

## 濒危植物秀丽兜兰叶绿体基因组特征与系统发育分析

范继征<sup>1</sup>, 李秀玲<sup>1\*</sup>, 李明智<sup>2</sup>, 卜朝阳<sup>1</sup>, 何荆洲<sup>1</sup>, 曾艳华<sup>1</sup>

1. 广西农业科学院花卉研究所, 广西南宁 530007; 2. 广州佰德生物科技有限公司, 广东广州 510555

**摘要:** 秀丽兜兰 (*Paphiopedilum venustum*) 具有较高的观赏价值和保护生物学价值, 野外资源濒临灭绝, 叶绿体基因组 (cpDNA) 具有基因组小、结构稳定、高度的保守性等优点被广泛用于植物系统发育和物种鉴定, 了解秀丽兜兰的叶绿体基因组结构, 对揭示兜兰属植物的系统进化关系具有重要意义。本研究通过二代高通量测序技术获得秀丽兜兰叶绿体全基因组序列, 利用 GeSeq、blast 和 hmmer 软件对叶绿体基因组进行注释, 采用 MISA、CodonW、fasttree 等生物信息学软件对其基因组结构、基因数目、序列重复、密码子偏好性和系统发育进行分析。结果表明: 秀丽兜兰叶绿体基因组结构保守, 具有典型的环状四分体结构, 一对反向重复区 (IRs) 将大单拷贝区 (LSC) 和小单拷贝区 (SSC) 分开, 全长 158 298 bp, GC 含量为 35.4%, 共注释得到 129 个基因, 包括 79 个蛋白编码基因, 38 个 tRNA 基因、8 个 rRNA 基因和 4 个假基因; 检测到 78 个 SSR 位点, 以单核苷酸重复和二核苷酸重复为主, 分别占 84.62% 和 10.26%, 无五核苷酸重复, 并且大多由 A 或 T 构成; 确定了高频密码子 32 个, 90.6% 都以 A 或 U 结尾, 编码亮氨酸 (Leu) 的密码子使用频率最高, 而编码半胱氨酸 (Cys) 的密码子使用频率最低。通过比较 18 个兰科植物叶绿体基因组, 构建系统进化树发现, 紫花杓兰和西藏杓兰 2 个杓兰属植物被单独聚成一支, 兜兰属被分成两大类, 其中秀丽兜兰 (*P. venustum*) 与同属的紫纹兜兰 (*P. purpuratum*) 亲缘关系密切。本研究筛选出的 SSR 和核苷酸多态性能够作为兜兰属植物种质鉴定和遗传多样性的分子标记, 可为珍稀濒危植物秀丽兜兰及其近缘种的种质鉴定、系统发育等提供理论依据, 为兜兰属植物种群恢复、生物多样性保护、创新利用及新品种选育奠定基础。

**关键词:** 秀丽兜兰; 叶绿体基因组; 特征分析; 系统发育分析

中图分类号: S 682.31 文献标识码: A

## Genomic Characteristics and Phylogenetic Analysis of Chloroplast of the Endangered Plant *Paphiopedilum venustum*

FAN Jizheng<sup>1</sup>, LI Xiuling<sup>1\*</sup>, LI Mingzhi<sup>2</sup>, BU Zhaoyang<sup>1</sup>, HE Jingzhou<sup>1</sup>, ZENG Yanhua<sup>1</sup>

1. Flower Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, Guangxi 530007, China; 2. Guangzhou Bio & data Technology Co., Ltd, Guangzhou, Guangdong 510555, China

**Abstract:** *Paphiopedilum venustum*, an endangered wild resources, has high ornamental value and conservation biological value. Chloroplast genome (cpDNA), small with stable structure and high conserved in nature, has been widely used in plant phylogeny and species identification. Understanding the chloroplast genome structure of *P. venustum* is of great significance in revealing the phylogenetic relationship of *Paphiopedilum*. In this study, the whole chloroplast genome of *P. venustum* was sequenced by the Illumina sequencing technology, the chloroplast genome was annotated using GeSeq, BLAST and hmmer, the genome structure, gene number, repeats, codon usage bias, phylogenetic development were analyzed using the bioinformatics softwares such as MISA, codonW and Fasttree. The chloroplast genome of *P. venustum* had a conserved quadripartite structure, with a pair of inverted repeats (IRs) separated a large single copy (LSC) region and a small single copy (SSC) region from the small single-copy region (SSC), with a total length of 158 298 bp.

收稿日期 2022-06-21; 修回日期 2022-08-03

基金项目 广西自然科学基金项目 (No. 2020GXNSFBA297136); 广西农业科学院科技发展基金项目 (桂农科 2021YT131); 广西农业科学院科技先锋队专项 (桂农科盟 202212-01)。

作者简介 范继征 (1982—), 女, 硕士, 高级农艺师, 研究方向: 兰科植物种质创新与新品种选育。\*通信作者 (Corresponding author): 李秀玲 (LI Xiuling), Email: congzxiao@163.com。

The guanine and cytosine (GC) content of chloroplast protein coding genes was 35.4%, and 129 genes were annotated, including 79 protein coding genes, 38 tRNA genes, 8 rRNA genes, and 4 pseudogenes. 78 simple sequence repeat (SSR) loci were detected, most of which were mononucleotide repeats and dinucleotide repeats, accounting for 84.62% and 10.26% respectively. There were no pentanucleotide repeats, and most of the repeats were composed of A or T base. 32 high-frequency codons were identified, and 90.6% of them ended in A or U, preferred codons ending in A or U. The codon encoding leucine (Leu) had the highest frequency of use, while the codon encoding cysteine (Cys) had the lowest frequency of use. Phylogenetic tree was constructed by comparing the chloroplast genomes of 18 orchidaceae plants. It was found that the two species of *Cypripedium calceolus* and *C. tibeticum* were separated from 16 species of *Paphiopedilum* and clustered into a single branch, and the genus of *Paphiopedilum* was divided into two groups, among which *P. venustum* and *P. purpuratum* were closely related. SSR and nucleotide polymorphisms can be used as molecular markers for germplasm identification and genetic diversity analysis of *Paphiopedilum*, which could provide a theoretical basis for germplasm identification and phylogeny of rare and endangered plant *Paphiopedilum* and its relatives, and lay a foundation for population restoration, biodiversity protection, innovative utilization and new variety breeding of *Paphiopedilum*.

**Keywords:** *Paphiopedilum venustum* (Wall. ex Sims) Pfitzer; chloroplast genome; characteristic analysis; phylogenetic analysis

**DOI:** 10.3969/j.issn.1000-2561.2023.06.003

秀丽兜兰 [*Paphiopedilum venustum* (Wall. ex Sims) Pfitzer] 隶属于兰科杓兰亚科兜兰属, 是第一个被人工栽培的兜兰属植物<sup>[1]</sup>, 主要分布于我国西藏东南部至南部, 以及孟加拉国、不丹、印度东北部、尼泊尔等国家和地区, 野外花期 1—3 月<sup>[2]</sup>。秀丽兜兰植株直立、丛生, 株高 7.0~9.0 cm, 叶片数 2~3 片, 长 7.5~12.0 cm, 宽 2.5~3.5 cm, 呈狭矩圆状椭圆形, 表面有网格斑, 背面有密集的紫色斑点, 花葶长 9.2~15.4 cm, 花朵横径 5.4~8.6 cm, 纵径 2.9~3.4 cm, 花朵色泽艳丽, 花期长, 花和叶均极具观赏价值。但由于人类活动以及自身生长要求和生长特点等原因, 秀丽兜兰已成为濒危物种(EN), 受到国际社会的严格保护<sup>[3-4]</sup>。

叶绿体是植物光合作用发生的细胞器, 具有一套完整的基因组<sup>[5]</sup>, 与核基因组等其他基因组相比, 叶绿体基因组具有高度的保守性以及自我复制机制、进化相对独立、基因组小、结构稳定、突变率低等特点<sup>[6-8]</sup>。随着二代测序技术的飞速发展, 通过测序组装和注释获得的完整的叶绿体基因组序列成本降低、速度加快<sup>[9]</sup>, 叶绿体基因组也越来越多的为进化分析、系统发育及分类鉴定等提供信息<sup>[10]</sup>。目前有关叶绿体基因组特征分析已被广泛应用于多种兜兰属植物, 有对单独一个种进行叶绿体基因组分析, 例如白旗兜兰 (*P. spicerianum*)<sup>[11]</sup>、白花兜兰 (*P. emersonii*)<sup>[12]</sup>、格丽兜兰 (*P. gratrixianum*)<sup>[13]</sup>、麻栗坡兜兰 (*P. malipoense*)<sup>[14]</sup>、紫纹兜兰 (*P. purpuratum*)<sup>[15]</sup>、飘带兜兰 (*P. parishii*)<sup>[16]</sup>、陈莲兜兰 (*P. tranlimianum*)<sup>[17]</sup>等,

也有对多个种进行比较分析<sup>[18-19]</sup>。现有研究结果表明兜兰属植物叶绿体基因组大小为 154 689~162 590 bp, 其结构与大多数陆生植物叶绿体结构相似, 为典型的四分结构, 包括大单拷贝区 (large single copy, LSC)、小单拷贝区 (small single copy, SSC) 和 2 个反向重复区 (inverted repeat, IR), 各区 GC 含量差异较大, 其中 IR 的 GC 含量较高, 不同的蛋白质编码在同义密码子的使用上有所不同, 偏好 A 或 U (T) 结尾的密码子。此外, 系统发育分析打破了形态学和细胞学等的局限性, 能够更好地反映物种之间的亲缘关系<sup>[11-19]</sup>。

有关秀丽兜兰的研究多集中在非共生萌发和繁殖技术<sup>[20]</sup>、细胞学研究<sup>[21-23]</sup>、病害鉴定<sup>[24]</sup>、分布地区和生长环境<sup>[25]</sup>等, 尚无秀丽兜兰叶绿体基因组的报道。本研究通过高通量测序方法获得秀丽兜兰完整的叶绿体基因组信息, 对其进行注释, 并对基因组中序列重复、SSR 位点及蛋白编码基因的密码子使用偏好性进行分析, 同时分析了系统发育关系。本研究结果可为秀丽兜兰及其他兜兰属植物种质资源鉴定、系统发育、遗传育种、种群恢复以及生物多样性保护等提供理论依据。

## 1 材料与方法

### 1.1 材料

供试材料为迁地保存在广西农业科学院花卉研究所兰科植物种质资源圃 (22°48'N, 108°22'E) 的秀丽兜兰, 取长势良好、无病虫害的新鲜幼嫩叶片擦拭干净, 用液氮速冻后于 -80 °C 保存备用。

## 1.2 方法

**1.2.1 DNA 提取、文库构建及测序** 取叶片 100 mg, 采用改良 CTAB 法<sup>[26]</sup>提取总 DNA, 样品基因组 DNA 质量检测合格后, 用超声波将合格的 DNA 片段化, 然后对片段化的 DNA 进行片段纯化、末端修复、3'端加 A、连接测序接头, 再用琼脂糖凝胶电泳进行片段大小选择, 进行 PCR 扩增形成测序文库, 并进行文库质检, 质检合格的文库用 BGISEQ-500 平台进行测序 (广州佰德生物科技有限公司)。

**1.2.2 叶绿体基因组的组装与注释** 对测序得到的原始数据 (raw data) 进行数据过滤, 去除其中的接头序列及低质量 reads 获得高质量的 clean data。使用 Bowtie2 v.2.3.4.3 软件<sup>[27]</sup>比对近缘物种叶绿体基因组数据库, 将比对上的 reads 数量最多的 3 个 CDS 序列作为项目样品的叶绿体基因组起始组装序列, 采用 NOVOPlasty v4.2.1 软件<sup>[28]</sup>组装叶绿体基因组。利用 GeSeq 和 blast 软件对叶绿体基因组 CDS 序列进行注释, 并利用 GeSeq 和 hmmer 软件对叶绿体基因组 rRNA 序列进行注释, 使用 tRNAscan-SE 和 ARAGORN 软件对叶绿体基因组 tRNA 序列进行注释, 并进行手动调整和确认获得最终注释结果, 最后采用 OrganellarGenomeDRAW 软件制作叶绿体基因组图谱<sup>[29-35]</sup>。将组装注释好的秀丽兜兰基因组序列上传至 GenBank 数据库, 获得登录号: MZ150831。

**1.2.3 叶绿体基因组 SSR 分析** 利用 MISA v1.01 (MicroSATellite identification tool) 软件对秀丽兜兰进行 cpSSR 的分析<sup>[36]</sup>。参数设置为单核苷酸 (mononucleotide)、二核苷酸 (dinucleotide)、三核苷酸 (trinucleotide)、四核苷酸 (tetranucleotide)、五核苷酸 (pentanucleotide) 和六核苷酸 (hexanucleotide) 的重复数阈值分别为 10、6、5、5、5 和 5。

**1.2.4 密码子使用分析** 计算秀丽兜兰 cpDNA 密码子的使用情况和相对密码子的偏好性 (relative synonymous codon usage, RSCU), RSCU 值则以 1 为限, 若 RSCU>1, 该密码子使用频率较高; 若 RSCU<1, 该密码子使用频率较低。

**1.2.5 系统发育关系分析** 从 NCBI 数据库中下载 16 个兜兰属和 2 个杓兰属植物, 利用 mafft 7.0 软件进行 cpDNA 序列比对, 使用 fasttree2.1.10 软件建立系统发育树, 选择 GTR+Gamma 模型, Shimodaira-Hasegawa 检测<sup>[37-38]</sup>。

## 2 结果与分析

### 2.1 秀丽兜兰叶绿体基因组的基本特征

秀丽兜兰叶绿体基因组具有典型的环状四分体结构, 包括 1 个大单拷贝区 (large single-copy region, LSC)、1 个小单拷贝区 (small single-copy region, SSC) 和 2 个反向重复区 (IR), 总长度为 158 298 bp, 其中 LSC、SSC 和 2 个 IR 长度分别为 87 775、949、34 787 bp, LSC、SSC 和 IR 区域的 GC 含量分别为 32.6%、25.1%和 39.1%。秀丽兜兰叶绿体基因组共含 129 个基因, 包括 79 个蛋白编码基因, 38 个 tRNA 基因、8 个 rRNA 基因和 4 个假基因 (2 个 ndhD、1 个 ndhJ、1 个 ndhK), 其中 22 个完整的基因位于 IR 重复区, 包括 10 个蛋白编码基因, 8 个 tRNA 和 4 个 rRNA, 1 个基因 rpl22 位于 LSC-IRB 连接处。IR 重复区域还含有 1 个假基因 ndhD 以及 rps12 基因的第 2 和第 3 个外显子, nps12 的第 1 个外显子位于 LSC 区域 (图 1)。秀丽兜兰叶绿体基因组注释基因信息见表 1。

### 2.2 秀丽兜兰叶绿体基因组简单序列重复

秀丽兜兰叶绿体基因组简单序列重复 (SSR) 信息见表 2, 从中共鉴定出 78 个 SSR 位点, 未发现五核苷酸重复, 最丰富的是单核苷酸重复, 有 66 个, 占总 SSRs 的 84.62%, 其次是二核苷酸重复 8 个, 占总 SSRs 的 10.26%, 三核苷酸重复和四核苷酸重复均为 1 个, 六核苷酸重复 2 个。其中 63 个单碱基重复都是由 A 或 T 构成, 因此, SSRs 的碱基组成偏向使用 A/T 碱基。

### 2.3 秀丽兜兰叶绿体基因组密码子偏好性

秀丽兜兰叶绿体基因组中各氨基酸的相对同义密码子使用频率 (relative synonymous code usage, RSCU) 分析结果见图 2 和表 3。从中可以看出, 高频密码子 (RSCU>1) 共有 32 个, 其中, 16 个以 U 结尾, 13 个以 A 结尾, 以 G 结尾的有 2 个, 以 C 结尾的密码子有 1 个, 说明秀丽兜兰叶绿体基因密码子偏好性以 U 和 A 结尾, 与较低的基因组和蛋白编码基因 GC 含量一致。

### 2.4 系统进化分析

基于叶绿体基因组序列对 18 个兰科植物 (含兜兰属物种 16 个、近缘属杓兰属物种 2 个) 进行系统发育分析, 并构建系统发育树 (图 3)。从图中可以看出, 18 个物种分为 3 大类, 其中, 紫花杓兰和西藏杓兰与其他物种遗传距离较远, 被单

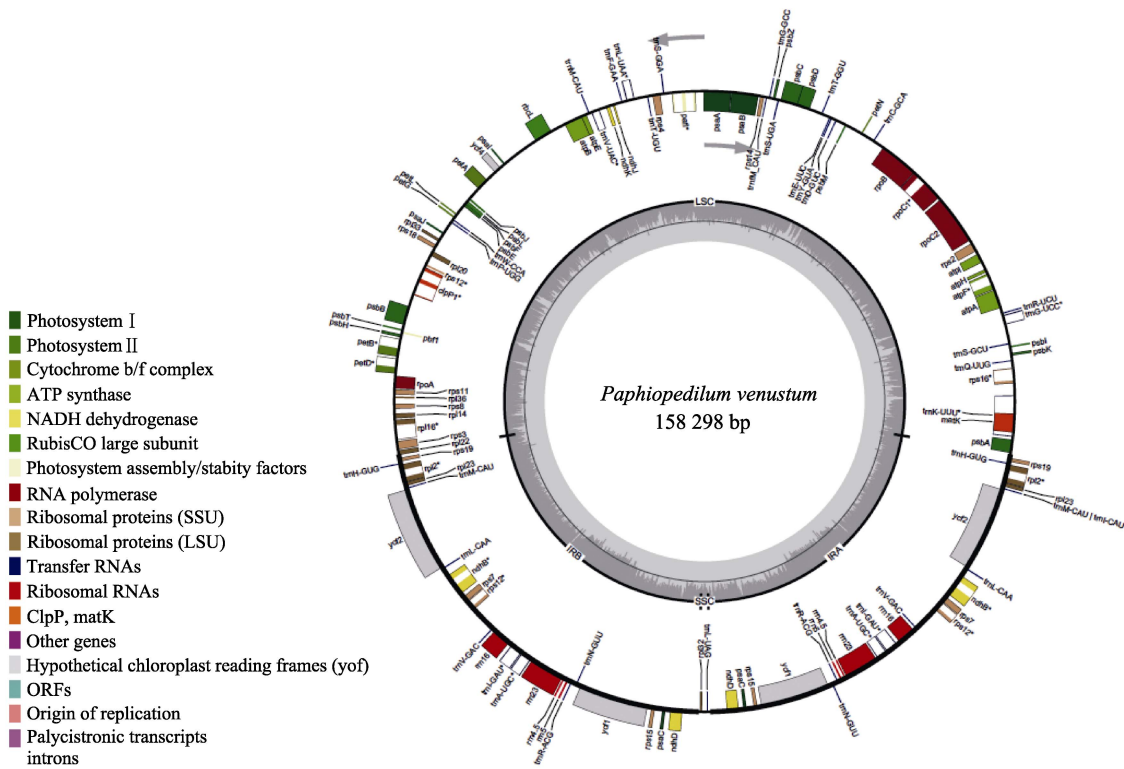


图 1 秀丽兜兰叶绿体基因组图谱  
Fig. 1 Gene map of chloroplast genome of *P. venustum*

表 1 秀丽兜兰叶绿体基因组注释基因列表  
Tab. 1 Genes present in chloroplast genome of *P. venustum*

功能 Function	家族名称 Family name	基因 Gene
光合系统基因	光系统 I 基因	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> <sup>*</sup> , <i>psaI</i> , <i>psaJ</i>
	光系统 II 基因	<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>psbT</i> , <i>psbZ</i> , <i>psbN</i>
	细胞色素 b/f 复合体基因	<i>petA</i> , <i>petB</i> , <i>petD</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	ATP 酶基因	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> , <i>atpH</i> , <i>atpI</i>
	NADH 氧化还原酶基因	<i>ndhB</i> <sup>*</sup> , <i>ndhD</i> <sup>*</sup> , <i>ndhJ</i> , <i>ndhK</i>
	二磷酸核酮糖羧化酶大亚基基因	<i>rbcL</i>
遗传系统基因	RNA 聚合酶亚基基因	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> , <i>rpoC2</i>
	核糖体蛋白小亚基基因	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> <sup>*</sup> , <i>rps8</i> , <i>rps11</i> , <i>rps12</i> <sup>*</sup> , <i>rps14</i> , <i>rps15</i> <sup>*</sup> , <i>rps16</i> , <i>rps18</i> , <i>rps19</i> <sup>*</sup>
	核糖体蛋白大亚基基因	<i>rpl2</i> <sup>*</sup> , <i>rpl14</i> , <i>rpl16</i> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> <sup>*</sup> , <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
	转运 RNA	<i>trnA_UGC</i> <sup>*</sup> , <i>trnC_GCA</i> , <i>trnD_GUC</i> , <i>trnE_UUC</i> , <i>trnF_GAA</i> , <i>trnM_CAU</i> , <i>trnG_GCC</i> , <i>trnG_UCC</i> , <i>trnH_GUG</i> <sup>*</sup> , <i>trnI_CAU</i> <sup>*</sup> , <i>trnI_GAU</i> <sup>*</sup> , <i>trnK-UUU</i> , <i>trnL_CAA</i> <sup>*</sup> , <i>trnL_UAA</i> , <i>trnL_UAG</i> , <i>trnM_CAU</i> , <i>trnN_GUU</i> <sup>*</sup> , <i>trnP_UGG</i> , <i>trnQ_UUG</i> , <i>trnR_ACG</i> <sup>*</sup> , <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> <sup>*</sup> , <i>trnV_UAC</i> , <i>trnW_CCA</i> , <i>trnY_GUA</i>
其他基因	核糖体 RNA	<i>rrn4.5</i> <sup>*</sup> , <i>rrn5</i> <sup>*</sup> , <i>rrn16</i> <sup>*</sup> , <i>rrn23</i> <sup>*</sup>
	翻译起始密码子	<i>infA</i>
	成熟酶基因	<i>matK</i>
	C 型细胞色素合成基因	<i>ccsA</i> <sup>*</sup>
	乙酰辅酶 A 羧化酶亚基	<i>accD</i>
	蛋白酶基因	<i>clpP</i>
假定叶绿体阅读框	<i>ycf1</i> <sup>*</sup> , <i>ycf2</i> <sup>*</sup> , <i>ycf3</i> , <i>ycf4</i>	

注：\*代表 IRS 区域有 2 个基因重复，粗体表示有 3 个拷贝。  
Note: \* means duplicated gene in IRs, bold means three gene copies.

表 2 秀丽兜兰 cpDNA 简单序列重复 (SSR) 信息

Tab. 2 cpDNA simple sequence repeat (SSR) information of *P. venustum*

类型 Type	重复序列 Repeat motif	数量 Number	比例 Percentage/%	合计占比 Total percentage/%
单核苷酸	A	27	40.91	84.62
	C	1	1.52	
	G	2	3.03	
	T	36	54.55	
二核苷酸	AT	2	25.00	10.26
	TA	6	75.00	
三核苷酸	TTA	1	100.00	1.28
四核苷酸	ATAA	1	100.00	1.28
五核苷酸				
六核苷酸	CCTCTT	1	50.00	2.56
	GGAAGA	1	50.00	
合计		78		100

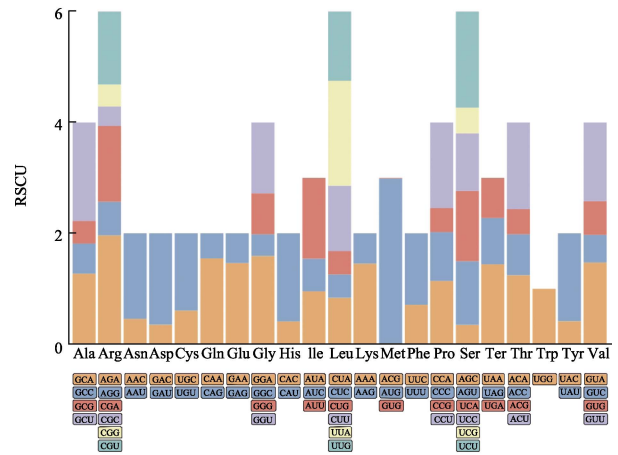


图 2 秀丽兜兰叶绿体基因组密码子 RSCU 分布情况  
Fig. 2 Relative synonymous codon usage (RSCU) distribution of genes in chloroplast genome of *P. venustum*

表 3 秀丽兜兰叶绿体各氨基酸相对同义密码子偏好性

Tab. 3 RSCU of amino acids in *P. venustum*

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	
丙氨酸 Ala	GCA*	401	1.28	半胱氨酸 Cys	UGC	93	0.61	
	GCC	170	0.54		UGU*	213	1.39	
	GCG	128	0.41	谷氨酰胺 Gln	CAA*	711	1.55	
精氨酸 Arg	GCU*	558	1.78		CAG	209	0.45	
	AGA*	520	1.96	天冬酰胺 Asn	AAC	301	0.46	
	AGG	161	0.61		AAU*	1008	1.54	
	甘氨酸 Gly	CGA*	362	1.37	天冬氨酸 Asp	GAC	189	0.35
		CGC	92	0.35		GAU*	877	1.65
CGG		106	0.40	谷氨酸 Glu	GAA*	1071	1.47	
CGU*		347	1.31		GAG	389	0.53	
酪氨酸 Tyr		GGA*	642		1.59	组氨酸 His	CAC	134
	GGC	157	0.39	CAU*	518		1.59	
	亮氨酸 Leu	GGG	297	0.74	赖氨酸 Lys	AAA*	1141	1.46
		GGU*	515	1.28		AAG	426	0.54
		脯氨酸 Pro	UAC	188	0.42	蛋氨酸 Met	ACG	2
UAU*			717	1.58	AUG*		591	2.98
苏氨酸 Thr			CUA	355	0.84	苯丙氨酸 Phe	GUG	2
	CUC		176	0.42	UUC		511	0.71
	CUG		177	0.42	丝氨酸 Ser	UUU*	923	1.29
	CUU*	500	1.18	AGC		117	0.35	
	UUA*	798	1.89	AGU*		381	1.15	
UUG*	527	1.25	UCA*	422		1.27		
终止密码子 Ter	CCA*	285	1.14	UCC*		346	1.04	
	CCC	220	0.88	UCG	154	0.46		
	CCG	109	0.44	UCU*	576	1.73		
	CCU*	385	1.54	异亮氨酸 Ile	AUA	675	0.96	
	色氨酸 Trp	ACA*	389		1.25	AUC	411	0.58
ACC		230	0.74		AUU*	1029	1.46	
ACG		142	0.46	缬氨酸 Val	UGG	437	1.00	
ACU*		487	1.56		GUA*	485	1.48	
终止密码子 Ter		UAA*	38	1.44	GUC	164	0.50	
	UAG	22	0.84	GUG	200	0.61		
	UGA	19	0.72	GUU*	466	1.42		

注：TER 表示终止密码子，\*为高频密码子。  
Note: TER means termination condon, \* means optimal codon.

独聚成一类；麻栗坡兜兰、硬叶兜兰等 5 个物种均归为一类，即宽瓣亚属；秀丽兜兰和带叶兜兰、紫纹兜兰等 11 个物种归为一类，属于兜兰亚属，且秀丽兜兰与紫纹兜兰关系最近。

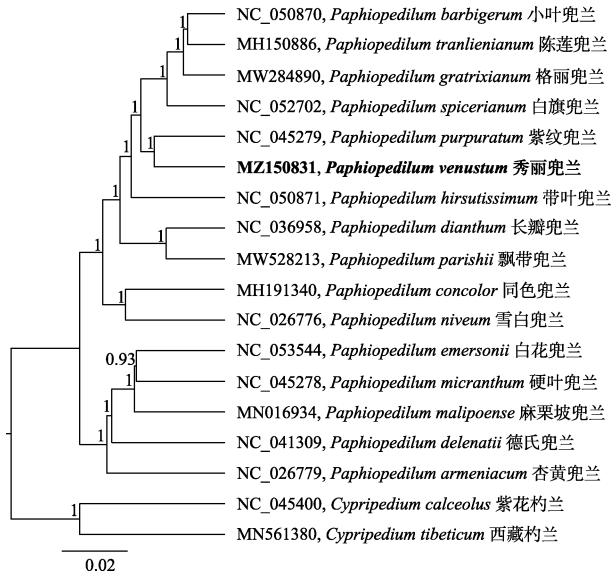


图 3 基于叶绿体基因组 18 种兰科植物构建的系统进化树  
Fig. 3 Phylogenetic tree based on 18 complete chloroplast genome sequences

### 3 讨论

兜兰是兰科 (Orchidaceae) 兜兰属 (*Paphiopedilum* Pfitzer) 植物的统称，因具有拖鞋状的唇瓣又被称为“拖鞋兰”或“仙履兰”，具有极高的观赏价值和商业价值<sup>[1, 39-40]</sup>。本研究采用第二代高通量测序技术对秀丽兜兰叶绿体基因组进行测序、组装和注释，获得了秀丽兜兰完整的叶绿体基因组。研究发现秀丽兜兰叶绿体基因组与前人<sup>[11-19]</sup>研究的兜兰属其他植物相似，具有典型的四分体环状结构，GC 含量远小于 AT 含量，其中 IR 区域 GC 含量最高，这与 rRNA 的 GC 含量较高，且全部定位在这些区域密切相关<sup>[41]</sup>。

简单序列重复 (SSR) 多存在于真核生物的细胞器基因组中，具有丰富的多态性，高度的重复性以及良好的可靠性和共显性，常被用于对种质资源进行遗传结构分析、物种鉴定和群体遗传学研究<sup>[19, 42]</sup>。本研究从秀丽兜兰叶绿体基因组中鉴定出 78 个 SSR 位点，单核苷酸有 66 个，占总 SSRs 的 94.62%，表明单核苷酸重复可能提供了更多的系统发育信息，63 个单碱基重复和其他碱基重复都是由 A 或 T 构成，A/T 重复类型占据主要地位，这与前人<sup>[18-19]</sup>对其他几种兜兰属植物研

究结果一致，得到的这些 SSR 序列可为兜兰属植物筛选 SSR 标记和指纹图谱研究提供依据。

在植物进化过程中，密码子的使用普遍表现出偏好性，自然选择、物种突变和遗传漂变是影响密码子偏好性的主要原因，密码子偏好性用相对同义密码子使用频率 (RSCU) 表示，若密码子无偏好性，则 RSCU 值为 1，如果该密码子比其他同义密码子来说出现更频繁，RSCU 值大于 1，反之亦然<sup>[43-45]</sup>。本研究秀丽兜兰叶绿体基因组 RSCU>1 的密码子共有 32 个，其中以 A 和 U 结尾的有 29 个，超过 90.6%，而 RSCU<1 的密码子则多以 C 或 G 结尾，表明秀丽兜兰叶绿体基因组偏向于使用以 A 或 U 结尾的同义密码子，与丁锐等<sup>[46]</sup>对杓兰属的研究结果一致，这可能与叶绿体基因组较低的 GC 含量有关，进一步表明密码子偏好性可能受基因组 GC 含量的影响<sup>[47]</sup>。根据秀丽兜兰叶绿体基因组的密码子偏好性可对目的基因进行优化，进一步提高基因转化和表达效率，这一研究结果可为秀丽兜兰遗传育种及叶绿体基因工程的外源基因高效、稳定表达提供科学依据。

从 18 种兰科植物叶绿体基因组系统发育分析发现，兜兰属和杓兰属是明显分开的两大类，兜兰属又分两大类，一类包含白花兜兰、硬叶兜兰、麻栗坡兜兰、德氏兜兰和杏黄兜兰；另一类包括小叶兜兰、陈莲兜兰、格丽兜兰、白旗兜兰、紫纹兜兰、秀丽兜兰、带叶兜兰、长瓣兜兰、飘带兜兰、同色兜兰和雪白兜兰，其中秀丽兜兰与紫纹兜兰亲缘关系密切。这一分类与其他学者利用形态学、细胞学等方法得到的分类结果相似<sup>[48-49]</sup>，与 SUN 等<sup>[18]</sup>和胡国家<sup>[19]</sup>通过叶绿体基因组序列建立的系统发育树基本一致，显示出叶绿体基因组序列在物种分类和进化研究中具有更高的分辨率和可靠性<sup>[50]</sup>，证明了植物叶绿体基因组系统发育树的构建在物种鉴定方面的准确性，同时还需要更加全面密集的取样，才能更准确地反映发育关系。

### 4 结论

秀丽兜兰叶绿体基因组为典型的四分结构，总基因组大小为 158 298 bp，总 GC 含量为 35.4%，含 129 个基因，共鉴定出 78 个 SSR 位点，碱基组成偏向使用 A/T 碱基，密码子偏好性以 U 和 A 结尾，系统发育分析显示秀丽兜兰与紫纹兜兰为姐妹种关系。研究结果可为兜兰属植物物种鉴定、

生物进化、系统发育、资源保护及分子生物学等研究提供理论参考。

## 参考文献

- [1] 罗毅波, 贾建生, 王春玲. 初论中国兜兰属植物的保护策略及其潜在资源优势[J]. 生物多样性, 2003(6): 491-498.  
LUO Y B, JIA J S, WANG C L. Conservation strategy and potential advantages of the Chinese *Paphiopedilum*[J]. Biodiversity Science, 2003(6): 491-498. (in Chinese)
- [2] 刘仲健, 陈心启, 陈利君, 雷嗣鹏. 中国兜兰属植物[M]. 北京: 科学出版社, 2009.  
LIU Z J, CHEN X Q, CHEN L J, LEI S P. The genus *Paphiopedilum* in China[M]. Beijing: Science Press, 2009. (in Chinese)
- [3] 杨颖婕, 黄家林, 胡虹, 张石宝. 中国兜兰属植物种质资源保护和利用研究进展[J]. 西部林业科学, 2021, 50(5): 108-112, 119.  
YANG Y J, HUANG J L, HU H, ZHANG S B. Progress on conservation and utilization of *Paphiopedilum* species in China[J]. Journal of West China Forestry Science, 2021, 50(5): 108-112, 119. (in Chinese)
- [4] 覃海宁, 杨永, 董仕勇, 何强, 贾渝, 赵莉娜, 于胜祥, 刘慧圆, 刘博, 严岳鸿, 向建英, 夏念和, 彭华, 李振宇, 张志翔, 何兴金, 尹林克, 林余霖, 刘全儒, 侯元同, 刘演, 刘启新, 曹伟, 李建强, 陈世龙, 金效华, 高天刚, 陈文俐, 马海英, 耿玉英, 金孝锋, 常朝阳, 蒋宏, 蔡蕾, 臧春鑫, 武建勇, 叶建飞, 赖阳均, 刘冰, 林秦文, 薛纳新. 中国高等植物受威胁物种名录[J]. 生物多样性, 2017, 25(7): 696-744.  
QIN H N, YANG Y, DONG S Y, HE Q, JIA Y, ZHAO L N, YU S X, LIU H Y, LIU B, YAN Y H, XIANG J Y, XIA N H, PENG H, LI Z Y, ZHANG Z X, HE X J, YIN L K, LIN Y L, LIU Q R, HOU Y T, LIU Y, LIU Q X, CAO W, LI J Q, CHEN S L, JIN X H, GAO T G, CHEN W L, MA H Y, GENG Y Y, JIN X F, CHANG Z Y, JIANG H, CAI L, ZANG C X, WU J Y, YE J F, LAI Y J, LIU B, LIN Q W, XUE N X. Threatened species list of China's higher plants[J]. Biodiversity Science, 2017, 25(7): 696-744. (in Chinese)
- [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.  
ZHANG Y J, LI D Z. Advances in phylogenomics based on complete chloroplast genomes[J]. Plant Diversity and Resources, 2011, 33(4): 365-375. (in Chinese)
- [6] 王玲, 董文攀, 周世良. 被子植物叶绿体基因组的结构变异研究进展[J]. 西北植物学报, 2012, 32(6): 1282-1288.  
WANG L, DONG W P, ZHOU S L. Structural mutations and reorganizations in chloroplast genomes of flowering plants[J]. Acta Botanica Boreali-Occidentalia Sinica, 2012, 32(6): 1282-1288. (in Chinese)
- [7] 谢海坤, 焦健, 樊秀彩, 张颖, 姜建福, 孙海生, 刘崇怀. 基于高通量测序组装‘赤霞珠’叶绿体基因组及其特征分析[J]. 中国农业科学, 2017, 50(9): 1655-1665.  
XIE H K, JIAO J, FAN X C, ZHANG Y, JIANG J F, SUN H S, LIU C H. Assembling and characteristic analysis of the complete chloroplast genome of *Vitis vinifera* cv. Cabernet Sauvignon from high-throughput sequencing data[J]. Scientia Agricultura Sinica, 2017, 50(9): 1655-1665. (in Chinese)
- [8] REN T, YANG Y, ZHOU T, LIU Z L. Comparative plastid genomes of *Primula* species: sequence divergence and phylogenetic relationships[J]. International Journal of Molecular Sciences, 2018, 19(4): 1050.
- [9] YANG Y C, ZHOU T, DUAN D, YANG J, FENG L, ZHAO G F. Comparative analysis of the complete chloroplast genomes of five *Quercus* species[J]. Frontiers in Plant Science, 2016, 7: 959.
- [10] LI C Y, ZHAO Y L, XU Z G, YANG G Y, PENG J, PENG X Y. Initial characterization of the chloroplast genome of *Vicia sepium*, an important wild resource plant, and related inferences about its evolution[J]. Frontiers in Genetics, 2020, 11: 73.
- [11] GE L P, TANG L, LI L, LUO Y. The complete chloroplast genome of an endangered orchid *Paphiopedilum spicerianum*[J]. Mitochondrial DNA Part B, 2020, 5(3): 3612-3613.
- [12] TANG F L, DENG L L, QIN H Z, SHI Y C. Complete chloroplast genome of *Paphiopedilum emersonii* (Orchidaceae)[J]. Mitochondrial DNA Part B, 2020, 5(3): 3518-3519.
- [13] HU C, JIANG K, ZENG X H, HUANG W C. The complete chloroplast genome sequence of a critically endangered orchid *Paphiopedilum gratixianum* (Orchidaceae)[J]. Mitochondrial DNA Part B, 2022, 7(4): 609-610.
- [14] LI L Q, HUANG J, CEHN L J, ZHANG Q Y, CHEN J B. The complete chloroplast genome of *Paphiopedilum malipoense* (Orchidaceae)[J]. Mitochondrial DNA Part B, 2019, 4(2): 2617-2618.
- [15] CHEN X D, PENG D H, LAN S R, CEHN J, FU W Q. The complete chloroplast genome sequence of *Paphiopedilum purpuratum* (Orchidaceae)[J]. Mitochondrial DNA Part B, 2019, 4(2): 3910-3911.
- [16] KAO H X, ZHAO Y X, YANG M Q, SUN Y J, CHENG J. The complete chloroplast genome sequences of an endangered orchid species *Paphiopedilum parishii* (Orchidaceae)[J]. Mitochondrial DNA Part B, 2021, 6(9): 2521-2522.
- [17] SUI C H, LIU S Y, LIU H M, LIU H Z. The complete

- chloroplast genome of *Paphiopedilum tranlimianum* (Orchidaceae)[J]. Mitochondrial DNA Part B, 2018, 3(2): 820-822.
- [18] SUN Y, ZOU P S, JIANG N N, FANG Y F, LIU G F. Comparative analysis of the complete chloroplast genomes of nine *Paphiopedilum* species[J]. Frontiers in Genetics, 2022, 12: 772415.
- [19] 胡国家. 兰科兜兰属与兰属植物比较叶绿体基因组学与系统发育研究[D]. 西安: 西北大学, 2020.  
HU G J. The complete chloroplast genomes of *Paphiopedilum* and *Cymbidium* (Orchidaceae) species: comparative genomic and phylogenetic analyses[D]. Xi'an: Northwest University, 2020. (in Chinese)
- [20] KAUR S, BHUTANI K K. Asymbiotic seed germination and multiplication of an endangered orchid-*Paphiopedilum venustum* (Wall. ex Sims.)[J]. Acta Societatis Botanicorum Poloniae, 2016, 85(2): 1-11.
- [21] KATAKI S K. Notes on systematics, cytology and culture of *Paphiopedilum venustum* (Lindl.) Pfitz[J]. Nelumbo-The Bulletin of the Botanical Survey of India, 1963, 5(1): 37-38.
- [22] RUTTER J C, WILLMER C M. A light and electron microscopy study of the epidermis of *Paphiopedilum* spp. with emphasis on stomatal ultrastructure[J]. Plant, Cell and Environment, 1979, 2(3): 211-219.
- [23] THORPE N. Accumulation of carbon-compounds in the epidermis of 5 species with either different photosynthetic systems or stomatal structure[J]. Plant, Cell and Environment, 1980, 3(6): 451-460.
- [24] BAG T K. Two new orchid hosts of *Sclerotium rolfsii* from India[J]. Plant Pathology, 2004, 53(2): 255.
- [25] GURUNG D B, GYELTSHEN N, TOBGAY K, DALSTRM S, WANGDI J, GHALLEY B B, CHAIDA L, PHUNTSO, GYELTSHEN N, DAWA K, WANGCHUK T, PRADHAN R, HOIJER T, GYEITSHEN C. Distribution and habitats of *Paphiopedilum* Pfitzer (Orchidaceae) known to occur in Bhutan[J]. Journal of Threatened Taxa, 2019, 11(9): 14101-14111.
- [26] DOYLE J J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue[J]. Phytochem Bull, 1987, 19: 11-15.
- [27] LANGMEAD B, SALZBERG S. Fast gapped-read alignment with Bowtie 2[J]. Nat Methods, 2012, 9: 357-359.
- [28] DIERCKXSENS N, MARDYLYN P, SMITS G. NOVO-Plasty: *de Novo* assembly of organelle genomes from whole genome data[J]. Nucleic Acids Research, 2017, 45: e18.
- [29] TILLICH M, LEHWARK P, PELLIZZER T, ULBRICHT-JONES E, FISCHER A, BOCK R, GREINER S. Ge-Seq-Versatile and accurate annotation of organelle genomes[J]. Nucleic Acids Res, 2017, 45: W6-W11.
- [30] JOHNSON M, ZARETSKAYA I, RAYTSELIS Y, MEREZHUK Y, MCGINNIS S, MADDEN T L. NCBI BLAST: a better web interface[J]. Nucleic Acids Research, 2008, 36: W5-W9.
- [31] EDDY S R. Accelerated profile HMM searches[J]. PLoS Computational Biology, 2011, 7(10): e1002195.
- [32] CHAN P P, LOWE T M. tRNAscan-SE: searching for tRNA genes in genomic sequences[J]. Gene Prediction, 2019, 1962: 1-14.
- [33] LASLETT D, CANBACK B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences[J]. Nucleic Acids Research, 2004, 32: 11-16.
- [34] KEARSE M, MOIR R, WILSON A, STONES-HAVAS S, CHEUNG M, STURROCK S, BUXTON S, COOPER A, MARKOWITZ S, DURAN C, THIERER T, ASHTON B, MEINTJES P, DRUMMOND A. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data[J]. Bioinformatics, 2012, 28: 1647-1649.
- [35] GREINER S, LEHWARK P, BOCK R. OrganellarGenomeDRAW (OGDRAW) Version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes[J]. Nucleic Acids Research, 2019, 47: W59-W64.
- [36] BEIER S, THIEL T, MNCH T, SCHOLZ U, MASCHER M. MISA-web: a web server for microsatellite prediction[J]. Bioinformatics, 2017, 33: 2583-2585.
- [37] KATO H, ROZEWICKI J, YAMADA K D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization[J]. Briefings in Bioinformatics, 2019, 20: 1160-1166.
- [38] PRICE M N, DEHAL P S, ARKIN A P. FastTree 2-Approximately maximum-likelihood trees for large alignments[J]. PLoS One, 2010, 5: e9490.
- [39] 曾宋君, 田瑞雪, 陈之林, 吴坤林, 段俊. 兜兰属植物杂交育种研究进展[J]. 热带亚热带植物学报, 2010, 18(4): 459-468.  
ZENG S J, TIAN R X, CHEN Z L, WU K L, DUAN J. Research progress on cross breeding of *Paphiopedilum*[J]. Journal of Tropical and Subtropical Botany, 2010, 18(2): 459-468. (in Chinese)
- [40] 尹玉莹, 房林, 李琳, 陈砚, 符稳群, 吴坤林, 曾宋君. 兜兰属植物花期调控研究进展[J]. 热带作物学报, 2022, 43(4): 769-778.  
YIN Y Y, FANG L, LI L, CHEN Y, FU W Q, WU K L, ZENG S J. Advances in flowering regulation of *Paphiopedilum*[J]. Chinese Journal of Tropical Crops, 2022, 43(4): 769-778. (in Chinese)

- [41] XIAO T W, XU Y, JIN L, LIU T J, YAN H F, GE X J. Conflicting phylogenetic signals in plastomes of the tribe *Lauraeae* (Lauraceae)[J]. PeerJ, 2020, 8: e10155.
- [42] 孙志轩, 敖平星, 毕玉芬, 赵雁. ‘德钦’紫花苜蓿叶绿体基因组序列及特征分析[J]. 草地学报, 2022, 30(2): 320-328.  
SUN Z X, AO P X, BI Y F, ZHAO Y. Complete chloroplast genomes sequence and characteristics analysis of *Medicago sativa* ‘Deqin’ [J]. Acta Agrestia Sinica, 2022, 30(2): 320-328. (in Chinese)
- [43] LIU Q P, XUE Q Z. Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species[J]. Journal of Genetics, 2005, 84(1): 55-62.
- [44] 李卫英, 辛静, 赵文植, 董章宏, 马路遥, 夏茂甜, 高洁, 辛培尧. 朱蕉叶绿体基因组特征与系统发育分析[J]. 热带作物学报, 2022, 43(4): 653-665.  
LI W Y, XIN J, ZHAO W Z, DONG Z H, MA L Y, XIA M T, GAO J, XIN P Y. Genomic characteristics and phylogenetic analysis of chloroplast of *Cordyline fruticosa*[J]. Chinese Journal of Tropical Crops, 2022, 43(4): 653-665. (in Chinese)
- [45] 张文娟. 基于密码子水平的生物信息学分析及进化研究[D]. 上海: 复旦大学, 2006.  
ZHANG W J. Codon analysis and its application in bioinformatics and evolutionary studies[D]. Shanghai: Fudan University, 2006. (in Chinese)
- [46] 丁锐, 胡兵, 宗小雁, 韩辰阳, 张丽杰, 陈旭辉. 杓兰叶绿体基因组密码子偏好性分析[J]. 林业科学研究, 2021, 34(5): 177-185.  
DING R, HU B, ZONG X Y, HAN C Y, ZHANG L J, CHEN X H. Analysis of codon usage bias in the chloroplast genome of *Cypripedium*[J]. Forest Research, 2021, 34(5): 177-185. (in Chinese)
- [47] 刘潮, 韩利红, 彭悦, 樊瑞卿, 王宇默. 黄丹木姜子叶绿体基因组特征分析[J]. 南方农业学报, 2022, 53(1): 12-20.  
LIU C, HAN L H, PENG Y, FAN R Q, WANG Y M. Characteristics of chloroplast genome of *Litsea elongata* (Wall. ex Nees) Benth. et Hook. f.[J]. Journal of Southern Agriculture, 2022, 53(1): 12-20. (in Chinese)
- [48] COX A V, PRIDGEON A M, ALBERT V A, CHASE M W. Phylogenetics of slipper orchids (Cypripedioideae, Orchidaceae): nuclear rDNA ITS sequences[J]. Plant Systematics and Evolution, 1997, 208: 197-223.
- [49] CRIBB P J. The genus *Paphiopedilum*[M]. Kota Kinabalu: National History Publications (Borneo) in Association with Royal Botanic Gardens, Kew. 1998.
- [50] ZHANG Y Y, TIAN Y L, TNG D Y P, ZHOU J B, ZHANG Y T, WANG Z W, LI P F, WANG Z S. Comparative chloroplast genomics of *Litsea* Lam. (Lauraceae) and its phylogenetic implications[J]. Forests, 2021, 12: 744.