

百合属药用植物叶绿体基因组密码子偏好性及系统发育研究

代国娜, 尚明越, 王嘉乐, 郑加梅, 廖彬彬, 刘颖琳*, 段宝忠*

大理大学药学院, 云南 大理 671000

摘要:目的 探究百合属 *Lilium* 药用植物叶绿体基因组密码子使用偏性及影响因素, 为百合属药用植物育种提供理论参考。方法 使用 CodonW1.4.2、CUSP、SPSS 等软件分析 9 种百合属药用植物的密码子偏好性。结果 9 种百合属植物密码子不同位点的 GC 含量均小于 0.5, 且 $GC1 > GC2 > GC3$; 其有效密码子数值均大于 45。共筛选到 21 个最优密码子, 其中有 20 个以 A/U 碱基结尾。PR2-plot 分析、ENC-plot 分析和中性绘图分析显示, 百合属叶绿体基因组密码子偏好性主要受自然选择影响, 同时也受突变因素影响。叶绿体蛋白编码序列的系统发育树和相对同义密码子使用度值的聚类分析结果, 均支持渥丹 *Lilium concolor* 从钟花组独立归并到卷瓣组。结论 9 种百合属植物叶绿体基因组密码子偏好性较弱, 自然选择是影响密码子偏好性形成的主要因素, 渥丹归入卷瓣组更合理。

关键词: 百合属; 叶绿体基因组; 密码子偏好性; 最优密码子; 自然选择

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2024)11-3835-10

DOI: 10.7501/j.issn.0253-2670.2024.11.025

Codon bias and phylogenesis analysis of chloroplast genome in medicinal plants of *Lilium*

DAI Guona, SHANG Mingyue, WANG Jiale, ZHENG Jiamei, LIAO Binbin, LIU Yinglin, DUAN Baozhong

College of Pharmaceutical Science, Dali University, Dali 671000, China

Abstract: Objective To explore the codon usage preferences and influencing factors of chloroplast (cp) genes in medicinal plants of the genus *Lilium* L. and provide a theoretical reference for breeding medicinal plants in *Lilium*. **Methods** CodonW1.4.2, CUSP, and SPSS software were employed to analyze the codon preferences of nine medicinal plants in *Lilium*. **Results** All of the GC content of different codon positions in nine *Lilium* species was less than 50%, with $GC1 > GC2 > GC3$. The effective number of codons (ENC) in the cp genomes of *Lilium* species is all more significant than 45. A total of 21 optimal codons were identified. Among them, 20 concluded with A/U bases. The PR2 plot analysis, ENC plot analysis, and neutral plot analysis demonstrated that the codon usage bias in the cp genomes of *Lilium* species is primarily influenced by natural selection and mutational factors. Furthermore, the phylogenetic tree of cp protein-coding sequences and the cluster analysis of relative synonymous codon usage (RSCU) values supported that *Lilium concolor* should be separated from the Sect. *Lophophorum* and merged into the Sect. *Sinomartagon*. **Conclusion** The codon preferences in the cp genomes of nine plants within the *Lilium* exhibit a relatively weak bias. Natural selection is the main factor affecting codon preference and placing *L. concolor* in the Sect. *Sinomartagon* is a more plausible classification.

Key words: *Lilium* L.; chloroplast genome; codon preference; optimal codon; natural selection

百合属 *Lilium* L. 为重要的药用类群, 其中卷丹 *L. lancifolium* Thunb.、百合 *L. brownii* F. E. Brown var. *viridulum* Baker. 及细叶百合 *L. pumilum* DC. 为《中国药典》2020 年版百合药材来源^[1]。现代研究表明, 百合含有酚酸、甾体皂苷和多糖类等化学成分,

具有抗炎、抗抑郁、降血糖和免疫调节等多种药理活性^[2-4], 在药品和食品领域市场需求量极大^[5-6]。由于百合为多年生植物, 种植周期长, 在生长过程中易受黑腐病、灰霉病和软腐病的侵害^[7]; 目前生产上虽已培育了部分百合品种, 但抗病性强的优质种质仍

收稿日期: 2023-12-09

基金项目: 云南省生物医药重大专项 (202002AA100007); 云南省兴滇英才支持计划 (YNWR-QNBJ-2020251)

作者简介: 代国娜 (1998—), 硕士研究生, 研究方向为中药资源与鉴定。E-mail: 1278804003@qq.com

*通信作者: 刘颖琳, 讲师, 研究方向为中药资源及品质评价。Tel: (0872)2257401 E-mail: lyldaliedu@163.com

段宝忠, 教授, 研究方向为中药资源与鉴定。Tel: (0872)2257411 E-mail: bzduan@126.com

然缺乏,极大制约了百合产业的发展。采用现代技术,选育高产、抗病性强的百合品种,是百合产业发展的重大需求。

叶绿体是光合作用和细胞质遗传的重要器官,也是植物生长能量提供和次生代谢物合成的重要场所^[8]。叶绿体中的密码子参与 mRNA 翻译成蛋白质的过程^[9],其使用偏好性可体现物种进化和突变模式,与基因表达水平息息相关^[10]。根据密码子偏好性优化外源基因,对培育优良新品种,提高植物产量和品质具有重要意义^[11-12]。

近年来,随着测序技术和基因工程的发展,根据叶绿体密码子偏好性培育优良植物品种引起了科学界的关注,一些高产、抗病性强的玉米 *Zea mays* L.^[13]、大豆 *Glycine max* (L.) Merr.^[14-15] 和水稻 *Oryza sativa* L.^[16-17] 品种已经被选育。目前,有关百合属叶绿体基因组的研究主要集中在系统发育、遗传多样性和分子鉴定方面^[18-20],尚未见百合密码子偏好性的研究,一定程度限制了百合优良种质的选育。鉴于此,本研究基于高通量测序技术,对百合属植物叶绿体基因组的密码子偏好性进行了分析,并对影响密码子偏好性的因素进行了探讨,以期研究百合属药用植物叶绿体基因组进化机制、遗传育种、资源鉴定和外源表达基因构建提供理论支撑。

1 材料

6个植物样本种球购自四川、河北和江西,种植于大理大学种质资源圃,2022年开花期经大理大学段宝忠教授鉴定为百合属植物渥丹 *L. concolor* Salisb.、麝香百合 *L. longiflorum* Thunb.、滇百合 *L. bakerianum* Coll. et Hemsl.、细叶百合 *L. pumilum* DC.、卷丹 *L. lancifolium* Thunb. 和百合 *L. brownii* var. *viridulum* Baker.。此外,从 NCBI 数据库下载东北百合 *L. distichum* Nakai et Kamibayashi. (MK493296)、马塘百合 *L. matangense* J. M. Xu. (MN745201) 和青岛百合 *L. tsingtauense* Gilg. (NC027675) 的叶绿体基因组序列用于分析。

2 方法

2.1 叶绿体基因组测序

利用 Illumina Novaseq6000 高通量测序平台对样本叶绿体基因组进行测序,结果上传至 NCBI 数据库 (<https://www.ncbi.nlm.nih.gov/>),登录号为 OR413692、OR413693、OR400157~OR400160。根据注释结果,从叶绿体基因组中提取出蛋白质编码序列 (coding DNA sequence, CDS),并选择长度大

于 300 bp、不重复、起始密码子为 ATG 和终止密码子为 TAA、TAG、TGA 的 CDS 进行后续分析。

2.2 密码子偏好性分析

使用 CodonW1.4.2 软件和在线程序 EMBOSS 计算 RSCU、ENC、密码子数 (N)、总 GC 含量 (GCall) 和密码子第 1 位 (GC1)、第 2 位 (GC2)、第 3 位 (GC3) 的 GC 含量,并使用 SPSS 26.0、Microsoft Excel 软件对各参数进行整理和绘图。

2.3 最优密码子分析

根据 ENC 值大小,将蛋白编码基因从高到低排序,取前 10% 和后 10% 的基因分别建立高表达和低表达的数据库。利用 CodonW 1.4.2 软件计算库中数据的 RSCU 值和 Δ RSCU 值 (Δ RSCU = RSCU 高表达 - RSCU 低表达)。其中, RSCU > 1 且 Δ RSCU \geq 0.08 的密码子为最优密码子^[21]。

2.4 PR2-plot 偏倚分析

碱基奇偶偏好性分析 (PR2-bias plot analysis, PR2-plot),用于分析各密码子第 3 位碱基 A、T 和 G、C 之间的使用频率是否均衡^[22]。使用 Microsoft Excel 软件,以 $G3/(G3+C3)$ 为横坐标 (X), $A3/(A3+T3)$ 为纵坐标 (Y),绘制散点图,每个点代表一个独立的基因,图中心点的位置表明密码子无使用偏好性 ($A=T, G=C$),其余点与中心点的矢量距离表示偏倚程度和方向^[23]。

2.5 ENC-plot 分析

有效密码子数 (ENC-plot) 分析可直观反映基因密码子使用模式,用于衡量突变及自然选择对密码子偏好性形成的影响程度^[24]。以 GC3 为横坐标 (X), ENC 为纵坐标 (Y) 绘制散点图,并根据公式绘制 ENC 期望值标准曲线,当基因的 ENC 值接近或分布在标准曲线上,表明该基因的密码子偏好性主要受突变压力影响,反之则主要受选择压力的影响^[25]。为更准确地衡量 ENC 理论值 (ENC_{exp}) 和实际值 (ENC_{Cobs}) 的差异大小,计算每个基因 ENC 比值,并做频率分布图。

$$ENC = 2 + GC3 + 29/[GC3^2 + (1 - GC3)^2]$$

$$ENC \text{ 比值} = (ENC_{exp} - ENC_{Cobs}) / ENC_{exp}$$

2.6 中性绘图分析

中性绘图分析是通过密码子 3 个位点碱基组成的相关性来判断影响密码子偏性因素的常用方法^[26]。同义密码子的突变位点通常为第 3 碱基,而非同义密码子的突变位点通常为第 1 或第 2 碱基^[27]。以 GC3 为横坐标 (X), GC12 (GC1 和 GC2 的平均值)

为纵坐标 (Y), 绘制散点图并进行回归分析, 得到回归系数、拟合函数和相关系数。当回归系数接近 1 时, 表明 GC1、GC2 和 GC3 的碱基组成相似, 密码子偏好性受突变的影响更多; 反之, 则表明 3 个位置碱基组成差异大, 密码子偏好性受自然选择的影响更多^[28]。

2.7 系统发育分析

从 NCBI 数据库中下载嘉兰 *Gloriosa superba* L. (NC030065) 和秋水仙 *Colchicum autumnale* L. (NC030064) 叶绿体基因组, 并提取蛋白编码序列, 以及 9 个百合属物种的蛋白编码序列, 共计 11 个物种用于系统发育分析。使用 Mafft v 7.49 对共有蛋白编码序列进行序列比对, 分别使用 RAxML (v.8.2.11) 和 MEGA v7.0 软件构建最大似然法 (maximum likelihood, ML) 和邻接法 (neighbor-joining, NJ) 系统发育树^[29]。

基于 RSCU 的聚类分析, 将 11 个物种的蛋白

编码序列按“2.1”项下的标准过滤后, 以每个物种的 RSCU 为变量 (剔除 AUG、UGG、UAA、UAG、UGA), 采用 SPSS 26.0 中进行聚类分析。

3 结果与分析

3.1 叶绿体基因的密码子偏好性分析

百合属植物叶绿体基因组中含有 82~84 条 CDS, 按“2.1”项下的标准筛选后, 每个物种各获得 48 条 CDS 用于密码子偏好性分析。如表 1 所示, 不同物种之间, 密码子的 GC1、GC2、GC3 和 GCall 含量较接近, 均小于 0.5, 且 GC1>GC2>GC3, 表明 CG 碱基在不同的位点中分布不均, 偏向在第 1、2 位出现, 9 种百合属植物密码子富含 A/U 碱基。此外, T3s 和 A3s 明显高于 G3s 和 C3s, 表明百合叶绿体编码基因序列组更偏向于以 A/T 结尾。分析显示所有物种的 ENC 值均大于 45, 提示百合属植物的密码子偏好性相对较弱。

表 1 百合属植物叶绿体基因组密码子偏好性基本参数

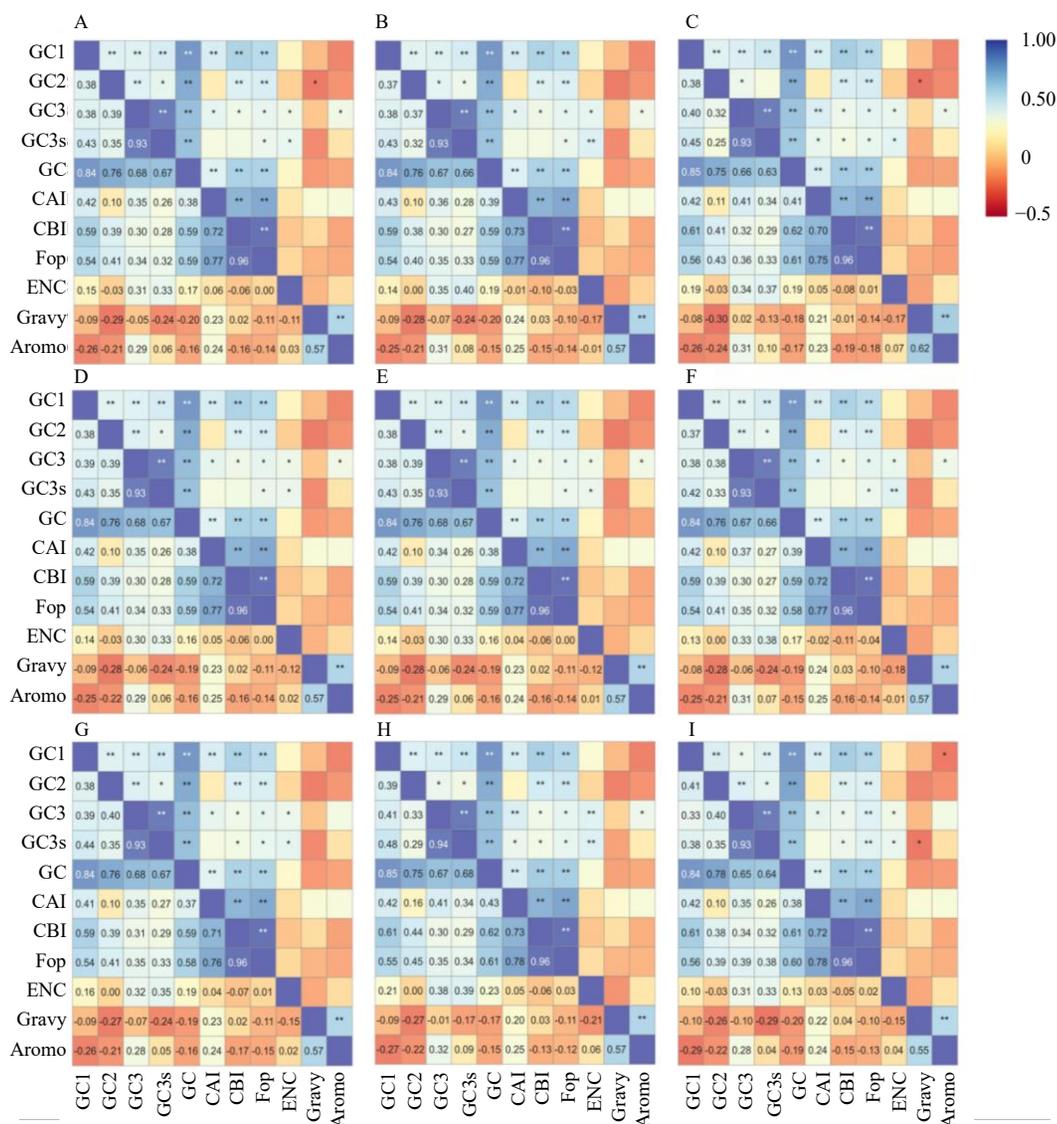
Table 1 Basic parameters of codon usage bias in chloroplast genome of *Lilium*

物种	GC1	GC2	GC3	GCall	A3s	T3s	G3s	C3s	ENC
渥丹	0.473 7	0.394 3	0.281 7	0.383 2	0.431 9	0.477 5	0.163 0	0.160 3	46.54
麝香百合	0.474 1	0.394 7	0.282 8	0.383 8	0.431 1	0.476 3	0.163 6	0.161 4	46.66
滇百合	0.472 7	0.394 1	0.282 8	0.383 2	0.433 6	0.475 2	0.162 3	0.162 4	46.67
细叶百合	0.473 8	0.394 6	0.281 6	0.383 3	0.431 8	0.477 5	0.162 9	0.160 3	46.54
卷丹	0.473 9	0.394 6	0.281 6	0.383 4	0.431 7	0.477 6	0.162 9	0.160 3	46.54
百合	0.474 0	0.394 7	0.282 6	0.383 8	0.431 0	0.476 7	0.163 3	0.161 4	46.64
东北百合	0.473 6	0.394 4	0.281 4	0.383 1	0.432 1	0.477 6	0.163 1	0.160 0	46.53
马塘百合	0.473 2	0.394 0	0.282 8	0.383 3	0.432 6	0.475 2	0.163 1	0.161 8	46.69
青岛百合	0.471 8	0.392 5	0.282 6	0.382 3	0.432 9	0.476 8	0.164 3	0.160 7	46.67

对 GC1、GC2、GC3、密码子适应指数 (codon adaptation index, CAI)、密码子偏好指数 (Codon bias index, CBI) 和密码子使用频率 (frequency of optional codons, FOP) 等参数进行相关性分析。结果显示 (图 1), 所有物种的 GC1、GC2 和 GC3 之间呈极显著或显著相关, 表明百合属植物密码子第 1、2 和 3 位的碱基组成相似。CAI、CBI 和 FOP 值与 GC、GC3 和 GC1 呈显著或极显著相关, 提示 GC 含量, 尤其是密码子第 1、3 位碱基的 GC 含量对百合属叶绿体基因组的表达有较大影响。ENC 值与 GC1 和 GC2 的相关性不显著, 与 GC3 和 GC3s 显著或极显著相关。此外, 渥丹和滇百合的蛋白质疏水性 (general average hydropathicity, Gravy) 与 GC2 呈显著相关性, 青岛百合的 Gravy 与 GC3s 呈显著相关性, 表明密码子第

2、3 位碱基组成对 Gravy 有一定的影响。除东北百合和青岛百合外, 其它 7 种百合的蛋白质芳香性 (aromaticity, Aromo) 与 GC3 呈显著相关性, 表明密码子第 3 位碱基组成对 Aromo 有一定的影响。

RSCU 是衡量密码子偏好性的重要参数, 当 RSCU>1 时, 表明该密码子偏好性强, 为高频密码子, 反之则偏好性弱, 为低频密码子^[30]。RSCU 值分析结果显示, 9 个物种的 RSCU 值较接近 (图 2), 所有物种 RSCU>1 的高频密码子均有 29 个, 占密码子总数的 49.15%, 其中以 A 结尾有 12 个, U 结尾 16 个, G 结尾 1 个; RSCU<1 的低频密码子有 30 个, 占比 50.85%, 其中以 C 结尾有 16 个, G 结尾 12 个, A 结尾 2 个。上述结果表明, 百合属植物高频密码子更偏向 A/U 结尾, 低频密码子更偏向以



A-渥丹; B-麝香百合; C-滇百合; D-细叶百合; E-卷丹; F-百合; G-东北百合; H-马塘百合; I-青岛百合; 下图同。
A-*L. concolor*; B-*L. longiflorum*; C-*L. bakerianum*; D-*L. pumilum*; E-*L. lancifolium*; F-*L. brownii* var. *viridulum*; G-*L. distichum*; H-*L. matangense*; I-*L. tsingtauense*.

图 1 百合属叶绿体基因组相关参数的相关性分析

Fig. 1 Correlation analysis of chloroplast genome related parameters in *Lilium*

G/C 结尾。

3.2 最优密码子分析

9种百合属植物的最优密码子如表2所示,在渥丹、麝香百合、滇百合、细叶百合、卷丹、百合、东北百合、马塘百合和青岛百合分别鉴定出11、12、12、11、12、14、12、15和14个最优密码子。除UUG外,其余均以A/U结尾,表明百合属最优密码子偏好使用A/U碱基结尾。此外,6个最优密码子(CGU、UCU、UUA、GCU、AAA、GGU)为9种植物共有,其中 Δ RSCU大于0.5的有2个,分别为编码精氨酸的CGU和丝氨酸的UCU。麝香百合和滇百合各有1个特有的最优密码子,分别是编码

精氨酸的AGA和丝氨酸UCA;马塘百合有3个特有的最优密码子,分别是编码异亮氨酸的AUU、天冬氨酸的GAU和组氨酸的CAU;渥丹和细叶百合中的最优密码子相同。

3.3 PR2-plot 偏倚分析

在突变压力的单独作用下,突变的随机性使密码子第3位碱基为A/T或C/G的概率相等,而当受到自然选择压力影响时,A/T或G/C的使用频率会出现不均等的现象。如图3所示,9种百合属植物的叶绿体编码基因不均匀的分布于4个区域内,大多远离中心,提示密码子第3位碱基的使用存在偏好性,自然选择是影响偏好性的主要因素。此外,

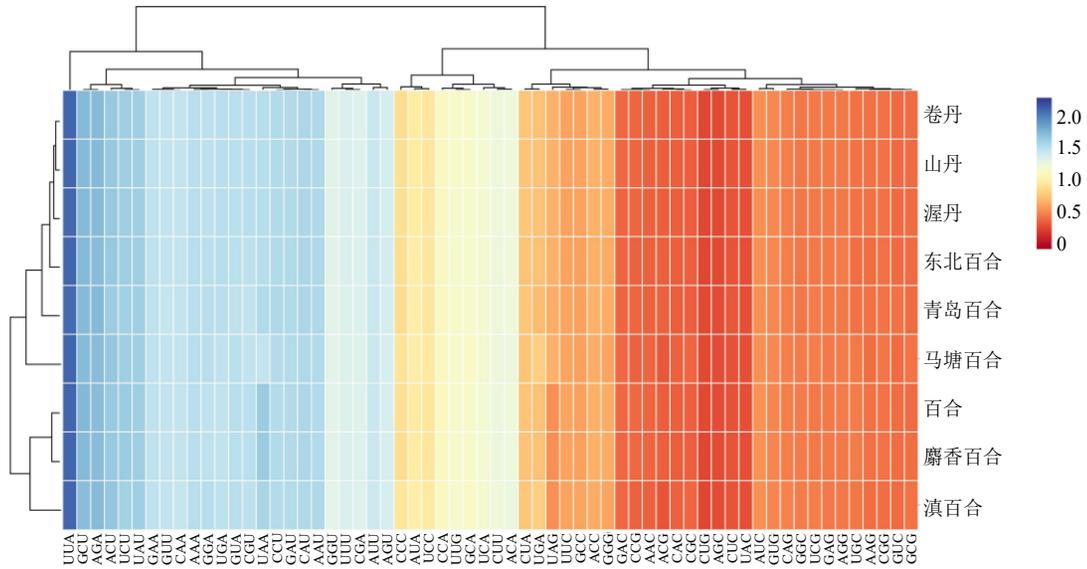


图 2 各物种叶绿体基因组 RSCU 值

Fig. 2 Chloroplast genome RSCU values of each specie

表 2 百合属植物叶绿体基因组最优密码子

Table 2 Optimal codons of chloroplast genomes in *Lilium*

氨基酸	密码子								
	渥丹	麝香百合	滇百合	山丹	卷丹	百合	东北百合	马塘百合	青岛百合
Arg	CGU*	CGU***							
	-	AGA*	-	-	-	-	-	-	-
Ser	UCU*	UCU***							
	AGU*	-	-	AGU*	AGU*	AGU*	AGU*	AGU*	AGU*
	-	-	UCA*	-	-	-	-	-	-
Thr	ACA*	ACA**	ACA*	ACA***	ACA**	ACA**	ACA**	-	ACA**
	-	ACU*	ACU*	-	-	-	-	ACU***	ACU*
Leu	UUA*	UUA**	UUA**	UUA*	UUA*	UUA*	UUA*	UUA*	UUA*
	UUG*	-	-	UUG**	UUG*	UUG**	UUG*	UUG*	UUG**
Ala	GCU*	GCU***	GCU**	GCU**	GCU*	GCU*	GCU*	GCU***	GCU*
Pro	CCU*	-	-	CCU*	CCU**	-	CCU**	-	CCU*
Lys	AAA*	AAA*	AAA**	AAA*	AAA*	AAA*	AAA*	AAA***	AAA*
Val	GUA*	-	-	GUA*	GUA*	GUA*	GUA*	GUA*	GUA**
	-	GUU**	GUU*	-	GUU*	GUU*	GUU*	GUU*	GUU*
Gly	GGU*	GGU*	GGU*	GGU*	GGU**	GGU*	GGU**	GGU**	GGU*
Gln	-	CAA*	CAA*	-	-	CAA*	-	-	-
Asn	-	AAU*	AAU*	-	-	AAU*	-	-	AAU*
Cys	-	-	-	-	-	UGU*	-	UGU*	-
Ile	-	-	-	-	-	-	-	AUU*	-
Asp	-	-	-	-	-	-	-	GAU*	-
His	-	-	-	-	-	-	-	CAU*	-
合计	11	12	12	11	12	14	12	15	14

*表示 $\Delta RSCU \geq 0.08$, **表示 $\Delta RSCU \geq 0.3$, ***表示 $\Delta RSCU \geq 0.5$ 。

* $\Delta RSCU \geq 0.08$; ** $\Delta RSCU \geq 0.3$, *** $\Delta RSCU \geq 0.5$ 。

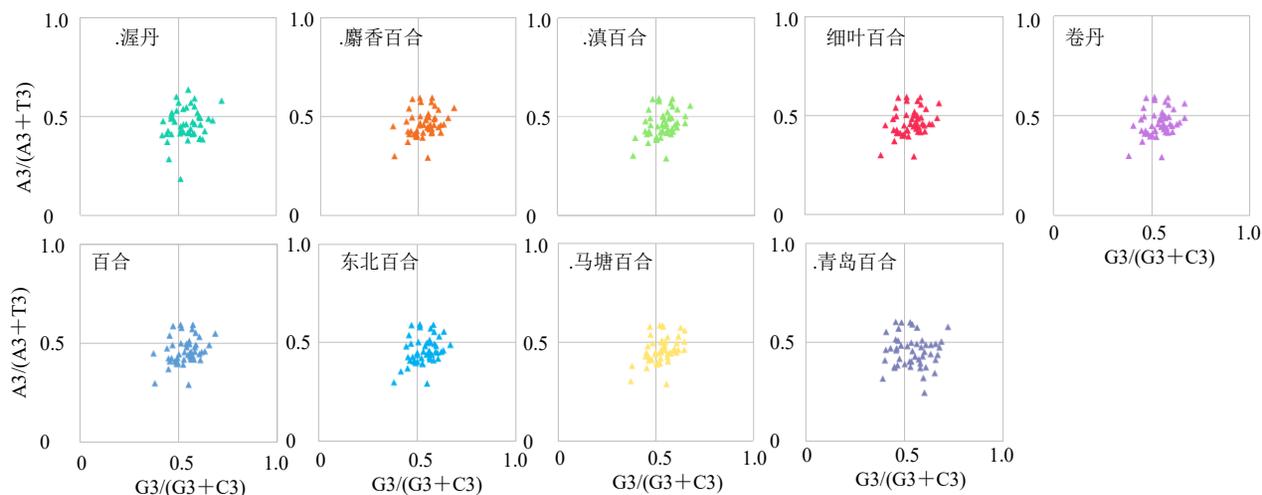


图3 PR2-plot 分析

Fig. 3 Analysis of PR2-plot

大部分基因分布于右下角,表明第3位碱基的使用频率为 $T > A, G > C$,即第3位密码子偏好使用 T/G 碱基。

3.4 ENC-plot 分析

ENC-plot 分析能表示 ENCexp 和 ENCobs 差异的大小,并反映突变或选择压力对密码子使用偏好性的影响^[31]。如图4所示,9个百合属植物叶绿体编码基因的 ENC-plot 图相似,大多数基因位于标准曲线下方,即大部分基因的 ENCobs 与 ENCexp 的偏差较大,表明百合属植物叶绿体基因组密码子使用偏好性受突变的影响较小,自然选择的影响更大。基因 ENC 比值频数分析的结果也支持上述观点,如图5可所示,ENC 比值分布在 $-0.05 \sim 0.05$ 的

基因有 11~14 个,占比 22.92%~29.17%,这些基因的 ENCobs 与 ENCexp 较为接近;比值在 $-0.05 \sim 0.05$ 段外的基因有 34~37 个,占比 70.83%~77.08%,这些基因的 ENCobs 与 ENCexp 的偏差较大,偏好性受突变的影响较小,自然选择的影响更大。

3.5 中性绘图分析

中性绘图分析可揭示 GC12 与 GC3 之间的相关性,并反映自然选择和突变压力对密码子偏性的影响程度^[32]。如图6所示,百合属植物密码子的 GC12 值在 0.34~0.56,GC3 值在 0.21~0.35,相关性系数为 0.430~0.473,表明 GC12 与 GC3 的相关性不显著,二者的进化模式存在一定差异。

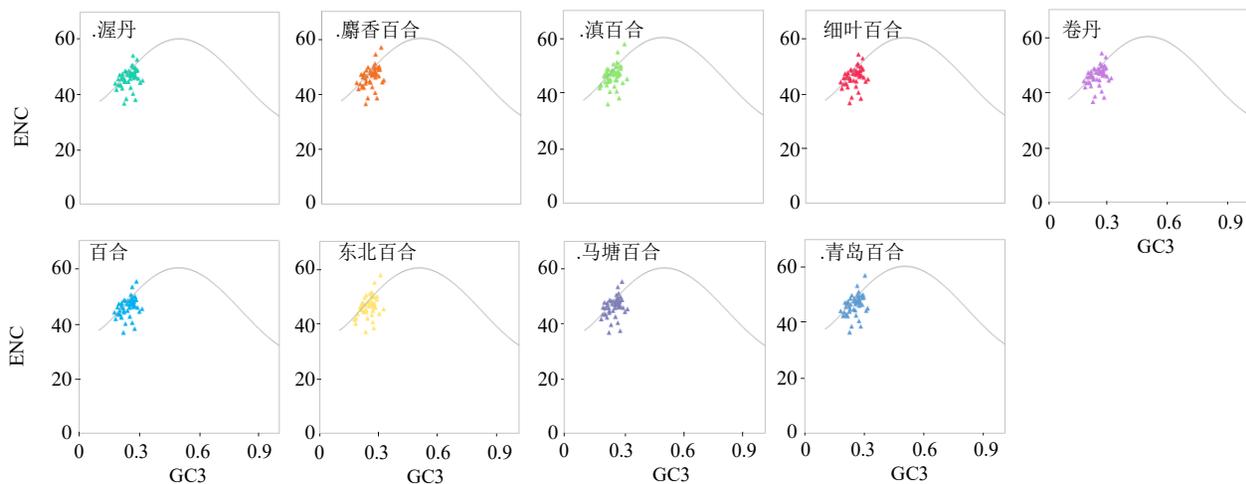


图4 ENC-plot 分析

Fig. 4 Analysis of ENC-plot

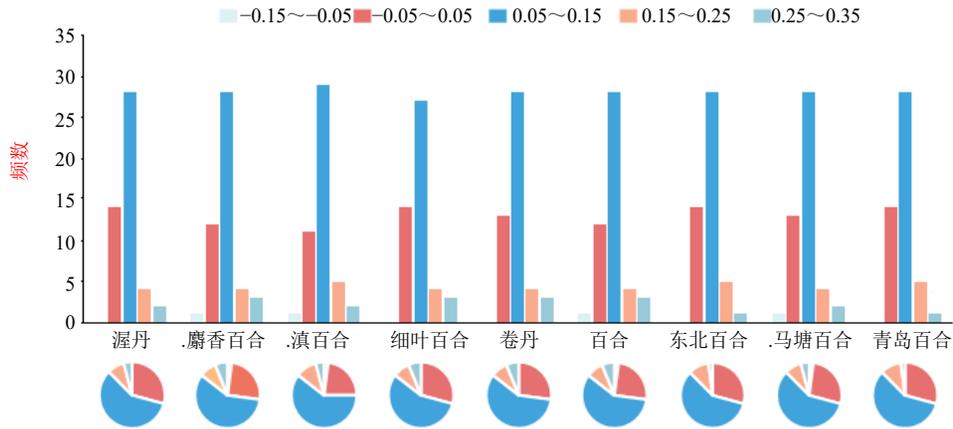


图5 ENC比值频率分布

Fig. 5 Frequency distribution of ENC ratio

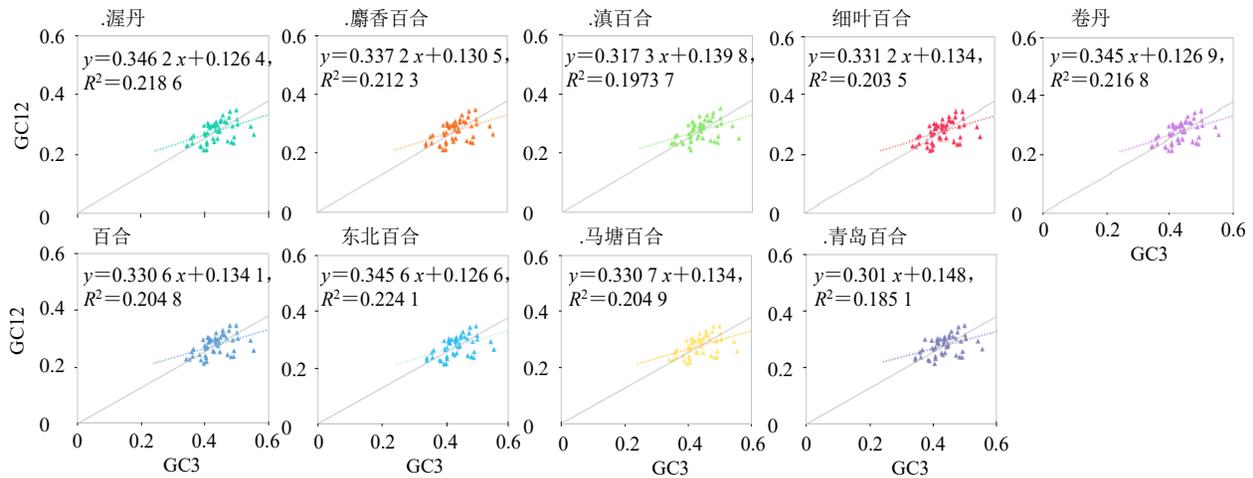


图6 中性绘图分析

Fig. 6 Analysis of neutrality plot

此外, 回归系数显示, 突变压力对密码子偏性的影响程度占 30.10%~34.62%, 自然选择占 69.90%~65.38%, 表明百合属叶绿体基因组密码子使用偏好性受突变压力的影响较小, 受自然选择的影响较大, 这与 ENC-plot 分析结果一致。

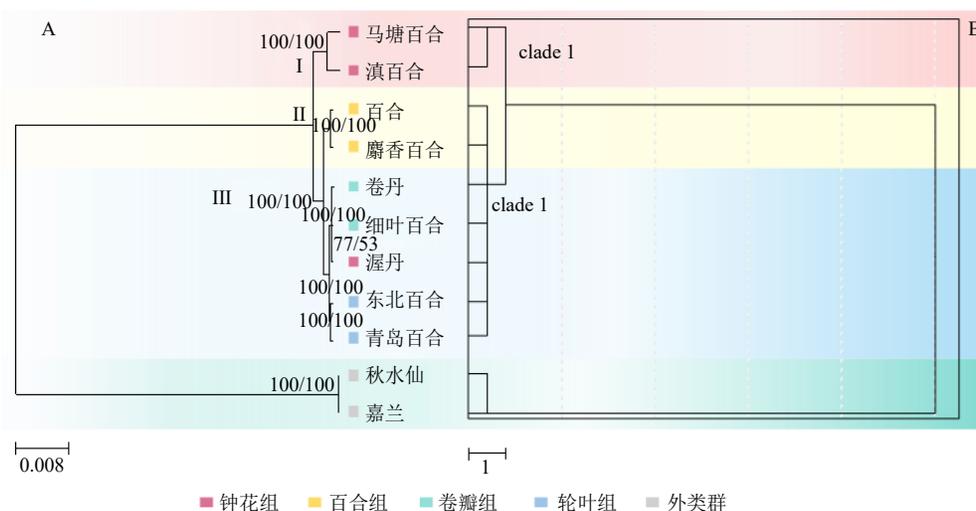
3.6 系统发育分析

基于百合属、秋水仙属和嘉兰属物种的共有 CDS 构建了 NJ 和 ML 树 (图 7), 发现它们的拓扑结构相似, 且具有较高的支持率。系统发育树中, 百合属物种主要被分成 3 个分支, 支持率均为 100%。分支 I 由马塘百合和滇百合组成; 分支 II 由百合和麝香百合组成; 分支 III 由卷丹、细叶百合、渥丹、东北百合和青岛百合组成。值得注意的是, 分支 III 中, 来自钟花组的渥丹与卷瓣组的细叶百合和卷丹构成了一个单系群, 支持率 100%, 并与轮叶组的东北百合和青岛百合形成了姊妹关系 (支持率

100%), 表明隶属于钟花组的渥丹与卷瓣组的关系较近。此外, RSCU 值聚类分析的结果与 CDS 系统发育树相比, 存在一定的差异, 他将百合属分为了 2 大类, 第一类包括马塘百合和滇百合, 第二类包括百合、麝香百合、卷丹、细叶百合、渥丹、东北百合和青岛百合。值得注意的是, 2 个树图均支持渥丹与卷瓣组关系更近, 这一观点与中国植物志将其归为钟花组的观点不一致。综上, CDS 系统发育树和 RSCU 值聚类分析均支持渥丹归入卷瓣组, 百合属叶绿体基因组密码子 RSCU 值可用于物种的系统发育探讨。

4 讨论

生物体在长期进化过程中, 往往会形成适应自身基因组环境的密码子使用模式^[33]。本研究发现, 百合属植物叶绿体基因组密码子偏好以 A/U 碱基结尾, 这一结论与先前对单子叶植物的研究相反,



A-基于 48 条共有 CDS 构建的 ML/NJ 树；B-基于密码子 RSCU 值聚类分析。
A-ML/NJ tree based on 48 shared CDS; B-Cluster analysis based on codon RSCU values.

图 7 百合属物种系统发育分析
Fig. 7 Phylogenetic analysis of *Lilium* species

即单子叶植物常以 G/C 结尾，而双子叶植物常以 A/U 结尾^[12,32,34]，形成这种差异的原因可能是由于不同科/属的单子叶植物的叶绿体基因组进化模式存在一定的差异，将来还需要通过大量的样本采集和测序，对单子叶和双子叶植物的密码子偏好性开展进一步的研究。ENC 值是衡量密码子偏好性的重要指标之一，当 ENC 值小于 35 时，密码子偏好性较强^[35]。百合属植物叶绿体基因的 ENC 值介于 46.54~46.69，表明它们的密码子偏好性较弱，与金银花 *Lonicera japonica* Thunb.^[36]、藿香 *Agastache rugosa* (Fisch. et Mey.) O. Ktze.^[37] 和北柴胡 *Bupleurum chinense* DC.^[38] 等大部分植物叶绿体基因组密码子的偏好性结果相似。

此外，中性绘图、ENC-plot 和 PR2-plot 分析的结果均表明，百合属叶绿体基因组密码子使用偏好性以自然选择为主，同时也受突变压力作用，与山楂属 *Crataegus* L.^[25]、绞股蓝属 *Gynostemma* Blume.^[39] 和鹅膏菌属 *Amanita* L.^[40] 密码子偏好性结果一致，与薏苡 *Coix lacryma-jobi* L.^[41]、橄榄 *Canarium album* (Lour.) DC.^[42] 等密码子偏好性受突变压力影响结论不一致，提示不同物种间密码子偏好性的主要影响因素有差异。最优密码子分析表明，渥丹和细叶百合的最优密码子相同，表明它们的密码子使用偏好性相似。之前的研究认为物种之间的亲缘关系越近，其密码子偏好性越相似^[43]，因此渥丹和细叶百合的亲缘关系较近。另外 9 种百合

属植物的最优密码子的组成有一定差异，但均偏向 A/U 结尾，与豇豆 *Vigna unguiculata* (L.) Walp.^[43] 和药用大黄 *Rheum officinale* Baill.^[44] 等类似。值得注意的是，系统发育分析发现，基于 CDS 构建的 ML 树、NJ 树与 RSCU 值聚类分析的结果均支持将渥丹从钟花组中独立，归并到卷瓣组中，这一结论与基于百合属叶绿体基因组^[45]、*rbcl*^[46] 和 *matK* 基因^[47] 系统发育分析结果一致，表明百合属密码子偏好性与物种的系统发育关系密切。此外，在 CDS 系统发育树中百合属被分为 3 个分支，而 RSCU 聚类分析中则被分为 2 个分支，其原因可能是 RSCU 聚类分析缺少了部分遗传信息，如无偏好的密码子和位点突变特征等^[34]。本研究首次探讨了 9 种百合属植物的叶绿体基因组密码偏好性及影响因素，为后续百合属叶绿体基因组外源的优化，以及高产、抗病性强的百合新品种培育提供了参考。

利益冲突 所有作者均声明不存在利益冲突

参考文献

[1] 粟倩, 吴萍, 夏伯候, 等. 百合化学成分及药理活性研究进展 [J]. 中国药学杂志, 2021, 56(11): 875-882.
[2] Zhou J, An R F, Huang X F. Genus *Lilium*: A review on traditional uses, phytochemistry and pharmacology [J]. *J Ethnopharmacol*, 2021, 270: 113852.
[3] 金唯唯, 金华燕, 费状丽, 等. 基于 HPLC 指纹图谱结合化学模式识别辨析补中益气产品质量标志物 [J]. 中草药, 2023, 54(7): 2114-2126.
[4] 何丹, 张海潮, 李世慧, 等. 百合化学成分、药理作用

- 及质量标志物的预测分析 [J]. 中华中医药学刊, 2022, 40(12): 205-212.
- [5] Wang P Y, Li J, Attia F A K, *et al.* A critical review on chemical constituents and pharmacological effects of *Lilium* [J]. *Food Sci Hum Wellness*, 2019, 8(4): 330-336.
- [6] 金月, 谢静, 辛二旦, 等. 基于 HPLC 指纹图谱和多成分定量评价不同品种百合鲜品质量 [J]. 中国野生植物资源, 2023, 42(5): 60-67.
- [7] 倪典墨, 李冉郡, 郭志静, 等. 百合科植物 *psbA* 基因密码子偏好性分析 [J]. 生物学杂志, 2020, 37(6): 68-72.
- [8] 叶友菊, 倪州献, 白天道, 等. 马尾松叶绿体基因组密码子偏好性分析 [J]. 基因组学与应用生物学, 2018, 37(10): 4464-4471.
- [9] 原晓龙, 李云琴, 张劲峰, 等. 乳油木叶绿体基因组密码子偏好性分析 [J]. 分子植物育种, 2020, 18(17): 5658-5664.
- [10] 吴巧雯. 透明颤菌血红蛋白基因的密码子优化、人工合成及其在烟草中的表达 [D]. 北京: 中国农业科学院, 2002.
- [11] 段祥光, 于涛, 曹萌, 等. 樟叶槭叶绿体基因组序列密码子偏性分析 [J]. 西北林学院学报, 2023, 38(1): 102-107.
- [12] Wang Y Z, Jiang D C, Guo K, *et al.* Comparative analysis of codon usage patterns in chloroplast genomes of ten *Epimedium* species [J]. *BMC Genom Data*, 2023, 24(1): 3.
- [13] 李圣彦, 郎志宏, 朱莉, 等. 利用密码子优化提高 *Bt cryIAh* 基因在转基因玉米(*Zea mays* L.)中的表达 [J]. 中国农业科技导报, 2011, 13(6): 20-26.
- [14] 寇莹莹, 宋英今, 杨少辉, 等. 植酸酶 *phyA* 基因的密码子优化及其在大豆中的表达 [J]. 作物学报, 2016, 42(12): 1798-1804.
- [15] 贾雪琦, 左永春. 大豆 ω -3 和 ω -6 脂肪酸脱氢酶基因密码子优化 [J]. 生物信息学, 2018, 16(1): 29-35.
- [16] 秦冠男, 翁绿水, 肖国樱. 水稻密码子优化基因 *Mat#* 增强草铵膦抗性 [J]. 农业生物技术学报, 2016, 24(10): 1457-1465.
- [17] 周宗梁. 水稻中 *cryIAh1* 基因密码子优化的研究 [D]. 哈尔滨: 东北农业大学, 2012.
- [18] 赖山潘, 沈呈娟, 周颂东, 等. 基于 3 种基因序列对百合属 *Leucolirion* 组的系统发育与种间关系研究 [J]. 西北植物学报, 2016, 36(8): 1541-1550.
- [19] 毕彧. 百合属的比较叶绿体基因组学研究 [D]. 长春: 吉林农业大学, 2017.
- [20] 陈名红, 熊华斌, 李成云. 分子标记在百合属植物遗传多样性研究中的应用 [J]. 生物技术通报, 2011(12): 65-69.
- [21] 舒军霞, 杨林, 周涛, 等. 刺柏属 4 种药用植物叶绿体基因组密码子偏好性分析 [J]. 中草药, 2022, 53(23): 7507-7515.
- [22] 赵月梅, 徐其碧, 杨贵清, 等. 艾纳香叶绿体基因组密码子使用偏性分析 [J]. 西部林业科学, 2023, 52(3): 55-62.
- [23] 肖明昆, 聂恺宏, 沈绍斌, 等. 复羽叶栎树叶绿体基因组密码子偏好性分析 [J]. 西南林业大学学报: 自然科学, 2023, 43(3): 56-63.
- [24] 王宇飞, 江媛, 杨成金, 等. 滇重楼叶绿体基因组特征及密码子偏好性分析 [J]. 分子植物育种, 2021, 19(22): 7448-7458.
- [25] 赵振宁, 孙浩田, 宋雨茹, 等. 山楂属植物叶绿体基因组特征与密码子偏好性分析 [J]. 江苏农业学报, 2023, 39(2): 504-517.
- [26] 马孟莉, 张薇, 孟衡玲, 等. 豆蔻属药用植物叶绿体基因组密码子偏性分析 [J]. 中草药, 2021, 52(12): 3661-3670.
- [27] 冯展, 江媛, 郑燕, 等. 肉苁蓉属植物叶绿体基因组密码子偏好性分析 [J]. 中草药, 2023, 54(5): 1540-1550.
- [28] 黄祥, 楚光明, 郑新开, 等. 睡莲属叶绿体基因组密码子偏好性及系统发育分析 [J]. 中国农业科技导报, 2022, 24(4): 75-84.
- [29] 江媛, 杨青淑, 王婧, 等. 毛重楼叶绿体基因组序列特征及其系统发育分析 [J]. 中草药, 2021, 52(13): 4014-4022.
- [30] 沈莲文, 田金红, 王玉昌, 等. 2 种玉兰属植物叶绿体基因组密码子偏好性分析 [J]. 西南林业大学学报: 自然科学, 2023, 43(2): 44-53.
- [31] 吴学俊, 梁湘兰, 易子群, 等. 湖北山楂叶绿体基因组密码子偏好性分析 [J]. 分子植物育种, 2023, 21(8): 2508-2516.
- [32] 李亚麒, 黄家雄, 娄予强, 等. 小粒咖啡铁皮卡叶绿体基因组密码子偏好性分析 [J]. 西北林学院学报, 2023, 38(2): 92-99.
- [33] 刘蒙蒙, 邢咏梅, 郭顺星. 基于转录组数据分析药用真菌猪苓密码子使用偏好性 [J]. 药学学报, 2020, 55(5): 1050-1055.
- [34] 季凯凯, 宋希强, 陈春国, 等. 木兰科叶绿体基因组的密码子使用特征分析 [J]. 中国农业科技导报, 2020, 22(11): 52-62.
- [35] 夏晞, 彭劲谕, 王大玮, 等. 3 种榕属叶绿体基因组密码子偏好性分析 [J]. 西北林学院学报, 2022, 37(5): 88-94.
- [36] 张加强, 刘慧春, 王杰, 等. 金银花大毛花叶绿体基因组密码子的偏好性分析 [J]. 浙江农业学报, 2023, 35(4): 821-830.
- [37] 梁湘兰, 覃逸明, 孙晓波, 等. 藿香叶绿体基因组密码子偏好性分析 [J]. 分子植物育种, 2023, 21(3): 809-

- 818.
- [38] Gao M Q, Huo X W, Lu L T, *et al.* Analysis of codon usage patterns in *Bupleurum falcatum* chloroplast genome [J]. *Chin Herb Med*, 2023, 15(2): 284-290.
- [39] Zhang Y R, Nie X J, Jia X O, *et al.* Analysis of codon usage patterns of the chloroplast genomes in the Poaceae family [J]. *Aust J Bot*, 2012, 60(5): 461.
- [40] Xu C, Dong J, Tong C F, *et al.* Analysis of synonymous codon usage patterns in seven different citrus species [J]. *Evol Bioinform Online*, 2013, 9: 215-228.
- [41] 张重菊, 赵懿琛, 赵德刚. 薏苡转录组中密码子使用偏好性分析 [J]. 种子, 2022, 41(4): 13-19.
- [42] 赖瑞联, 冯新, 陈瑾, 等. 橄榄转录组密码子使用偏好性及其影响因素 [J]. 核农学报, 2019, 33(1): 31-38.
- [43] 仇学文, 李丹, 甘玉迪, 等. 豇豆叶绿体基因组密码子使用偏好性分析 [J]. 核农学报, 2023, 37(6): 1118-1131.
- [44] 高梦琦, 邹建珍, 霍小位, 等. 基于转录组数据分析药用大黄的密码子使用偏好性 [J]. 中草药, 2021, 52(20): 6344-6349.
- [45] Lei F J, Zhang H R, Long Y P, *et al.* Characteristics and phylogenetic analysis of the complete chloroplast genome of *Lilium concolor* Salisb. (Liliaceae) from Jilin, China [J]. *Mitochondrial DNA B Resour*, 2022, 7(1): 30-31.
- [46] Duan Q, Liu F, Gui D P, *et al.* Phylogenetic analysis of wild species and the maternal origin of cultivars in the genus *Lilium* using 114 plastid genomes [J]. *Front Plant Sci*, 2022, 13: 865606.
- [47] Hayashi K, Kawano S. Molecular systematics of *Lilium* and allied genera (Liliaceae): phylogenetic relationships among *Lilium* and related genera based on the *rbcL* and *matK* gene sequence data [J]. *Plant Spec Biol*, 2010, 15(1): 73-93.

[责任编辑 时圣明]