b (×2)、 <i>trnR</i> -UCU、 <i>trnS</i> -GCU、 <i>trnS</i> -GGA、 <i>trnS</i> -UGA、 <i>trnT</i> -GGU、 <i>trnT</i> -UGU、 <i>trnV</i> -GAC ^b (×2)、 <i>trnV</i> -UAC ^{a*} 、 <i>trnW</i> -CCA、 <i>trnY</i> -GUA	73
核糖RNA基因		<i>rrn4.5</i> ^b (×2)、 <i>rrn5</i> ^b (×2)、 <i>rrn16</i> ^b (×2)、 <i>rrn23</i> ^b (×2)	
核糖体蛋白大亚基		<i>rpl2</i> ^{a*,b} (×2)、 <i>rpl14</i> 、 <i>rpl16</i> ^{a*} 、 <i>rpl20</i> 、 <i>rpl23</i> ^b (×2)、 <i>rpl32</i> 、 <i>rpl33</i> 、 <i>rpl36</i>	
核糖体蛋白小亚基		<i>rps2</i> 、 <i>rps3</i> 、 <i>rps4</i> 、 <i>rps7</i> ^b (×2)、 <i>rps8</i> 、 <i>rps11</i> 、 <i>rps12</i> ^{a**,b} (×2)、 <i>rps14</i> 、 <i>rps15</i> 、 <i>rps16</i> ^{a*} 、 <i>rps18</i> 、 <i>rps19</i>	
RNA聚合酶		<i>rpoA</i> 、 <i>rpoB</i> 、 <i>rpoC1</i> ^{a*} 、 <i>rpoC2</i>	
光合作用相关基因	ATP合酶亚基	<i>atpA</i> 、 <i>atpB</i> 、 <i>atpE</i> 、 <i>atpF</i> ^{a*} 、 <i>atpH</i> 、 <i>atpI</i>	45
	NADH脱氢酶亚基	<i>ndhA</i> ^{a*} 、 <i>ndhB</i> ^{a*,b} (×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>	
	细胞色素复合体b/f亚基	<i>petA</i> 、 <i>petB</i> ^{a*} 、 <i>petD</i> ^{a*} 、 <i>petG</i> 、 <i>petL</i> 、 <i>petN</i>	
	亚光系统 I	<i>psaA</i> 、 <i>psaB</i> 、 <i>psaC</i> 、 <i>psaI</i> 、 <i>psaJ</i>	
	亚光系统 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>	
其他基因	Rubisco大亚基	<i>rbcL</i>	
	乙酰CoA羧化酶	<i>accD</i>	5
	细胞色素C合成酶	<i>ccsA</i>	
	被膜蛋白	<i>cemA</i>	
	成熟酶	<i>matK</i>	
	蛋白酶	<i>clpP</i> ^{a**}	
未知功能基因	未知功能蛋白	<i>ycf1</i> 、 <i>Ψycf1</i> ^b 、 <i>ycf2</i> ^b (×2)、 <i>ycf3</i> ^{a**} 、 <i>ycf4</i> 、 <i>ycf15</i> ^b (×2)、 <i>Ψycf15</i> ^b (×2)、 <i>ycf68</i> ^{a*,b} (×2)	12
总数			135

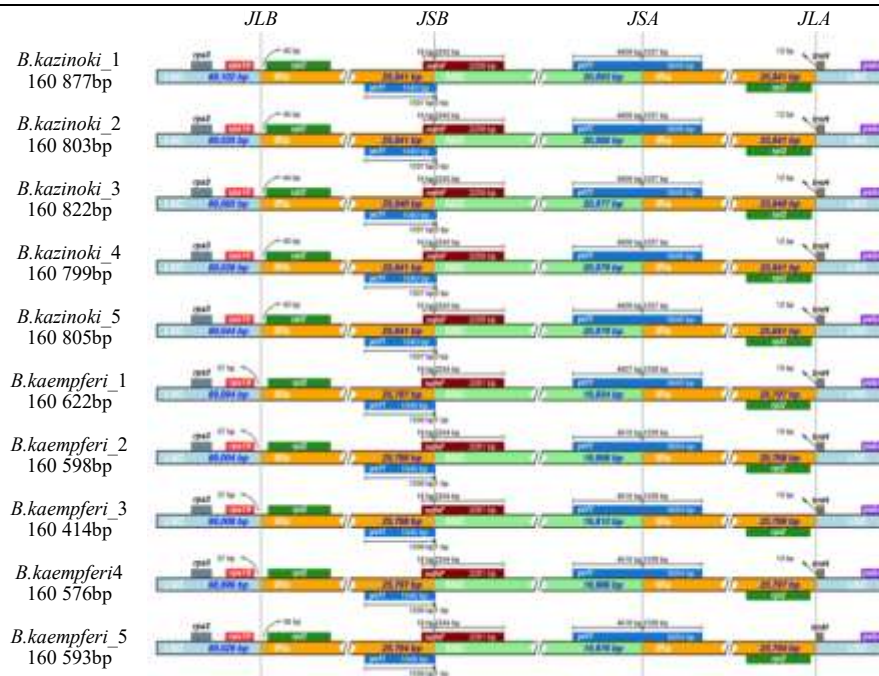


图 2 小构树和藤构叶绿体基因组边界分析

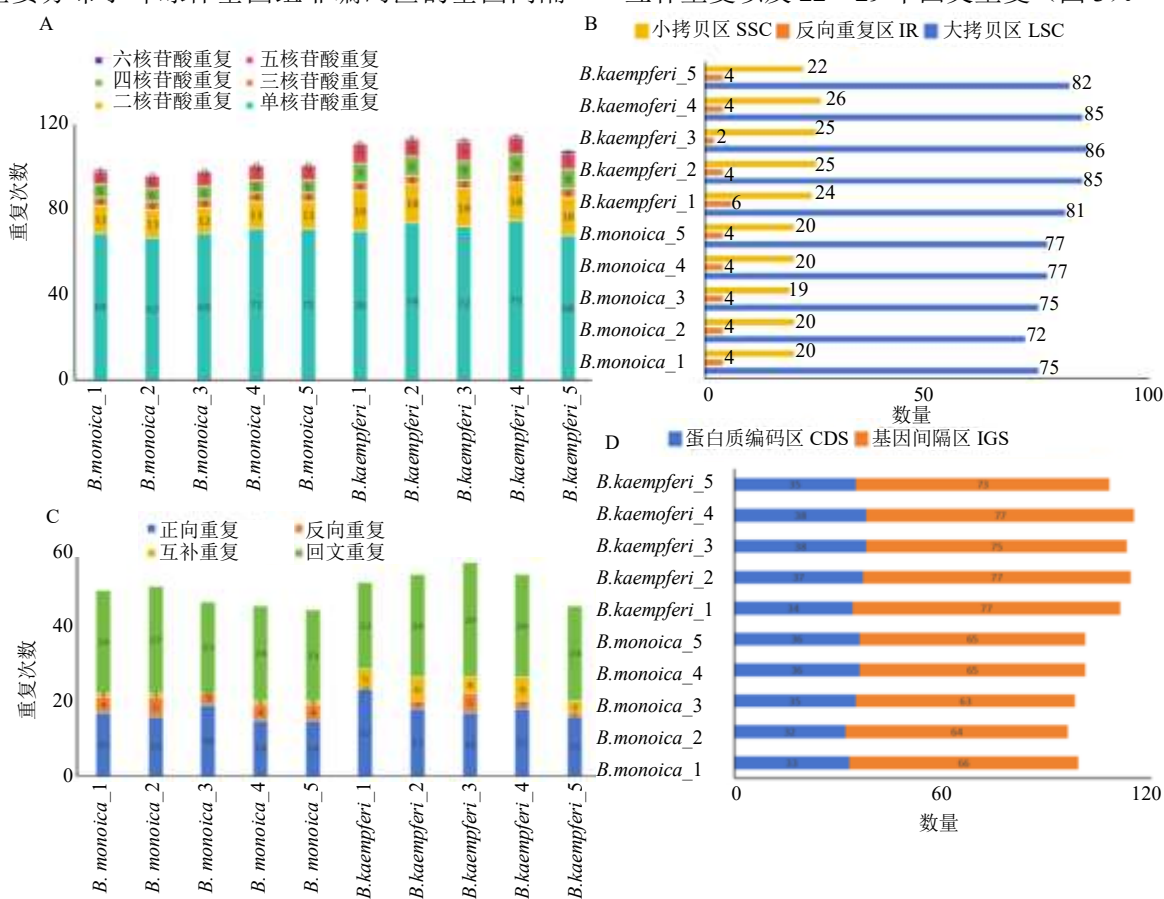
Fig. 2 Genomic boundary analysis of chloroplast genomes of *B. kazinoki* and *B. kaempferi*

检测到 1 个六核苷酸重复,而藤构中 *B.kaempferi_2*、*B.kaempferi_3*、*B.kaempferi_4* 和 *B.kaempferi_5* 均检测到 1~2 个六核苷酸重复。

小构树和藤构的 SSR 位点主要分布在 LSC 区,其次是 SSC 区,IR 区中最少。小构树的 SSR 位点有 72~77 个位于 LSC 区,19~20 个分布在 SSC 区,仅 4 个分布在 IR 区;藤构的 SSR 位点有 81~86 个位于 LSC 区,22~26 个分布在 SSC 区,仅 2~6 个分布在 IR 区。此外,小构树和藤构的 SSR 位点主要分布于叶绿体基因组非编码区的基因间隔

区。小构树中有 63~66 个位于基因间隔区,有 32~36 个位于编码区;藤构中有 73~77 个位于基因间隔区,有 34~38 个位于编码区。

利用 REPuter 对 5 个小构树个体和 5 个藤构个体叶绿体基因组进行重复序列分析。在小构树中鉴定到 42~48 个重复序列,包括 14~18 个正向重复;3~5 个反向重复;0~1 个互补重复以及 23~27 个回文重复。在藤构中鉴定到 43~54 个重复序列,包括 15~22 个正向重复;0~5 个反向重复;3~6 个互补重复以及 22~29 个回文重复(图 3)。



A-不同类型重复单元的数量; B-SSR 位点在叶绿体基因组 LSC、SSC 和 IR 的数量; C-不同类型重复序列的数量; D-SSR 位点在 CDS、IGS 区域的数量。

A-numbers of repetitive units in different types; B-numbers of SSRs in LSC (large single copy), SSC (small single copy), and IR (invert repeat) regions; C-numbers of repeat sequences in different types; D-numbers of SSRs in CDS (coding sequence) and IGS (intergenic spacers).

图 3 小构树和藤构叶绿体基因组的 SSR 分析和重复序列分析

Fig. 3 SSR analysis and repeat sequences analysis of the chloroplast genomes of *B. kazinoki* and *B. kaempferi*

3.4 核酸多态性分析

以 *B.monoica_1* 为参考序列,使用 mVISTA 软件对 5 个小构树个体和 5 个藤构个体叶绿体基因组进行序列多样性比较(图 4)。结果显示,种内序列相似性极高,种间序列相似度较高,10 条序列之间存在一定程度的差异。其中 LSC 区、SSC 区变异程

度明显高于 IR 区,非编码序列的变异高于编码区序列。

使用 DnaSP 6.0 软件分别计算了 5 个小构树个体和 5 个藤构个体叶绿体基因组序列和叶绿体基因组基因的 P_i 值(图 5)。小构树和藤构种内叶绿体基因组序列的 P_i 值变化范围分别为 0~0.021 和 0~0.026,

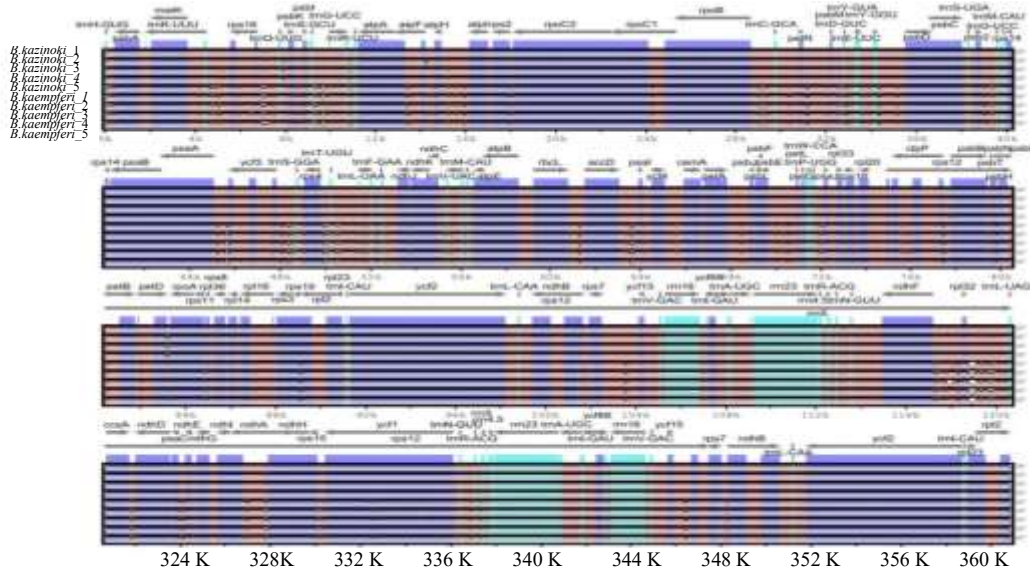
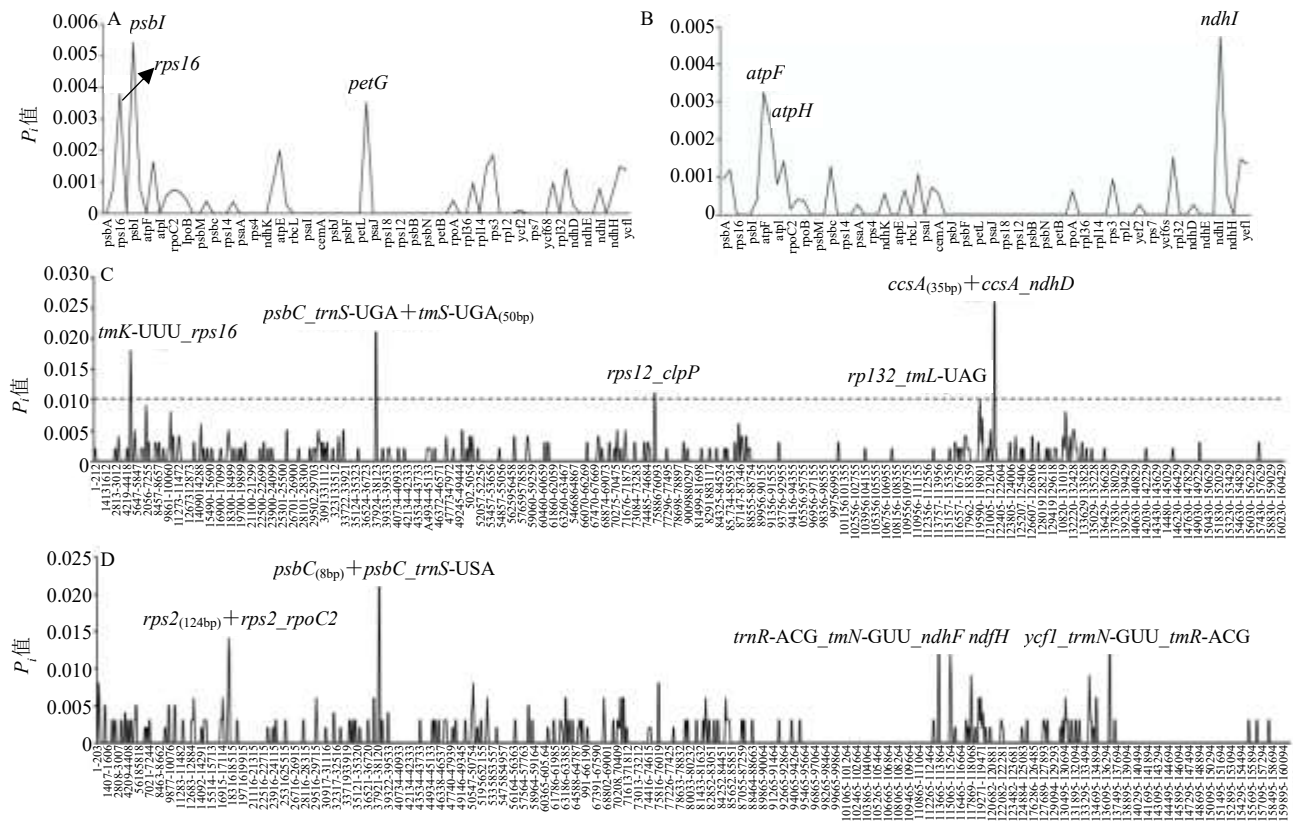


图4 小构树和藤构叶绿体基因组比较分析

Fig. 4 Comparative analysis of chloroplast genomes of *B. kazinoki* and *B. kaempferi*

A-小构树共有基因的核酸多态性分析; B-藤构共有基因的核酸多态性分析; C-小构树叶叶绿体基因组核酸多态性分析; D-藤构叶绿体基因组核酸多态性分析。

A-analysis of common genes nucleic acid polymorphism in *B. kazinoki*; B-analysis of common genes nucleic acid polymorphism in *B. kaempferi*; C-analysis of chloroplast genomic nucleic acid polymorphism in *B. kazinoki*; D-analysis of chloroplast genomic nucleic acid polymorphism in *B. kaempferi*.

图5 小构树和藤构叶绿体基因组和共有基因的核酸多态性分析

Fig. 5 Nucleic polymorphism analysis of chloroplast genomes and common genes for *B. kazinoki* and *B. kaempferi*

变异程度较高的位点主要位于 LSC 区和 SSC 区, IR 区相对保守, 这与图 4 结果一致。小构树和藤构种内分别筛选出 5 条 P_i 值大于 0.01 的高变异序列, 分别位于 LSC 区、SSC 区, 其中小构树中 *psbC* (8 bp) + *psbC_trnS-UGA* 区域的 P_i 值最高 (P_i 值为 0.21), 而藤构中 *ccsA* (35 bp) + *ccsA_ndhD* 区域 P_i 值最高 (P_i 值为 0.26)。小构树和藤构种内叶绿体基因组基因的 P_i 值变化范围分别为 0~0.004 73 和 0~0.005 41, 筛选出小构树中 *atpF*、*ndhF*、*ndhI* 基因和藤构中 *rps16*、*psbI*、*petG* 基因的 P_i 值大于 0.002。

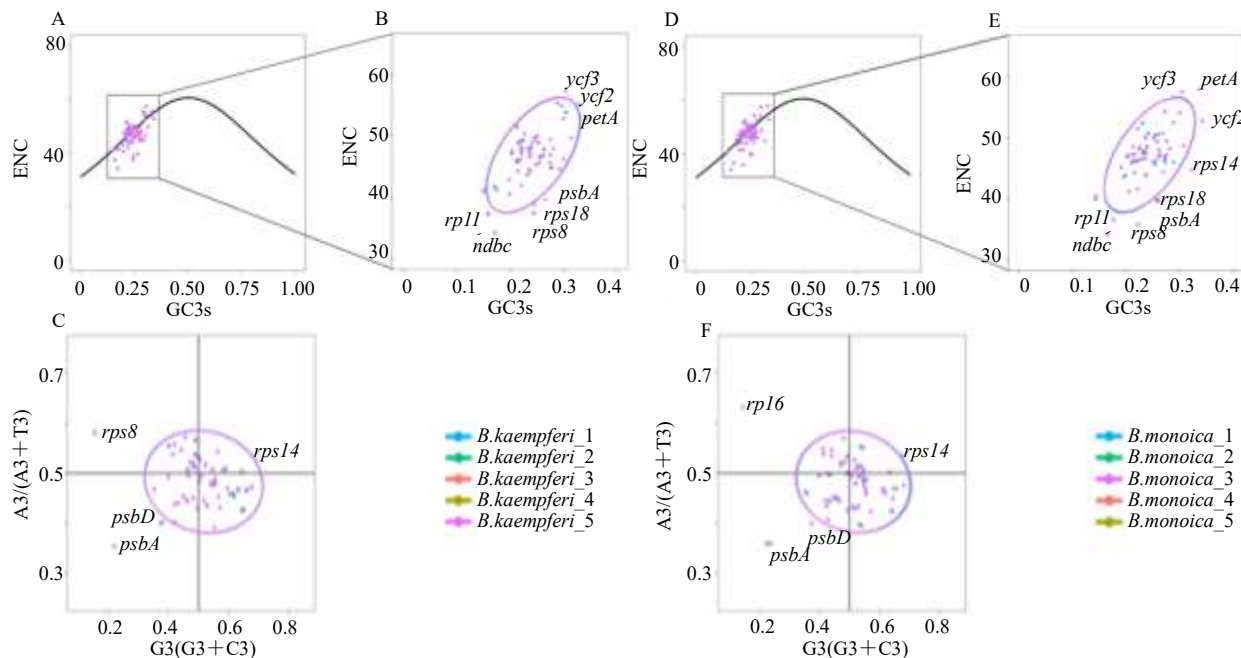
3.5 ENC-plot 分析

ENC 绘图分析 (图 6) 将 ENC 值和 GC3s 分布的关系可视化, 反映密码子偏好性受自然选择和突变的影响程度。结果显示, 5 个小构树和 5 个藤构的大部分基因未沿着标准曲线均匀分布, 小构树中有 51.9% 的基因分布在 -0.05~0.05, 与预期 ENC 值接近, 主要受突变的影响, 此区间以外 48.1% 的基因与预期 ENC 值存在一定差异, 主要受自然选择的影响; 藤构中有 50.8% 的基因分布在 -0.05~0.05, 与预期 ENC 值接近, 主要受突变的影响, 此区间以

外 49.2% 的基因与预期 ENC 值存在一定差异, 主要受自然选择的影响。说明小构树和藤构基因组的密码子偏好性主要受突变的影响, 而小构树受突变的影响更大。

3.6 PR2-plot 分析

PR2-plot 分析 (图 6) 以密码子第 3 位的核苷酸含量 A3、T3、C3、G3 为基础, 研究 A3/(A3+T3) 与 G3/(G3+C3) 的关系。以 A3 值等于 T3 值且 G3 值等于 C3 值的位置为中心, 通过各点与中心点的位置关系判断影响密码子的主要因素。PR2 分析结果显示, 5 个藤构和 5 个小构树叶绿体基因不均等的分布在 4 个区域内, 小构树中左上角区域共有 51 个基因, 左下角区域共有 69 个基因, 右上角区域共有 40 个基因, 右下角区域共有 100 个基因; 藤构中左上角区域共有 55 个基因, 左下角区域共有 65 个基因, 右上角区域和右下角区域与小构树相同分别有 40 个和 100 个基因。右下角区域基因分布较多, 即密码子第 3 碱基位置的碱基使用频率为 G>C、T>A, 其密码子偏好使用 T 和 G。小构树和藤构叶绿体基因组的密码子使用偏好性受



A-藤构叶绿体基因组密码子 ENC-plot 分析; B-藤构叶绿体基因组密码子偏离预期值的具体基因分析; C-藤构叶绿体基因组密码子 PR2-plot 分析; D-小构树叶绿体基因组密码子 ENC-plot 分析; E-小构树叶绿体基因组密码子偏离预期值的具体基因分析; F-小构树叶绿体基因组密码子 PR2-plot 分析。

A-ENC-plot analysis of chloroplast genome codons of *B. kaempferi*; B-analysis of specific genes with codon usage bias from the expected values in the chloroplast genome of *B. kaempferi*; C-PR2-plot analysis of chloroplast genome codons of *B. kaempferi*; D-ENC-plot analysis of chloroplast genome codons of *B. kazinoki*; E-analysis of specific genes with codon usage bias from the expected values in the chloroplast genome of *B. kazinoki*; F-PR2-plot analysis of chloroplast genome codons of *B. kazinoki*.

图 6 小构树和藤构叶绿体基因组密码子的 ENC-plot 分析和 PR2-plot 分析

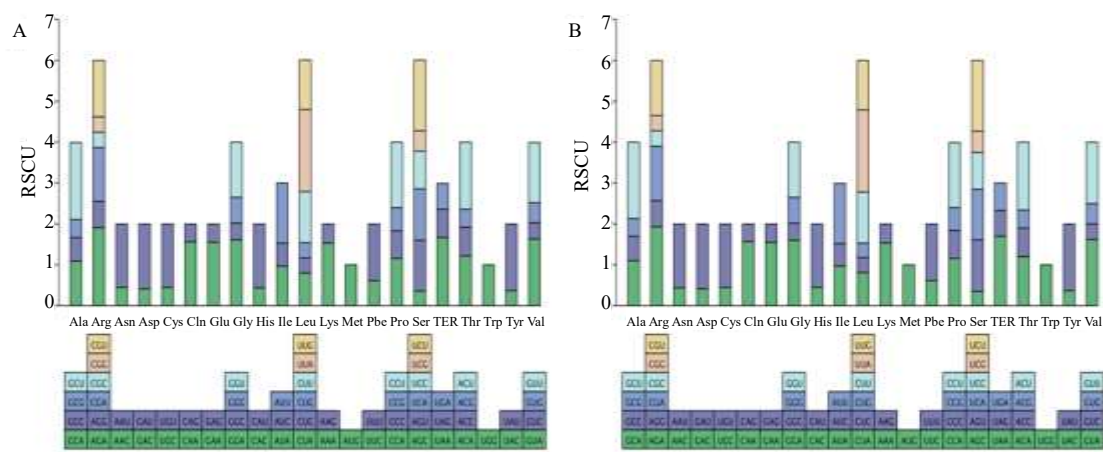
Fig. 6 ENC-plot analysis and PR2-plot analysis of chloroplast genome codons for *B. kazinoki* and *B. kaempferi*

自然选择和突变的双重影响，而基因偏向右下角区域的分布，也说明其受到突变的影响为主，自然选择的影响为辅。

3.7 最优密码子分析

分别以 5 个小构树个体和 5 个藤构个体中的相同基因为一个单位，将所有基因根据 ENC 值从高分到低分排列，选取两端占比为总数 5% 的基因建立高、低表达组。计算这 2 个基因库的密码子 RSCU 值和 ΔRSCU 值 ($\Delta\text{RSCU} = \text{RSCU}_{\text{高表达}} - \text{RSCU}_{\text{低表达}}$) [35]。RSCU 值大于 1 的密码子被定义为高频密码子 [36]，小构树和藤构中筛选出 30 个相同的高频密码子，其中 29 个以 A/U 结尾，仅 1 个以 G 结尾

(图 7)。小构树中 $\Delta\text{RSCU} \geq 0.08$ 的密码子有 28 个，有 10 个以 C 结尾，12 个以 G 结尾，4 个以 U 结尾，仅 2 个以 A 结尾；藤构有 $\Delta\text{RSCU} \geq 0.08$ 的密码子有 29 个，有 12 个以 C 结尾，11 个以 G 结尾，4 个以 U 结尾，仅 2 个以 A 结尾。其中 $\text{RSCU} > 1$ 且 $\Delta\text{RSCU} \geq 0.08$ 的密码子为最优密码子 [37]，小构树和藤构中筛选出 7 个共同的最优密码子，即 UUG、UCA、CCU、UAA、AAU、GAU 和 CGU，有 4 个密码子以 U 结尾，2 个密码子以 A 结尾，仅有 1 个密码子以 G 结尾，没有以 C 结尾的密码子 (表 4)。表明 5 个小构树和 5 个藤构的叶绿体基因组密码子偏好以 A/U 结尾。



A-小构树叶绿体基因组 RSCU 值；B-藤构叶绿体基因组 RSCU 值。
A-chloroplast genome RSCU value of *B. kazinoki*; B-chloroplast genome RSCU value of *B. kaempferi*.

图 7 小构树和藤构叶绿体基因组密码子分析

Fig. 7 Analysis of chloroplast genome codons of *B. kazinoki* and *B. kaempferi*

表 4 小构树和藤构叶绿体基因组密码子分析

Table 4 Codon analysis of chloroplast genomes for *B. kazinoki* and *B. kaempferi*

物种	密码子	氨基酸	RSCU	数量	RSCU高表达	数量	RSCU低表达	数量	ΔRSCU
小构树	Leu	UUG	1.21	2 297	1.53	525	1.03	30	0.50
	Ser	UCA	1.24	1 638	1.32	420	1.20	30	0.12
	Pro	CCU	1.59	1 734	1.42	280	0.89	20	0.53
	Ter	UAA	1.70	147	2.00	10	1.00	5	1.00
	Asn	AAU	1.56	4 208	1.52	900	1.18	50	0.34
	Asp	GAU	1.59	3 449	1.62	835	1.11	25	0.51
	Arg	CGU	1.35	1 390	0.74	155	0.46	15	0.28
藤构	Leu	UUG	1.21	2 291	1.52	520	0.88	25	0.64
	Ser	UCA	1.26	1 659	1.30	410	1.20	30	0.10
	Pro	CCU	1.60	1 739	1.44	285	0.89	20	0.55
	Ter	UAA	1.67	145	2.00	10	1.00	5	1.00
	Asn	AAU	1.55	4 205	1.52	895	1.18	50	0.34
	Asp	GAU	1.59	3 442	1.63	840	1.11	25	0.52
	Arg	CGU	1.38	1 431	0.76	160	0.63	20	0.13

3.8 系统发育分析

系统发育分析结果显示, 疏桑族内部构属、落叶花桑属和牛筋藤属各自构成了一个单系支。构属内部中 5 个藤构个体和 5 个小构树个体各自形成单系支, 其支持率均达 100%, 且与构树完全区分。与构树比, 小构树与藤构一起形成一个单系支, 具有更近的亲缘关系; 以枝长看, 与小构树比, 藤构和构树的亲缘关系更近。小构树中, *B.monoica*_1 (广西龙胜) 和 *B.monoica*_2 (贵州铜仁) 聚为一支, 支持率 100%;

*B.monoica*_3 (浙江宁波)、*B.monoica*_4 (安徽池州) 和 *B.monoica*_5 (江苏南京) 聚为另一支, 其中 *B.monoica*_4 和 *B.monoica*_5 又聚为次级分支, 支持率都为 100%。藤构中, *B.kaempferi*_5 (浙江宁波) 形成单独分支, 其余 *B.kaempferi*_2 (广西龙胜)、*B.kaempferi*_3 (贵州雷山) 和 *B.kaempferi*_4 (贵州铜仁) 等 3 个个体形成单系支 (支持率 100%), 并和 *B.kaempferi*_1 (广东龙胜) 个体构成姐妹支, 它们一起形成的单系支整体支持率为 84%。

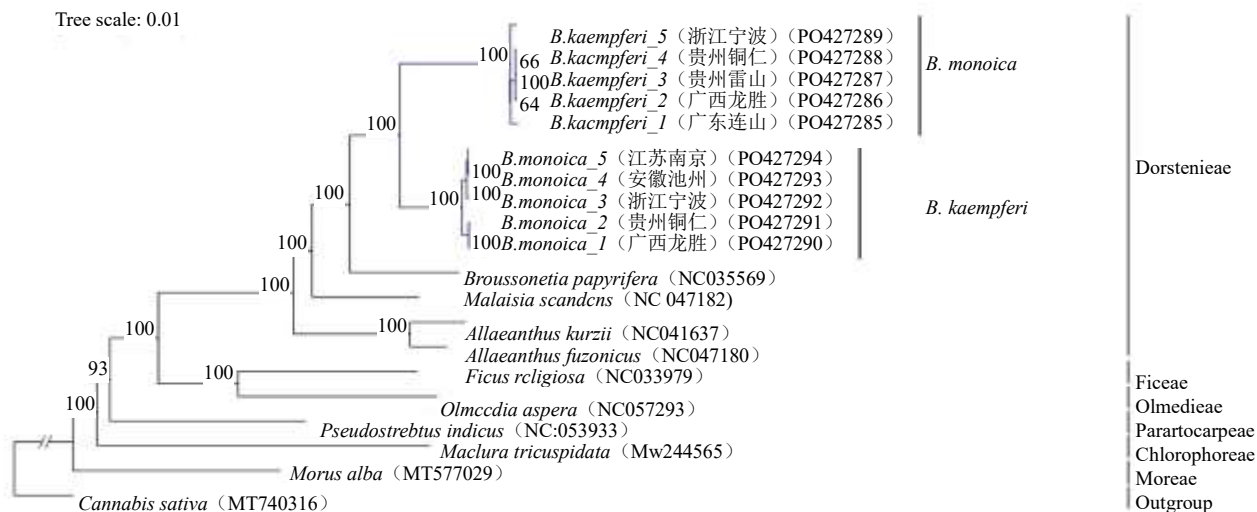


图 8 基于叶绿体全基因组构建的系统发育树

Fig. 8 Phylogenetic tree based on complete chloroplast genome

4 讨论

本研究通过对各 5 个产地的小构树和藤构个体进行高通量基因组重测序, 成功组装并注释了它们完整的叶绿体全基因组, 并进行比较分析。小构树和藤构的叶绿体基因组结构典型且保守, 序列全长分别为 160 799~160 877 bp 和 160 414~160 622 bp, 与已报道的构属植物叶绿体基因组相符^[21, 38]。这一发现表明, 尽管 2 个物种分布在不同地理区域, 叶绿体基因组在种内和种间的变异较少, 尤其是在小构树个体间的差异更为微小, 这也印证了叶绿体基因组在维持植物关键生命活动中的结构保守性。在叶绿体基因组的 GC 含量分析中, 小构树与藤构之间的 GC 含量非常接近, 尤其在 IR 区、LSC 区和 SSC 区的 GC 含量分布几乎一致, 这支持了叶绿体基因组在构属植物中序列组成高度保守的特性。叶绿体基因组的保守性主要是由于其在光合作用中的重要作用, 导致其在进化过程中受到强烈的功能约束, 变异速率较低^[39]。这种现象在构属植物中得到充分体现。此外, 小构树和藤构均注释出 135

个功能基因, 与已报道的构属植物叶绿体基因组注释结果一致^[21, 38]。IR 边界分析显示 10 个叶绿体基因组序列四分体的边界基因较为保守, 扩张收缩现象不明显。这种边界保守性是许多植物中叶绿体基因组的一般特征, 反映了叶绿体基因组在植物功能和结构上的净化选择压力。

SSR 是一段短而串联的 DNA 序列, 大小为 1~6 bp, 具有多态性高、覆盖整个基因组、稳定性好等特点, 可以有效反映种群内部的遗传多样性, 常作为分子标记在种群遗传结构分析、遗传多样性评估和物种鉴定中具有重要应用价值^[40-41]。在小构树和藤构的叶绿体基因组中分别鉴定出 98~101 个和 108~115 个 SSR 位点, 其中单核苷酸重复序列重复次数最多 (占 62.96%~70.41%), 且主要为 A/T 重复, 其次是二核苷酸重复序列 (占 12.24%~17.12%), 其余 SSR 也多由 A/T 碱基组成。这与榕属 *Ficus* L.^[42]、桑属 *Morus* L.^[43] 等桑科植物叶绿体基因组 SSR 主要由 A/T 碱基构成的研究结果一致。小构树中只有 *B.monoica*_1 中检测到 1 个六核苷酸

重复,而藤构中有 4 个个体检测到 1~2 个六核苷酸重复。此外,非编码区的 SSR 数量远多于编码区,这表明非编码区通常比编码区具有更高的变异率,反映出更高的遗传多样性,这可能是由于非编码区面临较小的选择压力^[44]。长重复序列在基因组结构变异和进化中起关键作用^[45],小构树和藤构叶绿体基因组中检测到 42~54 个 4 种类型的长重复序列,其中主要是正向重复(占 29.63%~44.90%)和回文重复(占 44.90%~56.25%),而反向重复和互补重复仅占总数的 6.82%~16.67%。正向重复和回文重复在长重复序列中占据主要比例,可能与叶绿体基因组的复制和修复机制有关。这些重复序列的存在可能影响基因组的重组和稳定性,进而影响基因组的结构和功能。本研究丰富了构属植物 SSR 位点和长重复序列研究,揭示了构属植物在遗传多样性、基因组结构保守性及进化机制等方面的重要特征。这不仅为构属植物的分子鉴定、群体遗传学研究提供了丰富的分子标记资源,也为进一步的系统发育和进化研究奠定了坚实的基础。

mVISTA 软件能够直观地展示不同样本或物种之间的基因组序列相似性和差异性^[46],DnaSP 6.0 帮助确定哪些基因或区域具有较高的变异性^[32]。通过 mVISTA 软件和 DnaSP 6.0 软件分析发现,小构树和藤构种内的叶绿体基因组序列高度相似,说明同种个体间基因组稳定性较高;而种间存在差异,支持它们在分类上的独立性。10 个个体的 IR 区较 LSC、SSC 区更为保守,非编码序列变异高于编码基因序列,这与 SSRs 位点分布远多于编码区的结果一致。此外,小构树中的 *psbC* (8 bp) + *psbC_trnS*-UGA 区域以及 *atpF*、*ndhF*、*ndhI* 基因和藤构中的 *ccsA* (35 bp) + *ccsA_ndhD* 区域以及 *rps16*、*psbI*、*petG* 基因变异度最大。这些变异位点可以开发为潜在的种内 DNA 条形码,辅助构属植物种内的种源鉴定,并在种群遗传研究中提供重要的分子依据。

ENC-plot 和 PR2-plot 分析的结果表明,小构树和藤构叶绿体基因组的密码子使用偏好性受突变和自然选择这 2 种机制的影响,但突变占主导地位。密码子偏好性分析分别得到小构树和藤构叶绿体基因组中相同的 30 个高频密码子,其中有 29 个以 A/U 结尾,有 1 个以 G 结尾,与大多数被子植物叶绿体基因组类似,都偏好使用 A/U 结尾的密码子^[47-49]。并且在小构树和藤构中筛选出了 7 个共同的最优密码子

(UUG、UCA、CCU、UAA、AAU、GAU 和 CGU)。由此阐明小构树和藤构叶绿体基因组密码子使用模式及偏好性。

为探讨小构树和藤构之间以及它们与桑科其他物种的系统发育关系,基于本研究获得的 10 个个体与桑科 6 个族中的 9 个个体叶绿体基因组构建了系统发育树。研究结果进一步支持了构属、牛筋藤属和落叶花桑属在桑科中的分类地位,并且揭示了小构树与藤构的亲缘关系较构树更近(图 8)。从形态上看,构属的圆柱形或球形雄性花序可以与落叶花桑属的细长柔荑花序和牛筋藤属的穗状花序或圆锥花序区分开来。在构属内部,小构树与藤构在形态特征上较构树更为相似,其中小构树和藤构均为灌木,而构树为乔木。本研究深化了桑科植物的系统发育认识,也展示了叶绿体基因组研究与传统形态学分类的互补性,为未来桑科植物群体分类、种群遗传多样性研究提供了新的数据和理论依据,尤其是在构属植物的系统进化方面,为相关研究提供了更多的基础资源。此外,小构树的 5 个个体根据地理位置呈现出一定的聚类现象:南部地区(广西龙胜、贵州铜仁)的个体聚为一支,而分布在长江下游北部地区(浙江宁波、安徽池州、江苏南京)的个体聚为另一支。藤构的个体也表现出相似的地理相关性,浙江宁波的 *B.kaempferi*_5 形成单独分支,而其余位于较南地区且地理位置相互接近的 4 个个体聚为另一支(虽然在系统发育树上支持率仅 84%)。这表明,小构树和藤构的亲缘关系与其地理分布,尤其是南北地理区划有显著关联,推测长江以南的山系或地理障碍在推动其种内亚支系分化中起到了重要作用。这种因地理隔离导致的种内亚支系分化现象在生物地理学和种群遗传结构研究中已有广泛印证^[50-51]。地理障碍,如山脉和河流,通常被认为是限制基因流动、促进局部种群分化的关键因素^[52-53]。在小构树和藤构的进化历程中,这类地理隔离效应可能是推动它们形成不同亚支的重要驱动力。因此,进一步结合地理和生态数据,探讨地理隔离对它们系统发育和种群结构的影响,将有助于加深对构属植物进化机制的理解。

本研究通过各 5 个产地的小构树和藤构个体叶绿体基因组进行测序、组装和注释,阐述了这 2 种构属植物的叶绿体基因组序列特征以及系统发育关系,对构属植物的遗传多样性以及系统发育研究进行补充,有助于后续构属植物的分子鉴定、系统

进化研究以及药用资源综合开发利用。本研究用小构树和藤构样本未涵盖更广泛的地理分布区域,且研究主要聚焦于叶绿体基因组的基因注释和变异分析,未涉及其表型变异、核基因组研究以及药用代谢组分的研究。未来研究应扩大地理采样范围,从更广泛的区域,特别是边缘分布区和地理隔离区收集小构树和藤构的种群样本,并加强小构树和藤构的核基因组和不同产地的表型变异研究,以深入探讨不同生态环境下的遗传多样性、适应性变异遗传机制及药用代谢组分特征与合成通路等。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Zhang Q, Sodmergen. Why does biparental plastid inheritance revive in angiosperms? [J]. *J Plant Res*, 2010, 123(2): 201-206.
- [2] Daniell H, Lin C S, Yu M, *et al.* Chloroplast genomes: Diversity, evolution, and applications in genetic engineering [J]. *Genome Biol*, 2016, 17(1): 134.
- [3] Zhang X, Gu C R, Zhang T X, *et al.* Chloroplast (Cp) transcriptome of *P. davidiana* Dode $\times$ *P. bolleana* Lauch provides insight into the Cp drought response and *Populus* Cp phylogeny [J]. *BMC Evol Biol*, 2020, 20(1): 51.
- [4] 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.
- [5] 张韵洁, 李德铎. 叶绿体系统发育基因组学的研究进展 [J]. *植物分类与资源学报*, 2011, 33(4): 365-375.
- [6] Bock R, Knoop V. *Genomics of chloroplasts and mitochondria* [M]. Dordrecht: Springer Netherlands, 2012.
- [7] Drouin G, Daoud H, Xia J N. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants [J]. *Mol Phylogenet Evol*, 2008, 49(3): 827-831.
- [8] Fu C N, Wu C S, Ye L J, *et al.* Prevalence of isomeric plastomes and effectiveness of plastome super-barcodes in yews (*Taxus*) worldwide [J]. *Sci Rep*, 2019, 9(1): 2773.
- [9] Liu C K, Lei J Q, Jiang Q P, *et al.* The complete plastomes of seven *Peucedanum* plants: Comparative and phylogenetic analyses for the *Peucedanum* genus [J]. *BMC Plant Biol*, 2022, 22(1): 101.
- [10] Chung K F, Kuo W H, Hsu Y H, *et al.* Molecular recircumscription of *Broussonetia* (Moraceae) and the identity and taxonomic status of *B. kaempferi* var. *australis* [J]. *Bot Stud*, 2017, 58(1): 11.
- [11] 郑汉臣, 黄宝康, 秦路平, 等. 构树属植物的分布及其生物学特性 [J]. *中国野生植物资源*, 2002, 21(6): 11-13.
- [12] 庞素秋, 黄宝康, 张巧艳, 等. 构树属植物的化学成分及药理作用研究进展 [J]. *药学服务与研究*, 2006, 6(2): 98-102.
- [13] 黄少鹏. 构树宁 E 抑制 LPS 诱导巨噬细胞炎症反应的新功能及机制研究 [D]. 广州: 广东药科大学, 2019.
- [14] 贾东辉, 杨雪莹. 构树叶中黄酮成分分析和抗氧化活性的测定 [J]. *职业与健康*, 2006, 22(17): 1352-1353.
- [15] 王亨, 杨雪莹, 何瑞, 等. 构树总黄酮对长波紫外线引起人角质形成细胞损伤的防护作用 [J]. *中华劳动卫生职业病杂志*, 2005, 23(6): 442-444.
- [16] 渠桂荣, 张倩, 李彩丽. 构树的药理与临床作用研究述略 [J]. *中医药学刊*, 2003, 21(11): 1810-1811.
- [17] Lin C N, Lu C M, Lin H C, *et al.* Novel antiplatelet constituents from Formosan moraceous plants [J]. *J Nat Prod*, 1996, 59(9): 834-838.
- [18] Ryu H W, Park M H, Kwon O K, *et al.* Anti-inflammatory flavonoids from root bark of *Broussonetia papyrifera* in LPS-stimulated RAW_{264.7} cells [J]. *Bioorg Chem*, 2019, 92: 103233.
- [19] Kuo W H, Liu S H, Chang C C, *et al.* Plastome phylogenomics of *Allaeanthus*, *Broussonetia* and *Malaisia* (Dorsteniaceae, Moraceae) and the origin of *B. $\times$ kazinoki* [J]. *J Plant Res*, 2022, 135(2): 203-220.
- [20] Zhou L N, Yuan L X, Li P, *et al.* Comparative analyses of plastomes in *Allaeanthus* and *Malaisia*: Structure, evolution, and phylogeny [J]. *Sci Rep*, 2024, 14(1): 22686.
- [21] Yang J H, Chu Q, Meng G, *et al.* The complete chloroplast genome sequences of three *Broussonetia* species and comparative analysis within the Moraceae [J]. *PeerJ*, 2022, 10: e14293.
- [22] Peng X J, Liu H, Chen P L, *et al.* A chromosome-scale genome assembly of paper mulberry (*Broussonetia papyrifera*) provides new insights into its forage and papermaking usage [J]. *Mol Plant*, 2019, 12(5): 661-677.
- [23] Hu Y M, Peng X J, Wang F F, *et al.* Natural population resequencing detects the genetic basis of local adaptation to low temperature in a woody plant [J]. *Plant Mol Biol*, 2021, 105(6): 585-599.
- [24] Hu Y M, Tang F, Zhang D, *et al.* Integrating genome-wide association and transcriptome analysis to provide molecular insights into heterophylly and eco-adaptability in woody plants [J]. *Hortic Res*, 2023, 10(11): uhad212.
- [25] Jin J J, Yu W B, Yang J B, *et al.* GetOrganelle: A fast and versatile toolkit for accurate *de novo* assembly of organelle genomes [J]. *Genome Biol*, 2020, 21(1): 241.
- [26] Shi L C, Chen H M, Jiang M, *et al.* CPGAVAS2, an integrated plastome sequence annotator and analyzer [J]. *Nucleic Acids Res*, 2019, 47(W1): W65-W73.
- [27] Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes [J]. *Nucleic Acids Res*, 2019, 47(W1): W59-W64.
- [28] Beier S, Thiel T, Münch T, *et al.* MISA-web: A web server

- for microsatellite prediction [J]. *Bioinformatics*, 2017, 33(16): 2583-2585.
- [29] Benson G. Tandem repeats finder: A program to analyze DNA sequences [J]. *Nucleic Acids Res*, 1999, 27(2): 573-580.
- [30] Amiryousefi A, Hyvönen J, Poczai P. IRscope: An online program to visualize the junction sites of chloroplast genomes [J]. *Bioinformatics*, 2018, 34(17): 3030-3031.
- [31] Frazer K A, Pachter L, Poliakov A, *et al.* VISTA: Computational tools for comparative genomics [J]. *Nucleic Acids Res*, 2004, 32(Web Server issue): W273-W279.
- [32] Rozas J, Ferrer-Mata A, Sánchez-DelBarrio J C, *et al.* DnaSP 6: DNA sequence polymorphism analysis of large data sets [J]. *Mol Biol Evol*, 2017, 34(12): 3299-3302.
- [33] 张洁, 陆嘉惠, 王倩倩, 等. 新疆产黄甘草叶绿体基因组特征及甘草属药用植物叶绿体基因组比较与系统发育分析 [J]. *药学报*, 2022, 57(5): 1516-1525.
- [34] Nguyen L T, Schmidt H A, von Haeseler A, *et al.* IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies [J]. *Mol Biol Evol*, 2015, 32(1): 268-274.
- [35] 唐晓芬, 陈莉, 李玉瑶. 密码子使用偏性量化方法研究综述 [J]. *基因组学与应用生物学*, 2013, 32(5): 660-666.
- [36] 林涛, 倪志华, 沈明山, 等. 高频密码子分析法及其在烟草密码子分析中的应用 [J]. *厦门大学学报: 自然科学版*, 2002, 41(5): 551-554.
- [37] Sharp P M, Li W H. The *Codon* adaptation index: A measure of directional synonymous *Codon* usage bias, and its potential applications [J]. *Nucleic Acids Res*, 1987, 15(3): 1281-1295.
- [38] Xu Z G, Yang G Y, Dong M, *et al.* The complete chloroplast genome of an economic and ecological plant, paper mulberry (*Broussonetia kazinoki* × *Broussonetia papyrifera*) [J]. *Mitochondrial DNA B Resour*, 2017, 3(1): 28-29.
- [39] Wicke S, Schneeweiss G M, DePamphilis C W, *et al.* The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function [J]. *Plant Mol Biol*, 2011, 76(3/4/5): 273-297.
- [40] Powell W, Morgante M, McDevitt R, *et al.* Polymorphic simple sequence repeat regions in chloroplast genomes: Applications to the population genetics of pines [J]. *Proc Natl Acad Sci USA*, 1995, 92(17): 7759-7763.
- [41] Yang H S, Li X P, Liu D J, *et al.* Genetic diversity and population structure of the endangered medicinal plant *Phellodendron amurense* in China revealed by SSR markers [J]. *Biochem Syst Ecol*, 2016, 66: 286-292.
- [42] Zhang Z, Zhang D S, Zou L, *et al.* Comparison of chloroplast genomes and phylogenomics in the *Ficus sarmentosa* complex (Moraceae) [J]. *PLoS One*, 2022, 17(12): e0279849.
- [43] Kong W Q, Yang J H. The complete chloroplast genome sequence of *Morus mongolica* and a comparative analysis within the Fabidae clade [J]. *Curr Genet*, 2016, 62(1): 165-172.
- [44] Shaw J, Lickey E B, Schilling E E, *et al.* Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III [J]. *Am J Bot*, 2007, 94(3): 275-288.
- [45] Asaf S, Khan A L, Khan M A, *et al.* Complete chloroplast genome sequence and comparative analysis of loblolly pine (*Pinus taeda* L.) with related species [J]. *PLoS One*, 2018, 13(3): e0192966.
- [46] De Carolis L, Marsico E, Arnaud V, *et al.* Assessing sound symbolism: Investigating phonetic forms, visual shapes and letter fonts in an implicit bouba-kiki experimental paradigm [J]. *PLoS One*, 2018, 13(12): e0208874.
- [47] Pechmann S, Frydman J. Evolutionary conservation of *Codon* optimality reveals hidden signatures of cotranslational folding [J]. *Nat Struct Mol Biol*, 2013, 20(2): 237-243.
- [48] Huang Y, Jin X J, Zhang C Y, *et al.* Plastome evolution of *Engelhardia* facilitates phylogeny of Juglandaceae [J]. *BMC Plant Biol*, 2024, 24(1): 634.
- [49] Li C L, Zhou L, Nie J B, *et al.* Codon usage bias and genetic diversity in chloroplast genomes of *Elaeagnus* species (Myrtiflorae: Elaeagnaceae) [J]. *Physiol Mol Biol Plants*, 2023, 29(2): 239-251.
- [50] Payacan C, Moncada X, Rojas G, *et al.* Phylogeography of herbarium specimens of asexually propagated paper mulberry [*Broussonetia papyrifera* (L.) L'Hér. Ex Vent. (Moraceae)] reveals genetic diversity across the Pacific [J]. *Ann Bot*, 2017, 120(3): 387-404.
- [51] Zhou L N, Yuan L X, Li P, *et al.* Comparative analyses of plastomes in *Allaeanthus* and *Malaisia*: Structure, evolution, and phylogeny [J]. *Sci Rep*, 2024, 14(1): 22686.
- [52] Tóth E G, Tremblay F, Housset J M, *et al.* Geographic isolation and climatic variability contribute to genetic differentiation in fragmented populations of the long-lived subalpine conifer *Pinus cembra* L. in the western Alps [J]. *BMC Evol Biol*, 2019, 19(1): 190.
- [53] Jiao X L, Wu L, Zhang D Z, *et al.* Landscape heterogeneity explains the genetic differentiation of a forest bird across the Sino-Himalayan mountains [J]. *Mol Biol Evol*, 2024, 41(3): msae027.