

• 药材与资源 •

棘豆属 9 种代表物种叶绿体基因组密码子使用偏好性的比较分析

张朋辉^{1,2}, 刘玉萍^{1,2,3}, 苏旭^{1,2,3*}, 靳佳瑞^{1,2}, 杨倩^{1,2}, 余明君^{1,2}, 曲荣举^{1,2}, 赵亚萍¹

1. 青海师范大学生命科学院, 青海 西宁 810008

2. 青海师范大学 青海省青藏高原生物多样性形成机制与综合利用重点实验室, 青海 西宁 810008

3. 青海师范大学高原科学与可持续发展研究院, 青海 西宁 810016

摘要: 目的 探究棘豆属 *Oxytropis* 物种密码子使用模式及影响因素。方法 基于棘豆属 9 个代表物种的叶绿体基因组数据, 采用 CodonW1.4.2、SPSS23 等软件分析密码子偏好性。结果 棘豆属 9 个代表物种密码子的不同位点 GC 含量差异较大, 均小于 50%, $GC_1 > GC_2 > GC_3$; 碱基组成与密码子适应指数 (codon adaptation index, CAI)、密码子偏好指数 (codon bias index, CBI)、密码子使用频率 (frequency of optional codons, FOP) 等密码子参数存在显著正相关; 棘豆属 9 个代表物种有效密码子数大于 35, 偏好性较弱; ENC-plot、PR2-plot 和中性绘图分析显示, 自然选择是影响棘豆属物种密码子偏好性的主要因素; 9 个代表物种鉴定到 23 个最优密码子, 均以 A/U 结尾; 系统发育和聚类分析表明, 棘豆属物种具有明显的属内分化, 黄花棘豆和甘肃棘豆的亲缘关系较近。结论 棘豆属 9 个代表物种的密码子偏好性较弱, 自然选择是影响棘豆属物种密码子偏好性的主要因素, 为棘豆属物种遗传进化研究提供了重要参考数据。

关键词: 叶绿体基因组; 有效密码子; 中性绘图; 最优密码子; 聚类分析

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2025)14-5169-10

DOI: 10.7501/j.issn.0253-2670.2025.14.020

Comparative analysis of codon usage bias in chloroplast genome of nine representative species of *Oxytropis*

ZHANG Penghui^{1,2}, LIU Yuping^{1,2,3}, SU Xu^{1,2,3}, JIN Jiarui^{1,2}, YANG Qian^{1,2}, YU Mingjun^{1,2}, QU Rongju^{1,2}, ZHAO Yaping¹

1. College of Life Sciences, Qinghai Normal University, Xining 810008, China

2. Key Laboratory of Biodiversity Formation Mechanism and Comprehensive Utilization of the Qinghai-Xizang Plateau in Qinghai Province, Qinghai Normal University, Xining 810008, China

3. Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining 810016, China

Abstract: Objective To explore the usage pattern and influencing factor of codons in *Oxytropis* species. **Methods** Based on the chloroplast genome data of nine representative species from *Oxytropis*, software such as CodonW1.4.2 and SPSS23 were used to analyze codon bias. **Results** The GC content of different codon sites from nine representative species of *Oxytropis* was significantly different, all of which were less than 50%, and $GC_1 > GC_2 > GC_3$. There was significantly positive correlation between the base composition and codon parameters such as codon adaptation index (CAI), codon bias index (CBI) and frequency of optional codons (FOP). The effective number of codons in nine representative species of *Oxytropis* was greater than 35, and the codon preference was weak. ENC-plot, PR2-plot and neutral plot analyses showed that natural selection was the main factor affecting codon preference of *Oxytropis*. A total of 23 optimal codons were identified from nine representative *Oxytropis* species, which ended in A/U. Phylogenetic and cluster analyses showed that the *Oxytropis* species had obvious intra-genus differentiation, and *Oxytropis ochrocephala* and *Oxytropis kansuensis* had the closest relationship. **Conclusion** The codon preference of nine representative *Oxytropis* species was weak, and natural selection was the main factor affecting the codon preference of *Oxytropis*. This study provides the important reference data for the genetics and evolution of the *Oxytropis* species.

Key words: chloroplast genome; effective codon; neutral plot; optimal codon; cluster analysis

收稿日期: 2025-02-03

基金项目: 青海省重大科技专项 (2023-SF-A5); 2023 年中央林业草原生态保护恢复资金野生动植物保护项目 (QHSY-2023-016); 青海省省财政林业改革发展资金林草新技术推广项目 (QSCZ-2023-001)

作者简介: 张朋辉 (2000—), 硕士研究生, 主要从事植物系统分类与分子生态。E-mail: zph200720@163.com

*通信作者: 苏旭, 教授, 博士生导师, 从事植物系统分类与分子生态研究。E-mail: xusu8527972@126.com

棘豆属 *Oxytropis* DC. 植物为豆科(Leguminosae Lindl.) 多年生草本或灌木, 全世界约有 300 种, 通常生长于山地草甸、大草原、石质斜坡、北极和高山冻土, 适应性强; 其中, 我国有 146 种, 主要分布于内蒙古和新疆的山地、荒漠以及青藏高原和西南横断山脉等地^[1]。已有研究表明, 棘豆属植物富含生物碱、黄酮、三萜皂苷类以及少量木脂素、甾体类等化合物^[2], 具有清热解毒、消炎镇痛、止血、生肌愈疮、抗肿瘤、抗菌等药理作用^[3]。

叶绿体是一种具有光合作用的半自主性细胞器, 为绿色植物和少量原生物所特有, 存在于植物细胞质体内, 常呈椭球形或片状, 因含叶绿素显绿色得名^[4]。叶绿体基因组是叶绿体中含有的遗传物质, 具有保守性、序列较短且简单、单亲遗传、易测序等特点, 因而成为植物遗传多样性研究和系统发育分析的理想材料^[5]。叶绿体基因组的研究始见于 1986 年 Shinozaki 和 Ohyama 分别对烟草 *Nicotiana tabacum* L.^[6] 和地钱 *Marchantia polymorpha* L.^[7] 叶绿体 DNA 的测序。随着高通量测序技术的迅速发展, 叶绿体基因组研究变得较为便捷容易, 相关研究成果涵盖了绿色植物的诸多类群, 如扇穗茅 *Littledalea racemosa* Keng^[8]、苦马豆 *Sphaerophysa salsula* (Pall.) DC.^[9]、葵花大蓟 *Cirsium souliei* (Franch.) Mattf.^[10]、蒺藜苜蓿 *Medicago truncatula* Gaertn.^[11] 等。并且, 由于物种间和物种内叶绿体基因组序列存在较大差异^[12], 因此, 叶绿体基因组测序也成为植物系统发育和遗传分化研究的一种重要分子标记方法。

密码子是 mRNA 上相邻的 3 个碱基序列, 生物体遗传信息传递过程中与 tRNA 上的反密码子互补编码氨基酸, 其中因简并性而编码同一氨基酸的密码子被称为同义密码子, 密码子偏好性是生物体中同义密码子不均等使用的现象^[11]。研究表明, 不同物种叶绿体基因组密码子使用模式具有较大差异, 单子叶植物往往偏好以 G/C 碱基结尾, 双子叶植物偏好以 A/U 碱基结尾^[12]。目前, 关于棘豆属植物叶绿体基因组的研究主要集中于序列特征^[13]和系统发育^[14-15]等领域, 然而却未见关于密码子使用模式的报道。据此, 利用高通量测序技术, 结合先前发表的叶绿体基因组数据, 本研究分析了棘豆属物种的密码子使用模式, 探究了密码子使用模式的影响因素, 以期棘豆属植物叶绿体基因组进化机制、系统发育关系、密码子优化和偏好机制研究提供参考数据。

1 材料

黄花棘豆 *O. ochrocephala* Bunge 健康植株采自中国青海海西蒙古族藏族自治州德令哈市 (N37°22'26.4", E97°22'15.6", alt. 2951.61 m), 经青海师范大学生命科学学院陈金元副教授鉴定为黄花棘豆。野外采集新鲜无病虫害叶片, 硅胶快速干燥后置于冰箱中保存, 用于 DNA 提取和测序。此外, 从 NCBI 下载二色棘豆 *O. bicolor* Bunge (MN255323.1)、*O. splendens* Douglas ex Hook. (MT409174)、*O. arctobia* Bunge (MT409175.1)、二型叶棘豆 *O. diversifolia* E. Peter (MT780271)、小花棘豆 *O. glabra* (Lam.) DC. (MW349014)、猫头刺 *O. aciphylla* Ledeb. (MW794135)、镰荚棘豆 *O. falcata* Bunge (OR491708.1) 和多枝棘豆 *O. ramosissima* Kom. (PP868418.1) 的叶绿体基因组序列用于联合分析。

2 方法

2.1 基因组测序和组装

采用改良的 CTAB 法从黄花棘豆干燥叶片中提取总基因组 DNA^[16]; 利用 Illumina Nova-Seq 6000 平台进行叶绿体基因组测序, 经过滤^[17]、组装^[18]和注释^[19-20]后将叶绿体基因组序列和注释信息提交至 GenBank (登录号 OR897029)。根据注释结果, 从叶绿体基因组中提取蛋白质编码序列 (coding DNA sequence, CDS), 选择长度大于 300 bp、不重复、起始密码子为 ATG 以及终止密码子为 TAA、TAG、TGA 的 CDS 序列进行后续分析。

2.2 相对同义密码子使用和有效密码子数 (effective number of codon, ENC) 分析

采用 Geneious Prime 2023.0.1 软件^[21]去除叶绿体基因组中重复基因及长度小于 300 bp 的 CDS 序列, 将符合条件的 CDS 序列导出为 fasta 格式; 利用 Codon W 1.4.2 软件^[22]计算不同个体 CDS 基因的 RSCU 值、ENC 值等基本参数。以 ENC 值为纵坐标, GC₃ 为横坐标, 依据公式^[23]绘制标准曲线和散点图, 根据基因的 ENC 值和标准曲线偏离程度判断密码子偏好性受到的影响。

$$ENC = 2 + GC_3s + 29/[GC_3s^2 + (1 - GC_3s)^2]$$

2.3 奇偶偏好性 (PR2-plot) 和中性绘图分析

奇偶偏好性分析 (PR2-bias plot analysis) 是为探究密码子第 3 位上 A 与 T、G 与 C 碱基分布是否均衡, 计算每个基因的 $A_3/(A_3 + T_3)$ 和 $G_3/(G_3 + C_3)$, 并分别以其为纵坐标和横坐标作图, 图中心点表示

A=T 且 C=G, 其余点为中心点向该点发出的矢量距离代表基因的偏倚程度和方向^[24]。中性绘图分析利用密码子 3 个位置的鸟嘌呤 (G) 和胞嘧啶 (C) 碱基相关性探究密码子偏性的影响因素, 统计密码子 GC₁、GC₂ 和 GC₃ 3 个位点的 GC 含量, 以 GC₃ 为横坐标, GC₁ 和 GC₂ 的平均值 GC₁₂ 为纵坐标制作散点图并进行回归分析, 图中每个点代表一个基因, 根据各点距对角线位置判断密码子使用模式是受突变还是自然选择影响^[25]。

2.4 最优密码子确定

以“2.2”项得到的 ENC 值为参考, 将棘豆属代表物种的蛋白质编码基因由低到高排序, 选取 ENC 值大小为前 10% 和后 10% 的基因各 5 个, 建立高表达和低表达的数据库; 利用 Codon W 1.4.2 软件分别计算 2 数据库基因的相对同义密码子使用频率 (relative synonymous codon usage, RSCU) 值, 计算高表达和低表达的差值, 以 $\Delta RSCU \geq 0.08$ 且 $RSCU > 1$ 作为最优密码子。

2.5 系统发育与 RSCU 聚类分析

从 NCBI 数据库下载甘肃棘豆 *O. kansuensis* Bunge (PQ790176.1) 的叶绿体基因组, 与上述棘豆属 9 个物种的叶绿体基因组数据共同用于系统发育分析; 利用 CPStools 软件^[26]提取棘豆属 10 个物种的共有 CDS 序列, 运用 MEGA11 中的 ClustalW 功能进行序列比对, 根据比对结果绘制邻接 (neighbor-joining, NJ) 树, Bootstrap 检验 1 000 次。将棘豆属代表物种的蛋白编码序列按“2.1”项标准过滤, 以每个物种各密码子 RSCU 值为变量 (剔除 AUG、UGG、UAA、UAG、UGA), 采用 SPSS23 进行聚类分析, 构建聚类树状图。

3 结果与分析

3.1 叶绿体基因组密码子使用偏好性分析

本研究结果表明, 棘豆属 9 个代表物种的叶绿

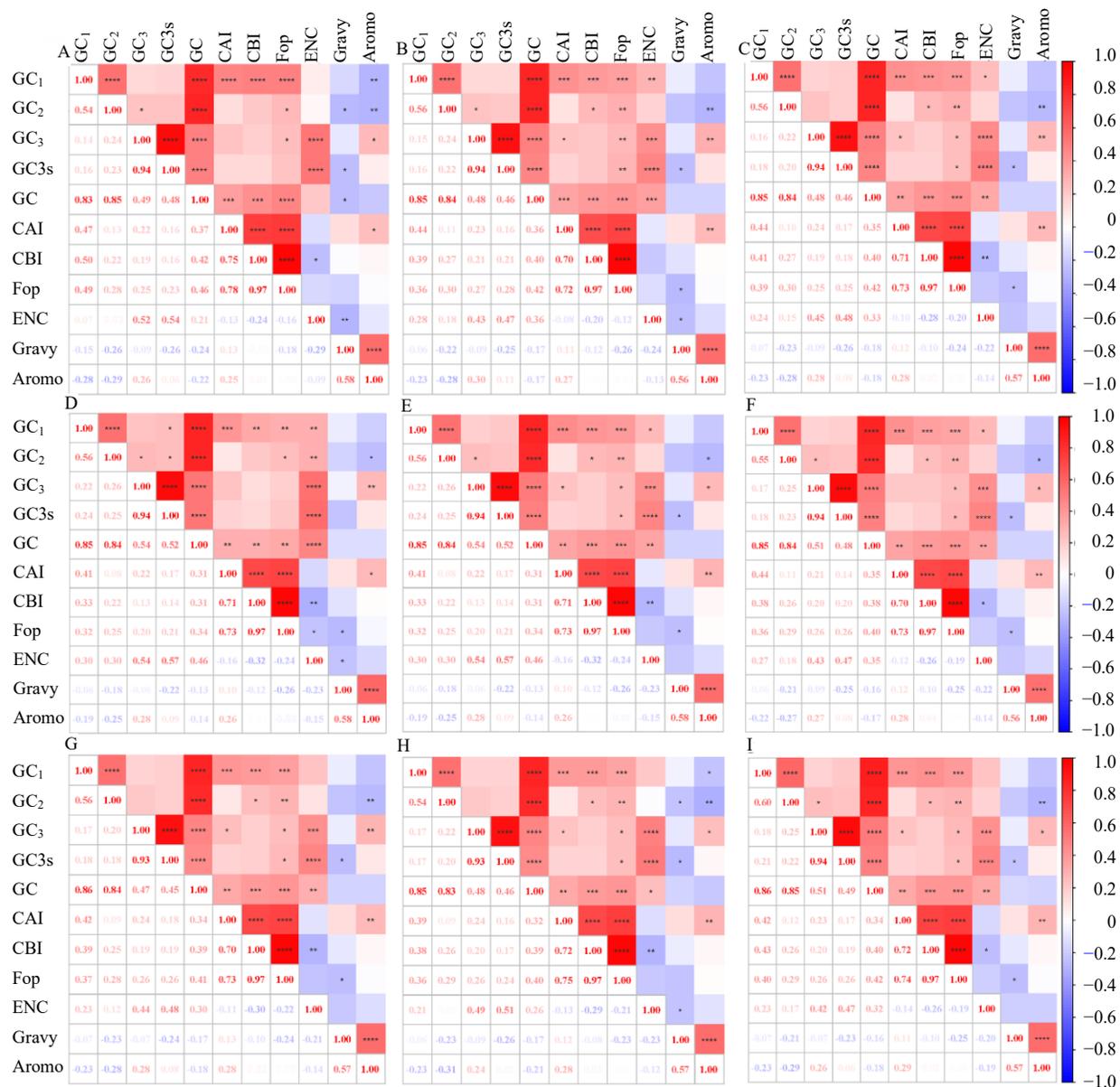
体基因组均含 76 条 CDS 序列, 依据“2.1”项标准筛选后每个物种获得 51~52 条用于密码子偏好性分析的 CDS 序列; 棘豆属 9 个代表物种整个编码区的 GC 含量 (GC_{all}) 处于 36.61%~36.87%, 3 个位点 (GC₁、GC₂ 和 GC₃) 的碱基含量分别为 45.19%~45.54%、37.13%~37.42%、27.26%~27.80% (表 1), 总体呈 GC₁>GC₂>GC₃ 的变化趋势, 表明棘豆属物种不同位点的 GC 含量接近且分布不均, 偏向出现在前 2 位碱基中; GC 碱基含量小于 50% 且 GC₃ 含量最低, 说明棘豆属 9 个代表物种的叶绿体密码子偏好使用 A/U 碱基, 多以 A/U 碱基结尾; 叶绿体基因组平均 ENC 值为 45, 证明棘豆属 9 个代表物种密码子偏好性较弱。

棘豆属 9 个代表物种叶绿体基因组 CDS 基因的 GC₁、GC₂、GC₃、ENC、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) 等参数的相关性分析结果显示, 所有物种的 GC₁、GC₂ 与 GC₃ 无明显相关性, GC₁ 与 GC₂ 间存在极显著相关, 且与第 3 位密码子的碱基组成差异较大 (图 1), 表明棘豆属 9 个代表物种的前 2 位密码子碱基组成相似, 其叶绿体基因组表达主要受第 1 位碱基 GC 含量的影响, 与第 2、3 位碱基的 GC 含量关联较小; 棘豆属 9 个代表物种叶绿体基因组的 ENC 值与 GC₁、GC₂ 无显著相关性, 而与 GC₃ 和 GC_{3s} 间具有显著相关性, 表明第 3 位密码子碱基组成改变可能是影响棘豆属密码子偏好性的主要因素; 黄花棘豆和镰荚棘豆的 GC₂ 与蛋白疏水性 (general average hydropathicity, Gravy) 呈显著负相关, 表明它们的叶绿体基因组密码子第 2 位碱基组成对 Gravy 具有一定影响; 蛋白质芳香性 (aromaticity, Aromo) 与 GC₂ 存在显著负相关, 与 GC₃ 呈显著正相关, 说明叶绿体基因组密码子第 2、3 位碱基对 Aromo 影响较大。

表 1 棘豆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 bias in chloroplast genome of *Oxytropis* plants

| 物种名 | 登录号 | GC ₁ /% | GC ₂ /% | GC ₃ /% | GC _{all} /% | ENC |
|------------------------------|----------|--------------------|--------------------|--------------------|----------------------|-------|
| 黄花棘豆 <i>O. ochrocephala</i> | OR897029 | 45.54 | 37.27 | 27.80 | 36.87 | 45.80 |
| 二色棘豆 <i>O. bicolor</i> | MN255323 | 45.21 | 37.19 | 27.42 | 36.61 | 44.98 |
| <i>O. splendens</i> | MT409174 | 45.20 | 37.15 | 27.48 | 36.61 | 45.02 |
| <i>O. arctobia</i> | MT409175 | 45.33 | 37.16 | 27.60 | 36.69 | 45.37 |
| 二型叶棘豆 <i>O. diversifolia</i> | MT780271 | 45.19 | 37.16 | 27.50 | 36.62 | 45.08 |
| 小花棘豆 <i>O. glabra</i> | MW349014 | 45.37 | 37.34 | 27.44 | 36.72 | 45.07 |
| 猫头刺 <i>O. aciphylla</i> | MW794135 | 45.22 | 37.13 | 27.47 | 36.61 | 45.06 |
| 镰荚棘豆 <i>O. falcata</i> | OR491708 | 45.30 | 37.19 | 27.53 | 36.67 | 45.03 |
| 多枝棘豆 <i>O. ramosissima</i> | PP868418 | 45.21 | 37.42 | 27.26 | 36.63 | 45.00 |



A-黄花棘豆; B-二色棘豆; C-*O. splendens*; D-*O. arctobia*; E-二型叶棘豆; F-小花棘豆; G-猫头刺; H-镰荚棘豆; I-多枝棘豆; 下同。

A-*O. ochrocephala*; B-*O. bicolor*; C-*O. splendens*; D-*O. arctobia*; E-*O. diversifolia*; F-*O. glabra*; G-*O. aciphylla*; H-*O. falcata*; I-*O. ramosissima*; same as below.

图1 棘豆属植物密码子基本参数相关性分析

Fig. 1 Correlation analysis of codon basic parameters of *Oxytropis* plants

3.2 相对同义密码子偏好性和有效密码子数绘图(ENC-plot)分析

RSCU 大小是衡量密码子偏好性程度的一个重要指标。若编码同一氨基酸的所有同义密码子被等量使用,期望频率是该密码子的观察频率(RSCU=1),表明该氨基酸对其同义密码子使用无偏好性;若观察频率高于期望频率(RSCU>1),说明氨基酸对该密码子使用偏好性强,为高频密码子;反之,则为低频密码子^[1]。本研究发现,剔除终止密码子和不表达密码子后,棘豆属

9个代表物种的RSCU值相差不大,即所有物种均有29个高频密码子,占密码子总数的44.62%,除密码子UUA的RSCU值大于2,UUG以G结尾外,其余高频密码子皆以A/U结尾且RSCU<2,其中以A结尾的有12个,以U结尾的有16个(图2);RSCU值小于1的低频密码子有30个,其中16个以C结尾,12个以G结尾,2个以A结尾(图2),这些结果表明棘豆属代表物种的高频密码子偏向以A/U结尾,低频密码子偏向以G/C结尾。

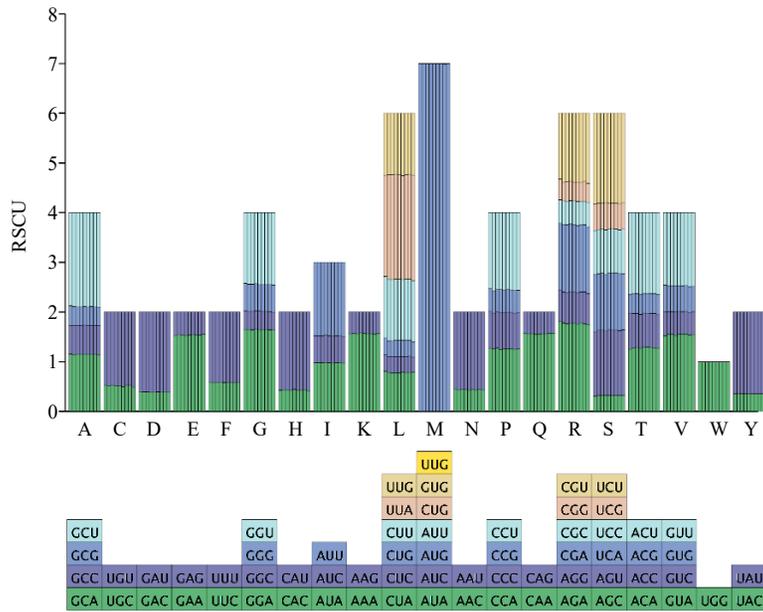


图2 棘豆属植物代表物种 RSCU 值

Fig. 2 RSCU values of representative *Oxytropis* plant species

ENC-plot通过对叶绿体基因组中各蛋白编码基因 ENC 值和 GC₃ 的关联分析, 根据基因散点与标准曲线的距离揭示影响物种密码子使用模式的主要因素。棘豆属 9 个代表物种蛋白编码基因

的 ENC-plot 图相似, 绝大部分基因落点于标准曲线下方, 仅少数基因位于标准曲线附近或上方; ENC 值介于 33.97~56.48, 大多数基因的 ENC > 45 (图 3), 表明棘豆属物种密码子的使用偏好受

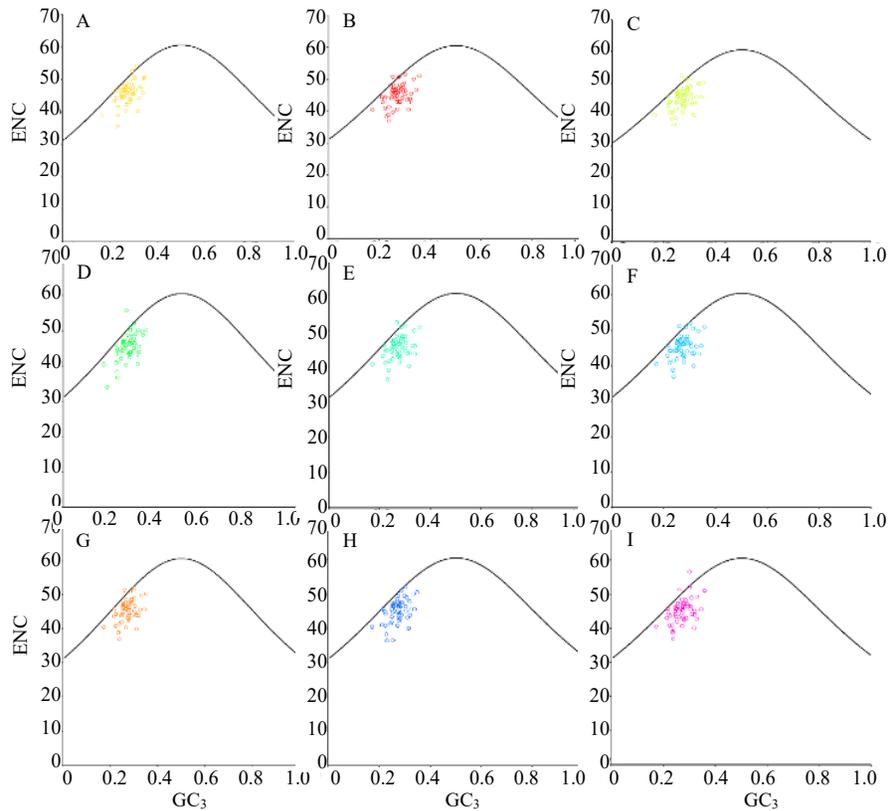


图3 棘豆属植物有效密码子数绘图分析

Fig. 3 ENC-plot analysis of *Oxytropis* plants

到自然选择影响较大，只有部分基因受到突变的影响。

3.3 奇偶偏好性绘图 (PR2-plot) 和中性绘图 (GC₃~GC₁₂) 分析

奇偶偏好性分析能够阐明物种密码子第3位碱基的偏倚程度。本研究结果表明，图中两轴均以0.5为中心点，将图分为4个象限，棘豆属9个代表物种的蛋白编码基因不均匀分布于4个象限中，以第

4象限较为集中(图4)，说明棘豆属物种密码子的第3位碱基使用存在偏倚，第3位碱基的使用频率为T>A，G>C，进一步证实棘豆属物种的密码子使用偏好性受自然选择和突变压力共同影响，尤其自然选择的影响较大。同时，对棘豆属9个代表物种蛋白编码基因进行中性绘图分析，发现该属物种的多数基因位于对角线两侧，仅少部分基因位于对角线上或沿对角线分布(图5)；GC₃的范围为

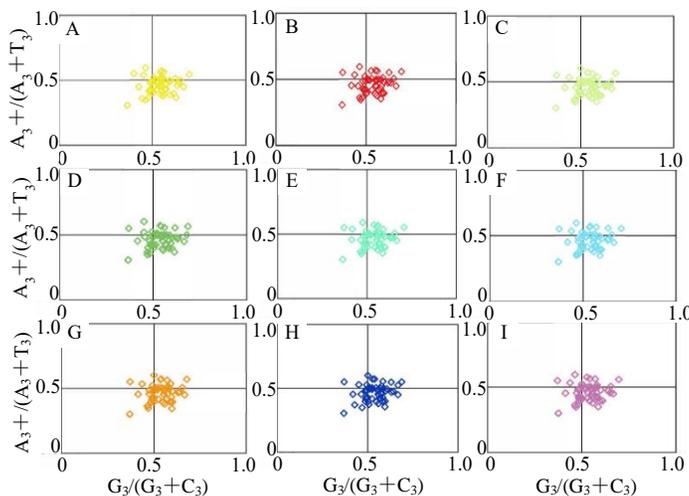


图4 棘豆属植物奇偶偏好性绘图分析

Fig. 4 PR2-plot analysis of *Oxytropis* plants

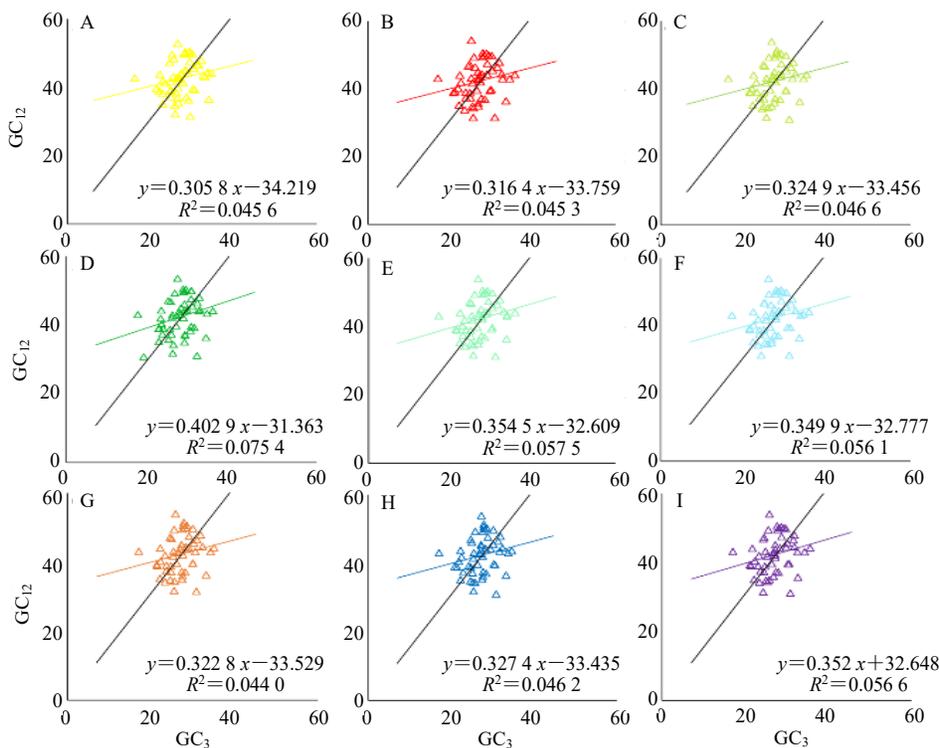


图5 棘豆属植物中性绘图分析

Fig. 5 Neutrality plot analysis of *Oxytropis* plants

16.26~35.64, GC₁₂介于 30.74~54.32, 回归系数在 0.305 8~0.402 9, GC₁₂与 GC₃不存在显著相关性(图 1), 表明两者的进化模式存在差异, 突变和自然选择对棘豆属物种的密码子使用偏好性均有影响, 并且自然选择产生的影响相对较大。

3.4 最优密码子确定

研究结果显示, 棘豆属黄花棘豆、二色棘豆、*O. splendens*、*O. arctobia*、二型叶棘豆、小花棘豆、猫头刺、镰荚棘豆和多枝棘豆中鉴定出的最优密码子数分别为 14、19、14、15、14、15、13、14 和 14 个, 最优密码子均以 A/U 结尾, 其中 12 个以 U 结尾、10 个以 A 结尾; 二色棘豆最优密码子多达 19 个, 猫头刺最优密码子最少、仅 13 个, 黄花棘豆、

O. splendens、二型叶棘豆、镰荚棘豆和多枝棘豆均含 14 个最优密码子(表 2)。更为重要的是, 8 个最优密码子为棘豆属 9 个代表物种所共有, 依次为 GCU、GGU、AAA、UUA、CGA、UCU、ACU 和 GUU, 分别编码丙氨酸(Ala)、甘氨酸(Gly)、赖氨酸(Lys)、亮氨酸(Leu)、精氨酸(Arg)、丝氨酸(Ser)、苏氨酸(Thr)和缬氨酸(Val); 其中, GCU、UUA 和 GUU 在所有物种中的 Δ RSCU 均大于 0.5(表 2)。此外, 二色棘豆有 2 个特有最优密码子, 分别是编码脯氨酸的 CCA 和编码丝氨酸的 UCA; 小花棘豆仅有一个特有最优密码子 CUU, 编码亮氨酸; 黄花棘豆和镰荚棘豆的最优密码子相同(表 2)。

表 2 棘豆属植物叶绿体基因组最优密码子分析
Table 2 Optimal codon analysis in chloroplast genome of *Oxytropis* plants

| 氨基酸 | 密码子 | | | | | | | | |
|-----|--------|--------|---------------------|--------------------|--------|--------|--------|--------|--------|
| | 黄花棘豆 | 二色棘豆 | <i>O. splendens</i> | <i>O. arctobia</i> | 二型叶棘豆 | 小花棘豆 | 猫头刺 | 镰荚棘豆 | 多枝棘豆 |
| Ala | GCU*** | GCU*** | GCU*** | GCU*** | GCU*** | GCU*** | GCU*** | GCU*** | GCU*** |
| Cys | UGU* | UGU** | UGU** | - | UGU* | UGU** | - | UGU** | UGU* |
| Asp | GAU* | GAU* | - | GAU** | - | GAU* | GAU* | GAU* | - |
| Glu | GAA** | GAA** | GAA** | GAA* | GAA** | - | GAA* | GAA** | GAA** |
| Phe | - | UUU*** | - | UUU* | - | - | - | - | - |
| Gly | - | GGA* | - | GGA** | - | - | - | - | - |
| His | GGU*** | GGU* | GGU*** | GGU* | GGU*** | GGU*** | GGU*** | GGU*** | GGU*** |
| Lys | - | CAU** | - | CAU** | - | - | - | - | - |
| Leu | AAA** | AAA** | AAA** | AAA* | AAA** | AAA* | AAA** | AAA* | AAA** |
| Pro | UUA*** | UUA** | UUA*** | UUA** | UUA** | UUA*** | UUA*** | UUA*** | UUA*** |
| | - | - | - | - | - | CUU* | - | - | - |
| Arg | - | CCA* | - | - | - | - | - | - | - |
| Gln | CCU** | - | CCU*** | - | CCU*** | CCU** | CCU** | CCU*** | CCU** |
| Thr | CAA* | CAA** | CAA** | - | CAA** | CAA* | CAA** | CAA** | CAA* |
| Ser | CGA** | CGA*** | CGA*** | CGA*** | CGA*** | CGA*** | CGA** | CGA*** | CGA*** |
| | CGU*** | - | CGU* | - | CGU* | - | CGU*** | CGU** | CGU* |
| | - | - | - | AGA** | - | AGA* | - | - | - |
| Val | - | UCA* | - | - | - | - | - | - | - |
| Tyr | UCU*** | UCU* | UCU*** | UCU*** | UCU*** | UCU*** | UCU*** | UCU*** | UCU*** |
| | ACU*** | ACU** | ACU*** | ACU*** | ACU*** | ACU*** | ACU*** | ACU*** | ACU** |
| | - | GUA* | GUA* | - | GUA* | - | - | - | GUA* |
| | GUU*** | GUU*** | GUU*** | GUU*** | GUU*** | GUU*** | GUU*** | GUU*** | GUU*** |
| | - | UAU* | - | UAU** | - | UAU* | - | - | - |
| 合计 | 14 | 19 | 14 | 15 | 14 | 15 | 13 | 14 | 14 |

*表示 Δ RSCU ≥ 0.08 , **表示 Δ RSCU ≥ 0.3 , ***表示 Δ RSCU ≥ 0.5 。

* means Δ RSCU ≥ 0.08 , ** means Δ RSCU ≥ 0.3 , *** means Δ RSCU ≥ 0.5 。

3.5 系统发育与 RSCU 聚类分析

基于棘豆属 10 个物种的共有 CDS 序列, 本研究采用邻接法(neighbor-joining, NJ)构建了棘豆属的系统发育树和密码子 RSCU 聚类树。NJ 树显示, 除

二型叶棘豆和猫头刺分别以 73%和 97%的支持率单独形成一支外, 其余分支的支持率均为 100%; 棘豆属物种属内遗传分化现象明显, 除多枝棘豆与小花棘豆、黄花棘豆与甘肃棘豆分别聚为一支外, 其余所有

物种单独形成一支。此外,密码子 RSCU 值聚类结果与 NJ 树具有大致相似的拓扑结构,不同之处主要体现在猫头刺和棘豆属 *O. splendens* 聚为一支,并与二

色棘豆和二型叶棘豆共同聚为一组,黄花棘豆与甘肃棘豆单独形成一支,并且两者的遗传距离相同,其余物种均单独聚成一支(图6)。

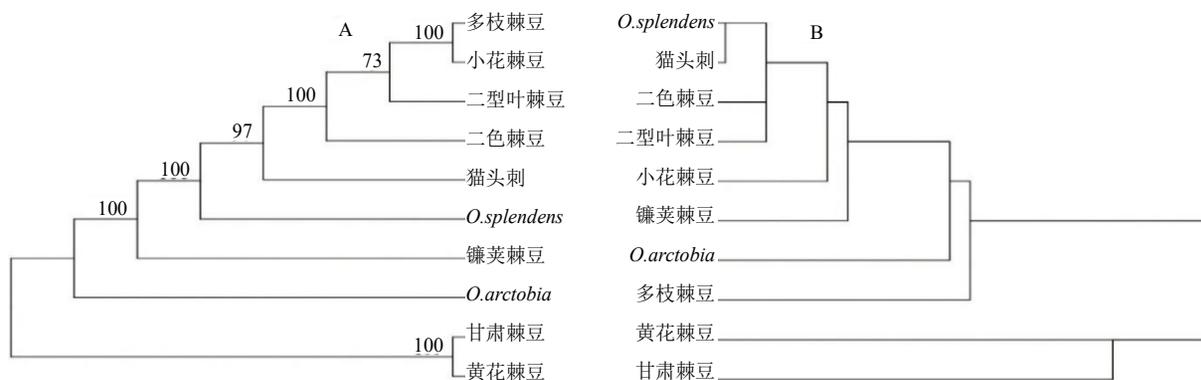


图6 基于共有 CDS 序列的 NJ 树 (A) 和基于 RSCU 值的聚类树 (B)

Fig. 6 NJ tree based on common CDS (A) and cluster tree based on RSCU values (B)

4 讨论

长期演化过程中,植物因适应不同的生长环境产生了各种各样的密码子使用模式和碱基组成^[27]。本研究通过对棘豆属 9 个代表物种叶绿体基因组数据的筛选和鉴定,获得 51~52 条蛋白编码基因的碱基组成和密码子使用模式,结果表明棘豆属不同物种间的碱基组成相似,总体呈 $GC_1 > GC_2 > GC_3$ 的变化趋势,总 GC 含量和密码子 3 个位置的 GC 含量均低于 50%,说明棘豆属物种偏向使用 A/U 碱基,多以 A/U 结尾。

相对同义密码子使用频度(RSCU)指一个密码子观察频率与期望频率之比,是衡量不同生物体内密码子使用偏好的标准。本研究结果表明,棘豆属所有参试物种的高频密码子和最优密码子偏向均以 A/U 碱基结尾,这与先前关于百合属 *Lilium* L.^[28]、甘草属 *Glycyrrhiza* L.^[29]、霸王属 *Sarcozygium* Bunge^[30]等物种的研究结果相同,说明棘豆属植物叶绿体基因组密码子使用模式较为保守。

ENC 能够反映密码子偏离随机选择程度,常介于 20~61。当 $ENC=20$ 时,说明每个氨基酸仅使用一个密码子,此时密码子偏性最大;当 $ENC=61$ 时,密码子被随机使用,没有偏倚;若 ENC 值小于或等于 35,基因具有较强的密码子偏好性,反之基因密码子偏性较弱^[23]。棘豆属 9 个代表物种 CDS 基因的 ENC 值处于 33.97~56.48,平均 ENC 值均大于或等于 45,表明棘豆属参试物种 CDS 基因的密码子偏好性较弱,这与先前已报道的诸多豆科植物

叶绿体基因组的密码子研究结果类似,如蒺藜苜蓿^[1]、苦豆子 *Sophora alopecuroides* L.^[31]、苦马豆^[32]等;同时,9 个代表物种 CDS 基因的第 3 位密码子碱基组成与前 2 位密码子相关性不显著,CAI、CBI 和 FOP 值与 GC_1 存在较显著相关性,ENC 值与 GC_3 相关性显著,蛋白疏水性和芳香性与 GC_2 和 GC_3 存在一定相关性,说明棘豆属参试物种的第 3 位密码子碱基组成与前 2 位密码子差异较大,密码子偏好性和功能表达与密码子 3 个位置的碱基组成密切相关,尤其是第 3 位密码子的碱基变化产生的影响最大。

先前研究认为,突变压力和自然选择是影响物种密码子偏好性的重要因素之一^[33-34],ENC 值是衡量物种自然选择与基因突变对同义密码子非均衡使用偏好程度的重要指标^[23]。本研究结果显示,棘豆属 9 个代表物种 CDS 基因的 ENC-plot 结果接近,绝大部分基因落点于标准曲线下方,说明棘豆属参试物种的实际 ENC 值与预期 ENC 值存在较大差异,表明其密码子使用偏好性主要受自然选择影响,仅部分受到基因突变的影响;PR2-plot 分析显示,密码子第 3 位碱基使用存在偏倚,总体使用频率为 $T > A$ 且 $G > C$,证明棘豆属参试物种的密码子使用偏倚性受到自然选择和突变压力的共同作用,尤以自然选择影响较大;中性绘图分析结果表明,棘豆属 9 个代表物种的回归系数均小于 0.5,进一步说明棘豆属物种密码子偏好性受到自然选择和基因突变的影响,自然选择影响仍是主导因素,

这与先前报道的绝大多数植物叶绿体基因组密码子使用偏好性结果类似^[35-37]，我们认为这可能是由于棘豆属物种的叶绿体基因组相对保守，不易受到基因突变的影响，密码子偏好性的形成主要受到长期的自然选择作用。

另外，基于密码子 RSCU 值构建的聚类树状图可以较好地反映物种间的遗传进化关系与密码子使用模式相关性，一定程度上能够印证系统发育分析的结果^[35]。本研究获得的聚类树状图与基于 CDS 序列构建的 NJ 树具有大致相同的拓扑结构，均呈现出明显的遗传分化现象，这与 Hu 等^[14]以苦豆子为外类群采用最大似然法和贝叶斯法构建的山羊豆族 (TRIB. Galegeae) 叶绿体全基因组系统发育树结果一致；同时，本研究得到的两种结果中黄花棘豆与甘肃棘豆均表现出较近的亲缘关系，这一结果与藏药中所记载的两者均为塞嘎的基原相符合^[2]，表明它们的密码子使用模式与其余参试物种间存在一定差异。本研究首次报道了棘豆属物种的叶绿体基因组密码子使用模式及影响因素，同时进行了系统发育和聚类分析，研究结果可为将来棘豆属种质资源保护利用和遗传改良提供理论基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中国植物志编辑委员会. 中国植物志 (第 42 卷) [M]. 北京: 科学出版社, 1998: 256.
- [2] 徐敏. 中国医学 R000134 中华本草藏药卷 [M]. 武汉: 湖北人民出版社, 2003: 605.
- [3] 马彦梅, 周文明, 杨新娟. 棘豆属植物化学成分和药理作用的研究进展 [J]. 西北林学院学报, 2005, 20(2): 167-170.
- [4] Neuhaus H E, Emes M J. Nonphotosynthetic metabolism in plastids [J]. *Annu Rev Plant Physiol Plant Mol Biol*, 2000, 51: 111-140.
- [5] Olmstead R G, Palmer J D. Chloroplast DNA systematics: A review of methods and data analysis [J]. *Am J Bot*, 1994, 81(9): 1205-1224.
- [6] Shinozaki K, Ohme M, Tanaka M, et al. The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression [J]. *EMBO J*, 1986, 5(9): 2043-2049.
- [7] Ohyama K, Fukuzawa H, Kohchi T, et al. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA [J]. *Nature*, 1986, 322: 572-574.
- [8] 刘玉萍, 吕婷, 朱迪, 等. 青藏高原特有种一藏扇穗茅叶绿体基因组测序及序列分析 [J]. 植物研究, 2018, 38(4): 518-525.
- [9] 苏丹丹, 刘玉萍, 刘涛, 等. 苦马豆叶绿体基因组结构及其特征分析 [J]. 植物研究, 2022, 42(3): 446-454.
- [10] 郑长远, 苏旭, 刘玉萍, 等. 葵花大蓟叶绿体基因组结构特征及系统发育分析 [J]. 药学报, 2022, 57(7): 2224-2233.
- [11] 杨国锋, 苏昆龙, 赵怡然, 等. 蒺藜苜蓿叶绿体密码子偏好性分析 [J]. 草业学报, 2015, 24(12): 171-179.
- [12] Li N, Li Y Y, Zheng C C, et al. Genome-wide comparative analysis of the Codon usage patterns in plants [J]. *Genes Genom*, 2016, 38(8): 723-731.
- [13] Liu S, Wei Y L, Si W, et al. Complete chloroplast genome sequence of *Oxytropis glabra* (Leguminosae) [J]. *Mitochondrial DNA B Resour*, 2021, 6(9): 2478-2479.
- [14] Hu X Y, Liu Y P, Chen J Y, et al. The complete chloroplast genome of *Oxytropis ochrocephala* Bunge 1874 (Fabaceae) and its phylogenetic analysis [J]. *Mitochondrial DNA B Resour*, 2024, 9(5): 641-646.
- [15] 高利霞. 基于 trnL-F 序列和 psbA-trnH 序列的 10 种棘豆属植物分子系统学研究 [D]. 呼和浩特: 内蒙古师范大学, 2013.
- [16] Doyle J J, Doyle J L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue [J]. *Phytochem Bull*, 1987, 19(1): 11-15.
- [17] Chen S F, Zhou Y Q, Chen Y R, et al. Fastp: An ultra-fast all-in-one FASTQ preprocessor [J]. *Bioinformatics*, 2018, 34(17): i884-i890.
- [18] Jin J J, Yu W B, Yang J B, et al. GetOrganelle: A fast and versatile toolkit for accurate *de novo* assembly of organelle genomes [J]. *Genome Biol*, 2020, 21(1): 241.
- [19] Tillich M, Lehwark P, Pellizzer T, et al. GeSeq-versatile and accurate annotation of organelle genomes Open Access [J]. *Nucleic Acids Res*, 2017, 45(W1): W6-W11.
- [20] Shi L C, Chen H M, Jiang M, et al. CPGAVAS2, an integrated plastome sequence annotator and analyzer [J]. *Nucleic Acids Res*, 2019, 47(W1): W65-W73.
- [21] Kearse M, Moir R, Wilson A, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data [J]. *Bioinformatics*, 2012, 28(12): 1647-1649.
- [22] Kim G B, Lim C E, Kim J S, et al. Comparative chloroplast genome analysis of *Artemisia* (Asteraceae) in east Asia: Insights into evolutionary divergence and phylogenomic implications [J]. *BMC Genomics*, 2020, 21(1): 415.
- [23] Wright F. The 'effective number of codons' used in a gene [J]. *Gene*, 1990, 87(1): 23-29.
- [24] Sueoka N. Near homogeneity of PR2-bias fingerprints in

- the human genome and their implications in phylogenetic analyses [J]. *J Mol Evol*, 2001, 53(4/5): 469-476.
- [25] Sueoka N. Directional mutation pressure and neutral molecular evolution [J]. *Proc Natl Acad Sci USA*, 1988, 85(8): 2653-2657.
- [26] Huang L J, Yu H X, Wang Z, *et al*. CPStools: A package for analyzing chloroplast genome sequences [J]. *iMetaOmics*, 2024, 1(2): e25.
- [27] Haig D, Westoby M. Genomic imprinting in endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis [J]. *Philosophical Transactions: Biological Sciences*, 1991, 333(1266): 1-13.
- [28] 代国娜, 尚明越, 王嘉乐, 等. 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 [J]. *中草药*, 2024, 55(11): 3835-3844.
- [29] 代江鹏, 蔡一鸣, 刘巧珍, 等. 甘草属 7 种植物叶绿体基因组密码子使用偏好性分析 [J]. *中草药*, 2023, 54(9): 2907-2916.
- [30] 周美, 王金涛, 庄重, 等. 4 种霸王属植物叶绿体基因组密码子偏好性分析 [J]. *西北植物学报*, 2024, 44(9): 1445-1455.
- [31] 胡夏宇. 苦豆子叶绿体基因组特征及其系统发育研究 [D]. 西宁: 青海师范大学, 2024.
- [32] 梁湘兰, 郭松. 苦马豆叶绿体基因组密码子偏好性分析 [J]. *西北林学院学报*, 2022, 37(2): 121-126.
- [33] Chen H X, Sun S C, Norenburg J L, *et al*. Mutation and selection cause codon usage and bias in mitochondrial genomes of ribbon worms (Nemertea) [J]. *PLoS One*, 2014, 9(1): e85631.
- [34] Trotta E. Selection on codon bias in yeast: A transcriptional hypothesis [J]. *Nucleic Acids Res*, 2013, 41(20): 9382-9395.
- [35] 杜雨, 李效雄, 贾西贝, 等. 枸杞叶绿体基因组密码子偏好性分析 [J]. *中草药*, 2024, 55(4): 1316-1325.
- [36] 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.
- [37] 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.

[责任编辑 时圣明]