

## • 药材与资源 •

## 黄花地黄叶绿体基因组序列特征及系统进化关系研究

赵俊禧, 石盼盼, 王小萱, 雷敏, 詹海仙\*, 杜晨晖, 廖雨婷, 尹凯  
山西中医药大学中药与食品工程学院, 山西 晋中 030619

**摘要:** 目的 以黄花地黄 *Rehmannia chrysantha* 为材料进行高通量测序, 明确黄花地黄叶绿体基因组的结构及其与近缘物种的系统发育关系。方法 使用 CodonW v1.3 软件对黄花地黄叶绿体进行密码子偏好性分析, 导入 IRscope、mVISTA 软件, 分析黄花地黄与其同属物种序列变异情况, 用 MEGA11 软件构建地黄属及其近缘属的系统发育树。结果 黄花地黄叶绿体基因组全长 153 789 bp, 是典型的环状双链四分体结构, 其中有 28 个密码子的使用频率较高, 中性绘图分析、ENC-plot 分析及 PR2-plot 分析表明黄花地黄密码子使用偏好性主要受自然选择影响。地黄属 IR 边界有不同程度的变化, 且主要集中于 JLA、JLB。rpl2、psbI、petN、psbZ、ycf1、ndhB 可分别作为鉴别地黄同属植物的 DNA 条形码序列。系统发育树表明黄花地黄与地黄亲缘关系较近。结论 地黄属叶绿体基因组蛋白编码基因倾向于使用 A/U 碱基编码的密码子, 自然选择是影响密码子偏好性的主要因素。研究结果可为地黄属的分子进化、系统发育及叶绿体基因工程提供参考。

**关键词:** 黄花地黄; 叶绿体基因组; 密码子偏好性; 序列变异; 系统发育

中图分类号: R283 文献标志码: A 文章编号: 0253-2670(2026)04-1450-10

DOI: 10.7501/j.issn.0253-2670.2026.04.021

## Sequence characterization and phylogenetic relationships of chloroplast genome of *Rehmannia chrysantha*

ZHAO Junxi, SHI Panpan, WANG Xiaoxuan, LEI Min, ZHAN Haixian, DU Chenhui, LIAO Yuting, YIN Kai  
College of Traditional Chinese Medicine and Food Engineering, Shanxi University of Chinese Medicine, Jinzhong 030619, China

**Abstract: Objective** High-throughput sequencing of *Rehmannia chrysantha* L. was performed to elucidate the structure of its chloroplast genome and its phylogenetic relationship with related species. **Methods** Codon preference analysis of chloroplasts of *R. chrysantha* was performed using CodonW v1.3 software. The codon usage bias was analyzed by using CodonW v1.3 software. Then, the obtained data were imported into IRscope and mVISTA to clarify sequence variations between *R. chrysantha* and related species. This study constructed a phylogenetic tree of the genus *Rehmannia* and its related genera on MEGA11 software. **Results** The chloroplast genome of *R. chrysantha* was a typical circular double-stranded molecule with a quadripartite structure and a total length of 153 789 bp. It has 28 high-frequency codons. Neutral plot analysis, ENC-plot analysis, and PR2-plot analysis all indicated that natural selection was the main cause of codon usage bias in *R. chrysantha*. It was also found that there were varying degrees of variation in the inverted repeat (IR) boundaries of the *Rehmannia* genus, and these changes were mainly concentrated in JLA and JLB. The genes rpl2, psbI, petN, psbZ, ycf1, and ndhB were identified as potential DNA barcodes for distinguishing closely related genus *Rehmannia*. In addition, phylogenetic analysis revealed a close evolutionary relationship between *R. chrysantha* and *R. glutinosa* L. **Conclusion** Protein-coding genes in the chloroplast genomes of genus *Rehmannia* exhibited a preferential use of codons ending in A/U, with natural selection being the primary factor influencing this codon usage bias. These findings will provide a foundation for future studies on molecular evolution, phylogeny, and chloroplast genetic engineering within this genus.

**Key words:** *Rehmannia chrysantha* L.; chloroplast genome; codon preference; sequence variation; phylogeny

收稿日期: 2025-10-03

基金项目: 山西省中药材产业技术体系(2026); 山西省科技创新人才团队专项(202304051001020); 山西中医药大学 2025 年研究生实践创新项目(X2025SJ036)

作者简介: 赵俊禧(2000—), 男, 天津市人, 硕士研究生, 从事中药资源开发与品质评价。E-mail: zhaojunxitj@126.com

\*通信作者: 詹海仙, 女, 教授, 硕士生导师, 从事药用植物资源研究。E-mail: zhan030006@126.com

地黄属 *Rehmannia* L. 主要分布于东亚地区, 目前在中国有地黄、天目地黄、湖北地黄、黄花地黄等 6 个物种分布, 其中, 地黄 *Rehmannia glutinosa* L. 为大宗中药, 生品具有清热凉血的功效<sup>[1]</sup>, 熟地具有滋阴补血的作用<sup>[2]</sup>, 其活性成分能缓解氧化应激和改善炎症<sup>[2]</sup>。地黄的近缘种中, 天目地黄在民间的应用价值与地黄相似, 化学成分也比较接近<sup>[3]</sup>。湖北地黄含有环烯醚萜苷类、苯乙醇苷类、糖类和紫罗兰酮类成分, 以全草入药<sup>[4]</sup>。据报道亲缘关系相近的植物通常具有相似的遗传关系和生理生化特征, 进而表现出含有相似化学成分的特点<sup>[5]</sup>。黄花地黄 *R. chrysantha* L. 为 2014 年发现的 1 个地黄属新种<sup>[6]</sup>, 被收录于《中国生物物种名录 2025 版》, 其主要特征为黄色花冠与绿色叶背, 而同属地黄为紫色叶背。

叶绿体是植物光合作用的重要场所, 与植物生长发育关系密切<sup>[7]</sup>。由于叶绿体基因组具有母系遗传、结构稳定、较小分子量等特点<sup>[8]</sup>, 被广泛应用于药用植物物种鉴定<sup>[9]</sup>、系统进化<sup>[10]</sup>、遗传多样性<sup>[11]</sup>及基因工程等研究。目前, 与黄花地黄同属的 5 个物种叶绿体基因组均有报道<sup>[12]</sup>, 基因组大小分别为地黄 (153 622 bp)、天目地黄 (154 055 bp)、湖北地黄 (153 089 bp)、裂叶地黄 (153 952 bp)、茄叶地黄 (153 989 bp)。上述 5 个地黄属植物叶绿体基因组均为环状四分体结构, 包括 1 个大单拷贝 (large single-copy, LSC)、1 个小单拷贝 (small single-copy, SSC) 以及 2 个反向重复序列 (inverted repeat sequence, IRs)。鉴于地黄属植物形态接近并具有较高的药用价值, 为了充分开发和利用地黄属资源, 本研究首次组装并注释了黄花地黄的叶绿体基因组, 丰富了地黄属的叶绿体基因组数据。同时, 结合已发表的地黄属 5 种植物的叶绿体基因组信息, 利用生物信息学手段对 6 个地黄属植物叶绿体的结构特征、重复序列分布、密码子偏好性、叶绿体基因组边界特征及系统发育关系进行分析, 研究结果将为地黄属植物的叶绿体遗传多样性、进化及系统发育研究提供参考。

## 1 材料

样品采集于天津市蓟州区 (40°5'25.97"N, 117°15'28.31"E), 经山西中医药大学杜晨晖教授鉴定为列当科地黄属黄花地黄 *R. chrysantha* L.。采集健康植株较新鲜叶片, 并用冰袋保存。后续研究的其他物种叶绿体基因组来源于 NCBI 数据库, 物种详情见表 1。

表 1 列当科 10 个物种的叶绿体基因组信息

Table 1 Chloroplast genomic information of 10 species of family Orobanchaceae

物种	Genbank 登录号
地黄 <i>R. glutinosa</i>	MG977439
天目地黄 <i>R. chingii</i>	KX426347
湖北地黄 <i>R. henryi</i>	KX636158
裂叶地黄 <i>R. piasezkii</i>	KX636160
茄叶地黄 <i>R. solanifolia</i>	KX636159
大黄花 <i>Cymbaria daurica</i>	NC_064104
光药大黄花 <i>C. mongolica</i>	NC_064388
肉苁蓉 <i>Cistanche deserticola</i>	KC_128846
黄花肉苁蓉 <i>C. phelypaea</i>	NC_025642
钟萼草 <i>Lindenbergia philippensis</i>	NC_022859

## 2 方法

### 2.1 DNA 提取与测序

使用 LyserPro 程序式冷冻组织破碎仪对黄花地黄叶片进行研磨, 采用改良 CTAB 方法<sup>[11]</sup>提取总 DNA, 1% 琼脂糖凝胶电泳和微量分光光度仪检测 DNA 质量。将合格的 DNA 样品, 送至深圳市惠通生物科技有限公司进行测序, 二代测序平台为 Illumina NovoSeq 6000。

### 2.2 叶绿体基因组组装、注释及物理图谱绘制

拼接软件选择对环状基因组拼接效果较好的 SPAdes (版本 3.14.1), 对 clean data 数据进行 de novo 组装, 设置 k-mer 21、45、65、85、105, 获得 graph 文件, 然后用 Bandage 软件可视化去除多余 contig 并编辑成环状, 叶绿体基因组注释采用 PGA 软件, 并手动矫正注释结果。将注释后的黄花地黄叶绿体基因组上传到 NCBI 数据库中, 获得登录号 (PX170676)。最后, 利用 OGDRAW 网站中的 Chloroplot<sup>[13]</sup>工具采用默认参数绘制黄花地黄叶绿体基因组圈图。

### 2.3 简单重复序列 (simple sequence repeat, SSR) 位点分析

应用在线工具 MISA 分析黄花地黄的 SSR, 设置参数: 单核苷酸序列为 8; 二、三核苷酸重复序列为 4; 四、五及以上核苷酸重复序列设为 3, 且每 2 个 SSRs 之间的距离为 100 bp<sup>[14]</sup>。

### 2.4 密码子偏好性分析

统计各基因编码区 GC 总含量 (GC<sub>all</sub>) 以及密码子第 1 位 (GC<sub>1</sub>)、第 2 位 (GC<sub>2</sub>)、第 3 位 (GC<sub>3</sub>) 的 GC 含量。用 CodonW v1.3 软件对黄花地黄叶绿体基因组密码子的相对同义密码子的使用情况进

行评价, RSCU 值 >1 的为高频使用密码子, RSCU 值 <1 的为低使用频率密码子<sup>[7]</sup>。

### 2.5 中性绘图、ENC-plot 绘图、PR2-plot 绘图分析

通过中性绘图、ENC-plot 和偏倚分析 (PR2-plot) 分析密码子偏好性。以所获得基因组的 GC<sub>12</sub> 含量为纵坐标, GC<sub>3</sub> 含量为横坐标, 对密码子基因进行中性绘图分析; ENC-plot 分析以 ENC 值为纵坐标, GC<sub>3</sub> 为横坐标绘制散点图, 同时通过绘制标准曲线分析突变和选择对密码子偏好性的影响, 标准曲线公式:  $ENC = 2 + GC_3 + 29/[GC_3^2 + (1 - GC_3)^2]$ ; 以  $G_3/(G_3 + C_3)$  为横坐标,  $A_3/(A_3 + T_3)$  为纵坐标绘制 PR2-plot 散点图, 利用中心点 (A = T, G = C) 分析碱基偏移的程度和方向。

### 2.6 序列比对分析

将测序所得黄花地黄叶绿体基因组与 5 个同属植物叶绿体注释序列列入 IRscope 软件, 分析叶绿体 IR 边界的收缩和扩张。同时, 用 mVISTA 比较黄花地黄和地黄属 5 种植物的叶绿体基因组差异。

### 2.7 系统发育分析

用 MEGA11 (最大似然法, bootstrap = 1 000)

软件分析 NCBI 中下载的地黄、天目地黄、湖北地黄、裂叶地黄、茄叶地黄以及同科植物大黄花、光药大黄花、肉苁蓉、黄花肉苁蓉、钟萼草的叶绿体基因组序列, 以及本研究的黄花地黄叶绿体基因组的序列并生成系统发育树。

## 3 结果与分析

### 3.1 黄花地黄叶绿体基因组基本特征

黄花地黄叶绿体基因组结构具有环状双链四分体结构 (图 1), 基因组全长为 153 789 bp, LSC 和 SSC 长度分别为 84 637 bp 和 24 137 bp, 1 对反向重复区 (IRa/IRb) 长为 17 608 bp。三段区域 GC 含量存在差异, IR (43.09%) > LSC (35.97%) > SSC (32.19%), 与大部分被子植物 GC 含量一致。黄花地黄叶绿体基因组注释结果 (表 2) 显示, 共注释到 103 个基因, 主要分为光合系统、自我复制、其他基因和功能未知基因 4 大类, rRNA 基因 4 个, tRNA 基因 30 个, CDS 基因 69 个。ndhA、ndhB、petB、petD、atpF、rpl16、rpl2、rps16、rpoC1、trnA-UGC、trnG-UCC、trnI-GAU、trnK-UUU、trnL-UAA、trnV-UAC 含有 1 个内含子, 而 rps12、clpP、ycf3 则含有 2 个内含子。

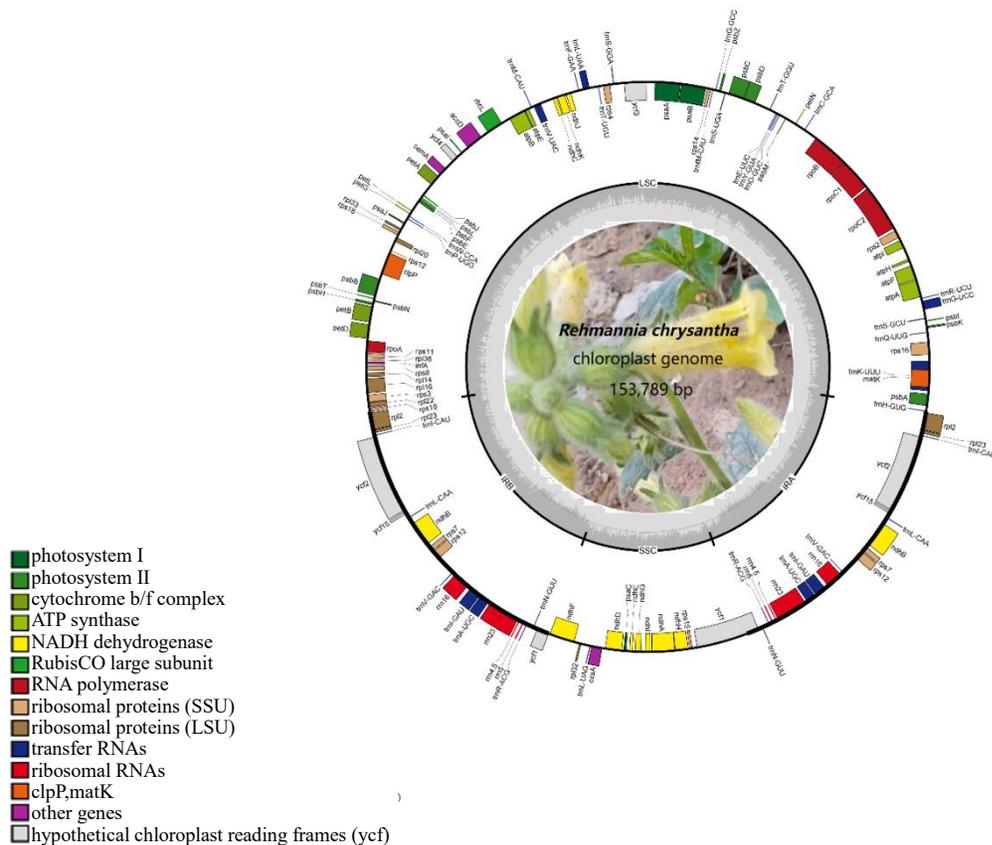


图 1 黄花地黄叶绿体基因组圈图

Fig. 1 Circularized map of chloroplast genome of *R. chrysantha*

表 2 黄花地黄叶绿体基因功能分类

Table 2 Chloroplast gene functional classification of *R. chrysantha*

基因功能	基因类型	基因名称
光合作用	光系统I	<i>psaA</i> 、 <i>psaB</i> 、 <i>psaC</i> 、 <i>psaI</i> 、 <i>psaJ</i>
	光系统II	<i>psbA</i> 、 <i>psbB</i> 、 <i>psbC</i> 、 <i>psbD</i> 、 <i>psbE</i> 、 <i>psbF</i> 、 <i>psbH</i> 、 <i>psbI</i> 、 <i>psbJ</i> 、 <i>psbK</i> 、 <i>psbL</i> 、 <i>psbM</i> 、 <i>psbN</i> 、 <i>psbT</i> 、 <i>psbZ</i>
	NADH脱氢酶	<i>ndhA</i> <sup>*</sup> 、 <i>ndhB</i> <sup>*(2)</sup> 、 <i>ndhC</i> 、 <i>ndhD</i> 、 <i>ndhE</i> 、 <i>ndhF</i> 、 <i>ndhG</i> 、 <i>ndhH</i> 、 <i>ndhI</i> 、 <i>ndhJ</i> 、 <i>ndhK</i>
	细胞色素b/f复合体	<i>petA</i> 、 <i>petB</i> <sup>*</sup> 、 <i>petD</i> <sup>*</sup> 、 <i>petG</i> 、 <i>petL</i> 、 <i>petN</i>
	ATP合成酶	<i>atpA</i> 、 <i>atpB</i> 、 <i>atpE</i> 、 <i>atpF</i> <sup>*</sup> 、 <i>atpH</i> 、 <i>atpI</i>
	Rubisco酶大亚基	<i>rbcL</i>
	自我复制	核糖体大亚基蛋白质
核糖体小亚基蛋白质		<i>rps11</i> 、 <i>rps12</i> <sup>**</sup> (2)、 <i>rps14</i> 、 <i>rps15</i> 、 <i>rps16</i> <sup>*</sup> 、 <i>rps18</i> 、 <i>rps19</i> 、 <i>rps2</i> 、 <i>rps3</i> 、 <i>rps4</i> 、 <i>rps7</i> (2)、 <i>rps8</i>
RNA聚合酶		<i>rpoA</i> 、 <i>rpoB</i> 、 <i>rpoC1</i> <sup>*</sup> 、 <i>rpoC2</i>
核糖体RNA		<i>rrn16</i> (2)、 <i>rrn23</i> (2)、 <i>rrn4.5</i> (2)、 <i>rrn5</i> (2)
转运RNA		<i>trnA</i> -UGC <sup>*</sup> (2)、 <i>trnC</i> -GCA、 <i>trnD</i> -GUC、 <i>trnE</i> -UUC、 <i>trnF</i> -GAA、 <i>trnG</i> -GCC、 <i>trnG</i> -UCC <sup>*</sup> 、 <i>trnH</i> -GUG、 <i>trnI</i> -CAU(2)、 <i>trnI</i> -GAU <sup>*</sup> (2)、 <i>trnK</i> -UUU <sup>*</sup> 、 <i>trnL</i> -CAA(2)、 <i>trnL</i> -UAA <sup>*</sup> 、 <i>trnL</i> -UAG、 <i>trnM</i> -CAU、 <i>trnN</i> -GUU(2)、 <i>trnP</i> -UGG、 <i>trnQ</i> -UUG、 <i>trnR</i> -ACG(2)、 <i>trnR</i> -UCU、 <i>trnS</i> -GCU、 <i>trnS</i> -GGA、 <i>trnS</i> -UGA、 <i>trnT</i> -GGU、 <i>trnT</i> -UGU、 <i>trnV</i> -GAC(2)、 <i>trnV</i> -UAC <sup>*</sup> 、 <i>trnW</i> -CCA、 <i>trnY</i> -GUA、 <i>trnY</i> -CAU
其他基因		成熟酶
	蛋白酶	<i>clpP</i> <sup>**</sup>
	外膜蛋白	<i>cemA</i>
	乙酰辅酶A羧化酶亚基	<i>accD</i>
	c型细胞色素合成基因	<i>ccsA</i>
未知功能基因	翻译起始因子	<i>infA</i>
	保守开放阅读框	<i>yef1</i> (2)、 <i>yef15</i> (2)、 <i>yef2</i> (2)、 <i>yef3</i> <sup>**</sup> 、 <i>yef4</i>

\*含有 1 个内含子; \*\*含有 2 个内含子。

\*gene with one intron; \*\*gene with two introns.

### 3.2 SSR 位点统计

应用 MISA 分析黄花地黄 SSR 位点,共检测到 180 个 SSR 位点(表 3),其中以单核苷酸最多,共有 132 个,其次是二核苷酸为 33 个,三核苷酸为 2

个,四核苷酸为 10 个,五核苷酸为 3 个,未发现六核苷酸重复位点。SSR 的类型以 A/T 为主,占比为 71.11%;其次为 AT/AT 和 AG/CT,分别占 8.88%、7.77%,其余序列占比较小。

表 3 黄花地黄叶绿体基因组的 SSRs 类型及数量

Table 3 SSRs types and quantities of chloroplast genome in *R. chrysantha*

核苷酸类型	SSR重复序列	重复序列个数	占比/%	合计占比/%
单核苷酸	A/T	128	71.11	73.33
	C/G	4	2.22	
二核苷酸	AC/GT	3	1.66	18.31
	AG/CT	14	7.77	
	AT/AT	16	8.88	
三核苷酸	AAG/CTT	1	0.55	1.10
	AAT/ATT	1	0.55	
四核苷酸	AAAC/GTTT	1	0.55	5.53
	AAAG/CTTT	3	1.66	
	AAAT/ATTT	2	1.11	
	AATC/ATTG	2	1.11	
	AATT/AATT	1	0.55	
	AGAT/ATCT	1	0.55	
五核苷酸	AAAAT/ATTTT	1	0.55	1.66
	AATCT/AGATT	2	1.11	

### 3.3 密码子偏好性分析

结果显示, 黄花地黄叶绿体基因组共有 20 863 个密码子, 其中有 28 个密码子使用度 (RSCU) 值大于 1, 34 个 RSCU 值小于 1, 有 2 个密码子 UGG、AUG 的密码子偏好性等于 1, 其中编码精亮氨酸 (Leu) 的密码子 UUA 的 RSCU 值最高 (1.91) (图 2-A)。GC<sub>all</sub> 及密码子 3 个位置的平均 GC 含量分别为 38.33%、47.07%、39.78%、28.13% (图 2-B),

GC<sub>3</sub> 的平均含量最低, 黄花地黄叶绿体基因组密码子偏好以 A/U 结尾。

### 3.4 中性绘图分析

中性绘图分析是评价密码子使用偏性的指标之一, 通过中性绘图分析地黄属 GC<sub>12</sub> 与 GC<sub>3</sub> 之间的相关性, 地黄属 6 个物种的叶绿体基因均位于对角线上方 (图 3), 且大部分基因与对角线距离较远。每个物种的回归系数均较小, 在 0.110~

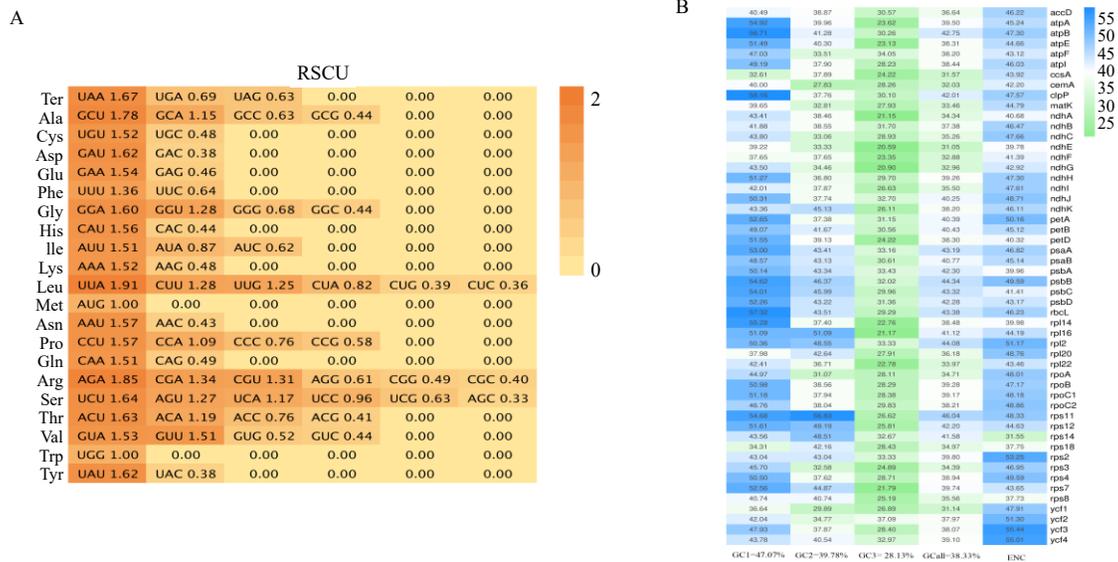


图 2 黄花地黄叶绿体基因组相对同义密码子使用度和 GC 含量 (A) 和 ENC (B) 值

Fig. 2 Relative synonymous codon usage and GCcontent (A) and ENC values (B) in chloroplast genome of *R. chrysanthra*

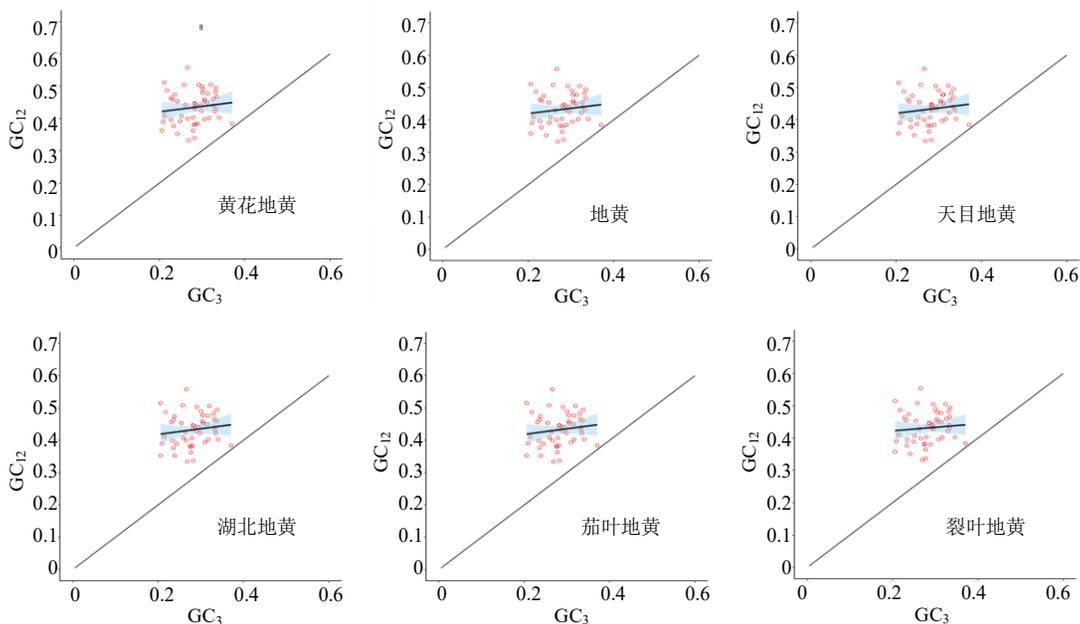


图 3 中性绘图分析

Fig. 3 Analysis of neutrality plot

0.174, 表明地黄属在进化过程中受突变的影响为 11.0%~17.4%, 而自然选择压力为 82.6%~89.0%。因此, 地黄属密码子偏好性受自然选择影响大。

### 3.5 ENC-plot 分析

ENC 与 GC<sub>3</sub> 关联分析显示, 以预期 ENC 值 (标准曲线) 为参考, 绝大部分基因的实际 ENC 值小于预期 ENC 值 (图 4), 说明黄花地黄及同属 5 个物种叶绿

体基因组密码子偏好性受自然选择影响; 同时, ENC 频数分布显示, 大部分基因的实际 ENC 比值分布于 0.15~0.50, 实际 ENC 值与预期 ENC 值相差较多, 离标准曲线较远, 说明地黄属密码子偏好性受自然选择影响大。

### 3.6 PR2-plot 分析

对地黄属植物叶绿体基因组进行 PR2-plot 分析。图 5 显示, 在以 0.5 为中心的 4 个象限各基因分布不均匀, 下方分布多于上方, 且右下角分布较多。说明

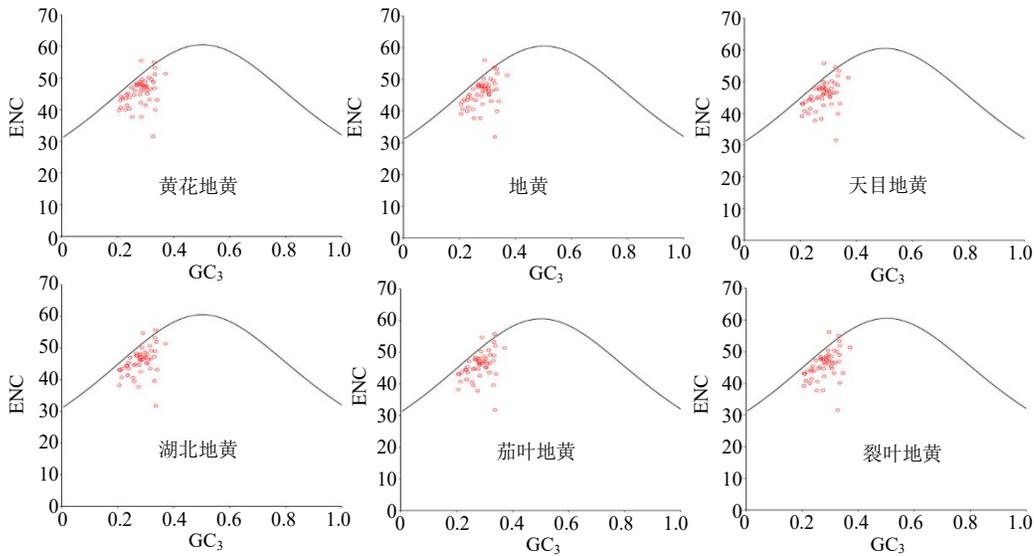


图 4 ENC-plot 绘图分析

Fig. 4 ENC-plot analysis

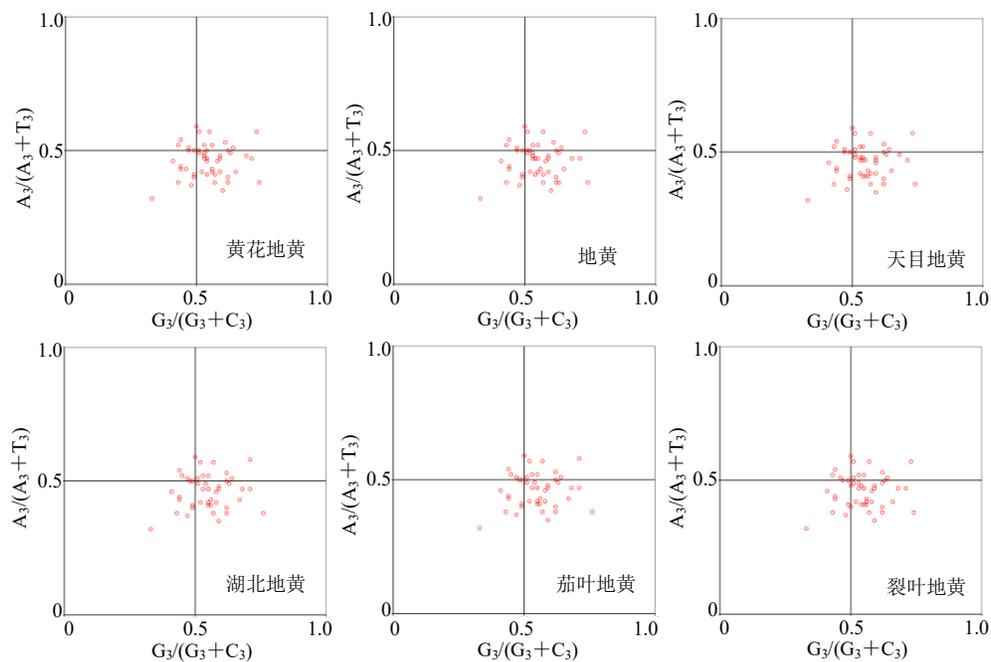


图 5 PR2-plot 分析

Fig. 5 PR2-plot analysis

地黄属物种叶绿体基因组的碱基使用频率 T>A、G>C，密码子第3位的碱基使用 T/G 的频率明显高于 A/C。进一步说明了地黄属植物叶绿体基因组密码子偏好性主要受选择影响，但同时也有其他因素的作用。

### 3.7 IR边界分析

分析黄花地黄与其同属植物叶绿体基因组 SC/IR 边界差异 (图 6)，黄花地黄和地黄 JLB 均位于基因 *rps19* 和 *rpl2* 之间的 IGS，而其余 4 个物种的 JLB 均位于 *rpl2* 基因中，但扩张程度稍有差异。地黄属物种的 JSB、JSA 分别位于 *ycf1* 基因 3'

端 3 bp 和 1 080 bp 处。而黄花地黄和地黄 JLA 分别位于基因 *rps12* 和 *trnH* 之间的 IGS、*rps12* 和 *psbA* 之间的 IGS，其他 4 个物种的 JLA 位于基因 *rps12* 中，只是扩张程度稍有差异。

### 3.8 叶绿体基因组比对分析

应用 mVISTA 在线工具，并选择检测基因重排和倒位的全局比对模式 (Shuffle-LAGAN)，对黄花地黄及同属 5 个物种：地黄、天目地黄、湖北地黄、茄叶地黄和裂叶地黄进行序列差异分析 (图 7)。在序列差异中，编码区差异小于非编码区域，大部分

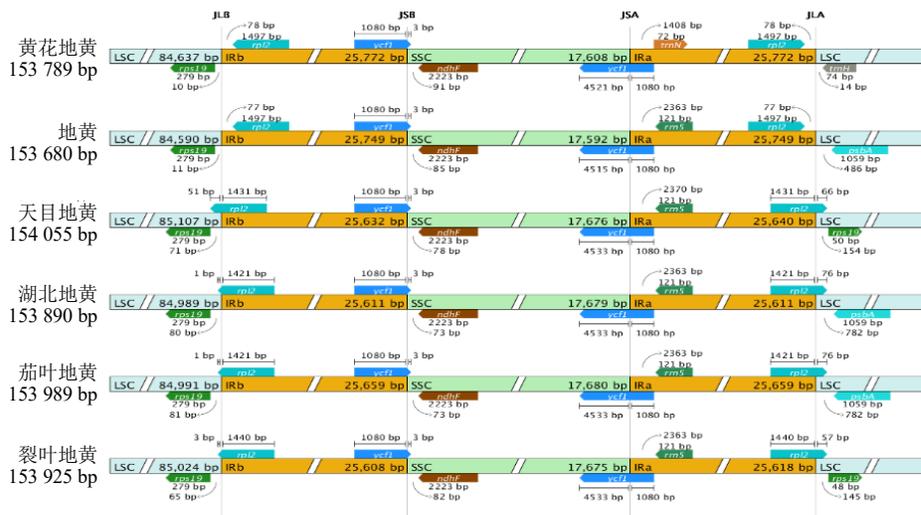
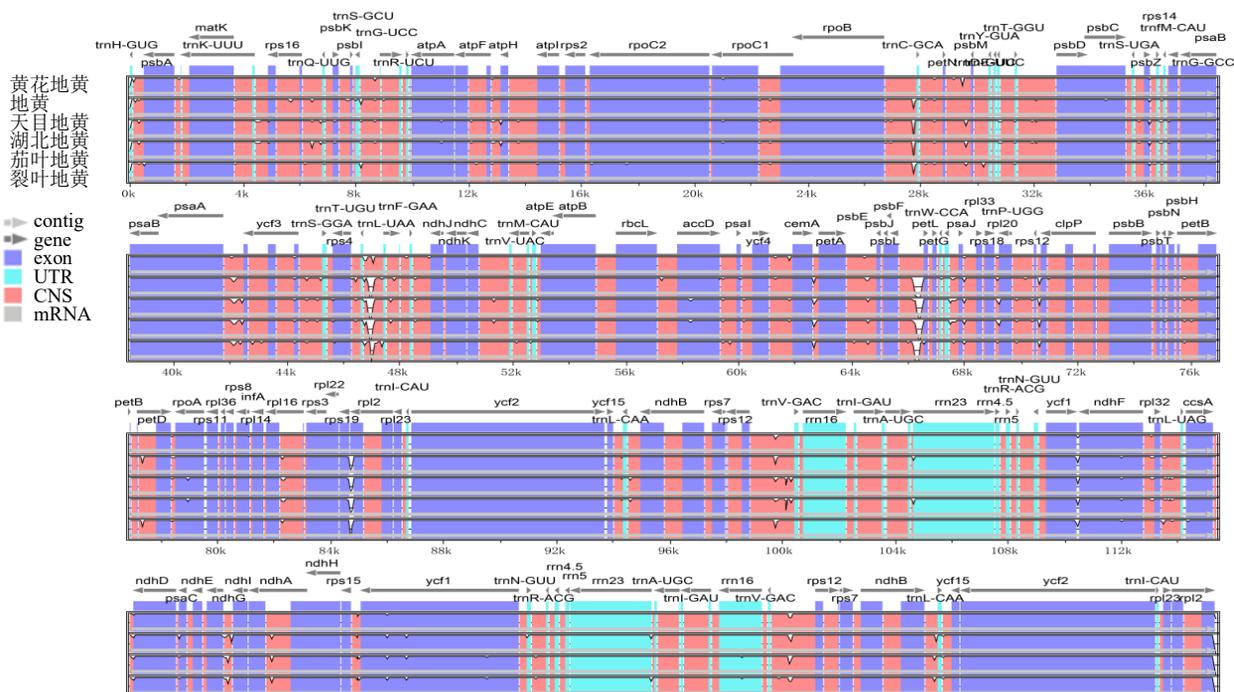


图 6 SC/IR 边界分析

Fig. 6 SC/IR boundary analysis



变异位点位于 IGS 内, 如 *rpl2*、*psbI*、*petN*、*psbZ*、*ycf1*、*ndhB*。

### 3.9 系统发育分析

为确定黄花地黄在地黄属的系统发育位置, 本研究选取地黄属 6 个物种: 黄花地黄、地黄、天目地黄、湖北地黄、茄叶地黄和裂叶地黄, 以及同科钟萼草属植物钟萼草, 肉苁蓉属肉苁蓉和黄花肉苁蓉, 大黄花属大黄花、光药大黄花构建系统发育树(图 8)。进化树主要分为 2 大支, 一支为大黄花和光药大黄花, 另一支为其余 9 个物种。第 2 大支又分为 2 小支, 一支为肉苁蓉属, 一支为地黄属和钟萼草。在地黄属中黄花地黄与地黄亲缘关系较近, 裂叶地黄和天目地黄遗传距离近, 而湖北地黄与茄叶地黄遗传距离短。

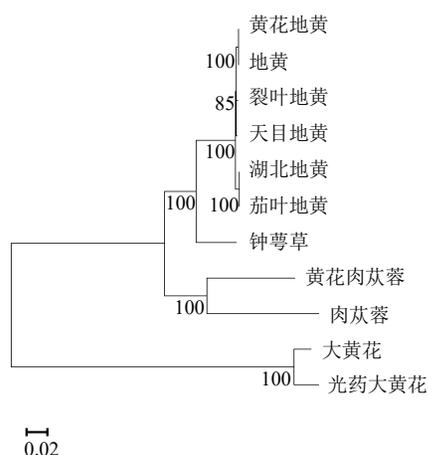


图 8 叶绿体基因组系统发育树

Fig. 8 Phylogenetic tree of chloroplast genome

### 4 讨论

叶绿体基因组是一个较保守的典型四分体结构序列, 基因之间相似性高, 差异主要源于基因间区。本研究首次完成黄花地黄叶绿体基因组的测序与解析。黄花地黄叶绿体基因组全长 153 789 bp, 为双链环状四分体, GC 含量为 37.92%, 其叶绿体基因组长度与地黄、天目地黄、湖北地黄、茄叶地黄、裂叶地黄相近。地黄属不同物种间在基因组大小和结构组成上表现出较高的保守性, 变异程度较小。

叶绿体基因组密码子使用偏好性与基因选择、遗传变异等有关, 同时也与物种间亲缘关系及生境状态密切相关<sup>[8]</sup>。有研究显示, 豆科的 27 个物种偏好使用的密码子中有 96.55% 以 U/A 结尾<sup>[15]</sup>。在橡胶树属物种中共鉴定出 30 个高频密码子, 其中以

U 或 A 结尾的密码子均占 42.28%, 以 G 结尾的密码子只有 1 个<sup>[16]</sup>。张菲菲对 2 个油茶属植物进行 RSCU 分析, 结果显示小果油茶有 28 种高频密码子, 其中以 U 或 A 结尾的分别有 12、13 个, 以 G 或 C 结尾的分别有 3、0 个; 越南油茶有 30 种高频密码子, 其中以 U 或 A 结尾的分别有 14、13 个, 以 G 或 C 结尾的分别有 3、0 个<sup>[17]</sup>。因此这两物种密码子偏好均以 A 或 U 结尾。黄花地黄有 28 个密码子 RSCU 值大于 1, 34 个 RSCU 小于 1, 有 2 个密码子 UGG、AUG 的密码子偏好性等于 1, 黄花地黄叶绿体基因组密码子偏好以 A/U 结尾。地黄、天目地黄、湖北地黄、裂叶地黄、茄叶地黄偏好使用的密码子均为 30 个, 其中仅有 1 个亮氨酸 Leu 的密码子以 G 结尾, 其余密码子均为 A/U 结尾。地黄属 6 个物种的偏好使用的密码子大多数以 A/U 结尾。

中性绘图分析、ENC-plot 分析和 PR2-plot 分析是评价密码子使用偏性的 3 个重要指标。在点地梅属中, 中性绘图结果显示, 点地梅属 10 个物种中叶绿体基因都分布于对角线上方, 且大部分基因离对角线较远, 各物种的回归系数都较小, 在 0.332 6~0.577 1。ENC-plot 图也显示还有许多基因分布在标准曲线上或下方附近, 比值频数分布图也显示有大量基因分布在 -0.05~0.05。因此, 在密码子偏性中, 突变压力也起了很大作用<sup>[10]</sup>。在黄芩属中 ENC 的值大于 35 的占多数, 表明多数密码子的偏性较弱, 作为密码子偏好性图像可视化的分析手段, ENC-plot 分析从 16 种黄芩属药用植物 ENC 和 GC<sub>3</sub> 的分布关系中发现, 自然选择压力对密码子使用偏好性的影响较大, 中性图分析的结果佐证了这一结论<sup>[8]</sup>。本研究中黄花地黄及同属 5 个物种的密码子偏好性分析显示, 中性绘图表明地黄属受自然选择压力为 82.6%~89.0%, 在 ENC-plot 分析中绝大部分基因的实际 ENC 值小于预期 ENC 值, 说明黄花地黄及同属 5 个物种叶绿体基因组密码子偏好性受自然选择影响, PR2-plot 分析也印证了其受选择的影响。

IR 区域的收缩和扩张会显著影响基因组的大小<sup>[19]</sup>, 在地黄与黄花地黄中, LSC/IRb 连接处位于 *rps19* 和 *rpl2* 之间, 而其余 4 个地黄属物种该边界全位于 *rpl2* 区域。地黄与黄花地黄的 *rpl2* 完全位于 IRb 区域, 而其他物种该基因在 IRb 区域的扩展长度可达 1 421~1 440 bp。地黄属物种的 SSC/IRb 边界位于 *ycf1* 区域, 且 *ndhF* 均完全位于 SSC 区域。

6个物种的SSC/IRa连接处均位于 $ycf1$ 基因中。黄花地黄和地黄LSC/IRa接合处分别位于 $rpl2$ 和 $trnH$ 基因重叠、 $rpl2$ 和 $psbA$ 基因重叠,而天目地黄、湖北地黄、裂叶地黄和茄叶地黄该连接处位于 $rpl2$ 基因内。

地黄属植物的SSC/IRb边界与SSC/IRa交界处,与无患子科植物基本一致均位于 $ycf1$ 基因中<sup>[20]</sup>,类似这种分布情况在美人蕉属的3个物种中也得到了印证<sup>[21]</sup>。地黄属物种的LSC/IRb边界分布比较特殊,大部分位于 $rpl2$ 基因中,而其他植物则一般位于 $rsp19$ 基因中,例如凤仙花属<sup>[22]</sup>、天南星属<sup>[23]</sup>。本属物种的LSC/IRa接合处位于 $rpl2$ 基因中,这与天南星属的位置相同,而悬钩子属该边界位于 $trnH$ 基因中<sup>[24]</sup>,槭属物种的该边界也位于 $trnH$ 基因<sup>[25]</sup>。由此可知,地黄属物种的JLB、JLA边界较为特殊与其他植物分布不同,而JSB、JSA边界与其他植物分布近似。

虽然DNA条形码已经普遍应用于药用植物的鉴别,ITS2、 $psbA$ 等序列作为DNA条形码十分成熟<sup>[23]</sup>,但基于黄花地黄叶绿体基因组与其同属植物的鉴别研究较少。对比叶绿体基因组序列可知, $rpl2$ 、 $psbI$ 、 $petN$ 、 $psbZ$ 、 $ycf1$ 、 $ndhB$ 可作为鉴别同属不同种地黄的候选序列。Xu等<sup>[24]</sup>以卵叶远志为参照序列,发现远志属9种植物的可变区域为 $ycf1$ 、 $ycf2$ 、 $ycf3$ 、 $ndhF$ 和 $ndhD$ 。由此可知, $rpl2$ 、 $psbI$ 、 $petN$ 、 $psbZ$ 、 $ycf1$ 、 $ndhB$ 序列在地黄属表现出高度的可变性,可作为DNA条形码的候选序列。

本研究首次解析了黄花地黄叶绿体基因组,综合分析了地黄属叶绿体基因组密码子使用偏性,结构和特征差异,筛选出了地黄属叶绿体基因组的差异片段,探讨了地黄属及其同科植物的系统进化关系,结果不仅丰富了地黄属叶绿体基因组数据信息,还为特异性DNA条形码的筛选及系统发育关系研究提供了依据。

**利益冲突** 所有作者均声明不存在利益冲突

#### 参考文献

- [1] 周添慧,周佳,吕思霖,等.基于“黑如漆、甘如饴”的地黄炮制过程质量评价研究[J].中草药,2025,56(7): 2344-2355.
- [2] 商岚清,程鑫雨,单鸿哲,等.地黄及其有效成分干预膝骨关节炎的作用机制研究进展[J].中草药,2025,56(1): 349-359.
- [3] 刘彦飞,史国茹,王欣,等.天目地黄化学成分研究[J].中草药,2016,47(11): 45-56.
- [4] 苗春妍,李铭铭,左鑫,等.湖北地黄CRISPR/Cas9基因编辑体系的建立[J].植物学报,2023,58(6): 905-916.
- [5] 王闰予,李国卫,索彩仙,等.不同基原忍冬属药材指纹图谱的建立[J].中草药,2021,52(14): 4353-4361.
- [6] Yang M Q, Wang Y L, Bai X R, et al. *Rehmannia chrysantha* (Rehmanniaceae), a new species from Inner Mongolia, northern China [J]. *Phytotaxa*, 2016, 265(2): 177.
- [7] 徐荣,宋文政,乔振升,等.富民枳叶叶绿体基因组密码子偏好性分析[J].中草药,2024,55(20): 7093-7100.
- [8] 才让扎西,胡夏宇,刘玉萍,等.苦豆子叶绿体基因组密码子偏好性分析及系统发育研究[J].广西植物,2025,45(7): 1295-1312.
- [9] 陈锦豪,程文萍,高静,等.基于叶绿体基因组种间高变区序列的黄精属药用植物分子鉴定[J].药学学报,2025,60(5): 1543-1554.
- [10] 郭玉朋,史明伟,马虎,等.点地梅属10种植物叶绿体基因组密码子偏好性分析[J].中草药,2025,56(4): 1355-1365.
- [11] 张丹,杜晨晖,裴香萍,等.党参线粒体和叶绿体微卫星标记的开发及应用[J].中国实验方剂学杂志,2021,27(23): 153-162.
- [12] Zeng S Y, Zhou T, Han K, et al. The complete chloroplast genome sequences of six *Rehmannia* species [J]. *Genes*, 2017, 8(3): 103.
- [13] 赵艳林,陈渊,沈晓婷,等.宽叶十万错叶绿体基因组比较分析和系统进化关系研究[J].中草药,2025,56(5): 1731-1746.
- [14] 李俊霖,郭淑红,张强,等.蒙古黄芪叶绿体全基因组特征解析及亲缘性分析[J].中草药,2024,55(7): 2366-2374.
- [15] Yang R, Xue Y, He X F, et al. Comprehensive analysis of chloroplast genomes in leguminous forage species: Codon usage, phylogenetic relationships, and evolutionary insights [J]. *Agronomy*, 2025, 15(4): 69-89.
- [16] Yang Y, Liu X Y, He L X, et al. Comparative chloroplast genomics and codon usage bias analysis in *Hevea* genus [J]. *Genes*, 2025, 16(2): 201.
- [17] 张菲菲,朱艳,韩长志.小果油茶与越南油茶中叶绿体基因组密码子偏好性[J].北京林业大学学报,2025,47(4): 37-49.
- [18] 荣正普,王杰敏,裴林,等.黄芩属药用植物叶绿体基因组密码子使用偏好性分析[J].中草药,2025,56(1): 269-281.
- [19] Liu H Z, Ye H, Zhang N Y, et al. Comparative analyses of chloroplast genomes provide comprehensive insights into the adaptive evolution of *Paphiopedilum* (Orchidaceae)

- [J]. *Horticulturae*, 2022, 8(5): 78-89.
- [20] Li J R, Wang H X, Wang L X, *et al.* Comprehensive analysis of the complete chloroplast genome of the cultivated soapberry and phylogenetic relationships of Sapindaceae [J]. *Ind Crops Prod*, 2025, 228: 120952.
- [21] Sun L H, Liu J X, Liu F Y, *et al.* Complete chloroplast genome sequences of three *Canna* species: Genome characterization, comparative analyses, and phylogenetic relationships within Zingiberales [J]. *Curr Issues Mol Biol*, 2025, 47(4): 222.
- [22] Yong Q Q, Li M J, Li Z, *et al.* Complete chloroplast genomes of 13 species of the *Impatiens* genus for genomic features and phylogenetic relationships studies [J]. *Sci Rep*, 2025, 15: 4258.
- [23] Tamboli A S, Youn J S, Kadam S K, *et al.* Chloroplast genome of *Arisaema takesimensis*: Comparative genomics and phylogenetic insights into the *Arisaema* [J]. *Biochem Genet*, 2025, doi:10.1007/s10528-025-11082-7.
- [24] Xu Y F, Li Y Q, Chen Y Z, *et al.* Comparative analysis of complete chloroplast genomes of *Rubus* in China: Hypervariable regions and phylogenetic relationships [J]. *Genes*, 2024, 15(6): 716.
- [25] Kadam S K, Tamboli A S, Youn J S, *et al.* Decoding the chloroplast genome of Korean endemic plant *Acer okamotoanum*: Comparative genomics, phylogenetic insights, and potential for marker development [J]. *Mol Biotechnol*, 2025, doi:10.1007/s12033-025-01383-y.

[责任编辑 时圣明]