

• 药材与资源 •

基于转录组数据分析药用大黄的密码子使用偏好性

高梦琦¹, 邹建珍¹, 霍小位², 李依民³, 张明英³, 高静³, 张岗^{3*}, 刘蒙蒙^{1*}

1. 河北大学中医学院, 河北 保定 071002

2. 河北大学药学院, 河北 保定 071002

3. 陕西中医药大学药学院 陕西省秦岭中草药应用开发工程技术研究中心, 陕西 西安 712046

摘要:目的 研究药用大黄 *Rheum officinale* 转录组编码序列的密码子使用特点及其影响因素, 为蒽醌类化合物异源合成的载体选择及药用大黄分子进化研究提供理论基础。方法 利用 perl 程序及 Codon W 软件分析 4733 条药用大黄转录组编码序列的密码子使用偏好性。结果 药用大黄转录组编码序列的 GC、GC3 平均含量分别为 45.6%、44.73%, GC12 与 GC3 存在显著正相关 ($r=0.215$, $P<0.001$); ENc-GC3 中性图及偏倚性分析结果显示, 大部分基因分布于远离期望曲线和平面中心点, 大部分基因分布偏离期望曲线和中心点。通过基因高表达筛选密码子方法确定了 29 个药用大黄的最优密码子, 大多数的最优密码子以 U 和 A 结尾。结论 突变压力在药用大黄转录组编码序列的密码子使用偏好性形成过程起到主要作用。

关键词: 药用大黄; 转录组; 密码子偏好性; 最优密码子; 突变压力

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2021)20-6344-06

DOI: 10.7501/j.issn.0253-2670.2021.20.024

Analysis of codon usage patterns in *Rheum officinale* based on transcriptome data

GAO Meng-qi¹, ZOU Jian-zhen¹, HUO Xiao-wei², LI Yi-min³, ZHANG Ming-ying³, GAO Jing³, ZHANG Gang³, LIU Meng-meng¹

1. College of Traditional Chinese Medicine, Hebei University, Baoding 071002, China

2. College of Pharmacy, Hebei University, Baoding 071002, China

3. Shaanxi Qinling Chinese Herbal Medicine Application Development Engineering Technology Research Center, College of Pharmacy, Shaanxi University of Chinese Medicine, Xi'an 712046, China

Abstract: Objective To study the characteristics of the codon usage patterns of the coding sequence of *Rheum officinale* based on transcriptome data. This study laid a foundation for future research on anthraquinone biosynthesis and molecular evolution in *R. officinale*. **Methods** Perl program and Codon W software were used to analyze the codon usage bias patterns of 4733 transcriptome coding sequences. **Results** The average contents of GC and GC3 of the coding sequence in *R. officinale* were 45.6% and 44.73%, respectively. There was a significant positive correlation between GC12 and GC3 ($r=0.215$, $P<0.001$). The results of ENc-GC3 and bias analysis showed that most genes were distributed far away from the standard curve and the center area of the axis, and a small number of genes distributed around the standard curve and the center area. Twenty-nine optimal codons were identified as being biased toward codons ending with U or A using the high expression superior codon analysis method. **Conclusion** Mutation pressure played a key role in the formation of codon usage bias in *R. officinale* coding sequences.

Key words: *Rheum officinale* Baill.; transcriptome; codon usage bias; optimal codons; mutation pressure

收稿日期: 2021-04-10

基金项目: 国家自然科学基金资助项目 (81973430); 河北省自然科学基金资助项目 (H2020201006); 河北大学医学学科培育项目 (2021B10); 陕西省重点产业创新链 (群) 项目 (2021ZDLSF04-01); 陕西省高校青年杰出人才支持计划项目; 陕西中医药大学学科创新团队项目 (2019-QN01)

作者简介: 高梦琦, 女, 硕士研究生, 研究方向为药用植物生物技术与分子生物学。E-mail: 438450837@qq.com

***通信作者:** 张岗, 男, 博士, 教授, 研究方向为中药资源与分子生物学。E-mail: jay_gumling2003@aliyun.com

刘蒙蒙, 男, 博士, 副教授, 研究方向为中药资源评价与利用。E-mail: 547717678@qq.com

药用大黄 *Rheum officinale* Baill. 属于蓼科大黄属多年生草本植物, 为《中国药典》2020 年版收录的大黄 3 种基原植物之一, 具有泻下攻积、清热泻火、凉血解毒、逐瘀通经、利湿退黄等功效^[1]。现代中药化学研究表明大黄中主要含有蒽醌类、酚类等物质, 其中蒽醌类化合物是主要的活性物质, 具有抗炎、抗肿瘤作用^[2-3]。目前药用大黄的资源主要来源于人工种植, 随着过度采挖, 野生资源已经濒临灭绝。因此, 开展蒽醌类化合物生物合成机制研究, 探索蒽醌类化合物生物合成路径, 是解决药用大黄野生资源濒危的思路之一。

遗传密码子是生物体 DNA 与蛋白之间信息传递的信息载体。大多数氨基酸不仅由一种密码子编码, 而是可以同时由 2~3 种密码子编码, 生物对于不同的密码子的使用具有一定偏好性^[4]。研究发现, 密码子使用偏好性受到 GC 含量 (鸟嘌呤和胞嘧啶含量)^[5]、基因表达水平^[6]、环境压力^[7]、突变压力^[8]、基因长度^[9]、蛋白结构^[10]及 tRNA 丰度^[11]的影响。随着高通量测序技术的发展, 药用植物的密码子使用偏好性研究有了显著的成果^[12-13], 但是目前还没有关于药用大黄密码子偏好性的研究。生物体选择最优密码子可以体现高基因的翻译效率及蛋白的表达量。因此, 研究密码子的使用偏好性, 有利于基因工程研究中密码子优化、最适外源宿主选择, 并对生物的系统进化研究具有重要意义。

本研究利用药用大黄的转录组数据对其编码序列进行密码子偏好性分析, 揭示药用大黄基因密码子使用偏好性的特征, 为药用大黄基因工程研究中进行密码子优化及选择最佳的外源载体提供理论基础。

1 数据

药用大黄转录组 (SRR10838870) 来源于本课题组前期选取药用大黄种苗根、根茎、样品, 利用 Illumina HiSeq 2500 测序平台进行转录组测序并拼接获得 Uingenex 序列, 利用 BLASTx 及 ESTScan 软件对上述 Uingenex 序列进行编码序列 (coding sequence, CDS) 分析。使用 perl 程序筛选编码序列中长度大于 300 bp 并且同时具有起始密码子及终止密码的序列, 用于药用大黄密码子使用偏好性分析, 最后共得到 4 733 条 CDS 进行分析。

2 方法

2.1 编码序列中碱基含量及中性绘图分析

使用 Codon W 软件分析药用大黄编码序列中

的有效密码子数 (effective number of codons, ENc)、密码子适应指数 (codon adaptation index, CAI) 及同义密码子相对使用度 (relative synonymous codon usage, RSCU), 计算序列中的 GC 含量, 利用编写的 perl 程序计算 GC1、GC2 及 GC3 的含量, 即密码子第 1、2、3 位的 GC 含量。GC12 为 GC1、GC2 的平均值。

中性绘图分析主要用于分析 GC12 与 GC3 之间的相关性, 以 GC12 为纵坐标, 以 GC3 为横坐标制作散点图, 并进行线性拟合回归分析。可用于评价自然选择压力及突变对密码子使用偏好性造成的影响。

2.2 GC3 关联及奇偶偏好性分析

ENc-GC3 关联分析主要用于分析密码子使用偏好性与碱基组成之间的关联。以 ENc 为纵坐标、GC3s 为横坐标绘制二维散点图, 并进行非线性拟合回归分析。如果基因的点分布在拟合曲线附近, 则表明密码子的使用偏好性主要是受到突变作用的影响; 如果基因点分布在拟合曲线下方, 则表明密码子的使用偏好性主要是受到选择压力的影响。奇偶偏好性分析以 A3/(A3+T3) 为纵坐标, 以 G3/(G3+C3) 为横坐标制作二维散点图, 以 (0.5, 0.5) 坐标为中心点, 从这个中心点制作的矢量表示了奇偶偏倚的程度和方向^[14]。

2.3 最优密码子分析

对于最优密码子的筛选, 本研究利用药用大黄编码书写的 CAI 值进行由高到低排列, 然后取前后各 5% 的编码序列, 分别代表高表达与低表达的基因序列。之后分析各个密码子的平均 RSCU 值。当密码子在高低表达基因中 RSCU 值的差异 > 0.3, 且在高表达基因中 RSCU ≥ 1、在低表达基因中的 RSCU ≤ 1, 即可被确定为最优密码子^[15-16]。

2.4 对应分析

对应分析是指利用多元统计分析的方法研究不同编码序列密码子的变异趋势。利用 Codon W 软件中对应分析功能, 将所有编码序列分布到 59 维向量空间, 进而获得密码子使用偏好性的特点。通过分析基因位置研究基因向量及基因间值的变异程度, 预测使密码子使用发生偏好性的原因^[17]。

3 结果

3.1 密码子 GC 含量组成分析

利用 Codon W 软件计算药用大黄编码序列中 GC 含量, 将药用大黄编码序列作为一个整体进行计算。结果如图 1 所示, GC 含量范围分布在 28.6%~

71.9%，主要分布在 40%~50%，GC 的平均含量为 45.6%。GC1、GC2、GC3 的平均含量分别为 48.63%、40.89%、44.37%。对比发现，GC2 的含量最小，GC1 与 GC3 的值相对比较接近，说明药用大黄编码序列更倾向于使用鸟嘌呤（G）和胞嘧啶（C）这 2 种碱基。

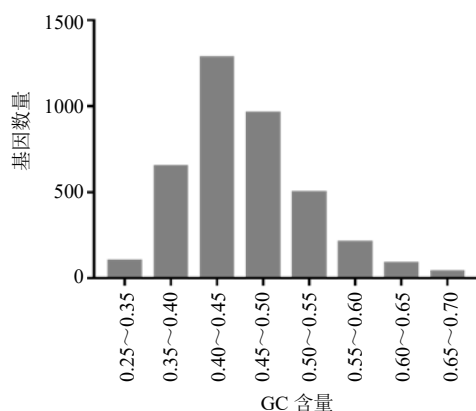


图 1 药用大黄编码序列中 GC 含量组成

Fig. 1 Distribution of GC content of coding sequences (CDS) in *R. officinale*

3.2 中性绘图分析

药用大黄编码序列的中性图分析如图 2 所示，其中 GC12 的数值范围为 0.15~0.87，GC3 的数值范围为 0.09~0.87。GC12 与 GC3 的相关性系数为 0.155，表明这两者之间的相关性比较弱，说明选择压力在药用大黄密码子使用偏好性的选择具有一定影响。

3.3 ENc-plot 绘图分析

ENc 的取值范围为 20~61，平均值为 47.58，其中 $ENc \leq 35$ 的基因有 386 个（8.15%）， $ENc = 61$

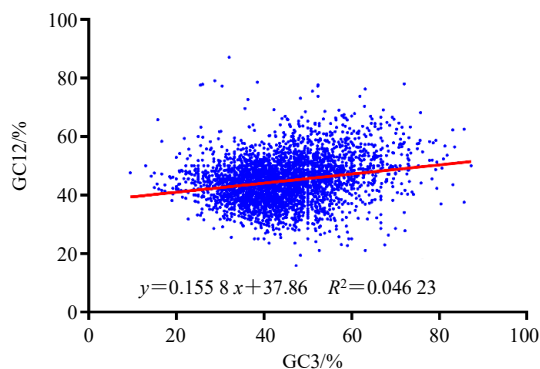


图 2 中性图分析

Fig. 2 Neutrality plot

的基因有 251 个（5.30%），说明药用大黄的编码序列除了少数密码子具有较强的偏好性，其整体的密码子使用偏好性较弱。以 ENc 为纵坐标、GC3s 为横坐标绘制散点图（图 3），图中各点代表基因分布情况。图 3 显示大部分基因远离期望曲线，说明密码子的偏好性主要来源于选择压力的影响，少部分基因均匀的分布在期望曲线两侧，说明部分基因收到突变压力的影响。

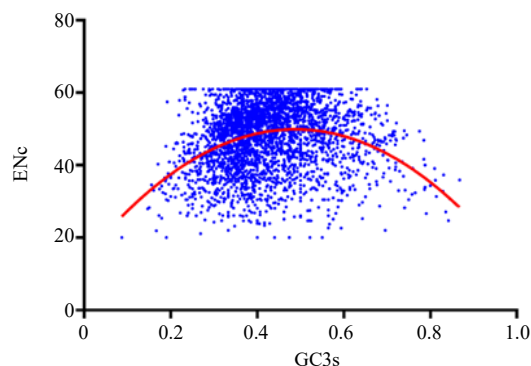


图 3 药用大黄转录组编码序列 ENC-GC3s 相关性分析

Fig. 3 Neutrality plot (ENC vs GC3)

利用 $(ENC_{exp} - ENC_{obs}) / ENC_{exp}$ 统计分析期望 ENc 与观测 ENc 之间的区别^[18]。如图 4 所示， $(ENC_{exp} - ENC_{obs}) / ENC_{exp}$ 的取值主要分布在 0~0.2，说明期望 ENc 与观测 ENc 两者之间的存在很小差异。进而说明在药用大黄编码序列密码子使用偏好性形成过程中，突变压力不是唯一的影响因素，但却是一个重要的影响因素。

3.4 对应性分析

对应性分析结果见图 5，GC 含量不同的基因被标记上不同的颜色，编码序列中 GC 含量大于 60%

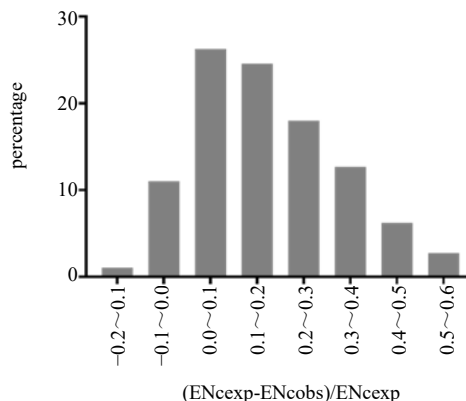
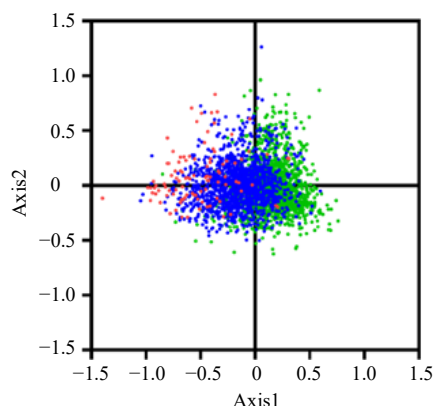


图 4 ENc 比值频率分布

Fig. 4 Frequency distribution of ENc ratio



红色-GC 含量大于 60% 的序列 蓝色-GC 含量 45%~60% 序列 绿色-GC 含量小于 45% 的序列
genes with GC content higher than 60%, within 45%—60% and lower than 45% were plotted as red, blue and green dots, respectively

图 5 对应性分析

Fig. 5 Correspondence analysis

的被标记为红色, GC 含量在 45%~60% 的被标记为蓝色, GC 含量小于 45% 的被标记为绿色。其中, GC 含量大于 60% 的编码序列分散的分布在坐标系中, 而 GC 含量小于 60% 的编码序列则在坐标系中分布比较集中。

此外, 选取 6 个编码序列相关的参数 (GC12、GC3、GCall、ENC、CAI 以及 Axis1) 进行相关性分析, 结果如表 1 所示, Axis1 与 GC12、GC3 这 2 个重要参数存在显著的负相关性 ($r=-0.418$, $r=-0.728$, $P<0.01$)。结合图 5 与表 1, 可以说明 GC 含量与 GC3 含量高的编码序列取向于分布在 Axis1 左侧。因此, 推测以 G/C 结尾的密码子应集中分布在 Axis1 负轴而以 A/U 结尾的密码子集中分布在 Axis1 正轴。

3.5 偏倚性分析

偏倚性分析结果如图 6 所示, 大部分药用大黄转录组编码序列样本点分布于中心点附近, 说明其

表 1 药用大黄编码序列对应参数的相关性分析

Table 1 Correlation analysis of *R. officinale* gene-related parameters

参数	GC3	GCall	ENC	CAI	Axis1
GC12	0.215**	0.858**	0.02	0.007	-0.418**
GC3		0.678**	0.01	0.015	-0.728**
GCall			0.08**	0.695**	0.011
ENC				0.070**	-0.019
CAI					-0.018

** $P<0.01$

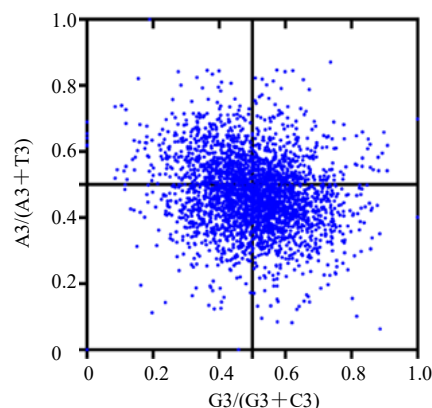


图 6 编码序列偏倚性分析

Fig. 6 Parity rule 2 (PR2) -bias plot

碱基中嘌呤与嘧啶出现的频率相对一致, 少部分编码序列偏离中心点较远, 进一步说明药用大黄密码子使用偏好性形成过程中主要受到突变压力的影响。

3.6 基因表达水平对密码子使用偏好性的影响

药用大黄编码序列的 CAI 的范围为 0.261~0.835, 表明药用大黄的基因表达水平比较高。同时, 对 CAI 与其他几个重要的参数进行了相关性分析, 包括 ENC、GC3、GC。结果如表 1 及图 7 所示, CAI 与 ENC 及 GC 之间都存在显著的正相关性 ($r=0.07$, $P<0.01$; $r=0.695$, $P<0.01$)。因此, 基因表达水平在药用大黄密码子使用偏好性形成过程中起到关键作用, 基因表达水平高并且基因中 GC 含量高, 密码子使用偏好性程度越高。

3.7 最优密码子分析

药用大黄编码序列高、低表达序列的 RSCU 值如表 2 所示。从中筛选到 29 个密码子作为药用大

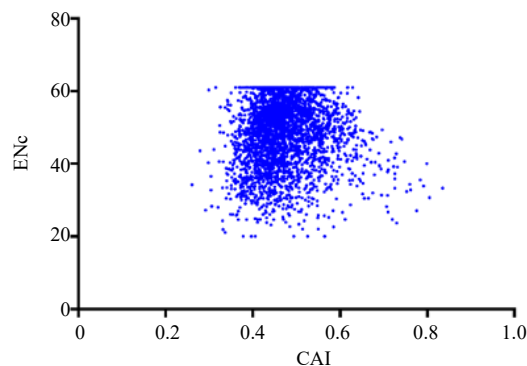


图 7 中性图分析

Fig. 7 Neutrality plot

表 2 药用大黄转录组高、低表达基因的 RSCU 比较分析

Table 2 Optimal codons of *R. officinale* genes based on the RSCU values

氨基酸	密码子	高表达 (数目)	低表达 (数目)	氨基酸	密码子	高表达 (数目)	低表达 (数目)
Phe	UUU*	1.34 (147)	0.51 (205)	Ser	UCU*	1.32 (132)	0.95 (300)
	UUC	0.66 (72)	1.49 (596)		UCC	0.23 (23)	2.00 (632)
Leu	UUA*	1.54 (158)	0.17 (43)		UCA*	2.07 (208)	0.62 (195)
	UUG*	1.23 (126)	0.87 (225)		UCG	0.07 (7)	1.12 (353)
	CUU*	1.37 (140)	0.97 (249)	Pro	CCU*	1.93 (170)	0.95 (392)
	CUC	0.23 (24)	2.55 (657)		CCC	0.36 (32)	1.03 (425)
	CUA*	1.19 (122)	0.42 (109)		CCA*	1.59 (140)	0.78 (324)
	CUG	0.44 (45)	1.03 (265)		CCG	0.11 (10)	1.25 (516)
Ile	AUU*	1.32 (207)	0.70 (175)	Thr	ACU*	1.49 (184)	0.73 (204)
	AUC	0.64 (100)	1.92 (478)		ACC	0.29 (36)	1.86 (523)
	AUA*	1.04 (163)	0.38 (95)		ACA*	1.99 (246)	0.43 (122)
Met	AUG	1.00 (438)	1.00 (584)		ACG	0.23 (28)	0.98 (275)
Val	GUU*	1.37 (166)	0.86 (245)	Ala	GCU*	1.40 (173)	0.87 (414)
	GUC	0.46 (56)	1.73 (491)		GCC	0.25 (31)	1.80 (852)
	GUA*	1.10 (133)	0.27 (78)		GCA*	2.29 (282)	0.43 (204)
	GUG	1.06 (128)	1.13 (322)		GCG	0.06 (7)	0.90 (426)
Tyr	UAU*	1.19 (147)	0.53 (111)	Cys	UGU*	1.31 (81)	0.51 (80)
	UAC	0.81 (101)	1.47 (311)		UGC	0.69 (43)	1.49 (235)
TER	UAA	1.52 (122)	1.20 (96)	TER	UGA	1.20 (96)	1.27 (102)
	UAG	0.29 (23)	0.52 (42)	Trp	UGG	1.00 (103)	1.00 (206)
His	CAU*	1.44 (123)	0.57 (159)	Arg	CGU	0.16 (10)	0.77 (154)
	CAC	0.56 (48)	1.43 (399)		CGC	0.00 (0)	1.71 (343)
Gln	CAA*	1.42 (274)	0.90 (274)		CGA	0.08 (5)	0.73 (146)
	CAG	0.58 (113)	1.10 (336)		CGG	0.05 (3)	1.00 (201)
Asn	AAU*	1.30 (293)	0.69 (219)	Ser	AGU*	1.55 (156)	0.38 (120)
	AAC	0.70 (157)	1.31 (413)		AGC	0.76 (76)	0.94 (299)
Lys	AAA*	1.05 (328)	0.60 (282)	Arg	AGA*	4.52 (275)	0.72 (144)
	AAG	0.95 (298)	1.40 (660)		AGG	1.18 (72)	1.08 (216)
Asp	GAU*	1.49 (349)	0.88 (437)	Gly	GGU*	1.25 (151)	0.86 (335)
	GAC	0.51 (118)	1.12 (554)		GGC	0.37 (45)	1.55 (607)
Glu	GAA*	1.38 (449)	0.65 (371)		GGA*	1.71 (207)	0.87 (341)
	GAG	0.62 (201)	1.35 (776)		GGG	0.68 (82)	0.72 (284)

*最优密码子

*optimal codon

黄编码序列的最优密码子, 在这些最优密码子中, 以 U (15 个) 及 A (13 个) 结尾的较多, 表明药用大黄转录组编码序列的最优密码子偏好以 U 或 A 结尾。

4 讨论

在生物长期的进化过程中, 形成了对密码子的使用的固定模式。今年来, 随着二代测序技术的发展, 探索天然产物合成路径并进行异源合成, 成为了研究热点。因此, 探索药用大黄基因密码子使用偏好性, 可为研究其密码子使用特点, 筛选最佳蛋白异源表达载体及密码子优化奠定理论基础。

本研究对药用大黄转录组数据中 4733 条完整编码序列的密码子使用偏好性进行分析, 结果显示, 药用大黄转录组编码序列的 GC、GC3 平均含量分

别为 45.6%、44.73%, 表明药用大黄转录组中的编码序列偏好使用富含 A、U 且以 A/U 结尾的密码子, 这与小麦 *Triticum aestivum* L.^[19]、水稻 *Oryza sativa* L.^[20]、玉米 *Zea mays* L.^[21] 密码子使用模式不一致, 这些植物的密码子富含 GC 碱基, 并且大多数密码子以 G/C 结尾, 药用植物川贝母转录组中的编码序列的 GC 和 GC3 平均含量都接近于 50%, 并没有表现出密码子使用偏好性^[15], 表明碱基组成在药用大黄响密码子偏好性形成过程中起到重要作用。

此外, 碱基突变和遗传选择也影响着密码子使用偏好性形成的重要因素。本研究中, 药用大黄转录组 GC12 与 GC3 含量具有显著相关性, ENc-plot 绘图及偏倚性分析分析结果表明, 大部分基因远离于标准曲线及坐标轴中心点, 基因的期望 ENc 值与观测 ENc

值具有较大差异,说明这部分基因密码子偏好性的形成是以遗传选择为主导因素;少数的基因分布于标准曲线附近及坐标轴中心点,基因的期望 ENc 值与观测 ENc 值基本保持一致,说明这部分基因密码子形成偏好性的形成是以突变为主导因素。

在构建蛋白表达载体时,不同的基因密码子使用偏好性差异很大,因此分析药用大黄的密码子使用特点对于载体的密码子优化具有重要意义。本研究筛选得到药用大黄 29 个最优密码子,大多数密码子以 A/U 结尾,从药用大黄编码序列中的 GC 含量可以看出,其偏好使用 A 或 U 碱基结尾的密码子。由此推断药用大黄基因的密码子偏好性的产生与其基因进化过程中 AU 到 GC 突变压力要比 GC 到 AU 突变压力高有关。

本研究对药用大黄转录组数据中编码序列的密码子使用特点进行了研究,探讨了影响其密码子使用偏好性形成的影响因素,并筛选出药用大黄基因的 29 个最优密码子,可为蒽醌类化合物的异源合成及其合成路径研究提供理论依据。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中国药典 [S]. 一部. 2020: 23.
- [2] Chiang J H, Yang J S, Ma C Y, *et al.* Danthron, an anthraquinone derivative, induces DNA damage and caspase cascades-mediated apoptosis in SNU-1 human gastric cancer cells through mitochondrial permeability transition pores and Bax-triggered pathways [J]. *Chem Res Toxicol*, 2011, 24(1): 20-29.
- [3] Ma J, Lu H, Wang S, *et al.* The anthraquinone derivative Emodin inhibits angiogenesis and metastasis through downregulating Runx2 activity in breast cancer [J]. *Int J Oncol*, 2015, 46(4): 1619-1628.
- [4] Liu X Y, Li Y, Ji K K, *et al.* Genome-wide *codon* usage pattern analysis reveals the correlation between *codon* usage bias and gene expression in *Cuscuta australis* [J]. *Genomics*, 2020, 112(4): 2695-2702.
- [5] Wang L, Roossinck M J. Comparative analysis of expressed sequences reveals a conserved pattern of optimal *codon* usage in plants [J]. *Plant Mol Biol*, 2006, 61(4/5): 699-710.
- [6] Paul P, Malakar A K, Chakraborty S. *Codon* usage and amino acid usage influence genes expression level [J]. *Genetica*, 2018, 146(1): 53-63.
- [7] Fabijanić M, Vlahoviček K. Big data, evolution, and metagenomes: predicting disease from gut microbiota *codon* usage profiles [J]. *Methods Mol Biol*, 2016, 1415: 509.
- [8] Camiolo S, Melito S, Porceddu A. New insights into the interplay between *codon* bias determinants in plants [J]. *DNA Res*, 2015, 22(6): 461-470.
- [9] Wada K, Wada Y, Ishibashi F, *et al.* *Codon* usage tabulated from the GenBank genetic sequence data [J]. *Nucleic Acids Res*, 1992, 20(Suppl): 2111-2118.
- [10] Duret L, Mouchiroud D. Expression pattern and, surprisingly, gene length shape *codon* usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis* [J]. *Proc Natl Acad Sci USA*, 1999, 96(8): 4482-4487.
- [11] Ikemura T. *Codon* usage and tRNA content in unicellular and multicellular organisms [J]. *Mol Biol Evol*, 1985, 2(1): 13-34.
- [12] 李滢, 匡雪君, 朱孝轩, 等. 长春花密码子使用偏好性分析 [J]. 中国中药杂志, 2016, 41(22): 4165-4168.
- [13] 赵春丽, 彭丽云, 王晓, 等. 莧菜 AtGAI 基因密码子偏好性与进化分析 [J]. 中国农业大学学报, 2019, 24(12): 10-22.
- [14] 秦政, 郑永杰, 桂丽静, 等. 樟树叶叶绿体基因组密码子偏好性分析 [J]. 广西植物, 2018, 38(10): 1346-1355.
- [15] 李滢, 匡雪君, 孙超, 等. 川贝母转录组密码子使用偏好性分析 [J]. 中国中药杂志, 2016, 41(11): 2055-2060.
- [16] Yang X, Luo X, Cai X. Analysis of *codon* usage pattern in *Taenia saginata* based on a transcriptome dataset [J]. *Parasit Vectors*, 2014, 7: 527.
- [17] 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.
- [18] Wright F. The 'effective number of codons' used in a gene [J]. *Gene*, 1990, 87(1): 23-29.
- [19] 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.
- [20] Liu Q P, Feng Y, Zhao X A, *et al.* Synonymous *codon* usage bias in *Oryza sativa* [J]. *Plant Sci*, 2004, 167(1): 101-105.
- [21] Fennoy S L, Bailey-Serres J. Synonymous *codon* usage in *Zea mays* L. nuclear genes is varied by levels of C and G-ending codons [J]. *Nucl Acids Res*, 1993, 21(23): 5294-5300.

[责任编辑 时圣明]