

• 农业生物技术 •

红边龙血树叶绿体基因组特征及其系统发育分析

王子豪, 郭佳乐, 范琪, 田泽园, 王雪晴, 郑维, 黄罗冬*

广西大学生命科学与技术学院, 广西 南宁 530004

王子豪, 郭佳乐, 范琪, 田泽园, 王雪晴, 郑维, 黄罗冬. 红边龙血树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.

WANG Zihao, GUO Jiale, FAN Qi, TIAN Zeyuan, WANG Xueqing, ZHENG Wei, HUANG Luodong. Characteristics of the chloroplast genome of *Dracaena marginata* and phylogenetic analysis[J]. Chinese Journal of Biotechnology, 2023, 39(7): 2926-2938.

摘 要: 红边龙血树(*Dracaena marginata*)是一种在全球广泛种植的龙血树属园艺植物,具有较高的观赏价值和药用价值。本研究首次利用高通量测序技术对红边龙血树叶片进行全基因组测序,组装得到完整的叶绿体基因组序列,并进行注释、序列特征比较和系统发育分析。结果表明,红边龙血树叶绿体基因组包含一个典型的四分体结构,长度为 154 926 bp,是目前已报道的龙血树属中叶绿体基因组最小的物种;共拥有 132 个基因,包含 86 个编码蛋白基因、38 个转运 RNA 基因和 8 个核糖体 RNA 基因;密码子偏好性分析发现存在偏好使用 A/U 碱基结尾的现象,整体上密码子偏好性较低;共鉴定出 46 个简单重复序列位点和 54 个长重复序列,分别在大单拷贝区与反向重复区有最大检出率;种间边界分析发现边界区域基因存在相对位置差异,扩张收缩情况总体较为相似;与近缘种进行系统发育分析,红边龙血树与细枝龙血树聚为一类,关系最近,符合形态学分类特征。对红边龙血树叶绿体基因组的解析为龙血树属植物的物种鉴定、遗传多样性和叶绿体转基因工程等提供了重要数据基础。

关键词: 红边龙血树; 叶绿体基因组; 简单重复序列位点分析; 系统发育; 比较基因组学分析

资助项目: 巴马县人才科技计划项目(202100180); 广西研究生教育创新计划项目(YCSW2022028)

This work was supported by the Bama County Program for Talents in Science and Technology, Guangxi, China (202100180) and by the Innovation Projects of Guangxi Graduate Research (YCSW2022028).

*Corresponding author. E-mail: ynhuangld@gxu.edu.com

Received: 2022-11-09; Accepted: 2023-01-09; Published online: 2023-05-25

Characteristics of the chloroplast genome of *Dracaena marginata* and phylogenetic analysis

WANG Zihao, GUO Jiale, FAN Qi, TIAN Zeyuan, WANG Xueqing, ZHENG Wei, HUANG Luodong*

College of Life Science and Technology, Guangxi University, Nanning 530004, Guangxi, China

Abstract: *Dracaena marginata* is a widely cultivated horticultural plant in the world, which has high ornamental and medicinal value. In this study, the whole genome of leaves from *D. marginata* was sequenced by Illumina HiSeq 4000 platform. The chloroplast genome were assembled for functional annotation, sequence characteristics and phylogenetic analysis. The results showed that the chloroplast genome of *D. marginata* composed of four regions with a size of 154 926 bp, which was the smallest chloroplast genome reported for *Dracaena* species to date. A total of 132 genes were identified, including 86 coding genes, 38 tRNA genes and 8 rRNA genes. Codon bias analysis found that the codon usage bias was weak and there was a bias for using A/U base endings. 46 simple sequence repeat and 54 repeats loci were detected in the chloroplast genome, with the maximum detection rate in the large single copy region and inverted repeat region, respectively. The inverted repeats boundaries of *D. marginata* and *Dracaena* were highly conserved, whereas gene location differences occurred. Phylogenetic analysis revealed that *D. serrulata* and *D. cinnabari* form a monophyletic clade, which was the closest relationship and conformed to the morphological classification characteristics. The analysis of the chloroplast genome of *D. marginata* provides important data basis for species identification, genetic diversity and chloroplast genome engineering of *Dracaena*.

Keywords: *Dracaena marginata*; chloroplast genome; simple sequence repeat site analysis; phylogeny; comparative genomics analysis

红边龙血树(*Dracaena marginata*)隶属于百合科(Liliaceae)龙血树属(*Dracaena*), 也称为千年木, 原产于马达加斯加和毛里求斯, 包含三色千年木、五彩千年木、七彩千年木等品种^[1]。红边龙血树茎秆细长, 一般在成熟后开始分枝, 其顶部为一簇边缘呈紫红色的细狭叶片, 整体姿态优雅美观, 被广泛应用于园林园艺、城市景观美化与室内陈设栽培。红边龙血树具有较强适应性, 特别是在季节性干旱及不规则降水环境下可长期生存, 因而在世界各地都有规模化引进种植。目前, 关于红边龙血树的研究多集中于室内外栽培和园艺研究相关, 而对于红边龙血树的物

种鉴定、遗传多样性和系统发育分析较少^[2-4]。

叶绿体是植物最重要的细胞器之一, 具有独立于核基因组之外、结构简单易于测序、在植物进化中高度保守等特点, 能为植物的进化分类与资源利用提供可靠理论依据^[5-6]。如基于叶绿体基因组可显著提高对栎属(*Quercu*)物种的辨别能力^[7]; 叶绿体基因组分析发现扁果草(*Isopyrum anemonoides*)与东北扁果草(*Isopyrum manshuricum*)虽在形态上较为相似但应分别独立为一个属^[8]等。同时, 以叶绿体转化为主的叶绿体基因工程, 在外源基因表达水平和转基因植物安全性有明显优势, 并能互补核转化在控制

转基因沉默和遗传稳定性方面的局限性^[9],在黄蒿和烟草中以叶绿体基因组为平台通过转基因工程已经实现了抗性性状的改良和产量的提升^[10-11]。可见叶绿体基因组的解析对植物遗传、农艺等多方面研究有重要的贡献。

在龙血树属物种叶绿体基因组解析方面,2019年首次报道6个龙血树属物种的叶绿体基因组,通过叶绿体基因组序列可将不同的物种进行准确的识别和分类鉴定^[12];随后,对加那利龙血树(*Dracaena draco*)叶绿体基因组的研究,揭示了其特殊的结构特征与在龙血树属内的进化地位^[13];2022年通过对索科特拉龙血树(*Dracaena cinnabari*)和山樱花龙血树(*Dracaena serrulata*)叶绿体基因组的比较分析,揭示了二者的分子特征及独立于其他龙血树属植物的系统发育关系^[14]。但由于龙血树属物种多样性较为丰富,对不同种类、形态特征和地理来源的龙血树物种研究较少,叶绿体基因组研究相对落后,截至目前仅有10种共12个龙血树叶叶绿体基因组完成测序,难以满足对龙血树属植物更深入的分子系统与基因工程研究需求,亟待补充与完善。

本研究首次对红边龙血树进行了叶绿体基因组测序分析,并与已报道的龙血树属物种叶绿体基因组进行了综合详细的比较,分析结果不仅有助于对红边龙血树进行分类鉴定和应用研究,对红边龙血树进行叶绿体基因组比较分析,同样丰富了龙血树属植物细胞器基因组信息库,以利于完善龙血树属起源与进化、品种改良以及种质资源开发利用的研究内容。

1 材料与方法

1.1 实验材料

本研究所用的红边龙血树种植于广西大学环境与资源微生物实验室苗圃(经度108.29,纬度22.85),供试材料生长健壮、长势良好。用

于比较分析的近缘种叶绿体序列均从NCBI(<https://www.ncbi.nlm.nih.gov/>)的GenBank数据库下载,所属龙血树物种包括2个被鉴定为剑叶龙血树(*Dracaena cochinchinensis*, MF943127; MN200195),2个被鉴定为海南龙血树(*Dracaena cambodiana*, MH293451; MN200194)以及被鉴定为长花龙血树(*Dracaena angustifolia*, MN200193)、细枝龙血树(*Dracaena elliptica*, MN200196)、河口龙血树(*Dracaena hokouensis*, MN200197)、矮龙血树(*Dracaena terniflora*, MN200198)、加那利龙血树(MN990038)、山樱花龙血树(MT408026)、金心香龙血树(*Dracaena fragrans*, MW123093)、索科特拉龙血树(OK235335)的叶绿体序列,共12个。其中GenBank登录号为MF943127的剑叶龙血树经分析认为鉴定有误,并非剑叶龙血树,文中名称以一种龙血树(*Dracaena* sp. MF943127)表示。

1.2 方法

1.2.1 基因组提取与测序

剪取新鲜嫩叶,利用十六烷基三甲基溴化铵法(cetyltrimethylammonium bromide, CTAB)提取叶片总DNA,经质量检测合格后构建350 bp文库,利用二代Illumina HiSeq4000平台进行双末端测序,测序由深圳华大科技有限公司完成,下机后对原始数据过滤,共得到48.7 Gb clean reads。

1.2.2 叶绿体基因组组装与注释

过滤后的数据使用GetOrganelle软件^[15]进行从头组装,经手动检查校正获得一条环状叶绿体基因组序列。使用CPGAVAS2(<http://47.90.241.85:16019/analyzer/home/>)在线注释工具进行在线注释^[16],以海南龙血树叶叶绿体基因组(Genbank登录号MN200194)的注释信息作为参考,并进行多轮手动矫正获得了完整的注释信息,上传NCBI的GenBank数据库获得登录号为OP756062。叶绿体基因组图谱利用

Chloroplast 在线工具(<https://irscope.shinyapps.io/Chloroplast/>)绘制^[17]。

1.2.3 叶绿体基因组特征分析

密码子偏好性分析使用 Codon W (v1.4.2) 软件对叶绿体基因组蛋白质编码区的密码子数量、密码子使用频率和同义密码子相对使用度(relative synonymous codon usage, RSCU)进行分析与统计。之后使用在线软件 REPuter (设置参数: maximum computed repeats 为 1 000, minimal repeat size 为 30, hamming distance 为 3)对红边龙血树长重复序列进行分析^[18], 通过 MISA 软件(默认参数)进行简单重复序列(simple sequence repeat, SSR)位点分析^[19]。进一步使用 IRscope 在线可视化工具, 对龙血树属叶绿体基因组四分体边界进行比较分析^[20]。

1.2.4 红边龙血树与其近缘物种系统发育分析

依据 NCBI 在线 BLAST 比对结果, 下载万年青属的开口箭(*Rohdea chinensis*, MH356725)、弯蕊开口箭(*Rohdea wattii*, MW822041)、黄精属的卷叶黄精(*Polygonatum cirrhifolium*, MZ150834)、多花黄精(*Polygonatum cyrtoneura*, MW248135)、长梗黄精(*Polygonatum filipes*, MZ150842), 所下序列与龙血树属 11 个物种 13 个叶绿体共同使用 Homblocks 软件进行多序列比对^[21], 使用 MEGA X 软件(预测模型为“GTR + G + I”, bootstrap value 1 000)对龙血树属植物以及近缘种构建叶绿体基因组系统发育树^[22]。

2 结果与分析

2.1 红边龙血树叶叶绿体序列特征分析

本研究所用的红边龙血树标本形态特征如图 1 所示, 其茎秆细长, 茎秆和叶片边缘呈红色或紫色, 剪取其健康的叶片, 提取 DNA 进行全基因组测序。红边龙血树叶叶绿体基因组大小为 154 926 bp, 总的 GC 含量为 37.54%, 具有典型

的四分体结构: 2 个反向重复区(inverted repeats, IR)将整个叶绿体基因组分隔为 1 个大单拷贝区(large single-copy, LSC)和 1 个小单拷贝区(small single-copy region, SSC)(图 2)。大单拷贝区和小单拷贝区长度分别为 83 883 bp 和 18 515 bp, GC 含量分别为 35.54%和 31.20%。2 个反向重复区大小均为 26 524 bp, GC 含量为 42.89%。

2.2 红边龙血树叶叶绿体基因类型分析

红边龙血树叶叶绿体基因组共注释出 132 个基因, 包括 86 个编码蛋白质基因、38 个转运 RNA (transfer RNA, tRNA)基因和 8 个核糖体 RNA (ribosomal RNA, rRNA)基因。进一步根据基因功能分为 4 类: 光合作用基因(photosynthesis gene)、自我复制基因(self-replication)、其他基因(other genes)和未知功能基因(unknown genes)。其中, 20 个基因存在多个拷贝(表 1), 23 个基因含有内含子, 包括 *pafl*、*clpP1*、*rps12* 基因有 2 个内含子(表 2)。

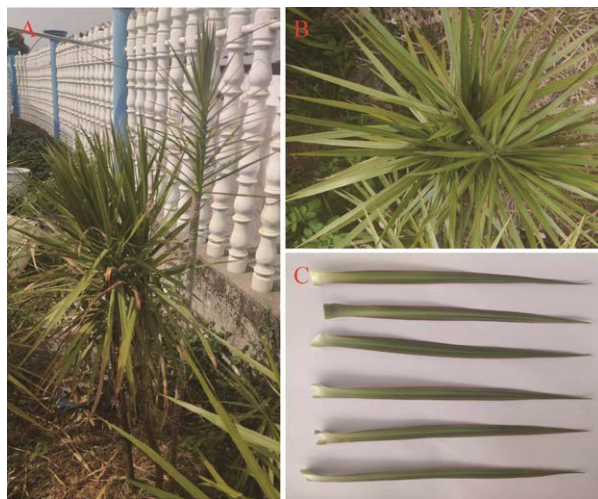


图 1 红边龙血树的植株和叶片形态特征 A: 红边龙血树全植株形态. B: 植株顶视图. C: 红边龙血树叶片

Figure 1 Morphological feature of *Dracaena marginata*. A: The whole plant morphology of *Dracaena marginata*. B: The top view of *Dracaena marginata*. C: The leaf appearance of *D. marginata*.

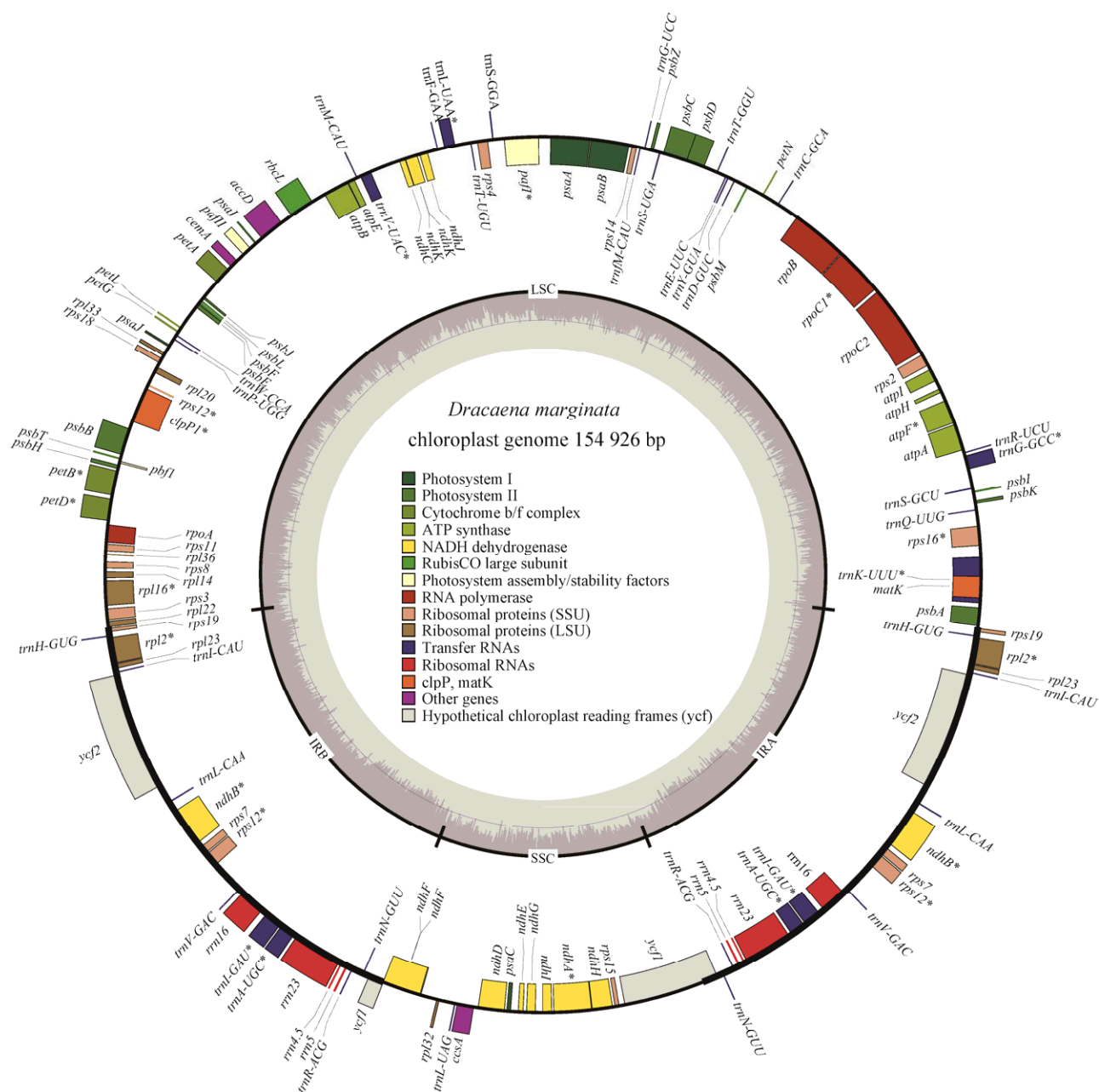


图2 红边龙血树叶绿体基因组图谱

Figure 2 The gene map of chloroplast genome of *Dracaena marginata*.

2.3 密码子偏好性分析

红边龙血树叶绿体基因组中有效密码子数 (effective number of codon, ENC) 为 55.59, 密码子的 GC 含量比例为 38.6%, 其中第 3 密码子是 GC 的比例为 36%, 密码子偏好性与龙血树

(GenBank 登录号 MF943127) 和海南龙血树最接近。在整个红边龙血树叶绿体基因组中, 与其他龙血树叶绿体基因组一样, 最常用到的密码子是编码苯丙氨酸(Phe)的 AAA (数量为 2 161), 而编码精氨酸(Arg)的 AGA 具有最高相对同义

密码子使用度(RSCU 值为 1.98), 且也是唯一高频使用的密码子(图 3)。此外, 还检测到 33 个 RSCU 值大于 1.00 的密码子(终止密码子除外), 它们当中有 27 个以 A 或 U 结尾, 仅有 6 个以 C 或 G 结尾, 这一结果与龙血树属其他物种相似, 表明在龙血树属的物种进化上叶绿体基因组密码子更倾向于以 A 或 U 结尾。

2.4 简单重复序列(SSR)与长重复序列分析

在红边龙血树叶体基因组中共预测到 54 个长重复序列(30–71 bp), 包含 23 个正向重复序列(forward repeat, F)与 31 个回文重复序列

(palindromic repeat, P), 定位于 LSC 区、SSC 区和 IR 区的占比分别为 38.9%、5.6%和 55.6% (图 4B)。在整个叶绿体基因组中未检测到反向重复序列(reverse repeat, R)与互补重复序列(complementary repeat, C)。此外, 还检测到 46 个简单重复序列(SSR), 包含 3 个单碱基重复序列与 3 个双碱基重复序列(图 4A)。其中定位于 LSC 区的位点有 30 个(占比 65.2%), 定位于 SSC 区的位点有 10 个(占比 21.7%), 定位于 IR 区的位点有 4 个(占比 8.7%)。在所有简单重复序列中, 单碱基重复序列(mono-SSRs)均为 A/T 重复类型, 其中, 重复单元

表 1 红边龙血树叶体基因组基因构成

Table 1 List of genes of *Dracaena marginata* chloroplast genome

Category of genes	Gene type	Group of genes
Photosynthesis gene	Rubisco large subunit	<i>rbcL</i>
	Subunits of ATP synthase	<i>atpA, atpB, atpE, atpF, atpH, atpI</i>
	Subunits of photosystem I	<i>psaA, psaB, psaC, psaI, psaJ</i>
	Subunits of photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbT, psbZ</i>
	Subunits of NADH-dehydrogenase	<i>ndhA, ndhB*, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
	Subunits of cytochrome b/f complex	<i>petA, petB, petD, petG, petL, petN</i>
Self-replication	DNA dependent RNA polymerase	<i>rpoA, rpoB, rpoC1, rpoC2</i>
	Large subunit of ribosome	<i>rpl14, rpl16, rpl2*, rpl20, rpl22, rpl23*, rpl32, rpl33, rpl36</i>
	Small subunit of ribosome	<i>rps11, rps12*, rps14, rps15, rps16, rps18, rps19*, rps2, rps3, rps4, rps7*, rps8</i>
	Ribosomal RNA	<i>rrn16*, rrn23*, rrn4.5*, rrn5*</i>
	Transfer RNA	<i>trnA-UGC*, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-GCC, trnG-UCC, trnH-GUG*, trnI-CAU*, trnI-GAU*, trnK-UUU, trnL-CAA*, trnL-UAA, trnL-UAG, trnM-CAU, trnN-GUU*, trnP-UGG, trnQ-UUG, trnR-ACG*, trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC*, trnV-UAC, trnW-CCA, trnY-GUA, trnY-M-CAU</i>
Other genes	Subunit of acetyl-CoA-carboxylase	<i>accD</i>
	c-type cytochrom synthesis gene	<i>ccsA</i>
	Envelop membrane protein	<i>cemA</i>
	Maturase gene	<i>matK</i>
Unkown genes	Conserved open reading frames	<i>ycf1*, ycf2*</i>

*: Two gene copies.

表 2 红边龙血树叶绿体多外显子的基因信息

Table 2 Information of gene with multiple exons of chloroplast genome of *Dracaena marginata*

Gene	Strand	Length of sequence (bp)				
		Exon I	Intron I	Exon II	Intron II	Exon III
<i>rps16</i>	LSC	35	868	218		
<i>atpF</i>	LSC	143	828	404		
<i>rpoC1</i>	LSC	431	719	1 625		
<i>pafl</i>	LSC	123	731	219	744	162
<i>rps12</i>	LSC-TR	25	543	231	—	113
<i>clpP1</i>	LSC	68	813	290	645	224
<i>petB</i>	LSC	5	763	641		
<i>petD</i>	LSC	7	747	474		
<i>rpl16</i>	LSC	8	1 017	398		
<i>rpl2</i>	IR	390	653	430		
<i>ndhB</i>	IR	776	700	755		
<i>rps12</i>	LSC-TR	113	—	231	543	25
<i>ndhA</i>	SSC	557	1 077	539		
<i>ndhB</i>	IR	776	700	755		
<i>rpl2</i>	IR	390	653	430		
<i>trnK-UUU</i>	LSC	36	2 567	39		
<i>trnG-GCC</i>	LSC	22	717	36		
<i>trnL-UAA</i>	LSC	16	545	49		
<i>trnV-UAC</i>	LSC	38	587	36		
<i>trnI-GAU</i>	IR	41	938	34		
<i>trnA-UGC</i>	IR	37	817	34		
<i>trnA-UGC</i>	IR	37	817	34		
<i>trnI-GAU</i>	IR	41	938	34		

—: Trans-splicing structure, without introns.

为 A 与重复单元为 T 的 SSR 位点均在重复频率为 10 时数目最多, 均为 13 个。双碱基重复序列(di-SSRs)均为 AT/TA 重复类型, 重复频率均为 6, 其中有一个 TA 重复位点位于 LSC 区, 2 个 AT 重复位点分别位于 LSC 区和 SSC 区。与其他龙血树叶绿体基因组相比, 红边龙血树叶绿体基因组长重复序列的类型与分布与细枝龙血树、龙血树(GenBank 登录号 MF943127)、长花

龙血树和金心香龙血树相似, 短重复序列的类型和分布与河口龙血树、细枝龙血树和龙血树(GenBank 登录号 MF943127)相似。13 个叶绿体基因组中, 海南龙血树(GenBank 登录号 MH293451)检测出最多的长重复序列位点(数目为 140), 山樱花龙血树次之(数目为 136), 但检测出数目最多的简单重复序列位点(数目为 63), 与 Ahmad 等^[14]的研究结果相似。

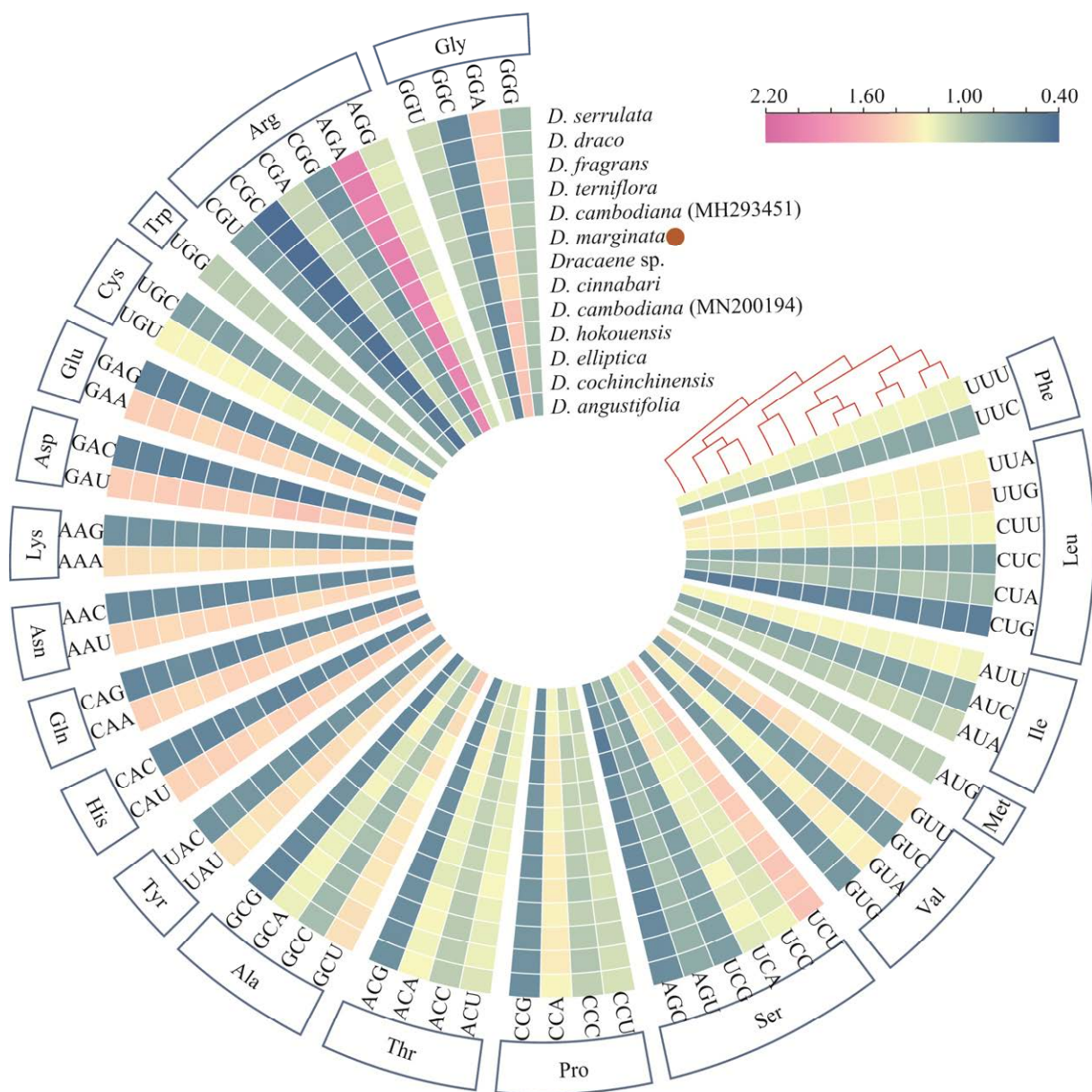


图3 13个龙血树叶叶绿体基因组蛋白编码基因密码子RSCU热图

Figure 3 Heatmap plot of codon RSCU of protein-coding genes in 13 *Dracaena* chloroplast genome.

2.5 红边龙血树叶叶绿体IR区边界结构差异分析

IR区通常高度保守,对维持叶绿体基因组的稳定有着重要作用^[23]。对叶绿体中IR区边界的结构分析发现13个龙血树叶叶绿体LSC区、SSC区、IR区大小分别为83 506–84 988 bp、18 402–18 657 bp、26 455–26 530 bp(图5)。

在13个龙血树叶叶绿体基因组中,红边龙血树的LSC区与SSC区最小。由两个IR区分隔单拷贝区所形成的边界上,*rpl22*与*rps19*基因分列在LSC/IRb(JLB)边界两侧,*ycf1*和*ndhf*基因分列或横跨于IRb/SSC(JSB)边界,*ycf1*基因横跨IRa/LSC(JSA)边界,*rps19*和*psbA*基因分列于SSC/Ira(JLA)边界。

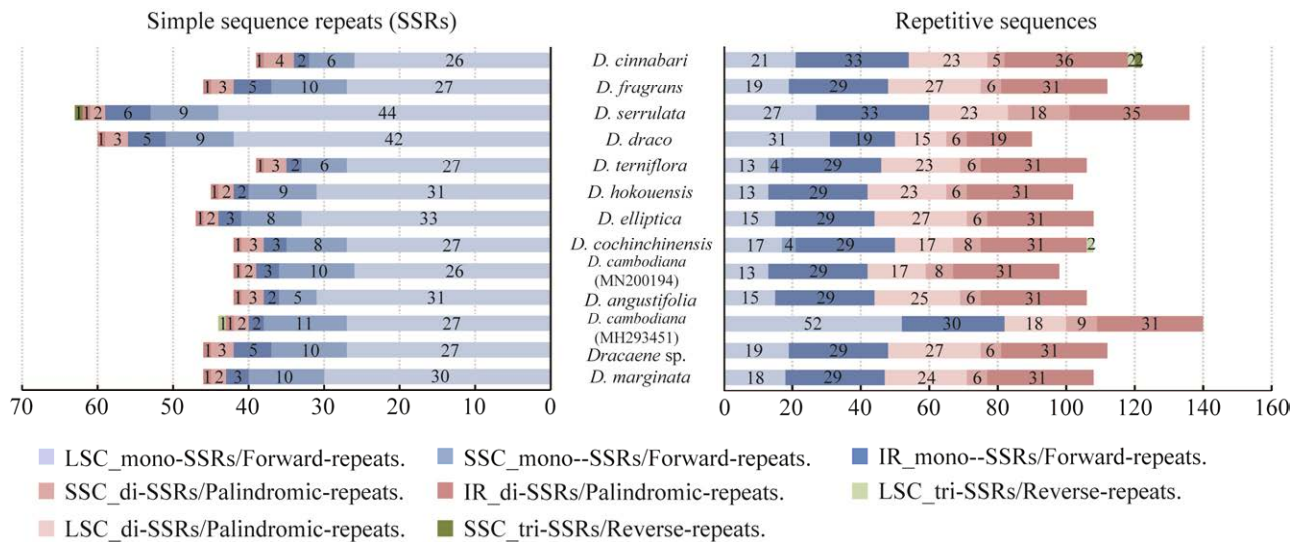


图 4 13 个龙血树叶叶绿体基因组微卫星重复序列(SSR)与长重复序列的类型与分布

Figure 4 The type and distribution of SSRs and repeats in 13 *Dracaena* chloroplast genomes.

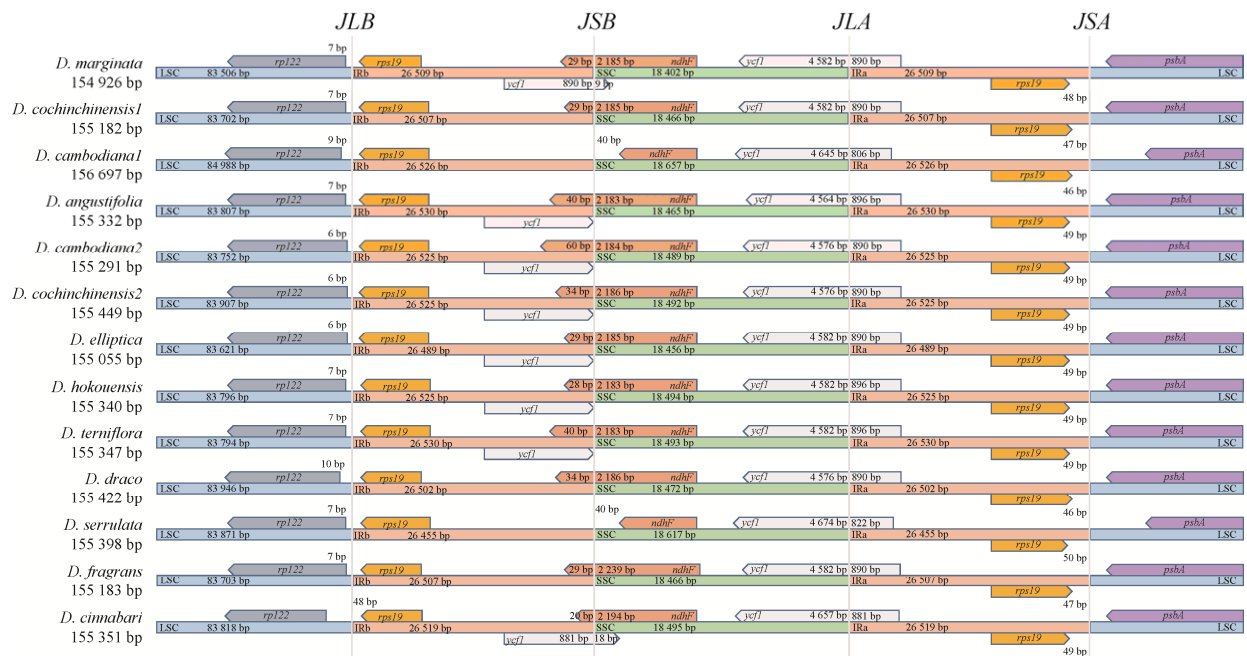


图 5 龙血树属 13 个叶绿体基因组 LSC、SSC 及 IR 区边界比较分析

Figure 5 LSC, SSC and IR boundaries of the chloroplast genomes in 13 *Dracaena* chloroplast genome.

红边龙血树 JSB 边界上的 *ycf1* 基因由 IRb 区跨越边界进入 SSC 区 9 个碱基, *ndhF* 基因由 SSC 区跨越边界进入 IRb 区 29 个碱基(图 5), 这与索科特拉龙血树 JSB 边界类型相似, 但是

JLB 边界上的 *rpl22* 基因仅在索科特拉龙血树中发生了位移。红边龙血树的 *ndhF* 基因也跨越 JSB 边界, 在 IRb 与 SSC 区长度分别为 29 bp 和 2 185 bp, 与龙血树(GenBank 登录号 MF943127)

和细枝龙血树的跨越结果完全一样; *ycf1* 基因跨越 JLA 边界在 SSC 与 IRa 区长度分别为 4 582 bp 和 890 bp, 与龙血树(GenBank 登录号 MF943127)、细枝龙血树和金心香龙血树的跨越结果完全一样。*ycf1* 和 *psbA* 基因在海南龙血树(GenBank 登录号 MH293451)和山樱花龙血树 JLA 与 JSA 边界上有着相似的位移情况, 体现出 IRa 区域向 SSC 区扩张但在 LSC 区方向上收缩, 红边龙血树的 JLA 与 JSA 边界与除这二者外其他龙血树 JLA 与 JSA 边界相似。

2.6 龙血树属植物系统发育分析

基于 13 个龙血树属植物与 2 个万年青属

植物、3 个芦笋属植物的叶绿体基因组构建系统发育树(图 6)。结果显示 13 个龙血树属植物均能很好地分类, 聚为一个大的分支。红边龙血树与细枝龙血树聚为一类, 两者亲缘关系最近, 相似度达到 99%。龙血树(GenBank 登录号 MF943127)和金心香龙血树聚为一类, 又与红边龙血树聚为一支。索科特拉龙血树与山樱花龙血树独立于其他龙血树单独聚为一类, 河口龙血树、矮龙血树、长花龙血树聚为一支, 加那利龙血树与两个海南龙血树及剑叶龙血树(GenBank 登录号为 MN200195)聚为一支。

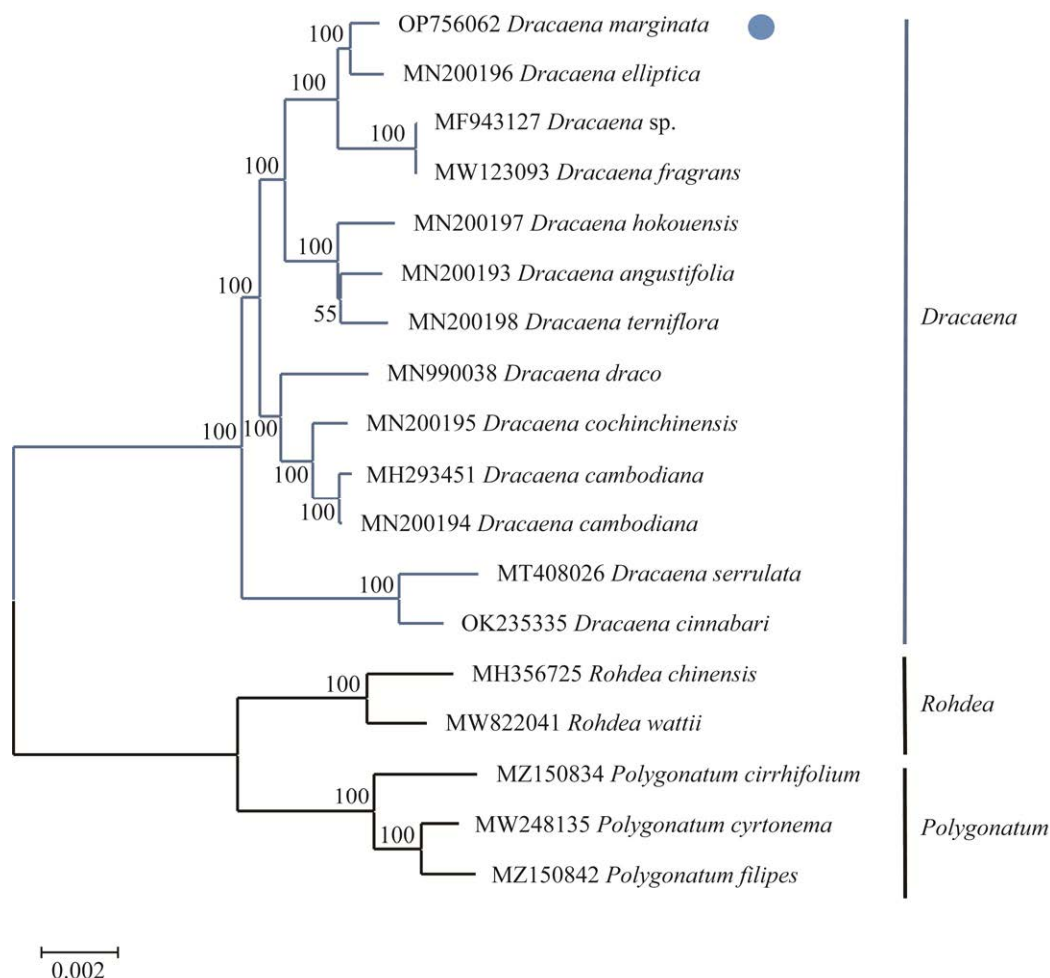


图 6 基于叶绿体全基因组构建的最大似然(ML)系统发育树

Figure 6 Maximum likelihood tree constructed by complete chloroplast genomes.

本研究构建的叶绿体基因组系统发育树拓扑结构与 Ahmad 等的结果一致^[14],通过对龙血树形态比对,剑叶龙血树与红边、细枝、金心香龙血树在形态上具有较大差异,而与海南龙血树形态接近,历来分子研究也支持剑叶龙血树与海南龙血树有较近的亲缘关系,剑叶龙血树与海南龙血树可能是进化过程中的姐妹种。值得注意的是,其中两个被命名为剑叶龙血树的叶绿体序列(GenBank 登录号分别为 MF943127 和 MN200195)并未聚在一起,显示出较远亲缘关系,综合分析认为 GenBank 登录号为 MF943127 的物种不应该是剑叶龙血树,上传的物种名鉴定错误,造成了分类学的混乱,建议予以更正。

3 讨论与结论

叶绿体基因组为母系遗传,保守性较高,尤其在龙血树属物种间,叶绿体结构以及基因组排列都高度保守。本研究首次组装获得了红边龙血树叶绿体基因组,全长 154 926 bp,深入分析了其叶绿体基因组结构特征,注释得到 132 个功能基因,与龙血树属其他叶绿体基因组的基因数量和种类相近,表明龙血树属植物的叶绿体基因组总体比较保守。

红边龙血树叶绿体密码子偏好性较低,与其他龙血树叶绿体使用密码子的频率有较高一致性。在 33 个 RSCU 值大于 1.00 的密码子中,有 27 个第 3 碱基为 A 或 U,表现出密码子以 A 或 U 结尾的偏好性。这种情况不仅出现在其他龙血树属物种中,在百合属^[24]和丝兰属^[25]等其他单子叶物种,山核桃属^[26]、山茶属^[27]和杨属^[28]等双子叶物种中都有所体现,原因可能是由于叶绿体基因组高 AT 含量所致。

检测 SSR 位点时设置较高长度下限(单核苷酸到六核苷酸的最小重复值分别设置为 10、

6、5、5、5、5),虽然有较低检出率,但可有效减少无效检出^[27]。在红边龙血树叶绿体基因组中共检测到 52 个 SSR 位点、108 个长重复序列位点(重复两次),其中 SSR 位点大多位于 LSC 区,而长重复序列位点则大多位于 IR 区。虽然检测 SSR 位点时参数设置不同,但本文龙血树属叶绿体基因组中,例如剑叶龙血树(GenBank 登录号 MN200195)、加那利龙血树和索科特拉龙血树等,SSR 位点类型与分布的研究结果与其他研究团队所报道结果相近^[12-14],而红边龙血树与龙血树(GenBank 登录号 MF943127)、细枝龙血树、金心香龙血树有着相似的重复序列类型及分布情况。进一步对 IR 边界进行研究,同样发现红边龙血树的 IR 边界基因跨越信息与龙血树(GenBank 登录号 MF943127)、细枝龙血树、金心香龙血树的跨越结果基本一致,体现了红边龙血树叶绿体结构与龙血树(GenBank 登录号 MF943127)、细枝龙血树、金心香龙血树有着高度相似性。

对龙血树属、万年青属和黄精属植物叶绿体基因组的聚类分析表明,红边龙血树与细枝龙血树聚为一支,关系最近,这与从重复序列及 IR 边界角度分析得出的二者近缘结果相一致。本研究使用叶绿体全基因组构建系统发育树的结果,支持山樱花龙血树与索科特拉龙血树单独聚为一支,加那利龙血树与海南龙血树、剑叶龙血树(GenBank 登录号 MN200195)聚为一支^[14],而与 Lu 等基于 *trnL-trnF*、*ndhF-rep132*、*trnQ-rps16* 等联合片段分析认为山樱花龙血树与加那利龙血树近缘的结果不相符^[29]。基于叶绿体全基因组建树可以充分利用保守的基因组中零散分布的多样性碱基,给予树状结构更高的支持度,比基于少量 DNA 片段构建的系统发育树有更高可信度。在龙血树属物种之间,仅通过少量基因进行鉴定存在一定误差。

综上所述,本研究获得了红边龙血树叶绿体基因组的序列、基因构成、基因组结构等基本特征,开展了多物种的比较叶绿体基因组分析,并利用叶绿体基因组序列构建了红边龙血树的系统发育关系,为未来红边龙血树的物种鉴定、进化地位、基因组研究、叶绿体转基因工程提供重要理论依据,同时也是对龙血树属植物细胞器的遗传多样性与综合开发利用研究的重要补充。

REFERENCES

- [1] 章玉平,陈丽云. 七彩千年木水培技术研究[J]. 广东农业科学, 2009, 36(10): 72-74.
ZHANG YP, CHEN LY. Study of hydroponic culture on *Dracaena marginata*[J]. Guangdong Agricultural Sciences, 2009, 36(10): 72-74 (in Chinese).
- [2] JUPA R, PLICHTA R, PASCHOVÁ Z, NADEZHDINA N, GEBAUER R. Mechanisms underlying the long-term survival of the monocot *Dracaena marginata* under drought conditions[J]. Tree Physiology, 2017, 37(9): 1182-1197.
- [3] HAMIDAH, PUTRI D, PURNOBASUKI H. Correlation of lead (Pb) content towards the number of stomata on the plant *Dracaena marginata* tricolor in some places of Surabaya City[J]. E3S Web of Conferences, 2020, 153: 02002.
- [4] VANLALRUATI, JAIN R, SINGH M, ANAND P, SINDHU SS, SANGWAN S. Effect of arbuscular mycorrhizal fungal consortia on growth, quality and nutrient uptake of *Dracaena* (*Dracaena marginata*)[J]. The Indian Journal of Agricultural Sciences, 2021, 91(1): 74-78.
- [5] ALJUHANI WS, ALJOHANI AY. Complete chloroplast genome of the medicinal plant *Cleome paradoxa* R.Br. Ex DC: comparative analysis, and phylogenetic relationships among the members of Cleomaceae[J]. Gene, 2022, 845: 146851.
- [6] JI YH, LIU CK, YANG ZY, YANG LF, HE ZS, WANG HC, YANG JB, YI TS. Testing and using complete plastomes and ribosomal DNA sequences as the next generation DNA barcodes in *Panax* (Araliaceae)[J]. Molecular Ecology Resources, 2019, 19(5): 1333-1345.
- [7] PANG XB, LIU HS, WU SR, YUAN YC, LI HJ, DONG JS, LIU ZH, AN CZ, SU ZH, LI B. Species identification of oaks (*Quercus* L., Fagaceae) from gene to genome[J]. International Journal of Molecular Sciences, 2019, 20(23): 5940.
- [8] 热伊汉古丽·图尔迪,慕丽红,田新民. 扁果草叶绿体基因组特征分析[J]. 生物工程学报, 2022, 38(8): 2999-3013.
TURDI R, MU LH, TIAN XM. Characteristics of the chloroplast genome of *Isopyrum anemonoides*[J]. Chinese Journal of Biotechnology, 2022, 38(8): 2999-3013 (in Chinese).
- [9] 苏涛,詹亚光,韩梅,郝爱平. 叶绿体基因工程: 一种植物生物技术的新方法[J]. 生物工程学报, 2005, 21(4): 674-680.
SU T, ZHAN YG, HAN M, HAO AP. Chloroplast genetic engineering: a new approach in plant biotechnology[J]. Chinese Journal of Biotechnology, 2005, 21(4): 674-680 (in Chinese).
- [10] KAUSHAL C, ABDIN MZ, KUMAR S. Chloroplast genome transformation of medicinal plant *Artemisia annua*[J]. Plant Biotechnology Journal, 2020, 18(11): 2155-2157.
- [11] 张剑锋,李阳,金静静,陈千思,谢小东,王中,罗朝鹏,杨军. 烟草叶绿体基因工程研究进展[J]. 烟草科技, 2017, 50(6): 88-98.
ZHANG JF, LI Y, JIN JJ, CHEN QS, XIE XD, WANG Z, LUO ZP, YANG J. Recent advances in tobacco chloroplast genetic engineering[J]. Tobacco Science & Technology, 2017, 50(6): 88-98 (in Chinese).
- [12] ZHANG ZL, ZHANG Y, SONG MF, GUAN YH, MA XJ. Species identification of *Dracaena* using the complete chloroplast genome as a super-barcode[J]. Frontiers in Pharmacology, 2019, 10: 1441.
- [13] CELIŃSKI K, KIJAK H, WILAND-SZYMAŃSKA J. Complete chloroplast genome sequence and phylogenetic inference of the canary islands dragon tree (*Dracaena draco* L.)[J]. Forests, 2020, 11(3): 309.
- [14] AHMAD W, ASAF S, KHAN A, AL-HARRASI A, AL-OKAISHI A, KHAN AL. Complete chloroplast genome sequencing and comparative analysis of threatened dragon trees *Dracaena serrulata* and *Dracaena cinnabari*[J]. Scientific Reports, 2022, 12: 16787.
- [15] JIN JJ, YU WB, YANG JB, SONG Y, DEPAMPHILIS CW, YI TS, LI DZ. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes[J]. Genome Biology, 2020, 21(1): 241.
- [16] SHI LC, CHEN HM, JIANG M, WANG LQ, WU X,

- HUANG LF, LIU C. CPGAVAS2, an integrated plastome sequence annotator and analyzer[J]. *Nucleic Acids Research*, 2019, 47(W1): W65-W73.
- [17] ZHENG SY, POCZAI P, HYVÖNEN J, TANG J, AMIRYOUSEFI A. Chloroplast: an online program for the versatile plotting of organelle genomes[J]. *Frontiers in Genetics*, 2020, 11: 576124.
- [18] KURTZ S, CHOUDHURI JV, OHLEBUSCH E, SCHLEIERMACHER C, STOYE J, GIEGERICH R. REPuter: the manifold applications of repeat analysis on a genomic scale[J]. *Nucleic Acids Research*, 2001, 29(22): 4633-4642.
- [19] BEIER S, THIEL T, MÜNCH T, SCHOLZ U, MASCHER M. MISA-web: a web server for microsatellite prediction[J]. *Bioinformatics*, 2017, 33(16): 2583-2585.
- [20] 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.
- [21] BI GQ, MAO YX, XING QK, CAO M. HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching[J]. *Genomics*, 2018, 110(1): 18-22.
- [22] NAKAMURA T, YAMADA KD, TOMII K, KATOH K. Parallelization of MAFFT for large-scale multiple sequence alignments[J]. *Bioinformatics*, 2018, 34(14): 2490-2492.
- [23] FU JM, LIU HM, HU JJ, LIANG YQ, LIANG JJ, WUYUN TN, TAN XF. Five complete chloroplast genome sequences from *Diospyros*: genome organization and comparative analysis[J]. *PLoS One*, 2016, 11(7): e0159566.
- [24] 毕戣. 百合属的比较叶绿体基因组学研究[D]. 长春: 吉林农业大学硕士学位论文, 2017.
- BI Y. Comparative chloroplast genomics of the genus *Lilium*[D]. Changchun: Master's Thesis of Jilin Agricultural University, 2017 (in Chinese).
- [25] 王飞, 辛雅莹, 董章宏, 赵文植, 李卫英, 马路遥, 夏茂甜, 辛培尧. 无刺龙舌兰叶绿体基因组特征及密码子偏好性分析[J]. *南方农业学报*, 2022, 53(4): 1030-1039.
- WANG F, XIN YX, DONG ZH, ZHAO WZ, LI WY, MA LY, XIA MT, XIN PY. Chloroplast genome characteristics and codon preference analysis of *Yucca treculeana*[J]. *Journal of Southern Agriculture*, 2022, 53(4): 1030-1039 (in Chinese).
- [26] SHEN JS, LI XQ, CHEN X, HUANG XL, JIN SH. The complete chloroplast genome of *Carya cathayensis* and phylogenetic analysis[J]. *Genes*, 2022, 13(2): 369.
- [27] 童一涵, 郑倩, 杜新明, 冯士令, 周莉君, 丁春邦, 陈涛. 多齿红山茶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.
- TONG YH, ZHENG Q, DU XM, FENG SL, ZHOU LJ, DING CB, CHEN T. Analysis on sequence characteristics of chloroplast genome of *Camellia polyodonta*[J]. *Journal of Plant Resources and Environment*, 2022, 31(5): 27-36 (in Chinese).
- [28] BU YF, WU XY, SUN N, MAN Y, JING YP. Codon usage bias predicts the functional *MYB10* gene in *Populus*[J]. *Journal of Plant Physiology*, 2021, 265: 153491.
- [29] LU PL, MORDEN CW. Phylogenetic relationships among dracaenoid genera (Asparagaceae: Nolinoideae) inferred from chloroplast DNA loci[J]. *Systematic Botany*, 2014, 39(1): 90-104.

(本文责编 郝丽芳)