

莧菜 *AtGAI* 基因密码子偏好性与进化分析

赵春丽 彭丽云 王晓 陈家兰 王乐 陈何 赖钟雄 刘生财*

(福建农林大学 园艺植物生物工程研究所,福州 350002)

摘要 为明确莧菜 *AtGAI* 基因密码子的使用特性,采用 CodonW、SPSS 软件及 EMBOSS 在线程序分析莧菜 *AtGAI* 基因密码子的偏好性,通过中性绘图、ENc 绘图和偏倚分析 *GAI* 基因密码子偏好性形成的可能原因。结果表明:1)莧菜 *AtGAI* 基因的有效密码子数(ENc)、密码子适应指数(CAI)和 GC 含量分别为 50.75、0.195 和 41.91%,表明其密码子偏好性较弱,偏好使用 A/T 结尾的密码子;2)CDS 序列及 RSCU 聚类分析表明,莧菜 *AtGAI* 基因与双子叶植物中石竹目含甜菜色素的植物偏好性最相近;3)密码子碱基成分和相关性分析发现,*GAI* 基因的密码子偏好性主要受碱基突变的影响;4)密码子使用频率比较分析发现,酵母真核表达系统更适合莧菜 *AtGAI* 异源表达,模式植物拟南芥、烟草、番茄和甜菜均可做为莧菜 *AtGAI* 基因的遗传转化受体,其中甜菜是其最理想的遗传转化受体。突变压力是 *GAI* 基因密码子偏好性形成的主要因素,莧菜 *AtGAI* 基因外源表达和遗传转化时需要根据宿主基因偏好性进行密码子优化和改造。

关键词 莧菜; *AtGAI* 基因; 密码子偏好性; 进化分析; 遗传转化受体

中图分类号 Q943.2; S636.4 **文章编号** 1007-4333(2019)12-0010-13 **文献标志码** A

Codon bias and evolution analysis of *AtGAI* in *Amaranthus tricolor* L.

ZHAO Chunli, PENG Liyun, WANG Xiao, CHEN Jialan, WANG Le, CHEN He,
LAI Zhongxiong, LIU Shengcai*

(Institute of Horticultural Plant Biological Engineering, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

Abstract To understand the codon bias of *AtGAI* in *Amaranthus tricolor* L., *AtGAI* gene codon bias was analyzed by CodonW, SPSS and EMBOSS online program. Neutral drawing, ENc plots and parity rule 2 plot were subsequently used for the exploration of possible factors that affect the formation of the bias. The optimal receptor system was explored by comparing the frequency of the *AtGAI* gene and the pattern biocodon usage. The results showed that: 1) The ENc, CAI, and GC content of *AtGAI* were respectively 50.75, 0.195, and 41.91%, indicating that the codon bias level of *AtGAI* in *A. tricolor* L. was low and biased toward the synonymous codons with A or T. 2) The CDS sequence and RSCU cluster analysis showed that the preference of *AtGAI* gene was the closest to that of Caryophyllales containing betalain in dicotyledonous plants. 3) The codon base composition and correlation analysis showed that the codon bias of *GAI* gene was mainly affected by the base mutation. 4) The codon usage frequency comparison and analysis showed that the yeast eukaryotic expression system was more suitable for heterologous expression of *AtGAI* gene. The model plants *Arabidopsis thaliana*, tobacco, tomato and beet can be used as genetic transformation receptors for *AtGAI* gene, but beet is the most ideal genetic transformation receptor. This study indicated that mutational stress was the main factor in the formation of *GAI* gene codon bias. It is necessary to optimize and modify *AtGAI* codon according to the codon bias of the host genome for heterologous expression and genetic transformation analysis.

Keywords *Amaranthus tricolor* L.; *AtGAI* gene; codon usage bias; evolution analysis; genetic transformation receptors

收稿日期: 2019-01-22

基金项目: 福建省自然科学基金(2018J01700); 福建农林大学科技创新专项基金项目(CXZX2017174, CXZX2016118, CXZX2017189); 福建省高原学科建设经费(102/71201801101)

第一作者: 赵春丽, 硕士研究生, E-mail: 1789308117@qq.com

通讯作者: 刘生财, 副教授, 主要从事园艺植物生物技术与品质代谢研究, E-mail: 1215698900@qq.com

密码子是 mRNA 上编码氨基酸的三联体核苷酸序列,是自然界中遗传信息准确识别和传递的基本载体^[1-2]。密码子编码的氨基酸中,除甲硫氨酸(Met)和色氨酸(Trp)之外,其他氨基酸均由 2 种或 2 种以上的密码子编码,这种编码同一种氨基酸的密码子称为同义密码子(Synonymous codon)^[3]。在长期进化过程中,物种对同义密码子的使用并不完全随机,往往会由于环境适应性改变或物种本身特殊的进化方式而倾向使用特定密码子,从而形成密码子使用偏好性(Codon usage bias, CUB)^[2],最优先被选用的密码子被称为最优密码子(optimal codon),一般出现在高表达基因中^[4]。密码子使用偏好性的评价参数主要是有效密码子数(Effective number of codons, ENc)、同义密码子相对使用度(Relative synonymous codon usage, RSCU)、不同位置的 G+C 含量和密码子适应指数(Codon adaptation index, CAI)等^[5]。密码子使用偏好性能够反映物种或基因的起源、进化规律和突变方式,对基因表达^[6]、基因功能预测^[7]、基因家族分化^[8]、蛋白质表达和蛋白质结构研究^[9]等具有重要参考价值。目前,主要有两种假说解释密码子使用偏好性,即“突变偏差与选择”假说^[10]和“突变选择漂移平衡”模型^[11]。在基因外源表达和遗传转化时,容易受物种间密码子偏好性差异制约而导致基因甲基化、低水平表达或基因沉默,往往需要根据目标基因密码子的特点选择使用偏性相近的宿主表达系统或对目标基因进行定向改造和优化^[12]。因此,密码子选择偏性分析能够为物种进化规律和基因功能的深入研究提供必要的理论指导。

GAI 属于 DELLA 蛋白家族,是 GAs 信号转导途径中的抑制因子^[13],当 GAs 信号存在时 GAI 被降解,从而解除 GAs 信号转导途径的阻遏状态,由 GAs 诱导的与生长相关的基因得以表达^[14]。Xie 等^[15]研究发现 DELLA 蛋白作为正调控因子介导 GAs 抑制的花青素合成。郑学立等^[16]和潘君飞等^[17]研究发现 GAs 抑制苋菜甜菜色素的合成,但其分子机制还不清楚。甜菜色素和花青素不能共存,但两者可能存在相同的调控因子^[18-20]。苋菜是石竹目苋科苋属一年生草本植物,茎和叶富含甜菜色素,有逐步取代甜菜作为甜菜色素提取的主要来源的趋势^[21]。甜菜色素是四大天然色素之一,具有较高的营养及药用价值^[22-23]。苋菜 *AtGAI* 基因属于 GAI 基因家族,可能通过响应 GAs 信号参与调控

苋菜幼苗生长及甜菜色素代谢^[24],但对于该基因最佳外源表达系统的选择以及调控甜菜色素代谢的具体分子机制等研究鲜有报道。本研究以苋菜 *AtGAI* 基因序列为材料,利用 CodonW 软件和 CHIPS 在线程序对该基因密码子使用偏好性进行分析,对苋菜 *AtGAI* 基因的功能及它的异源表达情况进行研究,旨在为深入研究苋菜 *AtGAI* 基因功能奠定基础。

1 材料与方 法

1.1 序列数据来源

苋菜 *AtGAI* (GenBank 登录号:MK049175)克隆于苋菜品种‘大红’(*A. tricolor* L. ‘Dahong’),全长为 2 168 bp,含有一个 1 818 bp 的开放阅读框(ORF),编码 605 个氨基酸。藜麦、甜菜、菠菜、葡萄、番茄和拟南芥等植物的 GAI 基因序列来源于 GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>),登录号见表 1。大肠杆菌(*Escherichia coli*)、酵母菌(*Saccharomyces cerevisiae*)、拟南芥(*Arabidopsis thaliana*)、烟草(*Nicotiana tabacum*)、番茄(*Solanum lycopersicum*)及甜菜(*Beta vulgaris* L.)等模式生物的基因组密码子偏好性数据来源于密码子使用数据库 Codon Usage Database(<http://www.kazusa.or.jp/codon/>)。

1.2 数据统计分析

参照李平等^[25]和孙晶等^[26]的方法采用 Codon W 软件和 CHIPS 在线程序(<http://emboss.toulouse.inra.fr/cgi-bin/emboss/chips>)对苋菜 *AtGAI* 基因及其他物种 GAI 基因的编码序列(Coding sequence, CDS)的有效密码子数(ENc)、密码子第 1、2 和 3 位上 G/C 含量(GC1、GC2 和 GC3)、密码子第三位上各个碱基含量(A3s、T3s、G3s 和 C3s)、密码子适应指数(CAI)和同义密码子相对使用度(RSCU)等密码子偏好性参数进行运算。采用 EMBOSS 在线程序中的 CUSP(<http://emboss.toulo use.inra.fr/cgi-bin/emboss/cusp>)分析基因密码子使用频率。使用 Excel 2016、Origin 2017 等软件对密码子相关参数进行重新统计,并进行热图、中性绘图、ENc 绘图和 PR2 分析。使用 SPSS 22.0 进行密码子偏好性参数的相关性分析,并对不同物种 GAI 基因的 RSCU 值进行密码子使用偏性聚类分析。使用 MEGA 5.0 邻接法(Neighbor-joining)构建不同物种 GAI 基因的 CDS 系统进化树。

2 结果与分析

2.1 苋菜 *AtGAI* 基因密码子使用偏好性分析

2.1.1 苋菜 *AtGAI* 基因 ENc、GC 及 CAI 分析

运用 CodonW 软件获得苋菜 *AtGAI* 基因的密码子适应指数(CAI),有效密码子数(ENc),运用 EMBOSS 在线网站获得各成员的总 GC 含量以及不同位置的 GC 含量。结果显示:苋菜 *AtGAI* 基因的 ENc 值为 50.75,较为接近 61,说明该基因编码氨基酸时对密码子的选择偏好性较弱,其表达水平可能相对较低^[27-28];苋菜 *AtGAI* 基因中密码子的 GC、GC1、GC2、GC3 含量分别为 41.91%、51.65%、41.91%和 32.18%,见表 1,说明苋菜 *AtGAI* 基因的密码子比较偏好以 A/T 结尾。苋菜 *AtGAI* 基因的 CAI 值为 0.195, ≤ 1.0 ,进一步表明苋菜 *AtGAI* 基因对密码子的选择偏好性较弱^[29]。

2.1.2 苋菜 *AtGAI* 基因 RSCU 分析

苋菜 *AtGAI* 基因的 RSCU 分析结果,图 1,偏好性较强的密码子有 28 个,其中以 A/T 结尾的密码子有 24 个,以 G/C 结尾的密码子有 4 个,说明苋菜 *AtGAI* 基因的密码子偏好以 A/T 结尾,与 GC 含量分析结果一致。在偏好性较强的密码子中,亮氨酸的密码子 UUA 的偏好性最强,其 RSCU 值达到了 2.21, CUU、UUG、CUG、CUC、CUA 密码子的 RSCU 值分别为 1.37、1.05、0.74、0.42 和 0.21,说明 UUA 是编码苋菜 *GAI* 蛋白质中亮氨酸残基的主要密码子,具有极强的使用偏好性。

此外,密码子 GUU、GCU 的 RSCU 值 > 2 ,具有较强的使用偏好性;密码子 ACG、CGC 的 RSCU 值为 0,说明苋菜 *AtGAI* 基因不使用这 2 个密码子;其余 27 种(RSCU < 1)密码子偏好性使用低于预期。一些密码子如, UUA、GUU 和 GCU 在同义密码子中具有优先选择性,使得苋菜 *AtGAI* 基因的 ENc 值偏离 61。

2.2 不同物种间 *GAI* 基因密码子使用偏好性比较

2.2.1 密码子偏好性相关参数比较

不同物种间 *GAI* 基因密码子选择偏好性相关参数,表 1。单子叶植物和双子叶植物间 *GAI* 基因的 CAI 值为 0.191~0.263,差异较小,而 ENc 和 GC 值却存在明显不同。双子叶植物中含甜菜色素的籽粒苋、甜菜、藜麦和菠菜及含花青素的拟南芥、番茄和毛白杨等的 *GAI* 基因与苋菜 *AtGAI* 基因的密码子偏好性参数相似,ENc 值在 40.00~58.00,

平均值为 51.98,说明双子叶植物 *GAI* 基因编码氨基酸密码子的偏好性较弱;而单子叶植物 *GAI* 基因的 ENc 值在 34.00~53.00,平均值为 42.64,除铁皮石斛和菠萝外,其它单子叶植物的 ENc 均分布于 35.00 左右,说明单子叶植物 *GAI* 基因的密码子偏好性较强,暗示这些基因的表达量可能较高。双子叶植物偏向使用以 A 或 T 结尾的密码子,而单子叶植物则偏向使用以 G 或 C 结尾的密码子^[30]。GC 含量分析发现,单子叶植物的 GC1、GC2、GC3 和 GC 含量均在 50%以上,说明单子叶植物中 *GAI* 基因对 G 或 C 具有更强的偏好性;双子叶植物中 GC1、GC2、GC3 和 GC 含量的均值依次为:54.79%、42.20%、48.32%和 48.63%,表明密码子第一位碱基偏好使用 G/C,第二和三位碱基偏好使用 A/T(U)。GC 含量为 48.63%,说明双子叶中 *GAI* 基因对 A 或 T 的偏好性更强。苋菜 *AtGAI* 基因的相关参数与双子叶植物类似,说明苋菜 *AtGAI* 基因具有双子叶植物密码子所具备的特性。

2.2.2 不同物种 *GAI* 基因 CDS 序列及 RSCU 值的相关聚类分析

通过邻接法对 21 个 *GAI* 基因进行系统进化树的构建(图 2)。21 个物种可聚为两大类:双子叶植物聚为一类,单子叶植物聚为一类。在双子叶植物中,苋菜 *AtGAI* 基因与石竹目含有甜菜色素的籽粒苋、甜菜、菠菜和藜麦的 *GAI* 基因聚为 1 支,其 ENc 值和 GC3 含量比较接近 ($50.11 \leq ENc \leq 54.48$, $41.91\% \leq GC3 \leq 48.00\%$),且籽粒苋与苋菜的亲缘关系最近;其他含花青素的植物聚为 1 个分支,其中蔷薇目的物种 ($46.64 \leq ENc \leq 54.83$, $54.19\% \leq GC3 \leq 58.61\%$)又单独聚为一个分支。在单子叶植物中,除铁皮石斛和菠萝外,小米、高粱和玉米的 ENc 值和 GC3 含量变化范围很小 ($34.68 \leq ENc \leq 38.15$, $84.78\% \leq GC3 \leq 90.54\%$)。综上所述,物种进化分类越近,其密码子使用偏好性越近。

使用 SPSS 22.0 对 21 个 *GAI* 基因的 RSCU 值进行聚类分析。同样将 21 个物种聚为两大类,但与 CDS 序列构建的进化树略有不同(图 3)。石竹目的植物与烟草、毛白杨、番茄和罂粟聚为一类,其中石竹目植物单独聚为一个支;蔷薇目的植物与单子叶植物及拟南芥、葡萄聚为一类,但蔷薇目植物与单子叶植物明显聚为两个单独的小支;与偏好性分析

表 1 不同物种间 *GAI* 基因密码子选择偏好性相关参数Table 1 Preference related parameters of *GAI* gene codons among different species

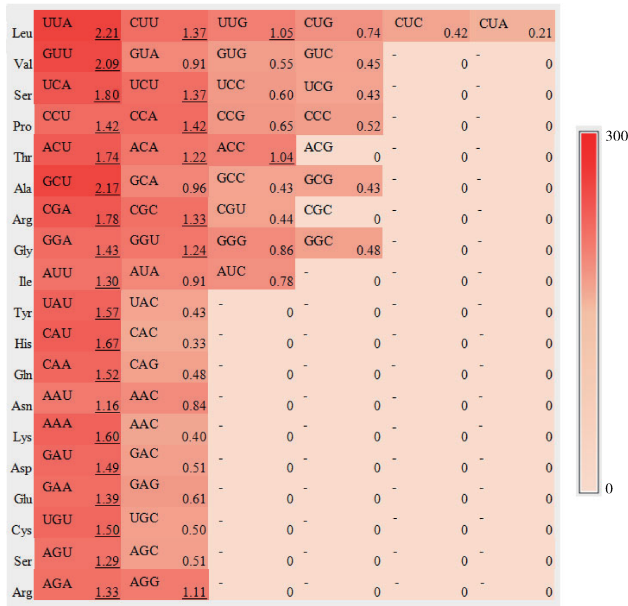
类别 classes	物种 Species	基因名 Gene name	基因登录号 GenBank accession number	G/C 含量/% G/C content			CAI	ENc	
				GC1	GC2	GC3			GC
	苋菜 <i>Amaranthus tricolor</i>	<i>AtGAI</i>	MK049175	51.65	41.91	32.18	41.91	0.195	50.75
	籽粒苋 <i>Amaranthus hypochondriacus</i>	<i>AhGAI</i>	AH004146	55.77	39.84	35.99	43.86	0.219	52.73
	甜菜 <i>Beta vulgaris</i>	<i>BvGAI</i>	XM_010683580.2	50.98	42.46	36.56	43.33	0.194	50.11
	藜麦 <i>Chenopodium quinoa</i>	<i>CqGAI-like</i>	XM_021871600.1	51.70	43.30	44.59	46.53	0.199	54.48
	菠菜 <i>Spinacia oleracea</i>	<i>SoGAI-like</i>	KX026951.1	53.73	44.81	45.45	48.00	0.202	54.47
双子叶植物 Dicyledon	拟南芥 <i>Arabidopsis thaliana</i>	<i>AtGAI</i>	NM_101361.3	58.80	38.76	47.94	48.50	0.249	49.73
	番茄 <i>Solanum lycopersicum</i>	<i>SlGAI</i>	NM_001247436.1	51.95	42.61	29.54	41.37	0.213	44.68
	毛白杨 <i>Populus tomentosa</i>	<i>PtGAI</i>	JX102471.1	55.79	41.06	45.03	47.30	0.250	56.81
	烟草 <i>Nicotiana tabacum</i>	<i>NtGAI-like</i>	XR_001656721.1	29.88	48.27	43.29	40.48	0.199	55.74
	罂粟 <i>Papaver somniferum</i>	<i>PsGAI1-like</i>	XM_026530997.1	55.73	43.01	21.53	40.03	0.214	40.69
	葡萄 <i>Vitis vinifera</i>	<i>VvGAI1</i>	HQ834311.1	60.91	42.64	64.30	55.95	0.228	51.67

表 1(续)

类别 classes	物种 Species	基因名 Gene name	基因登录号 GenBank accession number	G/C 含量/% G/C content					
				GC1	GC2	GC3	GC	CAI	ENc
	櫻桃 <i>Prunus avium</i>	<i>PaGAI</i>	XM_021956293.1	59.73	34.12	59.43	54.19	0.216	54.15
	水蜜桃 <i>Prunus persica</i>	<i>PpGAI</i>	XM_007214894.2	59.43	43.77	60.94	54.71	0.206	54.83
双子叶植物 Dicotyledon	梅花 <i>Prunus mume</i>	<i>PmGAI</i>	NM_001326504.1	60.37	43.98	61.20	55.18	0.207	54.12
	湖北海棠 <i>Malus hupehensis</i>	<i>MhGAI</i>	EF026101.1	60.09	41.94	72.14	58.06	0.251	48.19
	苹果 <i>Malus x domestica</i>	<i>MdGAI</i>	FJ535245.1	60.13	42.70	73.00	58.61	0.248	46.64
	铁皮石斛 <i>Dendrobium catenatum</i>	<i>DcGAI-like</i>	XM_020830052.1	63.93	47.49	67.33	59.59	0.225	50.23
	菠萝 <i>Ananas comosus</i>	<i>AcGAI</i>	LSRQ01008429.1	50.24	42.14	65.32	52.57	0.191	53.02
单子叶植物 Monocotyledon	高粱 <i>Sorghum bicolor</i>	<i>SbGAI</i>	XM_002450738.2	64.83	50.00	84.78	66.54	0.261	38.15
	小米 <i>Setaria italica</i>	<i>SiGAI</i>	XM_004980464.2	63.79	51.06	86.47	67.20	0.228	37.12
	玉米 <i>Zea mays</i>	<i>ZmGAI</i>	XM_008681950.2	64.51	50.39	90.54	68.48	0.263	34.68
均值 Average		—	—	56.58	43.07	54.89	51.61	0.223	50.64

注:GC1、GC2和GC3,分别表示密码子第1、2和3位上G/C含量;CAI,密码子适应指数;ENc,有效密码子数。下同。

Note:GC1,GC2 and GC3, respectively, indicate G/C content at positions 1,2 and 3 of the codon;CAI,Codon adaptation index;ENc, Effective codon number. The same below.

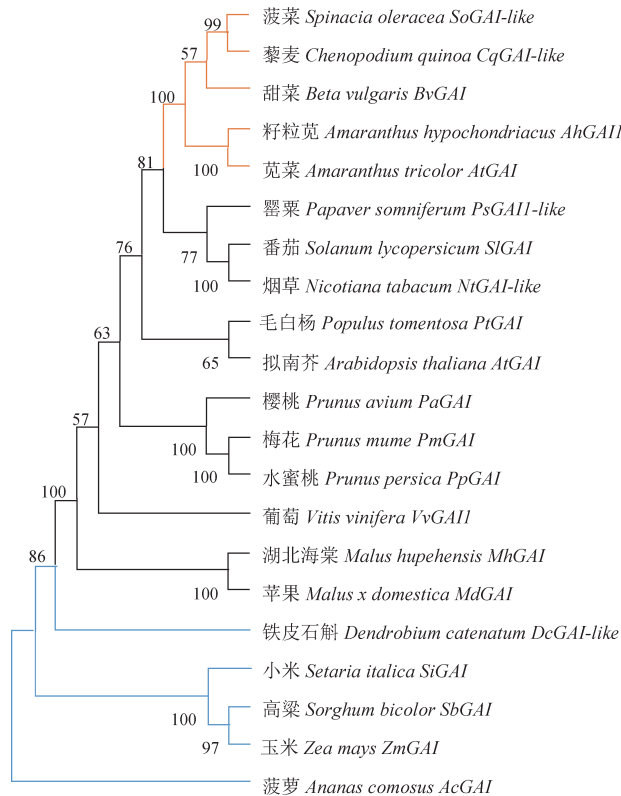


图中数字表示同义密码子相对使用度(RSCU);下划线表示苋菜 *AtGAI* 基因对该密码子使用频率较高。

The numbers in the figure indicate the relative synonymous codon usage (RSCU); Underlined value indicate that amaranth *AtGAI* gene uses this codon higher.

图 1 苋菜 *AtGAI* 基因同义密码子相对使用度分析

Fig. 1 Analysis of relative usage of synonymous codons of *AtGAI* in Amaranth



图中数字表示可信度。

The numbers in the figure indicate the credibility.

图 2 不同物种 *GAI* 相关基因系统进化树

Fig. 2 Construction of an evolutionary tree of *GAI* related gene system

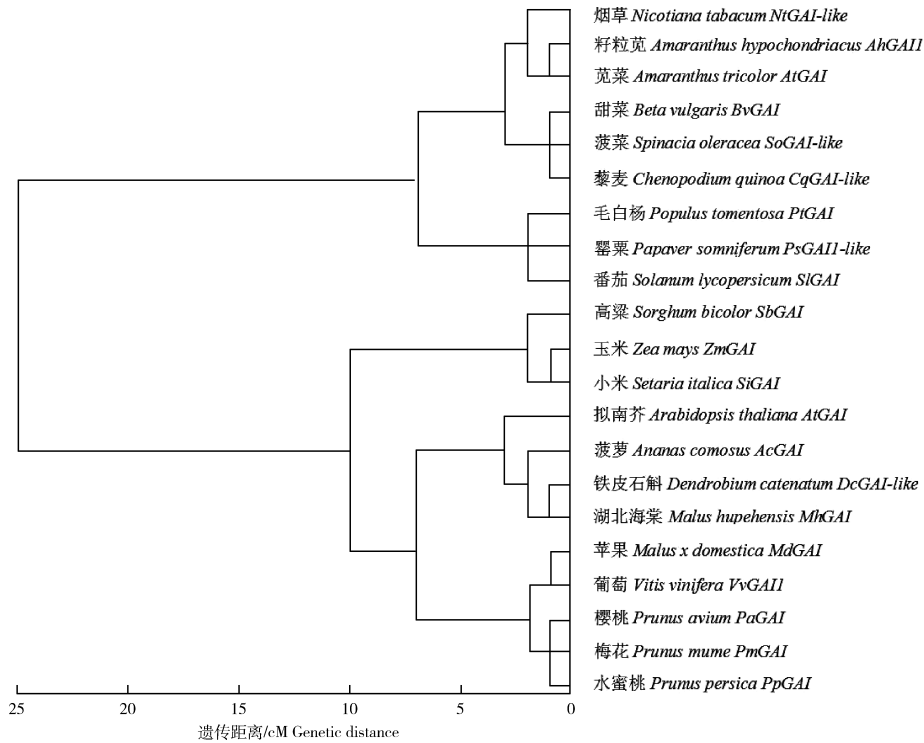


图3 GAI相关基因的RSCU值聚类分析

Fig. 3 RSCU cluster analysis of GAI related genes

结果类似。可见不同物种间GAI基因可能具有相似的密码子使用法则,但部分物种在长期进化过程中形成了一些独特的GAI基因密码子使用偏好性。密码子偏好性分类结果可能在很大程度上反映GAI基因特殊的进化规律,但采用CDS聚类更接近传统物种分类规律。

2.2.3 不同物种GAI基因中性分析

采用SPSS 22.0进行GC1、GC2、GC3、GC和ENc两两Pearson关联分析结果显示,GC与GC1、GC3呈极显著正相关($P < 0.01$),与GC2呈显著正相关($P < 0.05$),说明不同物种GAI基因第1、第2和第3位上的碱基组成较为相似(表2)。对不同物种GAI基因进行中性分析(图4),发现GC3的范围为21.53%~90.54%,GC12的范围为39.08%~57.45%,GC12的范围略小,使大部分基因分布在回归线左右两侧。GC3与GC12的线性归系数为0.166,极显著性相关系数为0.726($P < 0.01$),与Pearson分析一致,其碱基组成差异较小,主要受突变压力影响^[31-36]。苋菜AtGAI基因位于回归线的左侧,与双子叶植物GAI基因的分布距离较近,与单子叶植物GAI基因的分布距离较远,这与聚类分析结果相吻合。

2.2.4 不同物种GAI基因ENc分析

对不同物种GAI基因的ENc进行分析(图5),结果显示,罂粟、番茄、苹果、小米和玉米的基因位点分布在期望曲线的上面,其余基因位点分布在期望曲线下面,苋菜的基因位点离期望曲线最近。分布的趋势与标准曲线的趋势大致吻合,GC3值分布在0.20~0.95,说明GAI的密码子偏好性受碱基突变的影响较大,与中性绘图分析的结果一致。但是也有部分基因位点距离期望曲线较远,说明密码子的偏好性受碱基突变的影响外,也受其他因素(自然选择、基因长度和基因表达量等)的影响^[37-38]。

2.2.5 不同物种GAI基因PR2分析

由图6可知,PR2分析(PR2-bias plot analysis)不同物种GAI基因的A3/(A3+T3)或G3/(G3+C3)值都偏离了0.5,而当该比值偏离0.5时表明基因编码氨基酸时密码子偏好性除了碱基突变之外还受到其他因素的影响^[38]。从横坐标来看,大致均匀分布于0.5区域的左右两侧表明密码子在G和C之间没有显著的偏好性。从纵坐标来看,主要集中在<0.5的区域表明密码子在A和T之间具有一定的偏好性。由此推断自然选择等其他压力在不同物种GAI家族进化过程中也发挥一定作用。

表 2 *GAI* 基因家族成员密码子成分相关性分析

Table 2 Correlative analysis of codon components of *GAI* gene family members

参数 Parameter	GC1	GC2	GC3	GC	ENc
GC1	—				
GC2	0.072	—			
GC3	0.585**	0.492*	—		
GC	0.732**	0.490*	0.975**	—	
ENc	-0.467*	-0.563**	-0.433*	-0.522*	—

注：** 和 * 分别表示 $P < 0.01$ 和 $P < 0.05$ 水平上极显著差异和显著差异。

Note: ** and * indicate significant correlations at 0.01 and 0.05 levels, respectively.

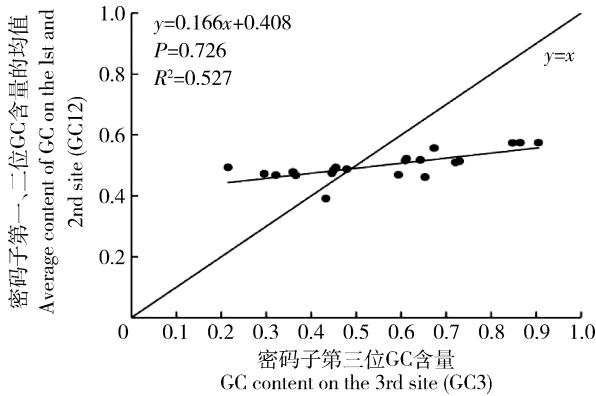


图 4 *GAI* 相关基因密码子使用偏好性中性绘图

Fig. 4 Neutrality plot on the codon usage bias of *GAI* related genes

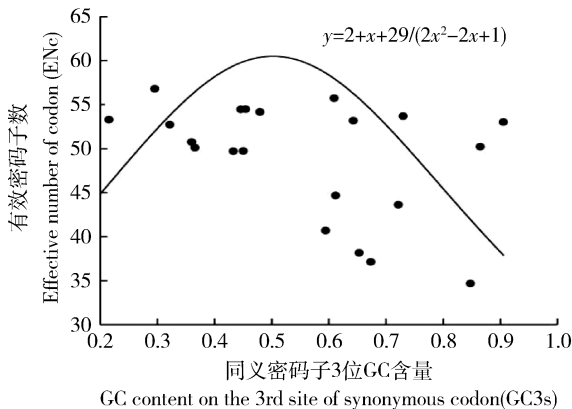
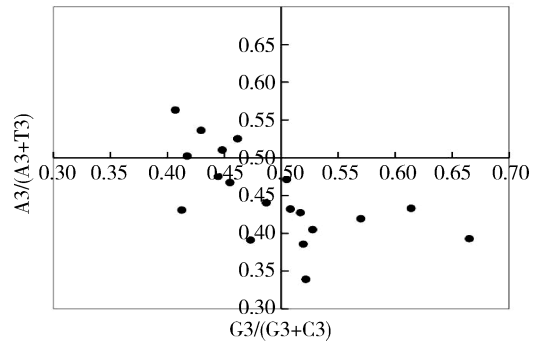


图 5 *GAI* 相关基因 ENc-plot 分布

Fig. 5 Effective number of codons (ENc) used in *GAI* related genes



A3、T3、G3、C3, 分别表示密码子第三位上各个碱基的含量。

A3, T3, G3, and C3 indicate the content of each base in the 3rd position of the codon, respectively.

图 6 奇偶偏好偏差分析

Fig. 6 Parity rule 2 (PR2) bias plot

基因通常需要借助遗传转化体系较为完善的受体进行异源表达。将苋菜 *AtGAI* 基因密码子使用频率与拟南芥、烟草、番茄、甜菜、大肠杆菌和酵母菌基因组密码子使用频率进行比较, 筛选密码子使用差异较大的个数, 见表 3。苋菜 *AtGAI* 基因与大肠杆菌、酵母菌的密码子使用频率差异性较大的个数分别为 28.0 和 9.0, 通常认为比值介于 0.5~2.0 时物种间的使用特性较一致, 否则说明密码子偏性差异较大^[27]。所以酵母菌更适合做苋菜 *AtGAI* 基因的真核表达载体, 若想选用大肠杆菌做原核表达载体, 需针对大肠杆菌的密码子偏好性, 优化苋菜 *AtGAI* 基因的密码子。苋菜 *AtGAI* 基因与拟南芥、烟草和番茄基因组密码子使用频率差异性较大的个数都为 10.0, 说明三者之间没有差异, 都可以作为苋菜 *AtGAI* 基因进行遗传转化的受体; 甜菜

2.3 苋菜 *AtGAI* 基因受体系统的选择

由于苋菜的遗传转化体系尚需完善, 苋菜中的

与苋菜亲缘关系较近,都具有甜菜色素代谢途径, 用频率的差异较大的个数为9.0,是苋菜瞬时表达的
其密码子使用频率与苋菜 *AtGAI* 基因的密码子使 的良好受体。

表3 不同物种 *GAI* 基因与常见模式生物基因组密码子使用偏性比较

Table 3 Comparison of codon usage bias of *GAI* related gene and common pattern organisms

密码子 Codon	氨基酸 Amimo acid	<i>AtGAI</i>	<i>AtGAI</i> / <i>Bv</i>	<i>AtGAI</i> / <i>At</i>	<i>AtGAI</i> / <i>Nt</i>	<i>AtGAI</i> / <i>Sl</i>	<i>AtGAI</i> / <i>Ec</i>	<i>AtGAI</i> / <i>Sc</i>
GCA	Ala	18.18	<u>0.46</u>	1.04	0.79	0.82	0.88	1.13
GCC	Ala	8.26	0.55	0.80	0.66	0.82	<u>0.32</u>	0.66
GCG	Ala	8.26	1.84	0.92	1.42	1.59	<u>0.26</u>	1.35
GCT	Ala	41.32	1.94	1.46	1.32	1.35	<u>2.65</u>	1.96
TGC	Cys	3.31	0.64	<u>0.46</u>	<u>0.46</u>	<u>0.49</u>	<u>0.48</u>	0.70
TGT	Cys	9.92	1.00	0.94	1.01	0.92	1.80	1.24
GAC	Asp	14.88	1.04	0.86	0.88	0.99	0.80	0.74
GAT	Asp	42.98	1.02	1.17	1.16	1.09	1.34	1.14
GAA	Glu	38.02	1.08	1.11	1.06	1.09	1.00	0.78
GAG	Glu	16.53	0.51	0.51	0.56	0.62	0.93	0.87
TTC	Phe	18.18	1.08	0.88	1.01	1.04	1.08	1.00
TTT	Phe	23.14	1.10	1.06	0.92	0.89	1.00	0.89
GGA	Gly	24.79	1.16	1.02	1.07	0.97	<u>2.75</u>	<u>2.27</u>
GGC	Gly	8.26	0.81	0.90	0.74	0.85	<u>0.30</u>	0.85
GGG	Gly	14.88	1.20	1.46	1.42	1.38	1.32	<u>2.48</u>
GGT	Gly	21.49	0.77	0.97	0.96	0.90	0.88	0.90
CAC	His	3.31	0.70	<u>0.38</u>	<u>0.38</u>	<u>0.42</u>	<u>0.34</u>	<u>0.43</u>
CAT	His	16.53	1.54	1.20	1.23	1.07	1.22	1.21
ATA	Ile	11.57	0.98	0.92	0.83	0.83	<u>2.14</u>	0.65
ATC	Ile	9.92	0.57	0.54	0.71	0.71	<u>0.41</u>	0.58
ATT	Ile	16.53	0.64	0.77	0.59	0.59	0.55	0.54
AAA	Lys	33.06	1.22	1.07	1.01	1.06	1.00	0.78
AAG	Lys	8.26	<u>0.22</u>	<u>0.25</u>	<u>0.25</u>	<u>0.27</u>	0.77	<u>0.27</u>
CTA	Leu	3.31	<u>0.34</u>	<u>0.33</u>	<u>0.35</u>	<u>0.33</u>	0.83	<u>0.25</u>
CTC	Leu	4.96	<u>0.46</u>	<u>0.31</u>	<u>0.40</u>	<u>0.44</u>	<u>0.45</u>	0.92
CTG	Leu	11.57	1.41	1.18	1.13	1.10	<u>0.23</u>	1.11
CTT	Leu	21.49	0.79	0.89	0.90	0.86	1.84	1.78
TTA	Leu	34.71	<u>3.10</u>	<u>2.73</u>	<u>2.59</u>	<u>2.41</u>	<u>2.50</u>	1.30
TTG	Leu	16.53	0.89	0.79	0.74	0.68	1.18	0.61
ATG	Met	29.75	1.30	1.21	1.19	1.20	1.10	1.42

表 3(续)

密码子 Codon	氨基酸 Amimo acid	<i>AtGAI</i>	<i>AtGAI</i> / <i>Bv</i>	<i>AtGAI</i> / <i>At</i>	<i>AtGAI</i> / <i>Nt</i>	<i>AtGAI</i> / <i>Sl</i>	<i>AtGAI</i> / <i>Ec</i>	<i>AtGAI</i> / <i>Sc</i>
AAC	Asn	26.45	1.66	1.27	1.48	1.53	1.24	1.06
AAT	Asn	36.36	1.39	1.63	1.30	1.19	1.96	1.00
CCA	Pro	18.18	0.80	1.12	0.92	0.95	<u>2.14</u>	1.00
CCC	Pro	6.61	0.86	1.25	1.00	1.16	1.14	0.97
CCG	Pro	8.26	1.80	0.96	1.65	1.80	<u>0.38</u>	1.56
CCT	Pro	18.18	0.78	0.97	0.97	0.95	<u>2.49</u>	1.34
CAA	Gln	36.36	2.20	1.86	1.76	1.73	<u>2.42</u>	1.32
CAG	Gln	11.57	0.79	0.76	0.77	0.83	0.39	0.96
AGA	Arg	9.92	0.67	0.52	0.62	0.60	<u>3.42</u>	<u>0.47</u>
AGG	Arg	8.26	0.65	0.75	0.68	0.69	<u>4.35</u>	0.90
CGA	Arg	13.22	1.11	<u>2.10</u>	<u>2.49</u>	<u>2.45</u>	<u>3.39</u>	<u>4.41</u>
CGG	Arg	9.92	<u>2.25</u>	<u>2.02</u>	<u>2.68</u>	<u>3.20</u>	1.57	<u>5.83</u>
CGT	Arg	3.31	0.51	<u>0.37</u>	<u>0.44</u>	<u>0.48</u>	<u>0.16</u>	0.51
AGC	Ser	9.92	1.21	0.88	0.99	1.07	0.62	1.02
AGT	Ser	23.14	1.35	1.65	1.74	1.52	<u>2.44</u>	1.63
TCA	Ser	34.71	<u>2.23</u>	1.90	1.97	1.68	<u>4.45</u>	1.85
TCC	Ser	11.57	1.27	1.03	1.13	1.17	1.30	0.81
TCG	Ser	8.26	1.62	0.89	1.56	1.48	0.95	0.97
TCT	Ser	26.45	1.12	1.05	1.32	1.25	<u>3.04</u>	1.13
ACA	Thr	11.57	0.70	0.74	0.66	0.65	1.41	0.65
ACC	Thr	9.92	0.96	0.96	1.02	1.15	<u>0.43</u>	0.79
ACT	Thr	16.53	0.86	0.94	0.81	0.83	1.82	0.81
GTA	Val	16.53	1.62	1.67	1.45	1.48	1.49	1.40
GTC	Val	8.26	0.66	0.65	0.74	0.82	0.55	0.71
GTG	Val	9.92	0.67	0.57	0.59	0.62	<u>0.39</u>	0.94
GTT	Val	38.02	1.33	1.40	1.42	1.36	<u>2.05</u>	1.73
TGG	Trp	11.57	0.88	0.93	0.95	0.86	0.76	1.12
TAC	Tyr	4.96	<u>0.31</u>	<u>0.36</u>	<u>0.37</u>	<u>0.40</u>	<u>0.41</u>	<u>0.34</u>
TAT	Tyr	18.18	1.17	1.25	1.02	0.98	1.10	0.96

注: *AtGAI*、*Bv*、*At*、*Nt*、*Sl*、*Ec*、*Sc* 分别表示苋菜 *AtGAI* 基因及甜菜、拟南芥、烟草、番茄、大肠杆菌、酵母菌基因组密码子使用频率;下划线表示密码子使用频率差异较大。

Note: *AtGAI*, *Bv*, *At*, *Nt*, *Sl*, *Ec*, and *Sc* represent the frequency of codon usage of *Amaranthus tricolor AtGAI* gene and *Beta vulgaris*, *Arabidopsis thaliana*, *Nicotiana tabacum*, *Solanum lycopersicum*, *Escherichia coli*, and *Saccharomyces cerevisiae* genome, respectively. The underline indicates that the codon usage frequency is quite different.

3 讨论

3.1 苋菜 *AtGAI* 基因密码子偏好性分析

双子叶植物偏向使用以 A 或 T 结尾的密码子,而单子叶植物偏向使用以 G 或 C 结尾的密码子^[30]。本研究发现,苋菜 *AtGAI* 基因密码子的 GC 含量、ENc 值和 CAI 值分别为 41.91%、50.75 和 0.195,说明苋菜 *AtGAI* 基因的密码子偏好以 A/T 结尾,选择偏好性较弱,表达水平可能较低。苋菜 *AtGAI* 基因的 RSCU 分析发现,某些密码子(UUA、GUU 和 GCU)具有极强的偏好性,使得 ENc 值偏离 61.00。不同物种间 *GAI* 基因密码子偏好性分析发现:单子叶植物的密码子以 G/C 结尾,与 Wang 等^[30]一致;双子叶植物的密码子多以 A/T 结尾,但葡萄、樱桃、水蜜桃、梅花、湖北海棠和苹果等偏向使用以 G/C 结尾的密码子,可能是由于基因的类型和功能不同造成的^[39],也可能是蔷薇科植物的 *GAI* 基因在系统进化和选择过程中逐渐偏好使用以 G/C 结尾的密码子,以更好的发挥其功能。苋菜 *AtGAI* 基因密码子偏好性相关参数与双子叶植物较为相似,且与石竹目含甜菜色素的植物最接近;与单子叶植物的相关性参数差异较大。以上分析说明苋菜 *AtGAI* 基因密码子偏好性符合双子叶植物对密码子的使用特性,且与亲缘关系较近的物种使用偏好性相似。

3.2 苋菜 *AtGAI* 基因聚类分析比较

亲缘关系较近的物种密码子偏好性较为一致^[40],但不同基因特殊的密码子偏好性导致这种聚类结果往往并不能准确反映物种亲缘关系及真实系统分类^[9]。密码子使用偏性聚类在较小的分类单元中可能提供较为可靠的分类依据,而在较大样本分类中,只能作为传统分类的补充^[41]。在本研究中,CDS 聚类分析和 RSCU 聚类分析存在一定差异。在 CDS 聚类分析中,将 21 个含 *GAI* 基因的物种分为单子叶植物和双子叶植物两大类,苋菜 *AtGAI* 基因归于双子叶植物一类。在 RSCU 聚类分析中,石竹目的植物与烟草、毛白杨、番茄和罂粟聚为一类,蔷薇目的植物与单子叶植物聚为一类,与偏好性分析结果类似。以上分析说明 CDS 聚类结果在亲缘关系鉴定中比 RSCU 聚类结果更为准确,而 RSCU 聚类分析可能在一定程度上反映 *GAI* 基因特殊的个体进化规律。因此,采用 CDS 聚类分析和 RSCU 聚类分析相结合更能提高分类结果的准确度及客观

说明物种的进化规律。

3.3 不同物种 *GAI* 基因密码子偏好性影响因素

密码子使用偏好性主要受突变压力和自然选择效应的影响。此外,密码子偏好性还与基因长度、基因碱基组成、GC 含量、基因表达水平、内含子结构、DNA 甲基化、蛋白疏水性、氨基酸保守性、tRNA 丰度及 mRNA 二级结构有关^[2-3]。本研究对不同物种 *GAI* 基因的密码子进行中性绘图、ENc 绘图及 PR2 偏好性分析,结果表明 *GAI* 基因的密码子偏好性主要受碱基突变的影响,但也受自然选择等其他因素的影响。

3.4 苋菜 *AtGAI* 基因异源表达和遗传转化受体系统的选择

苋菜 *AtGAI* 基因密码子使用频率研究结果表明,酵母真核表达系统更适用于苋菜 *AtGAI* 基因异源表达,但需要对部分密码子进行优化和改造^[42]。通过与不同模式植物对比发现,苋菜与拟南芥、烟草、番茄的偏性差异一样,与甜菜的偏性差异最小,且甜菜和苋菜亲缘关系较近又都含有甜菜色素代谢途径,故甜菜是苋菜 *AtGAI* 基因最理想的遗传转化受体。但目标基因能否高效表达还会受转化效率、miRNA 转录后水平调控以及 mRNA 的结构稳定性等诸多因素影响,如何实现苋菜 *AtGAI* 基因高效遗传转化仍有待深入研究。

参考文献 References

- [1] 陈哲,胡福初,王祥和,范鸿雁,张治礼. 菠萝密码子使用偏好性分析[J]. 果树学报,2017,8(8):946-955
Chen Z, Hu F C, Wang X H, Fan H Y, Zhang Z L. Analysis of codon usage bias of *Ananas comosus* with genome sequencing data[J]. *Journal of Fruit Science*, 2017, 8 (8): 946-955 (in Chinese)
- [2] 赖瑞联,冯新,陈瑾,韦晓霞,陈义挺,吴如健. 橄榄查尔酮异构酶基因 *CHI* 的密码子偏好模式[J]. 应用与环境生物学报, 2017, 23(5): 177-183
Lai R L, Feng X, Chen J, Wei X X, Chen Y T, Wu R J. Codon usage pattern of chalcone isomerase gene (*CHI*) in *Canarium album* (Lour) Raeusch[J]. *Chinese Journal of Applied and Environmental Biology*, 2017, 23(5): 177-183 (in Chinese)
- [3] 李蓉,谢析颖,王雪晶,苏立遥,林玉玲,郭容芳,陈裕坤,赖钟雄,徐涵. 兰科植物 *FNR* 基因的密码子偏好性分析[J]. 热带作物学报, 2018, 39 (6): 1137-1145
Li R, Xie X Y, Wang X J, Su L Y, Lin Y L, Guo R F, Chen Y K, Lai Z X, Xu H. Codon usage bias of Ferredoxin-NADP+ Oxidoreductase (*FNR*) in *Orchidaceae*[J]. *Chinese Journal of*

- Tropical Crops*, 2018, 39 (6): 1137-1145 (in Chinese)
- [4] Ghaemmaghami S, Huh W K, Bower K, Howson R W, Belle A, Dephore N, O'Shea E K, Weissman J S. Global analysis of protein expression in yeast[J]. *Nature*, 2003, 425(6959): 737-741
- [5] 吴宪明, 吴松锋, 任大明, 朱云平, 贺福初. 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
- Wu X M, Wu S F, Ren D M, Zhu Y P, He F C. The analysis method and progress in the study of codon bias[J]. *Hereditas*, 2007, 29(4): 420-426 (in Chinese)
- [6] Pek H B, Klement M, Ang K S, Chung B K, Ow D S, Lee D Y. Exploring codon context bias for synthetic gene design of a thermostable invertase in *Escherichia coli*. [J]. *Enzyme & Microbial Technology*, 2015, 75-76: 57-63
- [7] Pan L L, Wang Y, Hu J H, Ding Z T, Li C. Analysis of codon use features of stearoyl-acyl carrier protein desaturase gene in *Camellia sinensis* [J]. *Journal of Theoretical Biology*, 2013, 334(19): 80-86
- [8] Wei L, He J, Jia X, Qi Q, Liang Z S, Zheng H, Ping Y, Liu S Y, Sun J C. Analysis of codon usage bias of mitochondrial genome in *Bombyx mori* and its relation to evolution[J]. *BMC Evolutionary Biology*, 2014, 14(1): 262
- [9] 赖瑞联, 林玉玲, 钟春水, 赖钟雄. 龙眼生长素受体基因 *TIR1* 密码子偏好性分析[J]. *园艺学报*, 2016, 43(4): 771-780
- Lai R L, Lin Y L, Zhong C S, Lai Z X. Analysis of codon bias of auxin receptor gene *TIR1* in *Dimocarpus longan* [J]. *Acta Horticulturae Sinica*, 2016, 43(4): 771-780 (in Chinese)
- [10] Li W H. Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons[J]. *Journal of Molecular Evolution*, 1987, 24(4): 337-345
- [11] Bulmer M. The selection-mutation-drift theory of synonymous codon usage[J]. *Genetics*, 1991, 129(3): 897-907
- [12] Thanaraj T A, Argos P. Ribosome-mediated translational pause and protein domain organization[J]. *Protein Science A Publication of the Protein Society*, 1996, 5(8): 1594-1612
- [13] 李巍, 萧浪涛. *GAI/RGA* 蛋白家族的研究进展[J]. *生物技术通报*, 2007(5): 39-42
- Li W, Xiao L T. Research progress of *GAI/RGA* family proteins [J]. *Biotechnology Bulletin*, 2007 (5): 39-42 (in Chinese)
- [14] 王小青. 拟南芥 *GAI* 基因转化小桐子的研究[D]. 海口: 海南大学, 2010
- Wang X Q. Studies on transforming *GAI* Gene of *Arabidopsis thaliana* into *Jatropha curcas* [D]. Haikou: Hainan University, 2010 (in Chinese)
- [15] Xie Y, Tan H J, Ma Z X, Huang J R. DELLA proteins promote anthocyanin biosynthesis via sequestering MYB12 and JAZ suppressors of the MYB/bHLH/WD40 complex in *Arabidopsis thaliana* [J]. *Molecular Plant*, 2016, 9(5): 711-721
- [16] 郑学立. 苋菜甜菜红素合成的生理生化与分子生物学研究 [D]. 福州: 福建农林大学, 2016
- Zheng X L. Studies on physiology, biochemistry and molecular biology on betalain synthesis in *Amaranthus tricolor* [D]. Fuzhou: Fujian Agriculture and Forestry University, 2016 (in Chinese)
- [17] 潘君飞, 彭丽云, 赵春丽, 王晓, 张梓浩, 陈裕坤, 赖钟雄, 刘生财. 苋菜 *amaARF6* 基因克隆及其外源激素处理下在幼苗中响应分析[J]. *东北农业大学学报*, 2018, 49(5): 24-32
- Pan J F, Peng L Y, Zhao C L, Wang X, Zhang Z H, Chen Y K, Lai Z X, Liu S C. Cloning of *amaARF6* and its response to exogenous hormones treatment in amaranth seedlings [J]. *Journal of Northeast Agricultural University*, 2018, 49(5): 24-32 (in Chinese)
- [18] Sakuta M. Diversity in plant red pigments: Anthocyanins and betacyanins[J]. *Plant Biotechnology Reports*, 2014, 8(1): 37-48
- [19] Pavokovic D, Krsnikrasol M. Complex biochemistry and biotechnological production of betalains. [J]. *Food Technology & Biotechnology*, 2011, 49(2): 145-155
- [20] Stafford H A. Anthocyanins and betalains: Evolution of the mutually exclusive pathways[J]. *Plant Science*, 1994, 101(2): 91-98
- [21] Cai Y, Sun M, Corke H. HPLC characterization of betalains from plants in the amaranthaceae[J]. *Journal of Chromatographic Science*, 2005, 43(9): 454
- [22] Tanaka Y, Sasaki N, Ohmiya A. Biosynthesis of plant pigments: Anthocyanins, betalains and carotenoids [J]. *The Plant Journal*, 2008, 54(4): 733-749
- [23] 黄欢, 胡慧霞, 姜联合, 王亮生, 任红旭. 植物甜菜色素的研究进展[J]. *生物学通报*, 2014, 49(8): 1-4
- Huang H, Hu H X, Jiang L H, Wang L S, Ren H X. Research progress on plant betalain [J]. *Bulletin of Biology*, 2014, 49(8): 1-4 (in Chinese)
- [24] 赵春丽, 王晓, 潘君飞, 彭丽云, 赖钟雄, 刘生财. 苋菜 *AtGAI* 基因克隆及表达分析[J]. *西北植物学报*, 2019, 39(2): 199-209
- Zhao C L, Wang X, Pan J F, Peng L Y, Lai Z X, Liu S C. Cloning and expression analysis of *AtGAI* gene in *Amaranthus tricolor* L [J]. *Acta Botanica Boreali-Occidentalia Sinica*, 2019, 39(2): 199-209 (in Chinese)
- [25] 李平, 白云凤, 冯瑞云, 王原媛, 张维锋. 籽粒苋苹果酸酶 (NAD-ME) 基因密码子偏好性分析[J]. *应用与环境生物学报*, 2011, 17(1): 12-17
- Li P, Bai Y F, Feng R Y, Wang Y Y, Zhang W F. Analysis of codon bias of NAD-ME gene in *Amaranthus hypochondriacus* [J]. *Chinese Journal of Applied and Environmental Biology*, 2011, 17(1): 12-17 (in Chinese)
- [26] 孙晶, 何涛, 王闰兰, 彭小风, 淳泽. 铁皮石斛尿苷二磷酸葡萄糖焦磷酸化酶基因 (*UGP*) 的密码子偏好性分析[J]. *应用与环境生物学报*, 2014, 20(5): 759-766
- Sun J, He T, Wan R L, Peng X F, Chun Z. The codon usage bias of UDP-glucose pyrophosphorylase gene (*UGP*) in *Dendrobium officinale* [J]. *Chinese Journal of Applied and*

- Environmental Biology*, 2014, 20(5): 759-766 (in Chinese)
- [27] Sharp P M, Li W H. An evolutionary perspective on synonymous codon usage in unicellular organisms[J]. *Journal of Molecular Evolution*, 1986, 24(1-2): 28
- [28] Jiang Y, Deng F, Wang H, Hu Z H. An extensive analysis on the global codon usage pattern of baculoviruses[J]. *Archives of Virology*, 2008, 153(12): 2273-2282
- [29] Ma Q P, Li C, Wang J, Wang Y, Ding Z T. Analysis of synonymous codon usage in *FAD7* genes from different plant species[J]. *Genetics and Molecular Research*, 2015, 14(1): 1414-1422
- [30] 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
- [31] Chen S L, Lee W, Hottes A K, Shapiro L, Mcadams H H. Codon usage between genomes is constrained by genome-wide mutational processes[J]. *Proc Natl Acad Sci U S A*, 2004, 101(10): 3480-3485
- [32] Knight R D, Freeland S J, Landweber L F. A simple model based on mutation and selection explains trends in codon and amino- acid usage and GC composition within and across genomes[J]. *Genome Biology*, 2001, 2(4): 1-13
- [33] Goetz R, Fuglsang A. Correlation of codon bias measures with mRNA levels: Analysis of transcriptome data from *Escherichia coli* [J]. *Biochemical & Biophysical Research Communications*, 2005, 327(1): 4-7
- [34] Ingvarsson P K. Gene expression and protein length influence codon usage and rates of sequence evolution in *Populus tremula* [J]. *Molecular Biology and Evolution*, 2006, 24(3): 836-844
- [35] Sueoka N. Directional mutation pressure and neutral molecular evolution [J]. *Proceedings of the National Academy of Sciences*, 1988, 85(8): 2653-2657
- [36] 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
- [37] Zhang W J, Zhou J, Li Z F. Comparative analysis of codon usage patterns among mitochondrion, chloroplast and nuclear genes in *Triticum aestivum* L [J]. *Journal of Integrative Plant Biology*, 2007, 49(2): 246-254
- [38] Kawabe A, Miyashita N T. Patterns of codon usage bias in three dicot and four monocot plant species[J]. *Genes & Genetic Systems*, 2003, 78(5): 343-352
- [39] Liu Q, Dou S, Ji Z, Xue Q. Synonymous codon usage and gene function are strongly related in *Oryza sativa* [J]. *Biosystems*, 2005, 80(2): 123-131
- [40] Zhou H, Wang H, Huang L F, Naylor M, Clifford P. Heterogeneity in codon usages of sobemovirus genes [J]. *Archives of Virology*, 2005, 150(8): 1591-1605
- [41] Christianson M. Codon usage patterns distort phylogenies from or of DNA sequences[J]. *American Journal of Botany*, 2005, 92(8): 1221-1233
- [42] 晁岳恩, 吴政卿, 杨会民, 何宁, 杨攀. 11种植物 *psbA* 基因的密码子偏好性及聚类分析[J]. *核农学报*, 2011, 25(5): 927-932
- Chao Y E, Wu Z Q, Yang H M, He N, Yang P. Cluster analysis and codon usage bias studies on *psbA* genes from 11 plant species[J]. *Journal of Nuclear Agricultural Sciences*, 2011, 25(5): 927-932 (in Chinese)

责任编辑: 吕晓梅