

研究论文

Research Article

九翅豆蔻叶绿体基因组特征及其密码子偏好性分析

傅武祥^{2,3} 郭雨桐¹ 张雪梅² 马占霞^{1,3*}

1 滇西应用技术大学普洱茶学院, 普洱, 665000; 2 云南农业大学农学与生物技术学院, 昆明, 650201; 3 中国科学院西双版纳热带植物园, 勐仑, 666303

* 通信作者, mazhanxia@xtbg.ac.cn

摘要 为探究九翅豆蔻 (*Amomum maximum*) 叶绿体全基因组 (chloroplast DNA, cpDNA) 特征和密码子使用偏好性, 本研究利用 Illumina HiSeq 测序平台对九翅豆蔻叶绿体全基因组进行测序, 通过组装与注释得到叶绿体基因组全序列, 使用 MISA、IQTREE 和 CodonW 等软件对九翅豆蔻叶绿体全基因组的微卫星位点 (microsatellite site)、密码子偏好性及系统发育进行分析。结果表明, 九翅豆蔻完整叶绿体全基因组长度为 163 602 bp, 基因组平均 G+C 含量为 36.1%, 其大单拷贝 (large single copy, LSC) 区和小单拷贝 (small single copy, SSC) 区长度分别为 88 626 bp 和 15 426 bp, 两个反向重复 (inverted repeat, IR) 区长度为 29 775 bp。九翅豆蔻叶绿体基因组共编码 132 个基因, 包含 86 个蛋白编码基因、38 个 tRNA 基因和 8 个 rRNA 基因。叶绿体基因组序列上共检测到 132 个简单重复序列 (simple sequence repeat, SSR) 位点。其中单核苷酸、二核苷酸、三核苷酸、四核苷酸和五核苷酸数量分别为 65、35、5、24、3 个, 绝大多数属于单碱基重复的 A/U 类型。系统发育分析显示九翅豆蔻在豆蔻属 (*Amomum*) 植物中单独聚为一支。这为九翅豆蔻种质资源鉴定、评价及物种形成机制研究奠定了基础。

关键词 九翅豆蔻; 叶绿体基因组; 微卫星位点; 密码子偏好性; 系统发育分析

Analysis of Chloroplast Genome Characteristics and Its Code Usage Bias of *Amomum maximum*

FU Wuxiang^{2,3} GUO Yutong¹ ZHANG Xuemei² MA Zhanxia^{1,3*}

1 West Yunnan University of Applied Sciences College of Tea (Pu'er), Pu'er, 665000; 2 College of Agronomy and Biotechnology, Yunnan Agricultural University, Kunming, 650201; 3 Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, 666303

* Corresponding author, mazhanxia@xtbg.ac.cn

DOI: 10.13417/j.gab.043.000017

Abstract To explore the characteristics and codon usage bias of the whole chloroplast genome (chloroplast DNA, cpDNA) of *Amomum maximum*, this study used the Illumina HiSeq sequencing platform to sequence the whole chloroplast genome of *A. maximum*, and obtained the complete chloroplast genome sequence through assembly and annotation. The microsatellite sites, codon bias and phylogeny of the chloroplast genome of *A. maximum* were analyzed by software such as MISA, IQTREE and CodonW. The results showed that the complete chloroplast genome of *A. maximum* was 163 602 bp in length, with an average G+C content of 36.1%. The lengths of the large single copy (LSC) region and the small single copy (SSC) region were 88 626 bp and 15 426 bp, respectively, and the two inverted repeat (IR) regions were 29 775 bp. The chloroplast genome of *A. maximum* encoded 132 genes, including 86 protein-coding genes, 38 tRNA genes and 8 rRNA genes. A total of 132 simple sequence repeat (SSR) sites were detected on the chloroplast genome sequence. Among them, the numbers of mononucleotide, dinucleotide, trinucleotide, tetranucleotide and pentanucleotide were 65, 35, 5, 24 and 3, respectively, most of which belonged to the A/U type of single base repeat. Phylogenetic analysis showed that

基金项目: 本研究由滇西应用技术大学教学质量工程项目 (21JK16) 及云南省教育厅科学研究基金项目 (2022J1045) 共同资助。
引用格式: 傅武祥, 郭雨桐, 张雪梅, 等, 2024. 九翅豆蔻叶绿体基因组特征及其密码子偏好性分析. 基因组学与应用生物学, 43(1): 17-30. [FU W X, GUO Y T, ZHANG X M, et al., 2024. Analysis of chloroplast genome characteristics and its code usage bias of *Amomum maximum*. Genomics and Applied Biology, 43(1): 17-30.]

通信作者简介: 马占霞, 助理研究员, 药用植物功能基因研究

收稿日期: 2022-11-19; 接受日期: 2023-03-05

A. maximum was clustered as a separate branch in the genus *Amomum*. This laid a foundation for the research of identification, evaluation and species formation mechanism of *A. maximum* germplasm resources.

Keywords *Amomum maximum*; Chloroplast genome; Microsatellite sites; Code usage bias; Phylogenetic analysis

九翅豆蔻 (*Amomum maximum*) 属姜科 (Zingiberaceae) 豆蔻属 (*Amomum*) 多年生常绿丛生草本植物, 生于海拔 350~800 m 林中阴湿处, 广泛分布于我国云南南部和东南亚热带地区 (中国科学院昆明植物研究所, 1997)。作为传统药物, 九翅豆蔻的根状茎和果常被用于治疗消化系统疾病 (Chetia et al., 2014; Luo et al., 2014; Guo et al., 2015), 且被证明具有潜在的抗肿瘤、抗炎和祛虫活性 (Ji et al., 2019; Lu et al., 2022)。其花、果和嫩茎心等还可作为野生蔬菜食用 (Huong et al., 2015)。在傣医药资料中有相关记载, 九翅豆蔻被称作贺故 (音)、郭姑 (音) 等, 主治腹部胀痛、消化不良和肢体关节酸痛等 (林艳芳等, 2003; 卢传礼, 2017)。

姜科植物种间甚至属间大多外形相似, 其起源、进化和系统分类一直存在争议。不同研究人员从形态特征、理化特性、组织解剖和分子标记等多角度对姜科植物进行鉴定, 对姜科植物分类具有重要意义 (Huong et al., 2015)。尽管姜科植物形态特征有一定区别, 但对于根状茎或果实药材鉴定而言, 传统鉴定方法有其局限性。传统中药和民族药基源植物的鉴定主要利用性状鉴定, 如外观形状、色、气、味、质地等特征。虽然该方法可直接鉴定, 但是主观性强, 其准确性取决于鉴定者的经验, 且对加工炮制后的碎片或粉末难以鉴定 (廖彩丽等, 2013)。显微鉴定、理化鉴定和指纹图谱是较为客观的鉴定方法, 但因物种、环境或人为主观性等因素, 鉴定的重复性和稳定性较差 (Xu et al., 2017)。九翅豆蔻栽培条件要求较高, 自然结实率低, 造成市场上出现混淆品, 包括同属的草果 (*A. tsaoko*)、云南豆蔻 (*A. repoeense*) 等, 混淆品之间的药用成分和功效均存在一定差异, 影响临床用药安全及其资源的可持续利用 (林艳芳, 2003)。为了保证九翅豆蔻栽培选种正确, 临床用药的安全有效, 其准确鉴别的问题亟待解决。

在植物药材的分子鉴定方面, DNA 条形码是非常快速和准确的鉴定方法 (袁庆军等, 2012; 朱斌等, 2021), 在建立了正品物种的参考 DNA 序列数据库后, 利用保守的叶绿体基因 *matK* 和 *rbcL* 可以将一些关系较远的贗品快速地区别开来, 而一些关

系近缘、形态相近的混淆品则可依靠核基因 *ITS* (*ITS2*) 和 *psbA-trnH* 进行快速、准确的鉴别 (China Plant BOL Group et al., 2011; 宁淑萍等, 2018)。目前植物的 DNA 条形码的筛选工作主要集中在叶绿体基因组上, 因为叶绿体基因组包含大量的遗传信息, 为研究提供了一个较大的数据基础 (Zeng et al., 2018); 叶绿体 DNA 的核苷酸置换率适中, 在应用上很有价值; 叶绿体基因组的编码区和非编码区的分子进化速度差异显著, 适用于不同阶层的系统学研究 (Razi and Muneer, 2021); 叶绿体基因组大小适中, 便于测序; 各植物类群叶绿体基因组之间具有良好的共线性, 便于比较分析。因此, 叶绿体基因组在植物分类与物种水平的系统发育等研究中具有明显的优势, 并且发挥着重要的作用 (Xu et al., 2001; 钟志敏等, 2018)。

本研究基于高通量测序技术和生物信息学分析方法, 通过组装、注释对九翅豆蔻叶绿体基因组的组成、结构和系统发育等进行分析, 解析九翅豆蔻的叶绿体基因组结构, 以期为九翅豆蔻资源鉴定、合理保护和科学开发利用提供理论依据。

1 结果

1.1 九翅豆蔻叶绿体基因组基本特征

九翅豆蔻叶绿体全基因组与豆蔻属其他植物叶绿体基因组结构相似, 为共价闭合的双链环状分子, 全长 163 602 bp, 是典型的四分体结构, 其中包含一个大单拷贝区长度为 88 626 bp, 一个小单拷贝区长度为 15 426 bp, 两个反向重复区长度为 29 775 bp。全基因组的 G+C 含量为 36.1%, 其中 IR 区的 G+C 含量最高 (41.2%), LSC 区 (33.8%) 和 SSC 区 (29.8%) 的均较低 (图 1)。整个叶绿体基因组共包括 132 个基因 (其中重复基因有: *ndhB*、*rpl2*、*rpl23*、*rps12*、*rps7*、*rrn16*、*rrn23*、*rrn4.5*、*rrn5*、*trnA-UGC*、*trnH-GUG*、*trnI-CAU*、*trnI-GAU*、*trnL-CAA*、*trnN-GUU*、*trnR-ACG*、*trnV-GAC*、*yef2*), 包括 38 个 tRNA 基因、8 个 rRNA 基因和 86 个蛋白质编码基因。*rps12* 基因具有反式剪切结构, 其 5'端位于 LSC 区, 而 3'端位于 IR 区 (图 1)。



图 1 九翅豆蔻叶绿体基因组图谱

Figure 1 Chloroplast genome map of *Amomum maximum*

在蛋白质编码基因中，与光合作用相关的基因有 45 个，包括 5 个光合系统 I 基因、15 个光合系统 II 基因、6 个细胞色素 b/f 复合体编码基因、6 个 ATP 合酶亚基基因、12 个 NADH 脱氢酶基因、1 个二磷酸核酮糖羧化酶大亚基基因；自我复制相关基因除 rRNA 基因和 tRNA 基因外，还包括 15 个核糖

体小亚基基因、11 个核糖体大亚基基因和 4 个 DNA 依赖性 RNA 聚合酶基因；此外还有 6 个其他功能基因和 5 个功能未知基因。九翅豆蔻叶绿体基因组中含有内含子的基因共 19 个，其中包括 5 个 tRNA 基因和 13 个蛋白质编码基因，其中 *ndhB*、*rpl2*、*rps12*、*trnA-UGC* 和 *trnI-GAU* 中含有 2 个内含子（表 1）。

表 1 注释的九翅豆蔻叶绿体基因组基因列表

Table 1 Annotated gene list of chloroplast genome of *Amomum maximum*

基因类别 Category of gene	基因分组 Group of gene	基因名称 Name of gene	数量 Amount
自我复制相关基因 Self-replication related gene	核糖体 RNA rRNA	<i>rrn23</i> [*] 、 <i>rrn5</i> [*] 、 <i>rrn16</i> [*] 、 <i>rrn4</i> 、 <i>5</i> [*]	8
	转运 RNA tRNA	<i>trnH-GUG</i> [*] <i>trnI-CAU</i> [*] <i>trnI-GAU</i> [*] <i>trnL-CAA</i> [*] <i>trnL-UAG</i> <i>trnL-UAA</i> <i>trnV-GAC</i> <i>trnV-UAC</i> <i>trnA-UGC</i> [*] <i>trnR-ACG</i> [*] <i>trnR-UCU</i> <i>trnN-GUU</i> [*] <i>trnP-UGG</i> <i>trnW-CCA</i> <i>trnM-CAU</i> <i>trnF-GAA</i> <i>trnT-UGU</i> <i>trnT-GGU</i> <i>trnS-GGA</i> <i>trnS-UGA</i> <i>trnS-CGA</i> <i>trnS-GCU</i> <i>trnJ-M-CAU</i> <i>trnG-UCC</i> <i>trnE-UUC</i> <i>trnY-GUA</i> <i>trnD-GUC</i> <i>trnC-GCA</i> <i>trnQ-UUG</i> <i>trnK-UUU</i>	38
	核糖体小亚基 Small subunit of ribosome	<i>rps2</i> <i>rps3</i> <i>rps7</i> [*] <i>rps4</i> <i>rps8</i> <i>rps11</i> <i>rps12</i> [*] <i>rps14</i> <i>rps15</i> <i>rps16</i> <i>rps18</i> <i>rps19</i> [*]	15
	核糖体大亚基 Large subunit of ribosome	<i>rpl14</i> <i>rpl16</i> <i>rpl20</i> <i>rpl22</i> <i>rpl23</i> [*] <i>rpl32</i> <i>rpl33</i> <i>rpl36</i> <i>rpl2</i> [*]	11
	DNA 依赖性 RNA 聚合酶 DNA dependent RNA polymerase	<i>rpoA</i> <i>rpoB</i> <i>rpoC1</i> <i>rpoC2</i>	4
光合作用相关基因 Photosynthesis related gene	光系统 I 亚基 Photosystem I	<i>psaA</i> <i>psaB</i> <i>psaC</i> <i>psaI</i> <i>psaJ</i>	5
	光系统 II 亚基 Photosystem II	<i>psbA</i> <i>psbB</i> <i>psbC</i> <i>psbD</i> <i>psbE</i> <i>psbF</i> <i>psbH</i> <i>psbI</i> <i>psbJ</i> <i>psbK</i> <i>psbL</i> <i>psbM</i> <i>psbN</i> <i>psbT</i> <i>psbZ</i>	15
	细胞色素 b/f 复合体 Cytochrome b/f complex	<i>petA</i> <i>petB</i> <i>petD</i> <i>petG</i> <i>petL</i> <i>petN</i>	6
	ATP 合酶亚基 Subunits of ATP synthase	<i>atpA</i> <i>atpB</i> <i>atpE</i> <i>atpF</i> <i>atpH</i> <i>atpI</i>	6
	二磷酸核酮糖羧化酶大亚基 Large subunit of Rubisco	<i>rbcL</i>	1
	NADH 脱氢酶亚基 Subunits of NADH dehydrogenase	<i>ndhA</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>	12
其他功能基因 Other functional gene	成熟酶 Maturase	<i>matK</i>	1
	包膜蛋白 Envelope protein	<i>cemA</i>	1
	乙酰辅酶 A 羧化酶亚基 Subunit of acetyl-CoA-carboxylase	<i>accD</i>	1
	c-型细胞色素合成基因 c-type cytochrome synthesis	<i>ccsA</i>	1
	ATP-依赖的蛋白酶亚基 P ATP-dependent protease subunit P	<i>clpP</i>	1
	翻译起始因子 Translation initiation factor	<i>infA</i>	1
功能未知基因 Functionally unknown gene	开放阅读框 Open reading frame	<i>ycf1</i> <i>ycf2</i> [*] <i>ycf3</i> <i>ycf4</i>	5

注：* 指两次拷贝。

Note: * refers to two copies.

1.2 九翅豆蔻叶绿体全基因组密码子偏好性分析

九翅豆蔻叶绿体全基因组有效密码子数 (effective number of codon, N_c) 为 54.4, 偏好性较弱, 密码子的第三位上的 G+C 碱基含量 GC3s 为 33.7%, G+C 含量为 37.1%, 均小于 50%, 说明九翅豆蔻叶绿体基因组的密码子使用更倾向于 A 和 T 碱基。将九翅豆蔻叶绿体基因组四大区域的 A+T 含量进行统计, LSC 区、SSC 区、IR 区的 A+T 含量分别为

66.2%、70.2%、58.8%, 其中 IR 区的 A+T 含量最低是由于最为保守的 rRNA 基因全部分布在此区间。

九翅豆蔻叶绿体基因组相对同义密码子使用度 (relative synonymous codon usage, RSCU) 分析如表 2 所示, 在 64 种密码子中, RSCU 值大于 1 的密码子有 32 个, 其中以 A 或 U 碱基结尾的密码子有 29 个, 而以 C 或 G 碱基结尾的只有 3 个, 表明九翅豆蔻叶绿体基因组的密码子偏向以 A 或 U 碱基结尾。

表 2 九翅豆蔻叶绿体基因组密码子 RSCU 分析

Table 2 RSCU analysis of codon in chloroplast genome of *Amomum maximum*

氨基酸或终止密码子 Amino acids/ending codon	密码子及 RSCU 值 Codon and RSCU values	
色氨酸 Trp	UGG	1.00
甲硫氨酸 Met	AUG	1.00
苯丙氨酸 Phe	UUU	1.28 UUC 0.72
酪氨酸 Tyr	UAU	1.37 UAC 0.63
半胱氨酸 Cys	UGU	1.26 UGC 0.74
组氨酸 His	CAU	1.44 CAC 0.56
谷氨酰胺 Gln	CAA	1.38 CAG 0.62
天冬酰胺 Asn	AAU	1.41 AAC 0.59
赖氨酸 Lys	AAA	1.40 AAG 0.60
天冬氨酸 Asp	GAU	1.45 GAC 0.55
谷氨酸 Glu	GAA	1.41 GAG 0.59
异亮氨酸 Ile	AUU	1.19 AUC 0.68 AUA 1.12
终止子 TER	UAA	1.23 UAG 0.76 UGA 1.01
脯氨酸 Pro	CCU	1.17 CCC 0.94 CCA 1.26 CCG 0.63
苏氨酸 Thr	ACU	1.22 ACC 0.95 ACA 1.25 ACG 0.85
缬氨酸 Val	GUU	1.41 GUC 0.66 GUA 1.29 GUG 0.64
丙氨酸 Ala	GCU	1.37 GCC 0.86 GCA 1.21 GCG 0.56
甘氨酸 Gly	GGU	1.06 GGC 0.54 GGA 1.48 GGG 0.92
丝氨酸 Ser	UCU	1.49 UCC 1.09 UCA 1.28 UCG 0.71 AGU 0.86 AGC 0.75
亮氨酸 Leu	UUA	1.54 UUG 1.15 CUU 1.21 CUC 0.66 CUA 0.93 CUG 0.51
精氨酸 Arg	CGU	0.71 CGC 0.40 CGA 1.07 CGG 0.71 AGA 1.99 AGG 1.12

注: 加粗的数据表示 RSCU 大于 1。

Note: The bold data indicates that the RSCU is greater than 1.

1.3 九翅豆蔻叶绿体全基因组的 SSR 分析

通过 MISA 工具检测, 九翅豆蔻叶绿体基因组中共有 132 个简单重复序列 (simple sequence repeat, SSR) 位点 (表 3, 表 4), 表中 p1、p2、p3、p4、p5、p6 分别表示单碱基重复、二碱基重复、三碱基重复、四碱基重复、五碱基重复、六碱基重复, c 表示组合 SSR 位点, 即两个 SSR 之间相距不超过 100 bp, 其

中单碱基重复 65 个, 二碱基重复 35 个, 三碱基重复 5 个, 四碱基重复 24 个, 五碱基重复 3 个, 组合重复 15 个, 未检测到六碱基重复位点。SSR 几乎都集中在 LSC 区域, 这可能与 LSC 区域较长有关。九翅豆蔻叶绿体基因组单碱基重复最多且多为 A/T 重复, 表明九翅豆蔻 SSR 更偏好于使用 A 或 T 碱基。

表 3 九翅豆蔻叶绿体基因组中的 SSR 类型

Table 3 Type of SSR in the *Amomum maximum* chloroplast genome

序号 Number	SSR 类型 SSR type	碱基组成 Base composition	大小/bp Size/bp	起点 Origin point	终点 Terminal point	所在区域 Region
1	p2	(AT)5	10	2 661	2 670	LSC
2	p1	(T)10	10	4 393	4 402	LSC
3	c	(TA)4(<TA>TT)(TATT)2	20	4 878	4 897	LSC
4	p1	(A)10	10	5 242	5 251	LSC
5	p1	(T)11	11	5 635	5 645	LSC
6	p1	(A)10	10	6 199	6 208	LSC
7	c	(AT)6ttatt...atatgga(T)11	120	6 583	6 702	LSC
8	p1	(T)10	10	7 022	7 031	LSC
9	p1	(A)10	10	7 235	7 244	LSC
10	c	(C)10(A)10	20	7 488	7 507	LSC
11	p1	(T)10	10	8 677	8 686	LSC
12	p1	(T)10	10	10 453	10 462	LSC
13	p4	(TTAT)3	12	12 800	12 811	LSC
14	c	(T)10ccac(TTTA)3	26	13 730	13 755	LSC
15	p4	(AAAG)3	12	13 961	13 972	LSC
16	p1	(A)10	10	14 722	14 731	LSC
17	c	(ATATA)3aga...aggata(AT)6	125	14 879	15 003	LSC
18	c	(AATA)3atactat(AATA)3atactat(AATA)3	50	15 193	15 242	LSC
19	p1	(T)10	10	17 406	17 415	LSC
20	c	(T)10ctgtg...aata(T)12	117	19 527	19 643	LSC
21	p1	(A)10	10	19 775	19 784	LSC
22	p2	(AT)5	10	21 003	21 012	LSC
23	p1	(T)10	10	24 065	24 074	LSC
24	p2	(TA)8	16	28 527	28 542	LSC
25	p1	(A)11	11	28 929	28 939	LSC
26	c	(T)10c(A)11	22	29 442	29 463	LSC
27	p1	(A)10	10	29 811	29 820	LSC
28	p1	(T)10	10	30 781	30 790	LSC
29	p1	(T)10	10	32 883	32 892	LSC
30	c	(ATTT)3t. t(TA)5a(T)10a..a(TTAT)3 (ATTT)3tt. t(TA)5.a(T)10aa(TTAT)3	91	33 641	33 731	LSC
31	p3	(TCT)4	12	35 000	35 011	LSC
32	p2	(TA)6	12	38 532	38 543	LSC
33	p2	(AT)5	10	38 696	38 705	LSC
34	p1	(A)10	10	39 008	39 017	LSC
35	c	(TAG)4tccat...cttctg(C)11(A)10	62	39 373	39 434	LSC
36	p4	(ATTT)3	12	48 212	48 223	LSC

续表
Continuing table

序号 Number	SSR 类型 SSR type	碱基组成 Base composition	大小/bp Size/bp	起点 Origin point	终点 Terminal point	所在区域 Region
37	p1	(T)10	10	49 210	49 219	LSC
38	c	(TA) 5tc... ctat (TA) 6attt (TA) 5tcc... ctat (TA) 5t(TA) 6ataa... atgat (A) 10	123	49 751	49 873	LSC
39	p2	(TA)7	14	50 055	50 068	LSC
40	p2	(AT)8	16	50 337	50 352	LSC
41	p4	(CAAA)3	12	50 457	50 468	LSC
42	p1	(T)10	10	53 401	53 410	LSC
43	p2	(AT)6	12	54 179	54 190	LSC
44	p1	(A)11	11	54 480	54 490	LSC
45	p4	(ATTT)3	12	56 227	56 238	LSC
46	p1	(T)11	11	58 703	58 713	LSC
47	p1	(T)11	11	58 874	58 884	LSC
48	p1	(T)13	13	59 238	59 250	LSC
49	p1	(A)10	10	60 889	60 898	LSC
50	p5	(TATTA)3	15	61 247	61 261	LSC
51	c	(AT)5tgt... atatttttg(T)11	87	63 239	63 325	LSC
52	p2	(AT)6	12	63 436	63 447	LSC
53	p1	(A)10	10	65 183	65 192	LSC
54	p2	(AT)5	10	67 143	67 152	LSC
55	p4	(AGAA)3	12	67 498	67 509	LSC
56	p2	(TA)5	10	69 368	69 377	LSC
57	p4	(GAAA)3	12	69 487	69 498	LSC
58	p1	(T)12	12	69 729	69 740	LSC
59	p4	(CATA)3	12	70 676	70 687	LSC
60	p1	(T)11	11	70 842	70 852	LSC
61	p3	(TTC)4	12	71 892	71 903	LSC
62	c	(A)10tagaca... tct(AGA)4	104	72 676	72 779	LSC
63	p1	(C)11	11	73 385	73 395	LSC
64	p2	(TA)7	14	74 169	74 182	LSC
65	p1	(T)11	11	75 178	75 188	LSC
66	p4	(ATAA)3	12	75 375	75 386	LSC
67	p1	(A)11	11	75 713	75 723	LSC
68	p1	(T)10	10	75 921	75 930	LSC
69	p1	(T)10	10	76 863	76 872	LSC
70	p1	(A)11	11	79 274	79 284	LSC
71	p2	(TA)5	10	81 699	81 708	LSC
72	p1	(A)10	10	84 165	84 174	LSC

续表
Continuing table

序号 Number	SSR 类型 SSR type	碱基组成 Base composition	大小/bp Size/bp	起点 Origin point	终点 Terminal point	所在区域 Region
73	p3	(GGA)4	12	84 746	84 757	LSC
74	p4	(TTCT)3	12	86 987	86 998	LSC
75	p5	(TATAT)3	15	88 154	88 168	LSC
76	p1	(T)13	13	89 070	89 082	IR
77	p1	(T)11	11	89 297	89 307	IR
78	p2	(GA)5	10	94 650	94 659	IR
79	p1	(A)10	10	95 921	95 930	IR
80	p2	(TA)5	10	99 595	99 604	IR
81	p1	(T)11	11	104 195	104 205	IR
82	p1	(T)10	10	116 648	116 657	IR
83	p4	(ATTT)3	12	117 419	117 430	IR
84	p4	(GAAT)3	12	118 170	118 181	IR
85	p2	(AT)5	10	118 408	118 417	LSC
86	p1	(A)12	12	118 871	118 882	LSC
87	c	(A)12taa...t(A)12	67	119 004	119 070	LSC
88	p2	(AT)5	10	120 363	120 372	LSC
89	p1	(A)10	10	122 945	122 954	LSC
90	p4	(TTTA)3	12	123 092	123 103	LSC
91	p1	(A)14	14	125 385	125 398	LSC
92	c	(TA)5at...ag(TA)5a...ata(AT)5	132	125 938	126 069	LSC
93	p4	(TATT)3	12	128 206	128 217	LSC
94	p2	(TA)8	16	128 651	128 666	LSC
95	p2	(TA)5	10	130 435	130 444	LSC
96	p1	(A)12	12	130 553	130 564	LSC
97	p1	(T)10	10	121 253	121 262	LSC
98	p4	(AATT)3	12	133 439	133 450	IR
99	p4	(CATT)3	12	134 047	134 058	IR
100	p4	(AAAT)3	12	134 799	134 810	IR
101	p1	(A)10	10	135 572	135 581	IR
102	p1	(A)11	11	148 024	148 034	IR
103	p2	(AT)5	10	152 624	152 633	IR
104	p1	(T)10	10	156 299	156 308	IR
105	p2	(TC)5	10	157 570	157 579	IR
106	p1	(A)11	11	162 922	162 932	IR
107	p1	(A)13	13	163 147	163 159	IR

表 4 九翅豆蔻叶绿体基因 SSR 数量
Table 4 SSR number of chloroplast gene in *Amomum maximum*

核苷酸类型 Nucleotide type	重复序列 Repeat sequence	重复次数 Repeat times (number)					重复序列数量 Number of repeat sequence
单核苷酸 Single nucleotide	A/T	10(36)	11(16)	12(6)	13(3)	14(1)	62
	C/G	10(1)	11(2)				3
二核苷酸 Dinucleotide	AG/CT	5(2)					2
	AT/AT	5(20)	6(7)	7(2)	8(3)	9(1)	33
三核苷酸 Trinucleotide	AAG/CTT	4(3)					3
	ACT/AGT	4(1)					1
	AGG/CCT	4(1)					1
四核苷酸 Tetranucleotide	AAAC/GTTT	3(1)					1
	AAAG/CTTT	3(4)					4
	AAAT/ATTT	3(14)					15
	AATG/ATTC	3(2)					2
	AATT/AATT	3(1)					1
	ACAT/ATGT	3(1)					1
五核苷酸 Pentanucleotide	AATAT/ATATT	3(3)					3
总计 Total							132

注: 表格中括号内的数量, 表示该类型的重复序列在不同重复次数下统计得到的数量。

Note: The number in parentheses in the table represents the number of repeat sequences of this type calculated at different repetition times.

1.4 系统进化分析

基于九翅豆蔻等 14 种植物建立系统发育树, 包括姜科豆蔻属的海南砂仁 (*Amomum longiligulare*)、白豆蔻 (*Amomum krarvanh*) 和草果; 姜科姜属 (*Zingiber*) 的姜 (*Zingiber officinale*) 和蜂巢姜 (*Z. spectabile*); 姜科山柰属 (*Kaempferia*) 的山柰 (*Kaempferia galanga*) 和紫花山柰 (*Kaempferia elegans*); 姜科姜黄属 (*Curcuma*) 的温郁金 (*Curcuma*

wenyujin)、姜黄 (*Curcuma longa*) 和观音姜 (*Curcuma roscoeana*), 以及两个外类群: 闭鞘姜科 (*Costaceae*) 的闭鞘姜 (*Hellenia speciosa*) 和美叶闭鞘姜 (*Costus pulverulentus*), 望鹤兰科 (*Strelitziaceae*) 的旅人蕉 (*Ravenala madagascariensis*)。建立的系统发育树 (图 2) 表明, 九翅豆蔻与豆蔻属植物单独聚为一支, 说明其亲缘关系较近, 而姜科其他植物则聚为另一支, 表明它们具有最近共同祖先种。

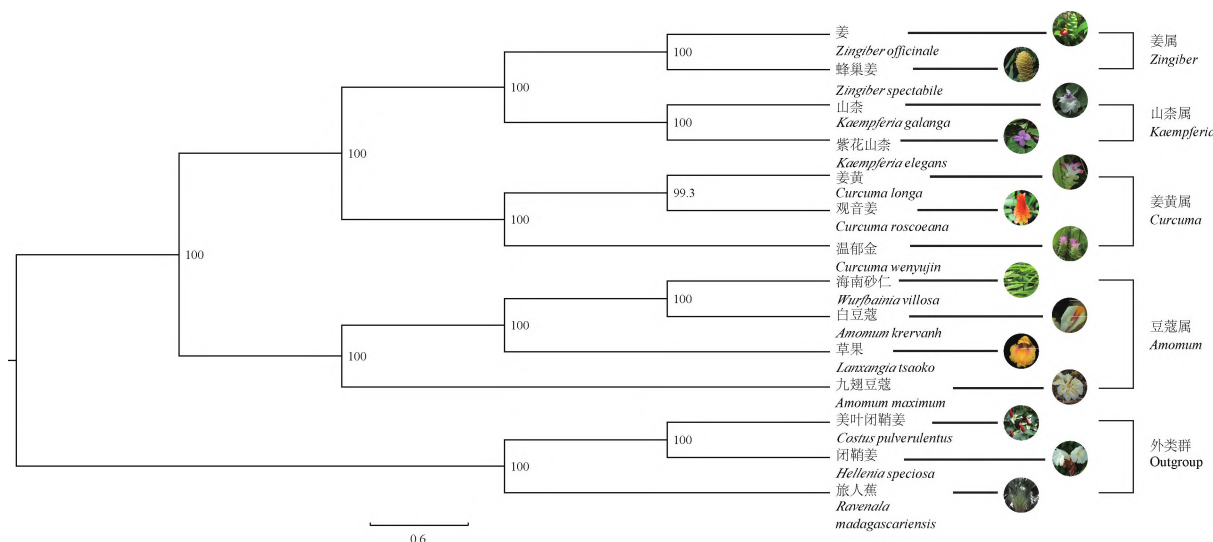


图2 基于14种植物叶绿体基因组构建的系统发育树

Figure 2 Phylogenetic tree based on chloroplast genomes of 14 plant species

2 讨论与结论

被子植物的叶绿体拥有相对独立的基因组，其基因组信息在被子植物中高度保守，同时叶绿体基因组拷贝数多，但其基因组分子量小，可通过一定技术快速且低成本地获得其全基因组信息，现在叶绿体基因组已成为探讨植物系统发育的重要依据 (Li et al., 2020)。本文通过高通量测序对九翅豆蔻叶绿体基因组进行研究分析，九翅豆蔻叶绿体全基因组长 163 602 bp，符合目前已发表的豆蔻属植物叶绿体基因组长度 162 766 ~ 164 069 bp (石林春等, 2010; 马孟莉等, 2021)，表明豆蔻属植物在叶绿体基因上有高保守性。已报道的豆蔻属叶绿体总基因数量在 133 ~ 135 个 (非重复蛋白编码基因在 79 ~ 80 个)，九翅豆蔻叶绿体基因组注释到 132 个基因 (38 个 tRNA 基因、8 个 rRNA 基因、86 个蛋白质编码基因)。豆蔻属物种间叶绿体基因组 G+C 含量相似，均在 36.0% ~ 36.1%，并且 LSC 区 (33.7% ~ 33.9%) 和 SSC 区 (29.8% ~ 30.1%) 的 G+C 含量显著低于 IR 区 (41.1%) 的，其主要原因与 8 个 G+C 含量较高的 rRNA 基因均分布在 IR 区有关。在编码蛋白的密码子中，部分密码子使用度相比其他同义密码子使用度高 (李娟和薛庆中, 2005)，九翅豆蔻编码蛋白的同义密码子 Nc 值为 54.4，说明九翅豆蔻叶绿体的密码子偏好性较弱，密码子的使用频率没有太大的差异。九翅豆蔻叶绿体基因组的 G+C 含量和 GC3s 含量分别为 37.1% 和 33.7%，占比较小，说明该叶绿体基因组在碱基组成上 A 和 T 的含量更多，且在 64 种密码子中，RSCU 值大于 1 的密

码子以 A 或 U 碱基结尾的密码子有 29 个，而以 C 或 G 碱基结尾的只有 3 个，密码子也主要以 A 或 U 碱基结尾。叶绿体中 SSR 拷贝数的变异是一种重要的分子标记，SSR 分子标记具有技术简单、稳定性好，以及共显性等特点，在遗传多样性分析、DNA 指纹图谱鉴定、基因定位和分子育种等方面应用广泛 (张立娜等, 2012; 李祥栋等, 2019)。

本研究利用 MISA 软件在单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸最小重复数设置为 10、5、4、3、3、3 的情况下共扫描到 132 个 SSR，其中多为由 A/T、AT/TA、AAAT/ATTT 和 AATT/AATT 组成，这表明九翅豆蔻叶绿体基因组中的 SSR 包含频繁的 A 或 T 重复。从系统发育分析看，姜科物种被分为 2 个大的类群，豆蔻属单独聚为一支，姜黄属、山柰属、姜属则聚为一支，其中姜属和山柰属在姜科中的亲缘关系较近；聚类图也清晰地展现了豆蔻属 4 个物种之间的亲缘关系，其中海南砂仁与白豆蔻亲缘关系较近，九翅豆蔻则在豆蔻属中单独成一支。此外，本研究也表明叶绿体基因组可有效区分姜科物种的系统发育关系，为深入研究姜科植物进化、遗传多样性和群体结构提供了参考。

3 材料与方法

3.1 九翅豆蔻植物材料和测序

九翅豆蔻植株来源于云南省西双版纳州中国科学院西双版纳热带植物园，采摘其新鲜幼嫩的叶片，经液氮速冻后，置于 -80 °C 冰箱保存备用。叶绿体基因组数据已信息上传到 NCBI 网站数据库 (索引

号: OQ448893)。

3.2 九翅豆蔻基因组 DNA 提取和测序

采用 TIANGEN 植物 DNA 提取试剂盒提取九翅豆蔻基因组总 DNA, 通过琼脂糖凝胶电泳和 NanoDrop2000 微量分光光度计检测 DNA 的纯度和浓度, 同时使用 Qubit 2.0 荧光定量仪检测 DNA 的浓度。随后使用 Agilent 2100 对文库的插入片段进行检测, 待插入片段大小符合预期, 使用 Q-PCR 方法对文库的有效浓度进行准确定量, 以保证文库质量。测序由中国科学院昆明植物研究所分子实验中心完成。

3.3 九翅豆蔻叶绿体基因组组装和注释

文库质检后利用二代测序技术依靠 Illumina HiSeq 测序平台对基因组测序, 最终得到 2 G raw data, Q30 大于 80%。分别使用软件 FastQC 和 Trimmomatic (Bolger et al., 2014) 对测序得到的 raw reads 进行质检和质控, 去除接头和低质量数据。以姜 (GenBank ID: NC_044775) 叶绿体基因组为参考基因组, 使用软件 Get Organelle v 1.7.5.2 (Jin et al., 2020) 进行组装, 工作原理如下: 调用 Bowtie2 依赖包, 利用 Seed 工具对 reads 进行 Mapping (Langmead and Salzberg, 2012); 通过 Contigs 的延伸抓取 reads; 调用 SPAdes 依赖包对 Contigs 进行组装获得 Graphs (Bankevich et al., 2012); 调用 BLAST 包对 Graphs 进行过滤 (Camacho et al., 2009), 获得 Complete fasta 文件及 gfa 文件; 在 Bandage 软件 (Wick et al., 2015) 中查看 gfa 文件是否闭合, 闭合即为目标植物的基因组草图。组装完成后使用 PGA 进行注释, 设置蛋白质编码基因 TBLASTN 最小识别可变阈值为 80%。初步注释结果使用 Geneious Prime 手动矫正, 利用 OGDRAW 软件 (Greiner et al., 2019) 对基因组数据进行基因组图谱绘制。

3.4 九翅豆蔻叶绿体基因组偏好性分析

利用 Codon W 1.4.2 软件分析九翅豆蔻叶绿体基因组的 RSCU、Nc、密码子数 (N) 等, 并对密码子总 G+C 含量进行分析。Nc 值通常在 20~60 之间, 数值越大其偏好性越弱。RSCU 值通常以 1 为界, 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; 若 RSCU=1, 则无偏好性。

3.5 九翅豆蔻叶绿体基因组 SSR 及重复序列的检测

利用 MISA 软件 (<http://pgrc.ipk-gatersleben.de/misa/misa.html>) 鉴定九翅豆蔻叶绿体基因组中的 SSR, 搜索参数设置为: 含有完全重复的单核苷酸最小重复数为 10, 二核苷酸最小重复数为 5, 三核苷酸最小重复数为 4, 四核苷酸、五核苷酸、六核苷酸最小重复数 3; 另外设置 2 个 SSR 之间的最小距离为 100 bp, 如果距离小于 100 bp, 则 2 个 SSR 被当作一个复合微卫星标记。

3.6 系统发育分析

从 NCBI 数据库下载已公布的姜科豆蔻属、山姜属、姜属、姜黄属和山柰属等共 13 个物种完整的叶绿体基因组序列, 同时以近缘的旅人蕉为外类群 (表 5), 利用软件 Phylosuite v 1.2.1 提取 83 个共有蛋白质编码基因序列, 利用 MAFFT-7.480 软件将提取的 83 条共有序列进行多重比对, 结果经手工检查与调整后用最大似然法 (maximum-likelihood method, ML) 对系统进化关系进行分析, ML 系统发育树用 IQTREE 软件生成, 选择最佳模型 GTR+F+R2, 自展值 ultrafast bootstrap (UFBoot) 为 1 000, 最后以 Figtree 进行系统发育树可视化分析 (Kalyanamoorthy et al., 2017; Xu et al., 2017; Hoang et al., 2018)。

表 5 13 种植物系统发育树的建立信息及其序列号

Table 5 Establishment information of 13 plants and their serial numbers in phylogenetic tree

物种 Species	GenBank 登录号 GenBank login ID	基因组 大小/bp Genome size/bp	总 G+C 含量/% Total G+C content/%	蛋白编码 基因数量 Number of protein coding gene	rRNA 基因 数量 Number of rRNA gene	tRNA 基因 数量 Number of tRNA gene	基因数量 Number of gene	假基因数量 Number of pseudogene
姜 <i>Zingiber officinale</i>	NC_044775	162 621	36.1	86	8	38	133	1
蜂巢姜 <i>Zingiber spectabile</i>	NC_020363	155 890	36.3	85	8	38	132	1

续表
Continuing table

物种 Species	GenBank 登录号 GenBank login ID	基因组 大小/bp Genome size/bp	总 G+C 含量/% Total G+C content/%	蛋白编码 基因数量 Number of protein coding gene	rRNA 基因 数量 Number of rRNA gene	tRNA 基因 数量 Number of tRNA gene	基因数量 Number of gene	假基因数量 Number of pseudogene
山柰 <i>Kaempferia galanga</i>	NC_040851	163 811	36.1	87	8	38	133	
紫花山柰 <i>Kaempferia elegans</i>	NC_040852	163 555	36.1	87	8	37	132	
姜黄 <i>Curcuma longa</i>	NC_042886	159 550	36.3	86	8	38	133	1
温郁金 <i>Curcuma wenyujin</i>	NC_045241	162 266	36.2	87	8	38	133	
观音姜 <i>Curcuma roscoeana</i>	NC_022928	159 512	36.3	86	8	38	133	1
白豆蔻 <i>Amomum krarvanh</i>	NC_036935	162 766	36.2	87	8	30	127	2
海南砂仁 <i>Amomum longiligulare</i>	NC_044774	164 068	36.0	87	8	38	133	
草果 <i>Amomum tsaoko</i>	MK926774	163 648	36.4	91	8	37	136	
闭鞘姜 <i>Hellenia speciosa</i>	OK641589	167 158	36.3	88	8	38	134	
美叶闭鞘姜 <i>Costus pulverulentus</i>	KF601573	163 239	36.0	86	8	38	132	
旅人蕉 <i>Ravenala madagascariensis</i>	NC_022927	166 170	37.3	83	8	38	133	4

作者贡献

傅武祥是本研究的实验设计和实验研究的执行人,完成数据分析和论文初稿的写作;郭雨桐和张雪梅参与实验设计和试验结果分析;马占霞是本研究的构思者及负责人,指导实验设计、数据分析和论文写作与修改。全体作者都已阅读并同意最终的文本。

参考文献

- 李娟,薛庆中,2005.拟南芥及水稻转录因子 MADS 密码子的偏好性比较.浙江大学学报(农业与生命科学版),31(5):513-517. [LI J, XUE Q Z, 2005. Comparison of MADS transcriptional factor on codon bias in *Arabidopsis* and rice. Journal of Zhejiang University (Agriculture and Life Sciences), 31(5): 513-517.]
- 李祥栋,石明,陆秀娟,等,2019.利用叶绿体基因组 SSR 标记揭示薏苡属种质资源的遗传多样性.华北农学报,34(S1):6-14. [LI X D, SHI M, LU X J, et al., 2019. Genetic diversity of germplasm resources in job's tears revealed by chloroplast genome SSR markers. Acta Agriculturae Boreali-Sinica, 34(S1): 6-14.]
- 廖彩丽,刘春生,张园园,等,2013.基于中药系统鉴别法的金铁锁及其混淆品的精确鉴别.中国中药杂志,38(8):1134-1137. [LIAO C L, LIU C S, ZHANG Y Y, et al., 2013. Accurate identification of *Psammosilene tunicoides* and its confused species by systematic identification method. China Journal of Chinese Materia Medica, 38(8): 1134-1137.]
- 林艳芳,依专,赵应红,2003.中国傣医药彩色图谱.昆明:云南民族出版社:178. [LIN Y F, YI ZH, ZHAO Y H, 2003. Chinese Dai Medicine Colorful Illustrations. Kunming: The Nationalities Publishing House of Yunnan: 178.]
- 卢传礼,2017.傣药九翅豆蔻化学成分和药理活性研究进展.亚太传统医药,13(14):50-52. [LU C L, 2017. Research progress on chemical constituents and pharmacological activities of Dai medicine Cardamom septempferi. Asia-Pacific Traditional Medicine, 13(14): 50-52.]
- 马孟莉,张薇,孟衡玲,等,2021.草果叶绿体基因组特征及

- 系统发育分析. 中草药, 52(19): 6023-6031. [MA M L, ZHANG W, MENG H L, et al., 2021. Characterization and phylogenetic analysis of the complete chloroplast genome of *Amomum tsaoko*. Chinese Traditional and Herbal Drugs, 52(19): 6023-6031.]
- 宁淑萍, 颜海飞, 郝刚, 等, 2008. 植物 DNA 条形码研究进展. 生物多样性, 16(5): 417-425. [NING S P, YAN H F, HAO G, et al., 2008. Current advances of DNA barcoding study in plants. Biodiversity Science, 16(5): 417-425.]
- 石林春, 宋经元, 陈士林, 等, 2010. 豆蔻属药用植物 DNA 条形码鉴定研究. 世界科学技术(中医药现代化), 12(3): 473-479. [SHI L C, SONG J Y, CHEN S L, et al., 2010. Identification of *Amomum* (Zingiberaceae) through DNA barcodes. World Science and Technology (Modernization of Traditional Chinese Medicine and Materia Medica), 12(3): 473-479.]
- 袁庆军, 张文婧, 姜丹, 等, 2012. 论中药分子鉴定的方法和原则. 植物分类与资源学报, 34(6): 607-613. [YUAN Q J, ZHANG W J, JIANG D, et al., 2012. On the methods and principles of molecular identification of Chinese herbs. Plant Diversity and Resources, 34(6): 607-613.]
- 张立娜, 曹桂兰, 韩龙植, 2012. 利用 SSR 标记揭示中国梗稻地方品种遗传多样性. 中国农业科学, 45(3): 405-413. [ZHANG L N, CAO G L, HAN L Z, 2012. Analysis of genetic diversity of *Japonica* rice Landrace in China with microsatellite marker. Scientia Agricultura Sinica, 45(3): 405-413.]
- 中国科学院昆明植物研究所, 1997. 云南植物志-第八卷-种子植物. 北京: 科学出版社: 410-411. [Kunming Institute of Botany, Chinese Academy of Science, 1997. Flora Yunnanica. Beijing: Science Press: 410-411.]
- 钟志敏, 赖小平, 黄松, 等, 2018. 基于 *yefI* 的姜科植物条形码鉴定及聚类分析. 中华中医药杂志, 33(9): 4089-4092. [ZHONG Z M, LAI X P, HUANG S, et al., 2018. Identification and clustering analysis of Zingiberaceae plant based on *yefI* barcode. China Journal of Traditional Chinese Medicine and Pharmacy, 33(9): 4089-4092.]
- 朱斌, 甘晨晨, 王洪程, 2021. 球花石斛 (*Dendrobium thyrsiflorum*) 叶绿体基因组特征及亲缘关系解析. 生物技术通报, 37(8): 13-22. [ZHU B, GAN C C, WANG H C, 2021. Characteristics of the complete chloroplast genome of *Dendrobium thyrsiflorum* and its phylogenetic relationship analysis. Biotechnology Bulletin, 37(8): 13-22.]
- BANKEVICH A, NURK S, ANTIPOV D, et al., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol., 19(5): 455-477.
- BOLGER A M, LOHSE M, USADEL B, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15): 2114-2120.
- CAMACHO C, COULOURIS G, AVAGYAN V, et al., 2009. BLAST+: architecture and applications. BMC Bioinformatics, 10: 421.
- CHEZIA M, GIRI B R, SWARGIARY A, et al., 2014. *Amomum maximum* Roxb (Zingiberaceae), a medicinal plant of Tripura, India: a natural anthelmintic? J. Advan. Micro. Res., 9(2): 148-153.
- CHINA PLANT BOL GROUP, LI D Z, GAO L M, et al., 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc. Natl. Acad. Sci. USA, 108(49): 19641-19646.
- GREINER S, LEHWARK P, BOCK R, 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res., 47(W1): W59-W64.
- GUO S S, YOU C X, LIANG J Y, et al., 2015. Essential oil of *Amomum maximum* Roxb. and its bioactivities against two stored-product insects. J. Oleo Sci., 64(12): 1307-1314.
- HOANG D T, CHERNOMOR O, VON HAESELER A, et al., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Oxford University Press, 35(2): 518-522.
- HUONG L T, DAI D N, THANG T D, et al., 2015. Volatile constituents of *Amomum maximum* Roxb and *Amomum microcarpum* C. F. Liang & D. Fang: two Zingiberaceae grown in Vietnam. Nat. Prod. Res., 29(15): 1469-1472.
- JI K L, FAN Y Y, GE Z P, et al., 2019. Maximumins A-D, rearranged labdane-type diterpenoids with four different carbon skeletons from *Amomum maximum*. J. Organ. Chem., 84(1): 282-288.
- JIN J J, YU W B, YANG J B, et al., 2020. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organellar genomes. Cold Spring Harbor Laboratory, 21(1): 241.
- KALYAANAMOORTHY S, MINH B Q, WONG T K F, et al., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods, 14(6): 587-589.
- LANGMEAD B, SALZBERG S L, 2012. Fast gapped-read alignment with bowtie 2. Nature Methods, 9(4): 357-359.
- LI D M, YE Y J, XU Y C, et al., 2020. Complete chloroplast genomes of *Zingiber montanum* and *Zingiber zerumbet*: genome structure, comparative and phylogenetic analyses. PLoS ONE, 15(7): e0236590.
- LU C L, WANG L N, LI Y J, et al., 2022. Anti-hyperglycaemic effect of labdane diterpenes isolated from the rhizome of *Amomum maximum* Roxb., an edible plant in Southwest China. Nat. Prod. Res., 36(10): 2570-2574.
- LUO J G, YIN H, FAN B Y, et al., 2014. Labdane diterpenoids from the roots of *Amomum maximum* and their cytotoxic eval-

- uation. *Helv. Chim. Acta*, 97(8): 1140-1145.
- RAZI K, MUNEEER S, 2021. Drought stress-induced physiological mechanisms, signaling pathways and molecular response of chloroplasts in common vegetable crops. *Crit. Rev. Biotechnol.*, 41(5): 669-691.
- WICK R R, SCHULTZ M B, ZOBEL J, et al., 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinform. (Oxf. Engl.)*, 31(20): 3350-3352.
- XU C, DONG W P, LI W Q, et al., 2017. Comparative analysis of six *Lagerstroemia* complete chloroplast genomes. *Front. Plant Sci.*, 8: 15.
- XU D H, ABE J, KANAZAWA A, et al., 2001. Identification of sequence variations by PCR-RFLP and its application to the evaluation of cpDNA diversity in wild and cultivated soybeans. *Theor. App. Genet.*, 102(5): 683-688.
- ZENG C X, HOLLINGSWORTH P M, YANG J, et al., 2018. Genome skimming herbarium specimens for DNA barcoding and phylogenomics. *Plant Methods*, 14: 43.

(责任编辑 谢为博)
(责任副主编 罗继景)