

## 黄槽毛竹叶绿体基因组及毛竹种下分类群的 叶绿体基因组序列比较

刘 萱, 邹龙海, 周明兵

(浙江农林大学 省部共建亚热带森林培育国家重点实验室/竹子研究院, 浙江 杭州 311300)

**摘要:** 【目的】对黄槽毛竹 *Phyllostachys edulis* f. *luteosulcata* 叶绿体基因组进行组装、注释和分析, 并与其他毛竹 *Ph. edulis* 种下分类群比较叶绿体遗传信息差异和系统进化关系。【方法】利用高通量二代测序数据组装了黄槽毛竹的叶绿体基因组序列并进行了基因结构注释, 通过生物信息学分析软件进行其组成、密码子偏好性、重复序列等分析。通过序列比对和系统进化分析, 比较不同毛竹种下分类群的系统进化关系和基因组序列差异。【结果】黄槽毛竹的叶绿体基因组是长度为 139 678 bp, 包含 132 个基因的双环 DNA; 包含蛋白质编码基因 85 个、核糖体 RNA (rRNA) 8 个和转运 RNA (tRNA) 39 个。该基因组的最优密码子偏好使用以 A/U 碱基结尾, 包含 49 个重复序列、55 个简单重复序列 (SSR) 位点, 其中简单重复序列最多的类型为 A/T。利用叶绿体基因组序列构建的系统发育分析显示: 黄槽毛竹与其他毛竹种下分类群共同组成单系分支, 且与毛竹原变种 *Ph. edulis* var. *pubescens* 亲缘关系最近。基于 7 个毛竹种下分类群的叶绿体基因组序列和编码基因特征分析显示: 毛竹种下分类群之间存在着编码基因数量和结构差异, 编码区和非编码区存在较低程度序列变异。【结论】首次对毛竹种下分类群的叶绿体基因组进行比较分析, 并揭示了这些种下分类群存在着一定程度的序列差异。这些变异资料可以用于毛竹种下分类群的鉴定比较。图 5 表 2 参 27

**关键词:** 黄槽毛竹; 叶绿体基因组; 系统进化; 序列变异

中图分类号: Q753; S795.7 文献标志码: A 文章编号: 2095-0756(2024)05-1037-10

## Chloroplast genome of *Phyllostachys edulis* f. *luteosulcata* and comparison of chloroplast genome sequence of subspecies of *Ph. edulis*

LIU Xuan, ZOU Longhai, ZHOU Mingbing

(State Key Laboratory of Subtropical Silviculture/Bamboo Industry Institute, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China)

**Abstract:** [Objective] This study aims to sequence, assemble, annotate, and analyze the chloroplast genome of *Phyllostachys edulis* f. *luteosulcata*. The research also involves comparing its chloroplast genetic information and phylogenetic relationships with those of other subspecies of *Ph. edulis* (moso bamboo). [Method] High-throughput sequencing data were used to assemble and annotate the complete chloroplast genome of *Ph. edulis* f. *luteosulcata*. Subsequently, we analyze the composition, codon preference, and repetitive sequences of the genome. Furthermore, sequence comparison and phylogenetic analysis were conducted to compare the phylogenetic relationships and genome sequence differences among different subspecies of moso bamboo. [Result] The chloroplast genome of *Ph. edulis* f. *luteosulcata* is a double-loop DNA of 139 678 bp in length containing 132 genes, including 85 protein-coding genes, eight ribosomal RNAs (rRNAs), and 39 transfer RNAs (tRNAs). The codon preference for this genome has an A/U base at the end. There are 49 repetitive

收稿日期: 2024-01-06; 修回日期: 2024-05-08

基金项目: 国家自然科学基金资助项目 (31870656); 浙江省自然科学基金重点项目 (LZ19C160002)

作者简介: 刘萱 (ORCID: 0000-0003-4546-9331), 从事毛竹生长发育研究。E-mail: liuxuan@stu.zafu.edu.cn。通信作者: 周明兵 (ORCID: 0000-0001-5674-4410), 教授, 博士, 从事毛竹生长发育研究。E-mail: zhoumingbing@zafu.edu.cn

sequences with the most common type being A/T and 55 Simple Sequence Repeat (SSR) sites. Phylogenetic analyses constructed using the chloroplast genome sequences showed that *Ph. edulis* f. *luteosulcata* is in a monophyletic branch together with other subspecies of *Ph. edulis* and is closely related to the original variety of *Ph. edulis* var. *pubescens*. The analysis of chloroplast genome sequence and coding gene characteristics showed that there were differences in the number and structure of coding genes and low degree of sequence variation among the subspecies of *Ph. edulis*. [Conclusion] This study is the first to comparatively analyze the chloroplast genomes of subspecies of *Ph. edulis* and reveals a degree of sequence variation in these subspecies of *Ph. edulis*. This variation information would be available for the identification and comparison of subspecies of *Ph. edulis*. [Ch, 5 fig. 2 tab. 27 ref.]

**Key words:** *Phyllostachys edulis* f. *luteosulcata*; chloroplast genome; phylogenetics; sequence variation

毛竹 *Phyllostachys edulis* 为竹亚科 Bambusoideae 刚竹属 *Phyllostachys* 的散生竹种, 广泛分布于东亚和东南亚, 其种植面积高达 460 多万  $\text{hm}^2$ <sup>[1]</sup>。由于生长速度快并且富含纤维素和半纤维素, 毛竹是中国重要的固碳植物和工业原料来源<sup>[2-3]</sup>。毛竹的种下分类群具有丰富的变型和栽培变型, 主要表现在秆型和秆色等方面的变异<sup>[4]</sup>。毛竹的秆色变异栽培种是优良的园林观赏竹类资源, 主要包括花毛竹 *Ph. edulis* f. *taokiang*、黄皮毛竹 *Ph. edulis* f. *holochrysa*、绿槽毛竹 *Ph. edulis* f. *bicolor*、黄槽毛竹 *Ph. edulis* f. *luteosulcata* 和绿槽龟甲竹 *Ph. edulis* f. *lücaoguijiazhu* 等。

高等植物的叶绿体是半自主的细胞器, 具有相对保守的遗传体系<sup>[5]</sup>。叶绿体基因组为母系遗传物质, 结构保守, 进化速率均衡, 分别可以应用于目级、科级、属级、种级的系统发育关系重建<sup>[6]</sup>。HUANG 等<sup>[7]</sup>通过叶绿体基因组系统进化分析, 将 11 个榕属 *Ficus* 物种根据不同亚属进一步划分为 3 个支系, 同时发现榕属与桑属 *Morus* 植物的属间亲缘关系。ZHANG 等<sup>[8]</sup>通过对 17 个孩儿参 *Pseudostellaria heterophylla* 品种的完整叶绿体基因组序列分析, 发现中国种群被聚类为单系群, 其中不开花的品种形成了独立的亚支系, 并且揭示了孩儿参的种内变异, 进一步支持了叶绿体基因组可以阐明近缘物种之间亲缘关系的观点。迄今为止, 绿槽毛竹<sup>[9]</sup>、厚壁毛竹 *Ph. edulis* f. *pachyloen*<sup>[10]</sup>、青龙竹 *Ph. edulis* f. *curviculmis*<sup>[11]</sup> 和龟甲竹 *Ph. edulis* f. *tubiformis*<sup>[12]</sup> 的叶绿体基因组已发布; 这些毛竹种下分类群的叶绿体基因组大小均为 139 678 bp, 包含 126~132 个基因, 其中包括 84~85 个蛋白质编码基因、8 个 rRNA 和 34~39 个 tRNA。黄槽毛竹是毛竹中秆色变异类型的重要栽培类型, 表现为茎秆节间槽部失绿而秆壁绿色, 具有较好的观赏价值<sup>[13]</sup>。黄槽毛竹的叶绿体基因组资料未见报道, 并且毛竹种下分类群的叶绿体基因组之间的比较缺乏研究。本研究将对黄槽毛竹的叶绿体基因组进行组装注释, 分析其基因组特征, 并比较与其他毛竹种下分类群之间的叶绿体遗传信息差异和系统进化关系。

## 1 材料与方法

### 1.1 材料

黄槽毛竹新鲜竹青采集于浙江农林大学东湖校区翠竹园 (30°26'N, 119°72'E)。黄槽毛竹取植物枝条标本 (编号 LIUX202301), 保存在浙江农林大学林业与生物技术学院植物标本室。

### 1.2 DNA 提取与测序

取 2 g 新鲜竹青进行液氮研磨, 随后转移到离心管, 配合天根 Plant Genomic DNA Kit (Cat.#DP305-03) 试剂盒和其建议的操作流程进行基因组 DNA 提取。对检验合格后的 DNA 样品进行二代测序文库构建。在 Illumina HiSeq 4000 平台测序 DNA 测序, 测序深度为 60×。原始数据可经国家基因库生命大数据平台 (CNGBdb) 的样本编号 CNS0314191 获取。

### 1.3 叶绿体基因组组装和注释

利用 fastp<sup>[14]</sup> 对二代测序数据进行过滤, 去除接头和切除低质量的数据, 具体参数设置为 “-w 16 -5 25 -3 25 --length\_required 50 --average\_qual 25 -z 9”; 通过 GetOrganelle<sup>[15]</sup> 对清洗后的数据进行叶绿体基因组进行组装; 参数设置为 “-F embplant\_pt -o output -R 10 -t 10 -k 21,45,85,105”。原始数据上传至美国

国家生物技术信息中心 (NCBI) GenBank (<https://www.ncbi.nlm.nih.gov>), 登录号为 OR597504。使用 CPGAVAS2 对叶绿体基因组进行注释。使用 CHLOROPLLOT 在线软件 (<https://irscope.shinyapps.io/Chloroplast/>) 绘制基因组图谱。

#### 1.4 叶绿体基因组密码子使用频率、简单重复序列 (SSR) 分析

利用 MISA<sup>[16]</sup> 在线软件检测简单重复序列, 相应参数设置为: 单核苷酸重复  $\geq 10$ , 二核苷酸  $\geq 5$ , 三核苷酸重复  $\geq 4$ , 四、五和六核苷酸重复  $\geq 3$ 。采用 REPuter 在线软件 (<https://bibiserv.cebitec.uni-bielefeld.de/reputer>) 验证黄槽毛竹的重复序列, 包括正向重复序列 (F)、反向重复序列 (R)、互补重复序列 (C) 以及回文重复序列 (P), 设置参数为 “Minimal repeat size = 15 bp, Edit distance  $\geq 3$  bp”。

#### 1.5 叶绿体基因组序列变异比较分析

边界收缩与扩张分析使用 Genepioneer 平台 (<http://cloud.genepioneer.com:9929/>) 的 CPJSdraw-边界图绘制-v1.0.0 功能进行绘制。利用 mVISTA 进行多序列变异分析, 比对运算模式设置为 Shuffle-LAGAN, 可视化参数设置为 Y 轴起始坐标为 0。核苷酸多样性分析使用 JSHYCloud 平台 (<http://cloud.genepioneer.com:9929>) 进行分析绘制。

#### 1.6 系统进化分析

序列比对使用软件 Mafft v7.490<sup>[17]</sup>, 参数设置为默认条件。系统发育进化关系重建分别采用了贝叶斯推理法和最大似然法。贝叶斯推理法所采用的软件为 MrBayes v3.2.7 (<https://mybiosoftware.com/mrbayes-3-1-2-bayesian-inference-phylogeny.html>), 参数设置为 “ngen=1 000 000, samplefreq=100, nchains = 4, temp = 0.1, burnin=2 500”。最大似然法采用了 IQ-TREE v2.0.7<sup>[18]</sup> 进行运算, 最佳进化模型为程序自动选择。构建系统发育树所用分类群及其序列数据登录号: 毛竹原变种 *Phyllostachys edulis* f. *pubescens* (HQ337796)、黄槽毛竹 (OR597504)、花毛竹 (SRR6705383)、青龙竹 (MW007169)、绿槽毛竹 (OM084949)、黄皮毛竹 (SRR6705382)、厚壁毛竹 (MN537809)、淡竹 *Ph. nigra* var. *henonis* (HQ154129)、箬竹 *Ph. glauca* (MT657329)、红壳雷竹 *Ph. incarnata* (OL457160)、桂竹 *Ph. reticulata* (MN537808)、麻竹 *Dendrocalamus latifloru* (FJ970916)、瓜多竹 *Guadua amplexifolia* (KM365071)、巴西玉米竺 *Raddia brasiliensis* (KJ870998) 和菝葜竹 *Olyra latifolia* (KF515509)。花毛竹和黄皮毛竹的基因组序列分别由 NCBI 获取的原始测序数据经 GetOrganelle 组装得到。

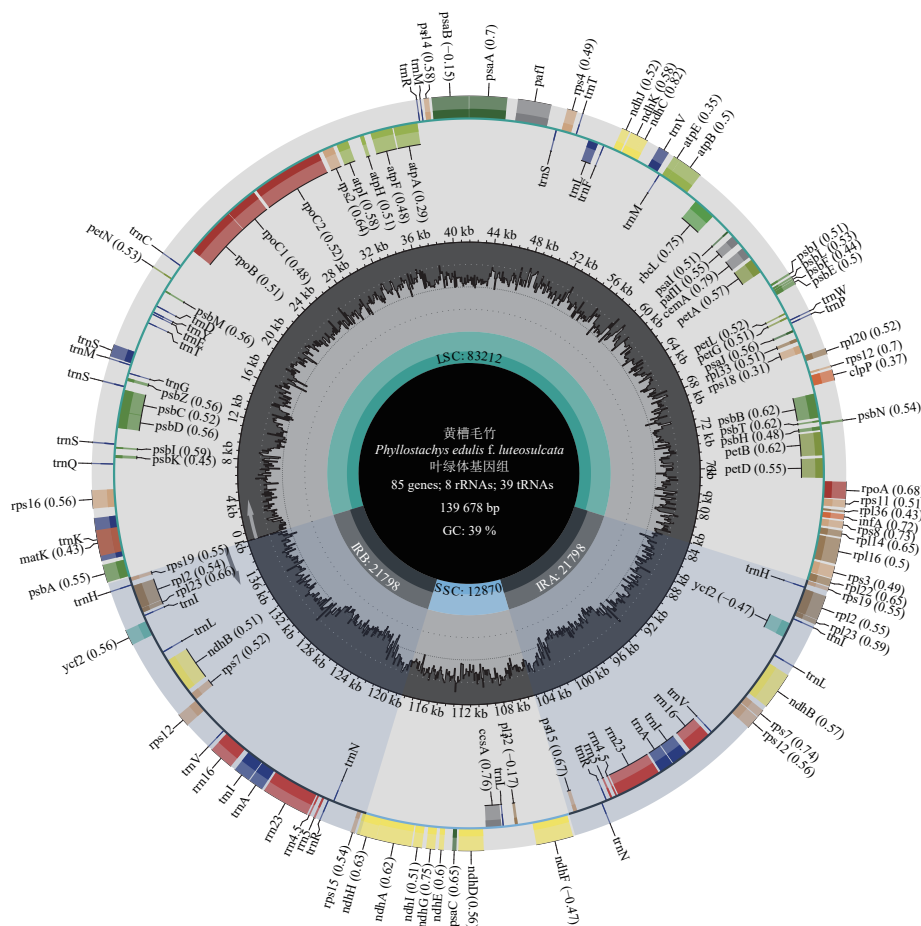
## 2 结果与分析

### 2.1 黄槽毛竹叶绿体基因组基本特征

黄槽毛竹绿色组织的二代测序获得 42.78 Gb 的原始数据, 包含了 285 198 507 对读段。原始数据被清洗后, 获得 283 052 833 对长度大于 50 bp 且平均碱基质量值大于 25 的读段。经 GetOrganelle 组装获得黄槽毛竹叶绿体基因组序列大小为 139 678 bp。该叶绿体基因组具有典型的环状四分体结构, 由大单拷贝区域 (LSR)、小单拷贝区域 (SSR)、2 个反向互补重复区域 (IRs) 等 4 个部分组成, 其中 LSR 长度为 83 212 bp, SSR 长度为 12 870 bp, 2 个 IRs (IRA 和 IRB) 长度为 43 596 bp (图 1)。整个叶绿体基因组中 GC 含量为 38.88%, LSC、SSC 和 IRs 区域 GC 含量分别为 36.97%、33.17% 和 44.22%, 其中 IRs 区域 GC 含量明显高于 2 个单拷贝区域。

### 2.2 黄槽毛竹叶绿体基因组的组成及功能

黄槽毛竹叶绿体基因组总共包含 132 个基因, 包括 85 个蛋白质编码基因、39 个 tRNA、8 个 rRNA (表 1)。根据生物学功能可以将 85 个蛋白质编码基因分为以下三大类: 与光合作用相关基因共 50 个; 与表达相关基因共 31 个; 其他基因共 4 个。该基因组中共 12 个基因具有内含子结构 [*ndhA*、*ndhB*(2)、*petB*、*petD*、*atpF*、*rpl16*、*rpl2*(2)、*ycf3*、*rps16* 和 *rpoC2*], 1 个蛋白质编码基因含有反式剪接基因 (*rps12*)。由于绿槽毛竹具有绿色的节间秆槽和黄色的节间秆壁, 与黄槽毛竹的秆色表型相反, 本研究比较了黄槽毛竹、绿槽毛竹和毛竹原变种的叶绿体基因组信息, 发现黄槽毛竹叶绿体基因组的蛋白质编码基因、tRNA 与毛竹原变种有所差异 (表 2), 相较于绿槽毛竹及毛竹原变种, 黄槽毛竹的蛋白质编码基因缺少 2 个 *ycf68*。黄槽毛竹叶绿体基因组的蛋白质编码基因数量也同毛竹原变种有所不同, 前者比后者多 1 个编码基因 *rps12*。



- 光系统 I (photosystem I) ■ 光系统 II (photosystem II) ■ 细胞色素 b/f 复合物(cytochrome b/f complex) ■ ATP 合成酶(ATP synthesis)
- NADH 脱氢酶(NADH dehydrogenase) ■ 二磷酸核酮糖羧化酶大亚基(RubisCO larg subunit) ■ RNA 聚合酶(RNA polymerase)
- 核糖体蛋白小亚基(small ribosomal protein) ■ 核糖体蛋白大亚基(large ribosomal protein) ■ clpP, matK, infA
- 假定开放性阅读框(hypothetical reading frame) ■ 转运 RNA(transfer RNA) ■ 核糖体 RNA(ribosomal RNA) ■ 其他(other)

环图中心标注物种名称、基因数量、基因组长度和 GC 含量。在第 1 个内圈标注了 SSC、IRA、IRB 和 LSC 的长度；第 2 内圈为各个区域的 GC 含量，用每个部分的阴影区域的比例标记，并标注 4 个区域的基因组位置。基因名称标记在最外圈（第 3 圈），基因按其功能分类用颜色标记。环图中的箭头表示内圈和外圈基因的转录方向。

图 1 黄槽毛竹叶绿体基因组图谱

Figure 1 Genome map of the chloroplast of *Ph. edulis f. luteosulcata*

### 2.3 黄槽毛竹的相对同义密码子使用频率 (RSCU)、重复序列和 SSR 分析

根据黄槽毛竹叶绿体基因组 RSCU(图 2A) 显示：共有 64 种密码子编码 20 种氨基酸，密码子中 UUA 的 RSCU 为最高值 (1.97)，CUG 为最低值 (0.33)。RSCU>1 的有 32 个，其中以 A/U 结尾的密码子有 30 个，提示密码子偏好使用以 A/U 结尾的密码子。重复序列按照不同的排列方式可分为正向重复序列、反向重复序列、回文重复序列和互补重复序列 4 种类型。黄槽毛竹叶绿体基因组中共检测出 49 个重复序列，其中 40 个为正向重复序列，9 个为回文重复序列，未发现反向和互补重复序列。所有重复序列长度为 30~100 bp，其中序列长度在 30~60 bp 的重复序列占比较高，为 85.7%。此外，在基因组中共检测出 55 个 SSR 位点 (图 2B)，包括 34 个单核苷酸重复序列、4 个二核苷酸重复序列、3 个三核苷酸重复序列、13 个四核苷酸重复序列和 1 个五核苷酸重复序列，其中简单重复序列最多的类型为 A/T，共 23 个。

### 2.4 黄槽毛竹与近缘分类群的叶绿体基因组序列分析

高等植物的叶绿体通常为典型的四分体结构，但是常常发生 SSC 和 LSC 的扩张和收缩，导致四分体的边界发生变化。对黄槽毛竹与其他毛竹变型等 7 个种下分类群的叶绿体基因组四分体边界进行了分析 (图 3A) 显示：黄槽毛竹叶绿体基因组四分体边界的扩张和收缩与其他分类群的高度一致，并且前者边界基因也与其他的高度统一。然而，黄槽毛竹、青龙竹和黄皮毛竹与其他毛竹种下分类群之间存在着边界编码基因编码长度不一致的情况，如前三者的 *rps19* 编码长度为 216 bp，其他毛竹种下分类群的该

表 1 黄槽毛竹叶绿体基因组的基因列表

Table 1 Gene of the *Ph. edulis* f. *luteosulcata* chloroplast genome

基因类别	基因分组	基因列表	
光合作用基因	ATP合酶亚基	<i>atpA</i> 、 <i>atpB</i> 、 <i>atpE</i> 、 <i>atpF*</i> 、 <i>atpH</i> 、 <i>atpI</i>	
	依赖ATP的Clp蛋白酶蛋白水解亚基	<i>clpP</i>	
	光合系统 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>	
	NADH脱氢酶亚基	<i>ndhA*</i> 、 <i>ndhB*</i> 、 <i>ndhB*</i> 、 <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b/f复合物亚基	<i>petA</i> 、 <i>petB*</i> 、 <i>petD*</i> 、 <i>petG</i> 、 <i>petL</i> 、 <i>petN</i>	
	光合系统 I 亚基	<i>psaA</i> 、 <i>psaB</i> 、 <i>psaC</i> 、 <i>psaI</i> 、 <i>psaJ</i>	
	光合系统 I 组件	<i>ycf2</i> 、 <i>pafI</i> ( <i>ycf3</i> )**、 <i>ycf2</i>	
	光合系统 II 组件	<i>pafII</i> ( <i>ycf4</i> )	
	二磷酸核酮糖羧化酶亚基	<i>rbcL</i>	
	表达相关基因	核糖体大亚基	<i>rpl14</i> 、 <i>rpl16*</i> 、 <i>rpl20</i> 、 <i>rpl22</i> 、 <i>rpl23</i> 、 <b><i>rpl23</i></b> 、 <i>rpl32</i> 、 <i>rpl33</i> 、 <i>rpl36</i> 、 <i>rpl2*</i> 、 <b><i>rpl2*</i></b>
依赖DNA的RNA聚合酶		<i>rpoA</i> 、 <i>rpoB</i> 、 <i>rpoC1</i> 、 <i>rpoC2*</i>	
核糖体小亚基		<i>rps2</i> 、 <i>rps3</i> 、 <i>rps4</i> 、 <i>rps7</i> 、 <b><i>rps7</i></b> 、 <i>rps8</i> 、 <i>rps11</i> 、 <i>rps12</i> 、 <b><i>rps12</i></b> 、 <i>rps14</i> 、 <i>rps15</i> 、 <b><i>rps15</i></b> 、 <i>rps16*</i> 、 <i>rps18</i> 、 <i>rps19</i> 、 <b><i>rps19</i></b>	
rRNA基因		<i>rrn16S</i> 、 <i>rrn23S</i> 、 <i>rrn4.5S</i> 、 <i>rrn5S</i> 、 <b><i>rrn5S</i></b> 、 <b><i>rrn4.5S</i></b> 、 <b><i>rrn23S</i></b> 、 <b><i>rrn16S</i></b>	
tRNA基因		<i>trnA-UGC*</i> 、 <b><i>trnA-UGC*</i></b> 、 <i>trnC-GCA</i> 、 <i>trnD-GUC</i> 、 <i>trnE-UUC</i> 、 <i>trnF-GAA</i> 、 <i>trnG-GCC</i> 、 <i>trnH-GUG</i> 、 <b><i>trnH-GUG</i></b> 、 <i>trnI-CAU</i> 、 <b><i>trnI-CAU</i></b> 、 <i>trnI-GAU*</i> 、 <b><i>trnI-GAU*</i></b> 、 <i>trnK-UUU*</i> 、 <i>trnL-CAA</i> 、 <b><i>trnL-CAA</i></b> 、 <i>trnL-UAA*</i> 、 <i>trnL-UAG</i> 、 <i>trnM-CAU</i> 、 <b><i>trnM-CAU</i></b> 、 <b><i>trnM-CAU</i></b> 、 <i>trnN-GUU</i> 、 <b><i>trnN-GUU</i></b> 、 <i>trnP-UGG</i> 、 <i>trnQ-UUG</i> 、 <i>trnR-ACG</i> 、 <b><i>trnR-ACG</i></b> 、 <i>trnR-UCU</i> 、 <i>trnS-CGA*</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> 、 <b><i>trnV-GAC</i></b> 、 <i>trnV-UAC*</i> 、 <i>trnW-CCA</i> 、 <i>trnY-GUA</i>	
其他基因		C型细胞色素合成基因	<i>ccsA</i>
		胞膜蛋白	<i>cemA</i>
		成熟酶	<i>matK</i>
		翻译起始因子	<i>infA</i>

说明：加粗代表多拷贝的基因；\*表示带1个内含子的基因；\*\*表示带2个内含子的基因。

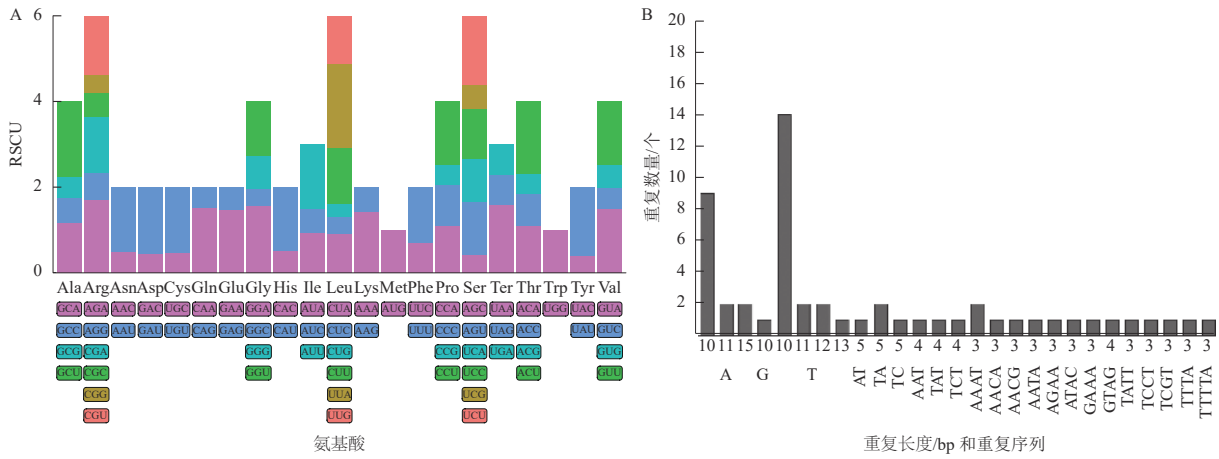
表 2 黄槽毛竹和绿槽毛竹、毛竹原变种的叶绿体基因差异

Table 2 Chloroplast gene differences among *Ph. edulis* f. *luteosulcata*, *Ph. edulis* f. *bicolor*, and *Ph. edulis* var. *pubescens*

项目	黄槽毛竹	毛竹	绿槽毛竹
蛋白编码基因数量	85	84	83
转运RNA数量	-	-	-
核糖体RNA数量	-	-	-
蛋白质编码基因差异	<i>rps12</i> (2)、 <i>ycf2</i> (2)	<i>rps12</i> 、 <i>ycf68</i> (2)	<i>rps12</i> 、 <i>ycf68</i>
转运RNA差异	<i>trnG-GCC</i> 、 <i>trnS-CGA</i>	<i>trnG-UCC</i> 、 <i>trnG-UCC</i>	<i>trnG-GCC</i> 、 <i>trnG-UCC</i>

说明：(2)代表拷贝数为2；-表示未检测到。

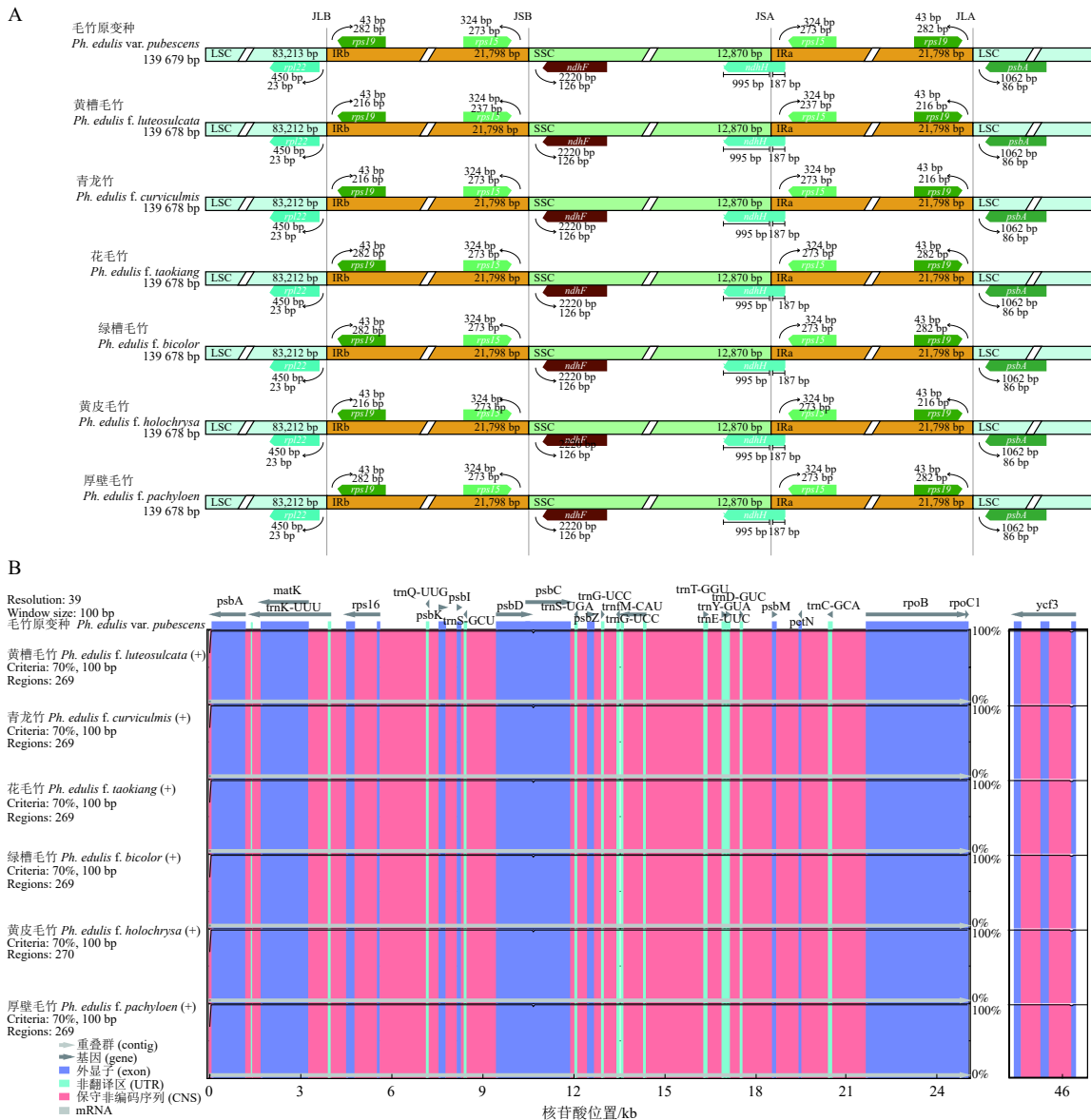
基因编码 282 bp。mVISTA 分析显示 (图 3B)：对比毛竹原变种，黄槽毛竹以及其他毛竹种下分类群的叶绿体基因组变异程度较低，仅在 *psbC*、*ycf3* 基因编码区与非编码区之间 (*psbA* 附近) 具有一定程度的序列差异。此外，基于 CPGview 鉴定黄槽毛竹、毛竹原变种和绿槽毛竹之间的顺式剪切基因，黄槽毛竹有 12 个，后两者有 11 个。其中黄槽毛竹较后两者多获得 *rpoC2* 的顺式剪切形式 (图 4A)。将核苷酸多样性 (Pi) 截断点设定为  $Pi \geq 0.04$ ，在 7 个毛竹种下分类群的基因组序列比对中发现了 3 个高变异区域



A. 黄槽毛竹叶绿体基因组 RSCU 分析; B. 黄槽毛竹叶绿体基因组中的重复序列长度与数量。

图 2 黄槽毛竹叶绿体基因组的重复分析

Figure 2 Repeat analysis of chloroplast genomes of *Ph. edulis f. luteosulcata*

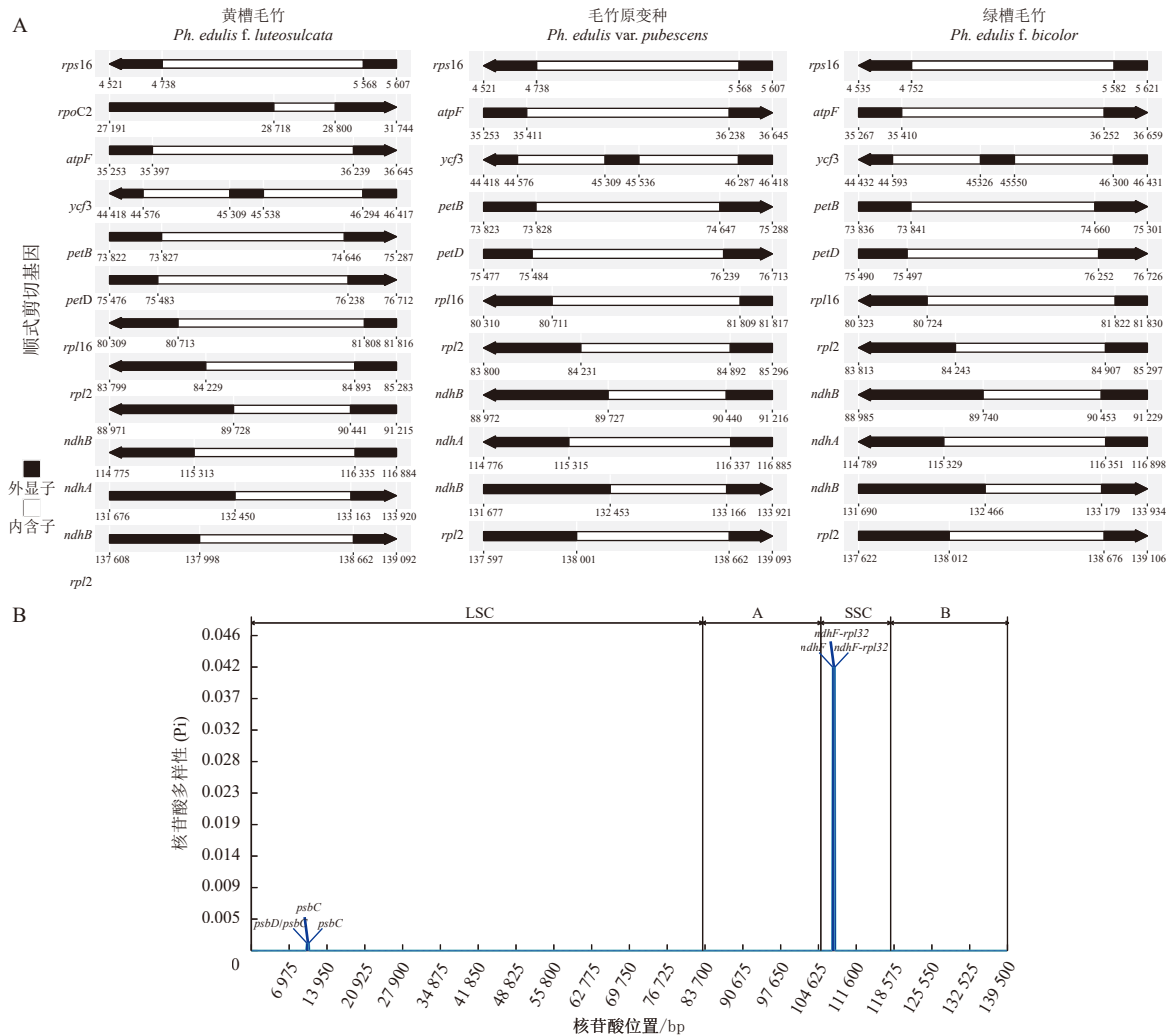


A. 叶绿体基因组四分体序列边界扩张与收缩; 箭头表示这些基因与 IR 边界的距离; B. 基于毛竹原变种序列为基准的叶绿体基因组序列 mVISTA 分析。

图 3 黄槽毛竹与其他毛竹变型的叶绿体基因组序列变异比较

Figure 3 Comparison of chloroplast genome sequence variation between *Ph. edulis f. luteosulcata* and other subspecies in moso bamboo

(图 4B)。这些高变异区均位于蛋白编码基因 *ndhF* 附近，主要表现在叶绿体基因组 *ndhF* 和 *rpl32* 之间的 SSC 区。这些变异区域将可以用于毛竹种下分类群的分子鉴定依据。上述序列数据表明：毛竹种下分类群之间的序列变异较低，但不同分类群之间仍具有一些共有和特异的变异特征。



A. 黄槽毛竹、毛竹原变种和绿槽毛竹之间的顺式剪切基因比较；B. 7 个毛竹种下分类群叶绿体基因组的核苷酸多样性分析。滑动窗口大小和滑动步长大小分别设定为 600 和 200 bp。

图 4 黄槽毛竹与其他毛竹变型的顺式剪切基因比较和叶绿体基因组核苷酸多样性 (Pi) 分析

Figure 4 Comparison of the *cis*-splicing genes and analysis of nucleotide diversity (Pi) of cp genomes between *Ph. edulis f. luteosulcata* and other subspecies in moso bamboo

## 2.5 系统进化分析

为了进一步了解黄槽毛竹在竹子分类群内进化关系，利用毛竹种下分类群 (毛竹原变种、绿槽毛竹、厚壁毛竹、青龙竹、黄皮毛竹、花毛竹和黄槽毛竹)，4 个刚竹属物种 (淡竹、箬竹、红壳雷竹和桂竹)，麻竹、瓜多竹、外类群草本竹巴西玉米竺和菝葜竹的叶绿体全基因组进行系统进化树的重建。贝叶斯推理法重建的系统进化树 (图 5) 显示：毛竹种下分类群为单系且毛竹分类群与淡竹为姐妹。毛竹分类群之中，黄槽毛竹与毛竹原变种为姐妹关系，两者与花毛竹和青龙竹组成的分支为姐妹关系。秆色变异表型与黄槽毛竹呈“反转”状态的绿槽毛竹和黄槽毛竹以及毛竹原变种之间，并没有组成一支，而是与黄皮毛竹组成姐妹类群。厚壁毛竹与本研究分析的其他 6 个毛竹种下分类群的为姐妹关系。

## 3 讨论

植物叶绿体基因组已被广泛应用于植物分类、分子进化以及系统发育等相关研究。竹亚科作为禾本科中唯一木质化结构的类群，其种群数量较多、分布较广且有性繁殖周期长，因此在系统分类学上造成

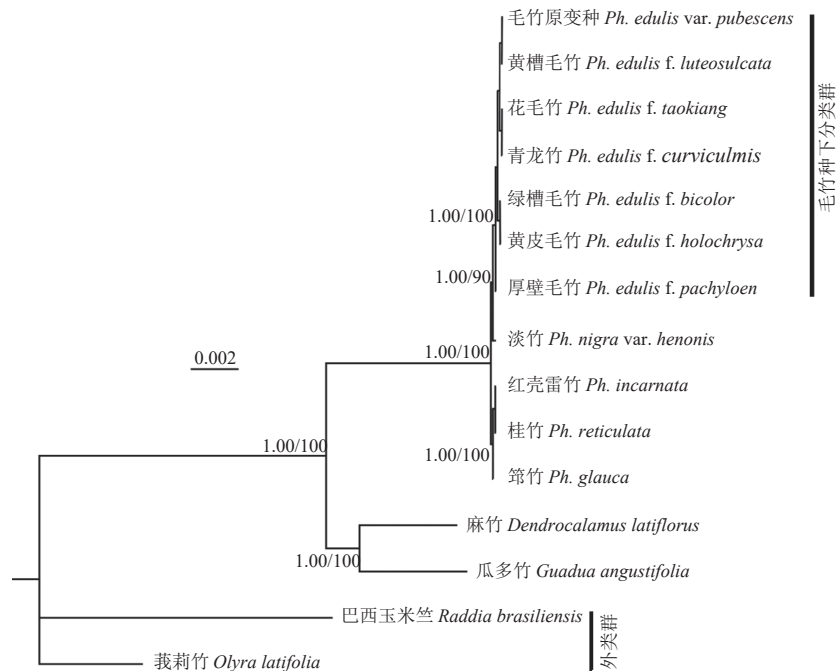


图5 基于叶绿体全基因组序列重建的黄槽毛竹在刚竹属中的系统进化关系

Figure 5 Phylogenetic relationship of *Ph. edulis f. luteosulcata* based on whole-chloroplast genome sequences

分类困难<sup>[19]</sup>。在分子水平上，黄槽毛竹所属的木本竹进化相对缓慢，较低的分子进化速度可能会使系统发育研究复杂化，因此对于这些较低阶元分类群，叶绿体系统发育基因组学是较好的解决方案<sup>[20]</sup>。叶绿体作为来源于蓝藻菌的独立细胞器<sup>[21]</sup>，具有较为保守的特性，叶绿体基因组在包含大量遗传信息的同时相对于核基因组和线粒体基因组序列大小适中，便于测序；其次，核酸取代率相对较慢，进化速率保守，为深层研究植物系统发育和物种鉴定提供基础<sup>[22]</sup>。

高等植物的叶绿体基因组通常为闭合的环形四分体结构，长度为108~165 kb，约包含80个编码基因<sup>[23]</sup>。本研究测序结果显示：黄槽毛竹的叶绿体基因组核苷酸序列为139 678 bp，为典型的四分体结构；这与竹子叶绿体基因组的长度接近且结构形式一致<sup>[9, 24]</sup>。黄槽毛竹叶绿体基因组共注释132个基因，包含85个蛋白质编码基因。这与系统发育分析所显示关系最近的毛竹原变种的编码基因数量差异较小，后者84个<sup>[19]</sup>。但最近报道的绿槽毛竹叶绿体基因组包含了85个蛋白编码基因<sup>[9]</sup>。本研究对黄槽毛竹、毛竹原变种和绿槽毛竹的叶绿体蛋白编码基因的结构分析发现：三者之间的*rpoC2*存在顺式剪切结构的差异；并且三者之间的tRNA和蛋白质编码基因类型也有所差异。虽然毛竹种下分类群的叶绿体基因组的四分体边界保持一致，但是边界的蛋白编码基因*rps19*的编码长度分为216和282 bp两大类。此外，叶绿体基因组的mVISTA分析也支持毛竹种下分类群之间也存在序列变异的区域，并且核苷酸多样性分析显示高变异区均位于叶绿体基因组SSC区的*ndhF*和*rpl32*之间。综合上述序列比较的信息，*rpoC2*的顺式剪切、*rps19*编码区长度以及高核酸多样性的*ndhF*和*ndhF-rpl32*区间等，可以用于毛竹种下分类群鉴定的潜在DNA片段。这些结果显示毛竹种下分类群的叶绿体基因组序列特征具有多样性。

重复序列广泛分布于叶绿体基因组中，能够通过特异结合蛋白质促使核酸形成复杂结构<sup>[7, 21]</sup>。本研究黄槽毛竹叶绿体基因组中的SSR分析显示：相对于二、三、四、五和六核苷酸重复相比，单核苷酸重复出现的频次更高，AT/TA以及TC是最常见的二核苷酸重复基序，而GC/CG较少或并不存在。这与VIEIRA等<sup>[25]</sup>研究的20种热带木本竹结果相同。密码子在叶绿体基因组中也起着关键作用，本研究相对同义密码子使用频率显示：RSCU值>1的有32个，其中以A/U结尾的密码子有30个，提示密码子偏好使用以A/U结尾的密码子，这与其他竹类叶绿体基因组的分析一致<sup>[26]</sup>。

叶绿体全基因组能够较好地鉴定种间及以上分类群，但偶见利用于种内群体的鉴定，如朴树*Celtis sinensis*<sup>[27]</sup>。本研究利用全基因组序列重建了7个毛竹种下分类群的系统发育关系。结果显示：毛竹种下分类群为单系(100/90，后验概率/最大似然法自展值)。然而，JING等<sup>[9]</sup>研究显示：毛竹种下分类群



与淡竹、箬竹和桂竹形成并系。本研究的系统发育分析显示: 这 3 个物种并不与毛竹分类群形成并系。这可能是由于采样类群不一致甚至序列比对的结果差异所导致。因此, 毛竹种下分类群具争议性的系统发育关系仍待更为全面和深入的研究。本研究重建的系统发育关系表明: 黄槽毛竹与毛竹原变种亲缘关系最近, 暗示两者有共同的起源; 与黄槽毛竹表型相反的绿槽毛竹与黄皮毛竹为姐妹关系, 表明两者有共同的最近祖先。目前, 通过叶绿体全基因组序列重建的进化树在毛竹种下分类群之间的系统发育关系鉴定分辨率较低, 仅能有效鉴定厚壁毛竹(厚壁毛竹+6 个分类群)和黄皮毛竹(黄皮毛竹+绿槽毛竹)的分支。

## 4 结论

黄槽毛竹的叶绿体基因组是长度为 139 678 bp 的双环 DNA, 包含 132 个基因。这些基因包括 85 个蛋白质编码基因、8 个核糖体 RNA(rRNA) 以及 39 个转运 RNA(tRNA)。该基因组偏好使用以 A/U 碱基结尾的密码子, 且简单重复序列最多的类型为 A/T。在系统发育分析方面, 黄槽毛竹与其他毛竹种下分类群共同构成了单系分支, 且与毛竹原变种具有最近的亲缘关系。种下分类群的叶绿体基因组比较分析发现毛竹种下分类群之间存在序列差异。

## 5 参考文献

- [1] ZHAO Hansheng, SUN Shuai, DING Yulong, *et al.* Analysis of 427 genomes reveals moso bamboo population structure and genetic basis of property traits [J/OL]. *Nature Communications*, 2021, **12**(1): 5466[2024-01-01]. doi: 10.1038/s41467-021-25795-x.
- [2] ZHAO Hansheng, WANG Jian, MENG Yufei, *et al.* Bamboo and rattan: nature-based solutions for sustainable development [J]. *The Innovation*, 2022, **3**(6): 37 – 38.
- [3] 郑钧, 吴仁武, 史琰, 等. 竹类植物的主要环境效应研究进展[J]. *浙江农林大学学报*, 2017, **34**(2): 374 – 380.  
ZHENG Jun, WU Renwu, SHI Yan, *et al.* Research progress on environmental effects of bamboo: a review [J]. *Journal of Zhejiang A&F University*, 2017, **34**(2): 374 – 380.
- [4] 史军义, 周德群, 马丽莎, 等. 中国竹类多样性及其重要价值[J]. *世界竹藤通讯*, 2020, **18**(3): 55 – 65, 72.  
SHI Junyi, ZHOU Dequn, MA Lisha, *et al.* Diversity and important value of bamboos in China [J]. *World Bamboo and Rattan*, 2020, **18**(3): 55 – 65, 72.
- [5] MAHAPATRA K, BANERJEE S, DE S, *et al.* An insight into the mechanism of plant organelle genome maintenance and implications of organelle genome in crop improvement: an update [J/OL]. *Frontiers in Cell and Developmental Biology*, 2021, **9**: 671698[2024-01-01]. doi: 10.3389/fcell.2021.671698.
- [6] HOLLINGSWORTH P M. Refining the DNA barcode for land plants [J]. *Proceedings of the National Academy of Sciences*, 2011, **108**(49): 19451 – 19452.
- [7] HUANG Yuying, LI Jing, YANG Zerui, *et al.* Comprehensive analysis of complete chloroplast genome and phylogenetic aspects of ten *Ficus* species [J/OL]. *BMC Plant Biology*, 2022, **22**(1): 253[2024-01-01]. doi: 10.1186/s12870-022-03643-4.
- [8] ZHANG Wujun, ZHANG Zhaolei, LIU Baocai, *et al.* Comparative analysis of 17 complete chloroplast genomes reveals intraspecific variation and relationships among *Pseudostellaria heterophylla* (Miq.) Pax populations [J/OL]. *Frontiers in Plant Science*, 2023, **14**: 1163325[2024-01-01]. doi: 10.3389/fpls.2023.1163325.
- [9] JING Wenxuan, HU Bo, WAN Rou, *et al.* A complete chloroplast genome of bamboo cultivar *Phyllostachys edulis* f. *bicolor* (Poaceae: Bambusoideae) [J]. *Mitochondrial DNA Part B*, 2023, **8**(4): 532 – 535.
- [10] HUANG Nianjun, LI Jiangping, YANG Guangyao, *et al.* Two plastomes of *Phyllostachys* and reconstruction of phylogenetic relationship amongst selected *Phyllostachys* species using genome skimming [J]. *Mitochondrial DNA Part B*, 2020, **5**(1): 69 – 70.
- [11] GAO Liqin, LI Yonglong, ZHANG Wengen, *et al.* The complete chloroplast genome of *Phyllostachys edulis* f. *curviculmis* (Bambusoideae): a newly ornamental bamboo endemic to China [J]. *Mitochondrial DNA Part B*, 2021, **6**(3): 941 – 942.
- [12] LIU Xinmiao, LIU Lei, LI Lubin, *et al.* The complete chloroplast genome of *Phyllostachys edulis* f. *tubiformis* (Bambusoideae): a highly appreciated type of ornamental bamboo in China [J]. *Mitochondrial DNA Part B*, 2022, **7**(1):

185 – 187.

- [13] 方伟. 中国经济竹类 [M]. 北京: 科学出版社, 2015.  
FANG Wei. *Chinese Economic Bamboo* [M]. Beijing: Science Press, 2015.
- [14] CHEN Shifu, ZHOU Yanqing, CHEN Yaru, *et al.* fastp: an ultra-fast all-in-one FASTQ preprocessor [J]. *Bioinformatics*, 2018, **34**(17): i884 – i890.
- [15] JIN Jianjun, YU Wenbin, YANG Junbo, *et al.* GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes [J/OL]. *Genome Biology*, 2020, **21**(1): 241 [2024-01-01]. doi: 10.1101/256479.
- [16] BEIER S, THIEL T, MÜNCH T, *et al.* MISA-web: a web server for microsatellite prediction [J]. *Bioinformatics*, 2017, **33**(16): 2583 – 2585.
- [17] NAKAMURA T, YAMADA K D, TOMII K, *et al.* Parallelization of MAFFT for large-scale multiple sequence alignments [J]. *Bioinformatics*, 2018, **34**(14): 2490 – 2492.
- [18] MINH B Q, SCHMIDT H A, CHERNOMOR O, *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era [J]. *Molecular Biology and Evolution*, 2020, **37**(5): 1530 – 1534.
- [19] ZHANG Yunjie, MA Pengfei, LI Dezhu. High-throughput sequencing of six bamboo chloroplast genomes: phylogenetic implications for temperate woody bamboos (Poaceae: Bambusoideae) [J/OL]. *PLoS One*, 2011, **6**(5): e20596 [2024-01-01]. doi: 10.1371/journal.pone.0020596.
- [20] 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.  
ZHANG Yunjie, LI Dezhu. Advances in Phylogenomics based on complete chloroplast genomes [J]. *Plant Diversity and Resources*, 2011, **33**(4): 365 – 375.
- [21] DOBROGOJSKI J, ADAMIEC M, LUCIŃSKI R. The chloroplast genome: a review [J/OL]. *Acta Physiologiae Plantarum*, 2020, **42**(6): 98 [2024-01-01]. doi: 10.1007/s11738-020-03089-x.
- [22] MA Pengfei, ZHANG Yuxiao, GUO Zhenhua, *et al.* Evidence for horizontal transfer of mitochondrial DNA to the plastid genome in a bamboo genus [J/OL]. *Scientific Reports*, 2015, **5**(1): 11608 [2024-01-01]. doi: 10.1038/srep11608.
- [23] GUISENGER M M, KUEHL J V, BOORE J L, *et al.* Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage [J]. *Molecular Biology and Evolution*, 2010, **28**(1): 583 – 600.
- [24] WYSOCKI W P, CLARK L G, ATTIGALA L, *et al.* Evolution of the bamboos (Bambusoideae; Poaceae): a full plastome phylogenomic analysis [J/OL]. *BMC Evolutionary Biology*, 2015, **15**(1): 50 [2024-01-01]. doi: 10.1186/s12862-015-0321-5.
- [25] VIEIRA L D N, DOS ANJOS K G, FAORO H, *et al.* Phylogenetic inference and SSR characterization of tropical woody bamboos tribe Bambuseae (Poaceae: Bambusoideae) based on complete plastid genome sequences [J]. *Current Genetics*, 2016, **62**(2): 443 – 453.
- [26] PEI Jialong, WANG Yong, ZHUO Juan, *et al.* Complete chloroplast genome features of *Dendrocalamus farinosus* and its comparison and evolutionary analysis with other Bambusoideae species [J/OL]. *Genes*, 2022, **13**(9): 1519 [2024-01-01]. doi: 10.3390/genes13091519.
- [27] 王国峥. 朴树叶绿体全基因组开发及三种生境类型下的朴树群体遗传学研究 [D]. 南京: 南京大学, 2020.  
WAN Guozheng. *Development of Complete Chloroplast Genome Sequence of Celtis sinensis (Cannabaceae) and Its Genetics Research under Three Habitats* [D]. Nanjing: Nanjing University, 2020.