

• 药材与资源 •

肉苁蓉属植物叶绿体基因组密码子偏好性分析

冯 展¹, 江 媛³, 郑 燕¹, 缪雨静², 黄林芳^{2*}, 罗光明^{1*}

1. 江西中医药大学药学院, 江西 南昌 330004

2. 中国医学科学院北京协和医学院 药用植物研究所, 北京 100193

3. 大理大学药学院, 云南 大理 671000

摘 要: **目的** 明确肉苁蓉属药用植物叶绿体基因组密码子使用偏性及影响因素。**方法** 以肉苁蓉 *Cistanche deserticola*、盐生肉苁蓉 *C. salsa*、沙苁蓉 *C. sinensis* 和管花肉苁蓉 *C. tubulosa* 为研究材料, 利用 CUSP 在线软件、CodonW 1.4.2、SPSS、Microsoft Excel 等软件分析其基因密码子使用偏性参数。**结果** 4 个肉苁蓉属植物叶绿体基因组的密码子使用模式相似, 密码子均以第 3 位碱基 A/T 结尾且密码子使用更偏向 A/T 碱基; 4 个物种的有效密码子数值 (effective number of codon, ENC) 均大于 35, 说明肉苁蓉属物种叶绿体基因密码子偏好性较弱。中性绘图分析、有效密码字数 (ENC-plot) 分析、奇偶偏好性 (PR2-plot) 分析和对应性分析的结果均说明自然选择是影响肉苁蓉属叶绿体基因组密码子偏好性的主要因素。利用同义密码子的相对使用频率 (relative synonymous codon usage, RSCU) 值筛选出肉苁蓉属 4 个物种共同拥有 4 个最优密码子。**结论** 对肉苁蓉属 4 个物种叶绿体基因组密码子使用偏性进行分析, 揭示了影响其密码子偏好性的影响因素, 为进一步探讨肉苁蓉属植物的系统进化、环境适应性和品种改良提供了理论基础。

关键词: 肉苁蓉属; 肉苁蓉; 盐生肉苁蓉; 沙苁蓉; 管花肉苁蓉; 叶绿体基因组; 密码子偏好性; 最优密码子

中图分类号: R286.12 **文献标志码:** A **文章编号:** 0253-2670(2023)05-1540-11

DOI: 10.7501/j.issn.0253-2670.2023.05.021

Codon use bias analysis of chloroplast genome of *Cistanche*FENG Zhan¹, JIANG Yuan³, ZHENG Yan¹, MIAO Yu-jing², HUANG Lin-fang², LUO Guang-ming¹

1. School of Pharmacy, Jiangxi University of Chinese Medicine, Nanchang 330004, China

2. Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China

3. School of Pharmacy, Dali University, Dali 671000, China

Abstract: Objective To clarify the codon usage bias and influencing factors of chloroplast genome in four medicinal plants of *Cistanche* spp. **Methods** The codon usage bias parameters of *Cistanche deserticola*, *C. salsa*, *C. sinensis* and *C. tubulosa* were analyzed by CUSP, CodonW 1.4.2, SPSS and Microsoft Excel. **Results** The chloroplast genomes of four *Cistanche* species had similar codon usage patterns, with the third base of the codons all ending in A/T and tending to use the A/T bases more favorably. The ENC (effective number of codon) values of the four species were all above 35, indicating that the codon preference of the chloroplast genes in *Cistanche* spp. is weak. The results of neutrality plot, ENC-plot, PR2-bias plot and correspondence analysis indicated that natural selection was the main factor influencing chloroplast genome codon usage bias in *Cistanche* spp. The RSCU (relative synonymous codon usage) values were used to identify four optimal codons shared by four species of *Cistanche* spp. **Conclusion** In this study, we analyzed the codon usage bias of chloroplast genomes of four species of *Cistanche* spp. and revealed the influencing factors affecting codon bias, which provided the corresponding theoretical basis for studying the systematic evolution, environmental adaptation and breed improvement of *Cistanche* spp. at molecular level.

收稿日期: 2022-08-09

基金项目: 国家自然科学基金资助项目 (82274045); 国家自然科学基金资助项目 (82073960); 国家自然科学基金资助项目 (U1812403-1); 国家外专局一带一路专项 (T2017052)

作者简介: 冯 展, 硕士研究生, 主要从事药用资源研究。Tel: 18280305532 E-mail: fengzhan1998@163.com

***通信作者:** 罗光明, 博士生导师, 教授, 主要从事中药鉴定及中药资源研究。Tel: (0791) 871189825 E-mail: jzlgm88@163.com

黄林芳, 博士生导师, 研究员, 主要从事中药鉴定及中药资源研究。Tel: 15801545922 E-mail: lfhuang@implad.ac.cn

Key words: *Cistanche Hoffmanns.* & Link; *Cistanche deserticola* Ma; *Cistanche salsa* (C. A. Mey.) G. Beck; *Cistanche sinensis* G. Beck; *Cistanche mongolica* Beck; chloroplast genome; codon usage bias; optimal codons

肉苁蓉属 *Cistanche* Hoffmanns. & Link 是列当科 (Orobanchaceae) 的多年生寄生草本植物, 主要分布于欧、亚洲, 中国主要有 4 种肉苁蓉属植物, 分布在内蒙古、宁夏、甘肃、青海以及新疆等地^[1]。本属植物在中国药用植物中占有及其重要的地位, 其中肉苁蓉 *Cistanche deserticola* Ma 的药用价值最高, 因其主要生长在沙漠地区, 素有“沙漠人参”的美誉。然而, 肉苁蓉属药用植物面临着植物分类混乱^[2]、市售品种混用^[3]的困境。

叶绿体是大多数绿色植物的光合作用场所, 参与发育和次级代谢活动^[4], 并协调细胞器和核基因组之间的基因表达^[5]。叶绿体具有自主遗传的基因组, 被广泛用于植物系统发育分析, 物种鉴定和遗传多样性表达等研究。近年来, 随着叶绿体基因组高通量测序技术逐渐成熟, 肉苁蓉属植物肉苁蓉、盐生肉苁蓉 *C. salsa* (C. A. Mey.) G. Beck、沙苁蓉 *C. sinensis* G. Beck、管花肉苁蓉 *C. tubulosa* Wight 等多种植物均已开展叶绿体测序研究, 并对其系统发育和遗传多样性进行了分析^[2, 6-7], 但尚未发现有关肉苁蓉属植物叶绿体基因组密码子偏好性方面的研究报道。

密码子 (codon) 又称遗传密码, 是核酸与蛋白质联系的桥梁, 是识别和传递生物遗传信息的重要载体, 在生物遗传和变异中至关重要的一部分^[8]。由于不同的物种其蛋白质翻译过程存在差异, 在翻译过程中倾向于使用一种或几种特定的同义密码子, 这种现象称为密码子偏好性 (codon usage bias, CUB)^[9], 密码子偏好性在 mRNA 翻译、DNA 转录、蛋白的结构、表达、功能和共翻译折叠等细胞代谢过程中发挥着重要作用^[10]。石岩硕等^[11]通过分析 4 种人参属 *Panax* Linn. 植物叶绿体基因组密码子偏好性, 推断出同属植物间亲缘关系更近; 宋芸等^[12]证明可基于密码子偏好性对 ICE1 基因进行优化, 使得其在低温胁迫下表达; 李显煌等^[13]发现密码子偏好性可反映出物种间的进化关系; 张俊焱等^[14]证明突变和自然选择共同影响獐牙菜 *Swertia bimaculate* (Sieb. et Zucc.) Hook. f. et Thoms. ex C. B. Clark 叶绿体基因组密码子偏好性, 为獐牙菜引种驯化提供依据。因此, 研究植物叶绿体基因组的密码子使用模式, 可为提高基因表达载体构建

效率、探讨物种进化关系、理解生物与环境适应的分子机制、改良植物品种等方面提供数据支持^[15]。

本研究基于高通量测序对肉苁蓉、盐生肉苁蓉、沙苁蓉、管花肉苁蓉四种肉苁蓉属植物的叶绿体基因组密码子的使用模式进行了分析, 确定 4 种肉苁蓉属的叶绿体基因组密码子偏好性及其影响因素, 为后续肉苁蓉属叶绿体基因组的基因表达检测、适应性及进化关系研究提供参考依据。

1 材料

肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉植株由不同地区采集 (表 1), 经北京协和医学院药用植物研究所黄林芳研究员鉴定为肉苁蓉属植物肉苁蓉 *C. deserticola* Ma、盐生肉苁蓉 *C. salsa* (C. A. Mey.) G. Beck、沙苁蓉 *C. sinensis* G. Beck 和管花肉苁蓉 *C. mongolica* Beck 的全株植物, 储存于北京协和医学院药用植物研究所植物标本室, 凭证标本为 CMPB16201、CMPB16202、CMPB16203 和 CMPB16204。叶绿体全基因组序列已上传至 GenBank, 收录号分别为 MN614127 (肉苁蓉)、MN614128 (盐生肉苁蓉)、MN614129 (沙苁蓉) 和 MN614130 (管花肉苁蓉)。肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉的叶绿体基因组大小分别为 109 495、111 710、111 500、75 375 bp, 分别包括 60、61、60、53 个蛋白质编码基因序列 (coding DNA sequence, CDS)。由于短序列无法准确地计算有效密码子数, 剔除了长度小于 300 bp 的 CDS^[16], 又因为终止密码子 UAA、UGA 和 UAG 不编码任何氨基酸, 并且 UGG 和 AUG 分别是色氨酸和蛋氨酸的唯一密码子, 这些密码子不存在偏好性^[17], 故选择的 CDS 以 ATG 为起始^[18], 以 TAA、TAG、TGA 为结尾, 最后为了降低结果的误差, 剔除了重复 CDS。处理后每个物种留存 18 条 CDS, 并用于后续分析。

表 1 肉苁蓉样品信息
Table 1 Information of samples

样品	采样地	采样时间
肉苁蓉	内蒙古阿拉善沙漠	2017-05
盐生肉苁蓉	新疆草原裕民	2017-05
沙苁蓉	宁夏银川	2018-04
管花肉苁蓉	新疆和田	2017-05

2 方法

2.1 密码子偏好性分析

将挑选出的 72 条 CDS 整合到一个 fasta 文件中,利用线上软件 CUSP(<https://www.bioinformatics.nl/cgi-bin/emboss/cusp>) 计算每个基因的编码区总 GC 含量 (GCall) 以及密码子第 1 位 (GC1)、第 2 位 (GC2)、第 3 位 (GC3) 的 GC 含量。利用 CodonW1.4.2 软件^[19]计算有效密码子数 (effective number of codon, ENC) 和同义密码子的相对使用频率 (relative synonymous codon usage, RSCU), 并进行对应性分析 (correspondence analysis, COA)。然后通过 Microsoft Excel 和 IBM SPSS19.0 软件对分析数据进行整理和绘图。

2.2 中性绘图分析

中性绘图分析是判断密码子偏好性影响因素的方法之一^[20]。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。以每个基因的 GC1 和 GC2 的平均值 (GC12) 为纵坐标、以 GC3 为横坐标在 Microsoft Excel 软件中绘制散点图,剖析 GC3 与 GC12 的相关性,其中每个散点则代表 1 个基因。如果所有的点都沿对角线分布,说明密码子 3 个位置的碱基没有明显差异,只受突变压力的影响;如果 GC12 和 GC3 变异的相关性很低,则表明自然选择是密码子使用模式的主要影响因素^[21]。

2.3 ENC-plot 分析

ENC 代表了密码子偏离随机选择的程度^[22],以 ENC 为纵坐标,GC3 为横坐标在 Microsoft Excel 软件中进行二维散点图的绘制,以公式 $ENC = 2 + GC_3 + 29/[GC_3^2 + (1 - GC_3)^2]$ 绘制标准曲线^[23],每个散点代表 1 个基因。散点图可反映出密码子使用偏好性和基因碱基组成之间的关系,如果密码子使用偏好性主要受到突变压力的影响,散点将位于标准曲线上或略低于标准曲线,相反,如果基因低于标准曲线,则主要受到选择和其他因素的影响^[24]。

2.4 PR2-plot 分析

PR2 偏倚分析 (PR2-bias plot analysis) 对密码子第 3 位碱基上的 A、T、C、G 含量进行分析,避免了密码子第 3 位碱基 A/T 和 C/G 之间的突变不平衡。分别以 $G_3/(G_3 + C_3)$ 为横坐标、以 $A_3/(A_3 + T_3)$ 为纵坐标在 Microsoft Excel 软件中制作散点图。通过判断散点与中心点 (A=T, C=G) 的矢量距离,来确定碱基偏移的程度和方向^[25]。

2.5 COA

COA 从多维空间中提取最具影响的轴和方向,被广泛用于研究基因间密码子使用变化的主要趋势^[26],通过分析 COA 轴分离出的基因能识别影响密码子使用偏性的主要原因^[27]。利用 CodonW1.4.2 软件进行 COA,根据肉苁蓉属四种植物的 RSCU 值,将所有基因密码子绘制在 59 维的向量空间中,以反映密码子使用变化的趋势。轴 1 (Axis 1) 的占比代表对密码子使用频率变化的最大影响,其余 58 个轴代表了逐渐减小的因素。根据基因在多维空间中的同义密码子使用情况,可以分析其主要变异来源^[28]。

2.6 最优密码子分析

将 4 种植物密码子的 ENC 值作为参考指标,取 ENC 值最高 (10%) 和最低 (10%) 的基因序列分别建库,作为高表达组和低表达组。利用 CodonW 1.4.2 软件计算 2 个表达库的 RSCU 值和 $\Delta RSCU$ 值 ($\Delta RSCU = RSCU_{\text{高表达}} - RSCU_{\text{低表达}}$)。其中, $\Delta RSCU \geq 0.08$ 的密码子作为高表达密码子, $RSCU > 1$ 的为高频密码子,同时满足 2 个条件的密码子为最优密码子^[29]。

3 结果与分析

3.1 叶绿体基因的密码子偏好性分析

肉苁蓉属 4 个物种叶绿体基因组的 GC 含量见图 1。肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉的 18 个叶绿体基因整个编码区的 GC 含量 (GCall) 范围在 36.60%~37.22%,叶绿体基因组第 1、第 2 和第 3 位密码子的 GC 含量分别为 44.03%~44.75%, 38.97%~39.82% 和 25.80%~27.08%。结果显示,4 个物种的平均 GC 含量均小于 50%,这说明 4 个物种的叶绿体基因组倾向于使用 A/T 碱基和 A/T 结尾的密码子。ENC 值从 20 (极偏好,每个氨基酸只使用 1 个密码子) 到 61 (无偏好,每个氨基酸均匀使用所有同义密码子),当 ENC 值 ≤ 35 时,认为密码子使用有非常显著的偏好^[22, 30],4 个物种的叶绿体基因组编码区的 ENC 值均大于 35,说明肉苁蓉属物种叶绿体基因密码子偏好性较弱。

4 个肉苁蓉属物种各叶绿体基因的 GC1、GC2、GC3、GCall、ENC 和密码子数 (codon counts, CC) 相关性分析见图 2。4 个物种的 GCall 均与 GC1 和 GC2 呈极显著相关 ($P < 0.01$),其中盐生肉苁蓉与管花肉苁蓉的 GC1 和 GC2 呈显著相关 ($P < 0.05$),表明这 2 个物种密码子第 1 位碱基和第 2 位碱基组成相似,

但与第 3 位碱基组成存在显著差别；肉苁蓉的 ENC 值与 GC3 呈显著相关，盐生肉苁蓉的 ENC 值与 GC1 和 GC2 呈显著相关，说明在肉苁蓉中，密码子第 3 位碱基的组成对密码子使用偏性有重要影响，在盐生肉苁蓉中密码子第 1、3 位碱基的组成对密码子使用偏性有重要影响，而在沙苁蓉和管花肉苁蓉中，碱基的组成对密码子使用偏性并无显著影响；4 个物种的基因密码子数 CC 与 ENC 值之间的相关系数介于 0.18（盐生肉苁蓉）~0.27（沙苁蓉），相关性均未达到显著水平，这说明 CC 对 ENC 的影响较弱，叶绿体基因的长度对密码子使用偏性无影响。

为了进一步探究密码子使用模式，分别计算四个物种的 RSCU 值（表 2），结果显示 4 个物种的 RSCU 值相差不大。RSCU>1 表明该密码子使用偏

性较高，反之，则为非偏好密码子。肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉中，RSCU>1 的氨基酸占比分别为 48.4%、50%、51.6%和 46.9%，其中以 A/U 结尾的氨基酸的个数分别为 27，27，28 和 26，这说明肉苁蓉属叶绿体基因组中 A 或 U 出现频率较高，是该属叶绿体基因组偏好密码子。RSCU<1 的氨基酸中，多数以 G/C 结尾，表明这些是肉苁蓉属叶绿体基因的非偏好密码子。

3.2 中性绘图分析

肉苁蓉属叶绿体基因中性绘图见图 3，肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉的 GC12 与 GC3 的相关系数分别 0.083、0.084、0.033 和 0.245，且相关性均未达到显著水平（ $P=0.05$ ），说明突变压力不会对对密码子使用偏性产生明显的影响；中性

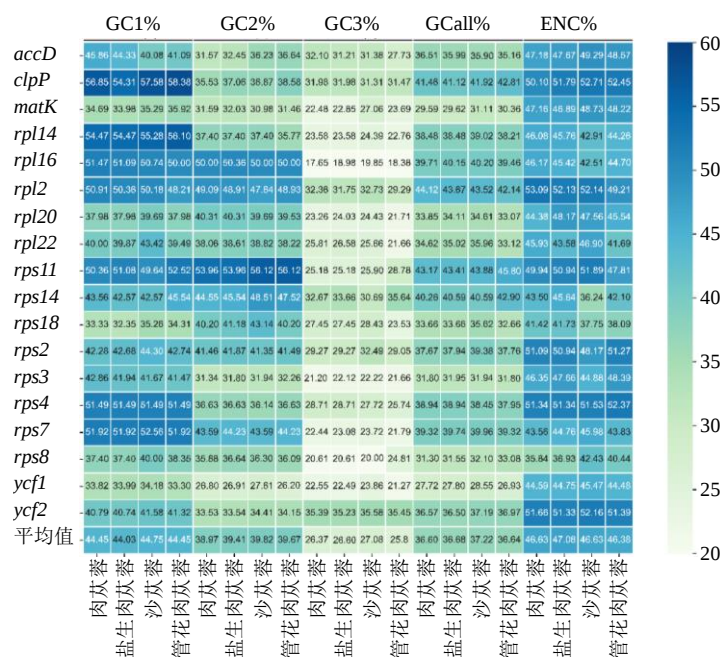


图 1 肉苁蓉属 4 个物种的叶绿体基因组密码子不同位置的 GC 含量

Fig. 1 GC content of different positions of codon in chloroplast genome of *Cistanche*

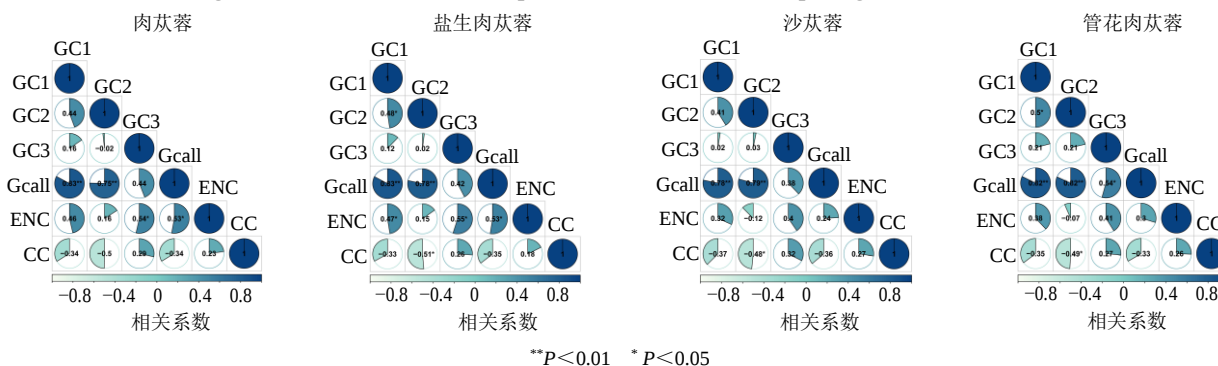


图 2 肉苁蓉属叶绿体基因组相关参数的相关性分析

Fig. 2 Correlation analysis of chloroplast genome-related parameters of *Cistanche*

表 2 肉苁蓉属叶绿体同义密码子的使用频率
Table 2 RSCU analysis of protein coding region in chloroplast of *Cistanche*

氨基酸 密码子		使用频率/%				氨基酸 密码子		使用频率/%			
		荒漠肉苁蓉	盐生肉苁蓉	沙苁蓉	管花肉苁蓉			荒漠肉苁蓉	盐生肉苁蓉	沙苁蓉	管花肉苁蓉
Phe	UUU	1.32	1.31	1.28	1.38	Tyr	UAU	1.38	1.34	1.40	1.45
	UUC	0.68	0.69	0.72	0.62		UAC	0.62	0.66	0.60	0.55
Leu	UUA	1.53	1.44	1.39	1.61	TER	UAA	1.37	1.30	1.18	1.36
	UUG	1.17	1.21	1.23	1.19		UAG	0.68	0.74	0.79	0.77
	CUU	1.19	1.22	1.24	1.15	His	CAU	1.41	1.33	1.37	1.42
	CUC	0.74	0.76	0.75	0.60		CAC	0.59	0.67	0.63	0.58
	CUA	0.82	0.87	0.87	0.93	Gln	CAA	1.41	1.37	1.39	1.39
	CUG	0.54	0.50	0.51	0.52		CAG	0.59	0.63	0.61	0.61
Ile	AUU	1.29	1.26	1.20	1.32	Asn	AAU	1.43	1.40	1.41	1.40
	AUC	0.69	0.69	0.71	0.64		AAC	0.57	0.60	0.59	0.60
	AUA	1.03	1.05	1.09	1.04	Lys	AAA	1.43	1.42	1.39	1.42
	AUG	1.00	1.00	1.00	1.00		AAG	0.57	0.58	0.61	0.58
Val	GUU	1.43	1.36	1.35	1.31	Asp	GAU	1.45	1.38	1.47	1.43
	GUC	0.76	0.73	0.71	0.67		GAC	0.55	0.62	0.53	0.57
	GUA	1.21	1.20	1.18	1.3	Glu	GAA	1.33	1.36	1.36	1.37
	GUG	0.60	0.71	0.75	0.72		GAG	0.67	0.64	0.64	0.63
Ser	UCU	1.51	1.46	1.54	1.50	Cys	UGU	1.25	1.25	1.16	1.21
	UCC	1.01	1.05	1.06	0.97		UGC	0.75	0.75	0.84	0.79
	UCA	1.10	1.27	1.31	1.32	TER	UGA	0.95	0.95	1.04	0.88
	UCG	0.77	0.72	0.64	0.64		UGG	1.00	1.00	1.00	1.00
Pro	CCU	1.20	1.12	1.18	1.23	Arg	CGU	0.72	0.68	0.66	0.72
	CCC	1.01	1.06	1.01	1.01		CGC	0.44	0.44	0.45	0.38
	CCA	1.15	1.17	1.19	1.17		CGA	1.04	1.00	1.01	0.97
	CCG	0.64	0.64	0.62	0.59		CGG	0.62	0.64	0.66	0.58
Thr	ACU	1.25	1.14	1.19	1.13	Ser	AGU	1.00	0.91	0.84	0.98
	ACC	0.96	1.03	1.01	0.99		AGC	0.62	0.59	0.61	0.58
	ACA	1.16	1.24	1.19	1.30	Arg	AGA	2.02	2.13	2.00	2.20
	ACG	0.64	0.60	0.61	0.58		AGG	1.16	1.1	1.21	1.14
Ala	GCU	1.37	1.31	1.19	1.33	Gly	GGU	0.99	1.02	0.94	0.96
	GCC	0.91	0.94	0.99	0.96		GGC	0.64	0.66	0.66	0.59
	GCA	1.11	1.12	1.17	1.07		GGA	1.44	1.38	1.42	1.36
	GCG	0.61	0.63	0.66	0.64		GGG	0.93	0.93	0.98	1.08

绘图斜率显示, 4 个叶绿体基因组密码子使用模式的突变压力占 4.44%~32%, 这说明 4 个叶绿体基因组密码子使用模式的突变压力占比小, 自然选择是影响密码子使用偏性的主要因素。

3.3 ENC-plot 绘图分析

肉苁蓉属 4 个物种的 ENC-plot 绘图见图 4, 大部

分基因与标准曲线有一定的距离, 只有小部分基因分布在标准曲线上, 这说明 ENC 的实际值与理论值存在一定的偏差且偏性较弱。大部分基因分布在标准曲线下方, 说明基因偏性主要受到自然选择的影响。

3.4 PR2-plot 绘图分析

肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁

蓉中各基因不均匀分布在 PR2-plot 绘图的 4 个平面区域中 (图 5), 结果显示, 大多数基因远离中心, 分布在右下角, 值得注意的是几乎没

有基因分布在左下角, 这表明 GC3 的使用主要受到自然选择影响, 且 4 种植物的 G3 存在 T/G 使用偏性。

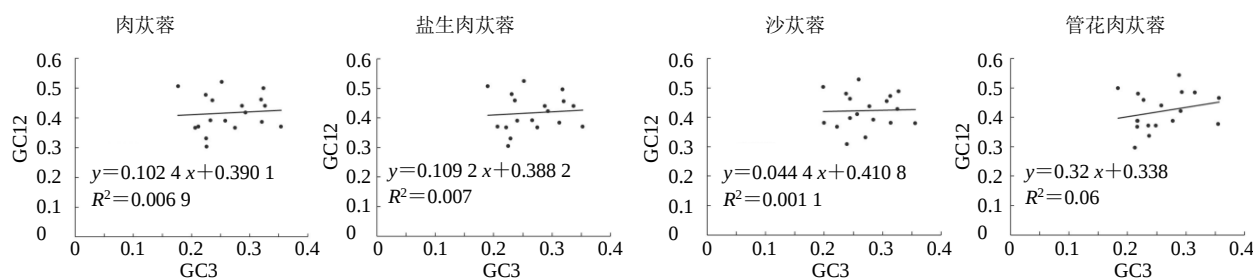


图 3 肉苁蓉属中性绘图分析

Fig. 3 Analysis of neutrality plot

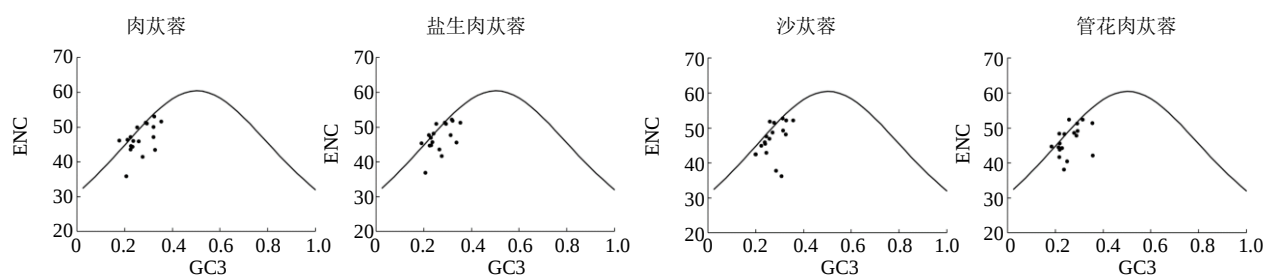


图 4 ENC-plot 分析

Fig. 4 ENC-plot analysis

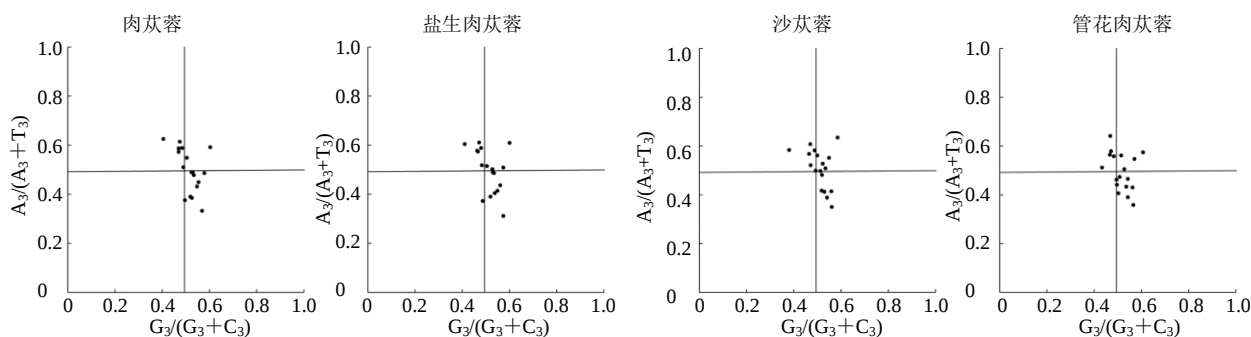


图 5 PR2-plot 分析

Fig. 5 PR2-plot analysis

3.5 对应性分析

为了探究肉苁蓉属植物叶绿体密码子变异源头, 对肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉叶绿体基因组密码子进行了基于 RSCU 值的对应性分析 (图 6), 前 4 轴的累计可解释变异分别占总变异的 52.97%、53.96%、53.29%和 52.18%, 第 1 轴的可解释变异分别占总变异的 17.89%、18.70%、17.39%和 17.57%, 第 2~4 轴的可解释变异分别为肉苁蓉 (12.87%、11.35%和 10.87%)、盐生肉苁蓉 (12.91%、12.00%和 10.34%)、沙苁蓉

(13.10%、12.07%和 10.73%) 和管花肉苁蓉 (12.12%、11.65%和 10.83%)。4 个物种的第 2、3、4 轴的可解释变异均小于第 1 轴, 因此第 1 轴对密码子偏好性的贡献率最大, 但第 1 轴仅代表肉苁蓉属叶绿体基因组中密码子使用偏性的部分变异, 这说明肉苁蓉属叶绿体的密码子使用偏性的影响因素并非一个, 还可能与突变、自然选择、基因长度和功能等因素有关。

3.6 最优密码子分析

肉苁蓉属 4 个物种的最优密码子分析见表 3。

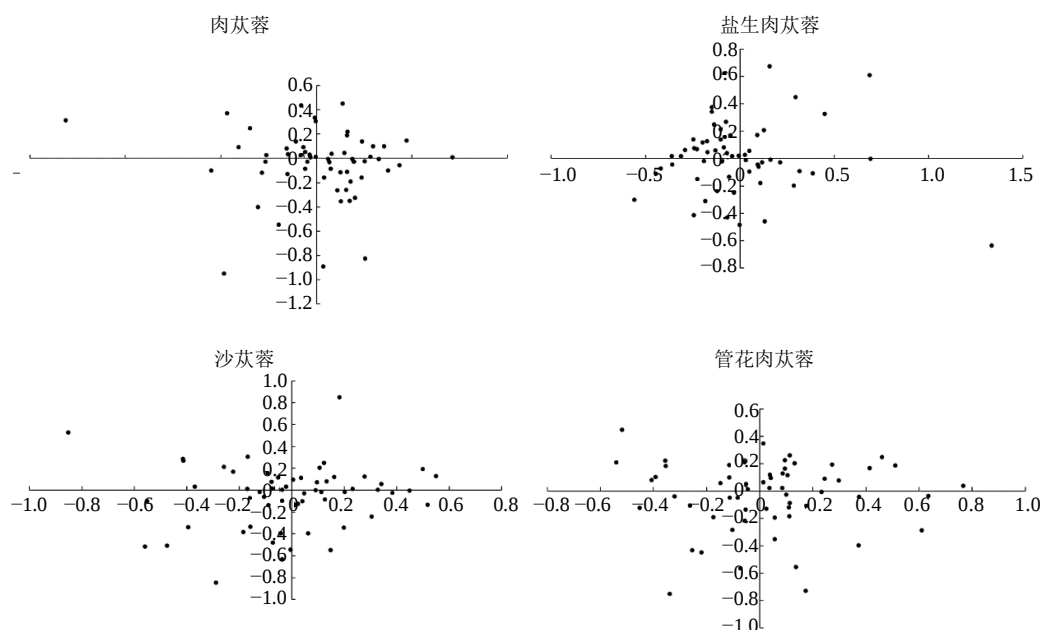


图6 对应性分析

Fig. 6 Corresponding analysis

表3 最优密码子分析

Table 3 Optimal codons in chloroplast genome of *Cistanche*

氨基酸	肉苁蓉				盐生肉苁蓉				沙苁蓉				管花肉苁蓉			
	密码子	低 RSCU	高 RSCU	Δ RSCU	密码子	低 RSCU	高 RSCU	Δ RSCU	密码子	低 RSCU	高 RSCU	Δ RSCU	密码子	低 RSCU	高 RSCU	Δ RSCU
Phe	<u>UUU</u>	2.00	1.14	-0.86	<u>UUU</u>	2.00	1.54	-0.46	<u>UUU</u>	1.75	1.09	-0.66	<u>UUU</u>	1.80	1.75	-0.05
	UUC***	0.00	0.86	0.86	UUC**	0.00	0.46	0.46	UUC***	0.25	0.91	0.66	UUC	0.20	0.25	0.05
Leu	<u>UUA</u>	2.50	1.42	-1.08	<u>UUA</u>	2.45	1.62	-0.83	<u>UUA</u>	1.83	0.92	-0.91	<u>UUA</u>	2.61	2.05	-0.56
	<u>UUG</u>	1.75	1.26	-0.49	<u>UUG</u>	1.64	0.97	-0.67	<u>UUG</u>	2.61	1.48	-1.13	<u>UUG</u>	1.57	0.68	-0.89
	<u>CUU</u>	1.00	0.95	-0.05	<u>CUU</u> **	1.09	1.46	0.37	<u>CUU</u> **	1.04	1.48	0.44	<u>CUU</u> **	1.04	1.36	0.32
	CUC***	0.25	1.26	1.01	CUC***	0.27	0.81	0.54	CUC*	0.26	0.54	0.28	CUC	0.26	0.27	0.01
	CUA**	0.50	0.95	0.45	CUA*	0.55	0.65	0.10	CUA***	0.26	0.95	0.69	CUA*	0.52	0.68	0.16
Ile	CUG*	0.00	0.16	0.16	CUG**	0.00	0.49	0.49	CUG***	0.00	0.62	0.62	CUG***	0.00	0.95	0.95
	<u>AUU</u>	1.34	1.18	-0.16	<u>AUU</u>	1.50	0.98	-0.52	<u>AUU</u>	2.14	1.32	-0.82	<u>AUU</u>	1.55	1.08	-0.47
	AUC*	0.52	0.79	0.27	AUC**	0.43	0.82	0.39	AUC***	0.00	0.75	0.75	AUC**	0.52	0.84	0.32
	<u>AUA</u>	1.14	1.03	-0.11	<u>AUA</u> *	1.07	1.20	0.13	<u>AUA</u>	0.86	0.93	0.07	<u>AUA</u> *	0.93	1.08	0.15
Met	AUG	1.00	1.00	0.00	AUG	1.00	1.00	0.00	AUG	1.00	1.00	0.00	AUG	1.00	1.00	0.00
Val	<u>GUU</u>	3.43	1.33	-2.10	<u>GUU</u>	3.43	1.68	-1.75	<u>GUU</u>	2.00	1.48	-0.52	<u>GUU</u>	4.00	1.75	-2.25
	GUC**	0.00	0.44	0.44	GUC*	0.00	0.26	0.26	GUC***	0.00	0.64	0.64	GUC***	0.00	0.50	0.50
	<u>GUA</u> ***	0.00	1.93	1.93	<u>GUA</u> ***	0.00	1.68	1.68	<u>GUA</u> ***	0.00	1.16	1.16	<u>GUA</u> ***	0.00	1.38	1.38
	GUG	0.57	0.30	-0.27	GUG	0.57	0.39	-0.18	GUG	2.00	0.72	-1.28	GUG**	0.00	0.38	0.38
Ser	<u>UCU</u>	2.40	0.91	-1.49	<u>UCU</u>	2.73	1.62	-1.11	<u>UCU</u> ***	0.90	1.69	0.79	<u>UCU</u>	2.33	1.83	-0.50
	<u>UCC</u> **	0.90	1.27	0.37	<u>UCC</u>	0.82	0.69	-0.13	<u>UCC</u>	1.20	1.07	-0.13	UCC	0.67	0.26	-0.41
	<u>UCA</u> ***	0.00	1.09	1.09	<u>UCA</u> ***	0.00	0.69	0.69	<u>UCA</u> **	0.90	1.30	0.40	<u>UCA</u> ***	0.33	1.83	1.50
	UCG	0.60	0.55	-0.05	UCG*	0.55	0.69	0.14	UCG	0.60	0.67	0.07	UCG	0.67	0.52	-0.15
Pro	<u>CCU</u> **	1.33	1.82	0.49	<u>CCU</u> ***	1.33	2.24	0.91	<u>CCU</u> ***	0.73	1.31	0.58	<u>CCU</u> *	1.71	1.80	0.09
	<u>CCC</u>	2.00	0.91	-1.09	<u>CCC</u>	2.00	0.64	-1.36	<u>CCC</u> *	0.73	0.91	0.18	<u>CCC</u>	1.71	0.80	-0.91
	<u>CCA</u> ***	0.00	0.91	0.91	<u>CCA</u> ***	0.00	0.96	0.96	<u>CCA</u>	2.18	1.07	-1.11	<u>CCA</u> ***	0.00	1.20	1.20
	CCG	0.67	0.36	-0.31	CCG	0.67	0.16	-0.51	CCG**	0.36	0.71	0.35	CCG	0.57	0.20	-0.37

续表 3

氨基酸	肉苁蓉				盐生肉苁蓉				沙苁蓉				管花肉苁蓉			
	密码子	低 RSCU	高 RSCU	Δ RSCU	密码子	低 RSCU	高 RSCU	Δ RSCU	密码子	低 RSCU	高 RSCU	Δ RSCU	密码子	低 RSCU	高 RSCU	Δ RSCU
Thr	<u>ACU</u>	1.45	0.76	-0.69	<u>ACU</u>	1.09	0.35	-0.74	<u>ACU</u> ***	0.36	1.62	1.26	<u>ACU</u>	1.45	0.75	-0.70
	ACC	1.82	1.33	-0.49	<u>ACC</u>	2.18	1.57	-0.61	<u>ACC</u>	2.18	0.58	-1.60	ACC	1.82	1.25	-0.57
	<u>ACA</u> ***	0.73	1.52	0.79	<u>ACA</u> ***	0.73	1.57	0.84	<u>ACA</u> *	1.09	1.19	0.10	<u>ACA</u> ***	0.73	1.25	0.52
	ACG**	0.00	0.38	0.38	ACG***	0.00	0.52	0.52	ACG*	0.36	0.62	0.26	ACG***	0.00	0.75	0.75
Ala	<u>GCU</u>	3.43	0.63	-2.80	<u>GCU</u>	3.33	0.75	-2.58	<u>GCU</u>	2.29	1.68	-0.61	<u>GCU</u>	3.33	0.91	-2.42
	GCC***	0.00	1.05	1.05	GCC***	0.00	1.00	1.00	GCC***	0.00	0.64	0.64	GCC***	0.00	1.27	1.27
	<u>GCA</u> ***	0.57	1.26	0.69	<u>GCA</u> ***	0.67	1.63	0.96	<u>GCA</u>	1.71	1.28	-0.43	<u>GCA</u> ***	0.67	1.45	0.78
	GCG***	0.00	1.05	1.05	GCG***	0.00	0.63	0.63	GCG**	0.00	0.41	0.41	GCG**	0.00	0.36	0.36
Tyr	<u>UAU</u>	2.00	1.33	-0.67	<u>UAU</u>	2.00	1.29	-0.71	<u>UAU</u>	2.00	1.59	-0.41	<u>UAU</u>	2.00	1.38	-0.62
	UAC***	0.00	0.67	0.67	UAC***	0.00	0.71	0.71	UAC**	0.00	0.41	0.41	UAC***	0.00	0.63	0.63
TER	<u>UAA</u>	1.50	0.00	-1.50	<u>UAA</u>	1.50	1.50	0.00	<u>UAA</u> ***	1.50	3.00	1.50	<u>UAA</u>	1.50	1.50	0.00
	UAG***	1.50	3.00	1.50	UAG	1.50	1.50	0.00	UAG	1.50	0.00	-1.50	UAG	1.50	0.00	-1.50
His	<u>CAU</u>	2.00	1.67	-0.33	<u>CAU</u>	2.00	1.69	-0.31	<u>CAU</u>	1.50	1.45	-0.05	<u>CAU</u>	2.00	1.50	-0.50
	CAC**	0.00	0.33	0.33	CAC**	0.00	0.31	0.31	CAC	0.50	0.55	0.05	CAC***	0.00	0.50	0.50
Gln	<u>CAA</u>	1.60	1.00	-0.60	<u>CAA</u>	1.60	1.54	-0.06	<u>CAA</u>	1.50	1.43	-0.07	<u>CAA</u> **	1.56	1.89	0.33
	CAG***	0.40	1.00	0.60	CAG	0.40	0.46	0.06	CAG	0.50	0.57	0.07	CAG	0.44	0.11	-0.33
Asn	<u>AAU</u> **	1.29	1.69	0.40	<u>AAU</u> *	1.38	1.56	0.18	<u>AAU</u> *	1.43	1.55	0.12	<u>AAU</u>	1.23	1.00	-0.23
	AAC	0.71	0.31	-0.40	AAC	0.63	0.44	-0.19	AAC	0.57	0.45	-0.12	AAC*	0.77	1.00	0.23
Lys	<u>AAA</u>	1.68	1.52	-0.16	<u>AAA</u>	1.70	1.66	-0.04	<u>AAA</u>	1.60	1.33	-0.27	<u>AAA</u>	1.74	1.73	-0.01
	AAG*	0.32	0.48	0.16	AAG	0.30	0.34	0.04	AAG*	0.40	0.67	0.27	AAG	0.26	0.27	0.01
Asp	<u>GAU</u>	1.60	1.17	-0.43	<u>GAU</u>	1.20	1.18	-0.02	<u>GAU</u> ***	0.67	1.63	0.96	<u>GAU</u>	1.67	1.65	-0.02
	GAC**	0.40	0.83	0.43	GAC	0.80	0.82	0.02	GAC	1.33	0.37	-0.96	GAC	0.33	0.35	0.02
Glu	<u>GAA</u>	1.60	1.22	-0.38	<u>GAA</u>	1.40	1.33	-0.07	<u>GAA</u>	1.60	1.30	-0.30	<u>GAA</u> **	1.27	1.76	0.49
	GAG**	0.40	0.78	0.38	GAG	0.60	0.67	0.07	GAG**	0.40	0.70	0.30	GAG	0.73	0.24	-0.49
Cys	<u>UGU</u>	2.00	1.71	-0.29	<u>UGU</u>	2.00	1.20	-0.80	<u>UGU</u>	2.00	1.31	-0.69	<u>UGU</u>	2.00	0.67	-1.33
	UGC*	0.00	0.29	0.29	UGC***	0.00	0.80	0.80	UGC***	0.00	0.69	0.69	UGC***	0.00	1.33	1.33
TER	UGA	0.00	0.00	0.00	UGA	0.00	0.00	0.00	<u>UGA</u>	0.00	0.00	0.00	UGA***	0.00	1.50	1.50
Trp	UGG	1.00	1.00	0.00	UGG	1.00	1.00	0.00	UGG	1.00	1.00	0.00	UGG	1.00	1.00	0.00
Arg	CGU***	0.60	1.36	0.76	CGU***	0.58	1.20	0.62	CGU	1.16	0.73	-0.43	CGU***	0.56	1.94	1.38
	CGC**	0.20	0.55	0.35	CGC**	0.19	0.60	0.41	CGC	0.39	0.24	-0.15	CGC***	0.19	0.88	0.69
	<u>CGA</u>	2.60	0.82	-1.78	CGA	2.52	0.75	-1.77	<u>CGA</u>	1.55	1.18	-0.37	CGA	2.25	1.41	-0.84
	CGG*	0.00	0.14	0.14	CGG*	0.00	0.15	0.15	CGG**	0.19	0.61	0.42	CGG**	0.00	0.35	0.35
Ser	AGU	2.10	1.45	-0.65	AGU	1.91	1.85	-0.06	AGU	2.40	1.02	-1.38	AGU	2.00	1.30	-0.7
	AGC***	0.00	0.73	0.73	AGC**	0.00	0.46	0.46	AGC*	0.00	0.25	0.25	AGC*	0.00	0.26	0.26
Arg	<u>AGA</u> **	2.00	2.32	0.32	<u>AGA</u> **	2.13	2.55	0.42	<u>AGA</u> **	1.74	2.07	0.33	<u>AGA</u>	2.25	0.88	-1.37
	<u>AGG</u> *	0.60	0.82	0.22	<u>AGG</u> *	0.58	0.75	0.17	<u>AGG</u> *	0.97	1.18	0.21	<u>AGG</u>	0.75	0.53	-0.22
Gly	GGU	2.00	1.49	-0.51	<u>GGU</u>	2.00	1.25	-0.75	GGU**	0.44	0.75	0.31	GGU	1.23	0.92	-0.31
	GGC*	0.00	0.09	0.09	GGC**	0.00	0.31	0.31	GGC	0.44	0.41	-0.03	GGC	0.92	0.31	-0.61
	<u>GGA</u>	1.67	1.40	-0.27	<u>GGA</u>	1.67	1.65	-0.02	<u>GGA</u>	2.22	1.87	-0.35	<u>GGA</u> ***	1.54	2.15	0.61
	GGG***	0.33	1.02	0.69	GGG**	0.33	0.78	0.45	GGG*	0.89	0.97	0.08	<u>GGG</u> **	0.31	0.62	0.31

下划线表示 RSCU 值>1 的密码子; *表示 Δ RSCU ≥ 0.08 , **表示 Δ RSCU ≥ 0.3 , ***表示 Δ RSCU ≥ 0.5 ; 加粗标记为最优密码子

Underline: Codon usage with the value of RSCU>1; *: Δ RSCU ≥ 0.08 , **: Δ RSCU ≥ 0.3 , ***: Δ RSCU ≥ 0.5 ; Optimal codons are marked in overstriking

结果显示,肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉各有 10、11、13 和 12 个。其中肉苁蓉与盐生肉苁蓉最为相似,共有 9 个同样的密码子(图 7),4 个物种共有的密码子有 4 个(GUA、UCA、CCU 和 ACA)。在共同的最优密码子中, Δ RSCU 均大于 0.5 的仅有 1 个,为 GUA。

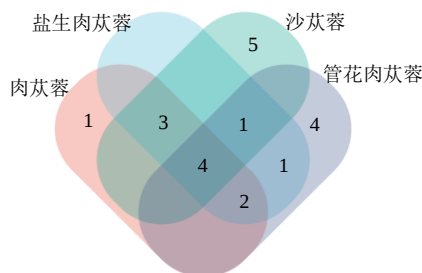


图 7 最优密码子个数韦恩图

Fig. 7 Venn diagram of optimal codons

4 讨论

密码子偏好性指的是在蛋白质的翻译过程中不同同义密码子使用频率的高低现象,是生物基因组进化的一个重要特征,密码子偏好性与 GC 含量、tRNA 丰度、基因表达水平和蛋白质结构等因素相关^[15],研究植物密码子的偏性对深入探究分子进化和外源蛋白表达有着及其重要的意义^[31-32]。在影响密码子偏好性的诸多因素中,突变和自然选择为主导因素,自然选择使得植物在翻译中优先编码最优密码子,突变会导致植物中部分非偏好性密码子的存在^[33-34],所以在此次研究中,密切关注了基因突变压力和自然选择压力。

本研究中肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉叶绿体基因组中 GCall 为 36.60%~37.22%,GC1、GC2 和 GC3 的含量分别为 44.03%~44.75%,38.97%~39.82%和 25.80%~27.08%,含量均在 50% 以下,表明 4 种肉苁蓉属植物的叶绿体基因第 3 位碱基多以 A/T 结尾或更偏向于 A/T 碱基,这与前人报道的大戟科^[32]、茄科^[35]、人参属^[11]、豆蔻属^[36]等植物叶绿体基因组研究一致,说明不同物种间叶绿体基因密码子使用偏好性存在着一定的相似性。

密码子第 3 位碱基的变化通常不会导致编码氨基酸的改变,因此研究密码子第 3 位碱基组成对研究密码子偏好性具有重要意义^[37]。对肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉叶绿体基因组的第 3 位碱基上的 A、T、C、G 含量进行 PR2 分析,结果显示: $T > A$ 、 $G > C$,这与小麦 *Triticum aestivum*

Linn.^[38]、剑麻 *Agave sisalana* Perr. ex Engelm.^[24]、黄芩 *Scutellaria baicalensis* Georgi^[39]、陆地棉 *Gossypium hirsutum* Linn.^[40]等植物表现出的 $T > A$ 、 $C > G$ 和胡杨 *Populus euphratica* Oliv.^[41]、禾本科(Poaceae Barnhart)植物^[42]表现出的 $A = T$ 、 $G = C$ 的结果不一致。由此可见,不同植物的叶绿体基因组中碱基组成情况存在明显差异,这导致了密码子偏好性有所不同。ENC 值是判断判断密码子偏好性强弱的标准,当 ENC 低于 35 时,密码子使用偏性较大,当 ENC 高于 35 时,偏性较小。在肉苁蓉属的四种植物种,叶绿体基因组编码区的 ENC 值均大于 35,说明肉苁蓉属物种叶绿体基因密码子偏好性较弱。在相关性分析中,4 个物种的 GCall 与 GC1 和 GC2 均呈极显著相关,除此之外,其余数值的相关性各有不同,由于这 4 种肉苁蓉属植物的生境不尽相同,可以推测生长环境的不同可能会导致不同物种间的 GC 差异。

中性绘图分析、ENC-plot 分析、PR2-plot 的结果均说明肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉叶绿体基因组密码子更偏向于受自然选择的影响,这与之前报道的豆蔻属 *Amomum* Roxb.^[36]、含笑属 *Michelia* Linn.^[43]和人参属^[11]植物结果一致,但蒺藜苜蓿 *Medicago truncatula* Gaertn.^[17]、拟南芥 *Arabidopsis thaliana* (L.) Heynh. 和杨树 *Populus przewalskii* Maxim.^[41]等植物的叶绿体基因组密码子偏好性则主要受突变的影响,这表明不同植物的密码子偏好性是多因素综合的结果。

最优密码子分析表明,肉苁蓉、盐生肉苁蓉、沙苁蓉和管花肉苁蓉的最优密码子数为 10、11、13 和 12 个,共有的最优密码子数个数为 4。值得注意的是,肉苁蓉与盐生肉苁蓉共享 9 个相同的最优密码子,是肉苁蓉属 4 个物种中最为相近的 2 个物种,可以推断出在肉苁蓉属内,肉苁蓉与盐生肉苁蓉具有相近的密码子使用偏好性,亲缘关系更为接近,这与缪雨静等^[44]和杨俏俏等^[2]通过肉苁蓉属的系统发育分析所得结论一致。

本研究首次基于叶绿体基因组密码子偏好性对国产肉苁蓉属进行了分析,系统性的阐明了影响密码子偏好性的因素,揭示了影响其密码子偏好性的因素与自然选择压力相关,筛选出了最优密码子,并且发现了属内肉苁蓉与盐生肉苁蓉亲缘关系更为接近。为后续的肉苁蓉属物种基因工程研究、资源鉴定、外源表达基因构建等提供了科学依据和理论支撑。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中国科学院中国植物志编辑委员会. 中国植物志(第四十二卷, 第一分册) [M]. 北京: 科学出版社, 1993: 26-29.
- [2] 杨俏俏. 中国肉苁蓉属植物叶绿体全基因组解析、比较基因组及系统发育研究 [D]. 北京: 北京协和医学院, 2019.
- [3] 张欣华, 高寒, 王嘉, 等. 内蒙古药品评价性抽验中肉苁蓉质量分析 [J]. 中国药业, 2022, 31(7): 85-88.
- [4] Wicke S, Schneeweiss G M, DePamphilis C W, et al. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function [J]. *Plant Mol Biol*, 2011, 76(3/4/5): 273-297.
- [5] Woodson J D, Chory J. Coordination of gene expression between organellar and nuclear genomes [J]. *Nat Rev Genet*, 2008, 9(5): 383-395.
- [6] Xu W Q, Chen H M, Tian L X, et al. Extensive gene loss in the plastome of holoparasitic plant *Cistanche tubulosa* (Orobanchaceae) [J]. *Mitochondrial DNA B Resour*, 2020, 5(3): 2679-2681.
- [7] Li X, Zhang T C, Qiao Q, et al. Complete chloroplast genome sequence of holoparasite *Cistanche deserticola* (Orobanchaceae) reveals gene loss and horizontal gene transfer from its host *Haloxylon ammodendron* (Chenopodiaceae) [J]. *PLoS One*, 2013, 8(3): e58747.
- [8] 郑元庭. 多功能密码子分析优化平台的构建及应用研究 [D]. 杭州: 浙江大学, 2020.
- [9] Hershberg R, Petrov D A. Selection on codon bias [J]. *Annu Rev Genet*, 2008, 42: 287-299.
- [10] Lyu X L, Yang Q, Zhao F Z, et al. Codon usage and protein length-dependent feedback from translation elongation regulates translation initiation and elongation speed [J]. *Nucleic Acids Res*, 2021, 49(16): 9404-9423.
- [11] 石岩硕, 赵永星. 人参属植物叶绿体基因组密码子偏好性分析 [J]. 分子植物育种, 2022, 20(19): 6350-6361.
- [12] 宋芸, 贾孟君, 陈亮, 等. 忍冬 ICE1 基因密码子偏好性分析及受体系统选择 [J]. 植物生理学报, 2020, 56(11): 2459-2468.
- [13] 李显煌, 杨生超, 辛雅萱, 等. 灯盏花叶绿体基因组密码子偏好性分析 [J]. 云南农业大学学报: 自然科学, 2021, 36(3): 384-392.
- [14] 张俊焱, 曾阳, 李锦萍, 等. 2 种药用獐牙菜叶绿体基因组密码子偏好性分析 [J]. 中国中医药信息杂志, 2022, 29(1): 96-102.
- [15] 吴宪明, 吴松峰, 任大明, 等. 密码子偏性的分析方法及相关研究进展 [J]. 遗传, 2007, 29(4): 420-426.
- [16] Rosenberg M S, Subramanian S, Kumar S. Patterns of transitional mutation biases within and among mammalian genomes [J]. *Mol Biol Evol*, 2003, 20(6): 988-993.
- [17] 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.
- [18] 侯哲, 姜晓鸣, 李昂, 等. 星毛唐松草叶绿体基因组的测序及其密码子偏好性分析 [J]. 西南林业大学学报: 自然科学版, 2020, 18: 1-13.
- [19] Sharp P M, Li W H. Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for 'rare' codons [J]. *Nucleic Acids Res*, 1986, 14(19): 7737-7749.
- [20] 陆奇丰, 骆文华, 黄至欢. 两种梧桐叶绿体基因组密码子使用偏性分析 [J]. 广西植物, 2020, 40(2): 173-183.
- [21] Sueoka N. Directional mutation pressure and neutral molecular evolution [J]. *Proc Natl Acad Sci USA*, 1988, 85(8): 2653-2657.
- [22] Wright F. The 'effective number of codons' used in a gene [J]. *Gene*, 1990, 87(1): 23-29.
- [23] Romero H, Zavala A, Musto H. Codon usage in *Chlamydia trachomatis* is the result of strand-specific mutational biases and a complex pattern of selective forces [J]. *Nucleic Acids Res*, 2000, 28(10): 2084-2090.
- [24] 金刚, 覃旭, 龙凌云, 等. 剑麻叶绿体基因组编码序列密码子的使用特征 [J]. 福建农林大学学报: 自然科学版, 2018, 47(6): 705-710.
- [25] Sueoka N. Translation-coupled violation of Parity Rule 2 in human genes is not the cause of heterogeneity of the DNA G+C content of third codon position [J]. *Gene*, 1999, 238(1): 53-58.
- [26] Lloyd A T, Sharp P M. Evolution of codon usage patterns: The extent and nature of divergence between *Candida albicans* and *Saccharomyces cerevisiae* [J]. *Nucleic Acids Res*, 1992, 20(20): 5289-5295.
- [27] Grantham R, Gautier C, Gouy M, et al. Codon catalog usage is a genome strategy modulated for gene expressivity [J]. *Nucleic Acids Res*, 1981, 9(1): r43-r74.
- [28] Wang H C, Hickey D A. Rapid divergence of codon usage patterns within the rice genome [J]. *BMC Evol Biol*, 2007, 7(1): S6.
- [29] Xiang H, Zhang R Z, Butler R R 3rd, et al. Comparative analysis of codon usage bias patterns in microsporidian genomes [J]. *PLoS One*, 2015, 10(6): e0129223.
- [30] Jiang Y, Deng F, Wang H L, et al. An extensive analysis on the global codon usage pattern of baculoviruses [J]. *Arch Virol*, 2008, 153(12): 2273-2282.
- [31] Sharp P M, Matassi G. Codon usage and genome evolution [J]. *Curr Opin Genet Dev*, 1994, 4(6): 851-860.
- [32] Wang Z J, Xu B B, Li B, et al. Comparative analysis of

- codon usage patterns in chloroplast genomes of six Euphorbiaceae species [J]. *PeerJ*, 2020, 8: e8251.
- [33] Shields D C, Sharp P M. Synonymous codon usage in *Bacillus subtilis* reflects both translational selection and mutational biases [J]. *Nucleic Acids Res*, 1987, 15(19): 8023-8040.
- [34] Plotkin J B, Kudla G. Synonymous but not the same: The causes and consequences of codon bias [J]. *Nat Rev Genet*, 2011, 12(1): 32-42.
- [35] Zhang R Z, Zhang L, Wang W, *et al.* Differences in codon usage bias between photosynthesis-related genes and genetic system-related genes of chloroplast genomes in cultivated and wild *Solanum* species [J]. *Int J Mol Sci*, 2018, 19(10): 3142.
- [36] 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.
- [37] 吴妙丽, 陈世品, 陈辉. 竹亚科叶绿体基因组的密码子使用偏性分析 [J]. *森林与环境学报*, 2019, 39(1): 9-14.
- [38] Zhang W J, Zhou J, Li Z F, *et al.* Comparative analysis of codon usage patterns among mitochondrion, chloroplast and nuclear genes in *Triticum aestivum* L [J]. *J Integr Plant Biol*, 2007, 49(2): 246-254.
- [39] 王文斌, 于欢, 邱相坡. 黄芩叶绿体基因组重复序列及密码子偏好性分析 [J]. *分子植物育种*, 2018, 16(8): 2445-2452.
- [40] 尚明照, 刘方, 华金平, 等. 陆地棉叶绿体基因组密码子使用偏性的分析 [J]. *中国农业科学*, 2011, 44(2): 245-253.
- [41] Zhou M, Long W, Li X. Patterns of synonymous codon usage bias in chloroplast genomes of seed plants [J]. *For Stud China*, 2008, 10(4): 235-242.
- [42] 张月荣. 禾本科叶绿体基因组密码子使用模式及紫茎泽兰叶绿体 RNA 编辑分析 [D]. 杨凌: 西北农林科技大学, 2013.
- [43] 周涛, 杨林, 舒军霞, 等. 3 种含笑属植物叶绿体基因组密码子偏好性分析 [J]. *西部林业科学*, 2022, 51(3): 91-100.
- [44] Miao Y, Chen H, Xu W, *et al.* Structural mutations of small single copy (SSC) region in the plastid genomes of five *Cistanche* species and inter-species identification [J]. *BMC Plant Biol*, 2022, 22(1): 412.

[责任编辑 时圣明]