

膜荚黄芪法呢基焦磷酸合酶基因密码子偏性分析

郎红, 吴松权*

延边大学农学院, 吉林 延吉 133000

摘要: 目的 法呢基焦磷酸合酶(farnesyl pyrophosphate synthase, FPS)是调控黄芪甲苷生物合成途径关键酶,为该基因外源表达选择合适的宿主,进而为提高黄芪甲苷产量提供理论依据。方法 以长白山膜荚黄芪为植物材料,克隆了FPS基因的编码序列。运用EMBOSS及Codon W等在线软件分析了FPS基因的密码子偏好性,并将其与玉米、青蒿等7种植物FPS基因及大肠杆菌基因组密码子偏好性进行比较。结果 膜荚黄芪FPS基因偏好于以A或T碱基结尾的密码子,与大肠杆菌*Escherichia coli*基因组共有22个密码子存在差异。结论 进一步提高膜荚黄芪FPS基因在大肠杆菌中的表达水平,需对其密码子进行优化。

关键词: 膜荚黄芪; 法呢基焦磷酸合酶; 密码子偏好性; 基因表达; 黄芪甲苷

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2015)20-3071-08

DOI: 10.7501/j.issn.0253-2670.2015.20.019

Analysis on codon bias of farnesyl pyrophosphate synthase gene in *Astragalus membranaceus*

LANG Hong, WU Song-quan

College of Agricultural, Yanbian University, Yanji 133000, China

Abstract: Objective Farnesyl pyrophosphate synthase (FPS) is the key enzyme in biosynthesis pathway of astragaloside IV. The purpose of the experiment was to provide the theory basis for selecting appropriate expression systems and regulating the content of astragaloside IV. **Method** FPS gene coding sequence was cloned based on *Astragalus membranaceus* from Changbai Mountain. Synonymous codons usage of FPS gene was analyzed by EMBOSS and Codon W programs and compared with the genome of other seven plants, such as *Zea mays* and *Artemisia apiacea*, and *E. coli*. **Results** FPS gene of *A. membranaceus* was bias toward the codon with A and T at the third codon position and there are 22 codons showing the significant differences between FPS gene of *A. membranaceus* and *E. coli* genome. **Conclusion** The codons need to be optimized to improve the expression level of FPS gene in *E. coli*.

Key words: *Astragalus membranaceus* (Fisch.) Bunge; farnesyl pyrophosphate synthase; codon bias; gene expression; astragaloside IV

密码子是核酸与蛋白质间遗传信息传递的桥梁,每个氨基酸至少由一个密码子编码,编码同一种氨基酸的密码子为同义密码子。在物种进化过程中,每一种基因都会为了适应环境而形成同义密码子使用不均衡的现象,这种现象称之为密码子偏好性,使用频率较高的密码子为最优密码子。分析基因对同义密码子使用情况,选择具有相似密码子使用偏好的宿主,对提高外源基因在宿主中的表达具有重要意义^[1]。

黄芪来源于豆科植物膜荚黄芪 *Astragalus membranaceus* (Fisch.) Bunge 或蒙古黄芪 *Astragalus membranaceus* (Fisch.) Bunge var. *mongholicus* (Bunge) Hsiao 干燥根^[2],是多年生药用草本植物。黄芪甲苷是黄芪主要活性成分之一,具有抗炎、抗病毒、抗氧化、

抗衰老及强心降压等多种药理功效。目前主要采用黄芪皂苷中的黄芪甲苷作为评价黄芪药材质量优劣的标准^[3]。法呢基焦磷酸合酶(farnesyl pyrophosphate synthase, FPS)是黄芪甲苷生物合成途径中的限速步骤,也是调控叶绿素、植物生长调节剂、甾醇等类异戊二烯化合物合成的关键酶^[4-6]。因此,通过外源调控FPS来增加黄芪甲苷的生物合成具有重要的意义。本实验以长白山膜荚黄芪为植物材料,克隆了FPS基因的编码序列,运用EMBOSS及Codon W等在线软件分析了该基因同义密码子的使用特性,并与大肠杆菌基因组密码子的使用进行相似性比较,为FPS基因在异源载体中高效表达及调控黄芪甲苷的生物合成提供理论依据。

收稿日期: 2015-04-24

作者简介: 郎红(1990—),女,硕士研究生,从事种质资源的研究。

*通信作者 吴松权,副教授。Tel: (0433)2435633 E-mail: arswsq@ybu.edu.cn

1 材料与方法

1.1 材料

黄芪药材经延边大学农学院植物学教研室石铁源教授鉴定延边地区膜荚黄芪 *Astragalus membranaceus* (Fisch.) Bunge, 克隆的 FPS 基因 (基因登录号 KP233831) CDS 全长为 1 029 bp, 共编码 343 个氨基酸。玉米、青蒿、长春花等植物的 FPS 完整编码区序列均来源于 GenBank 数据中心, 序列名称及相应的登录号见表 1。

表 1 FPS 基因编码序列来源

Table 1 Sources of coding sequences of FPS gene

物种	基因登录号
玉米 <i>Zea mays</i>	L39789.1
青蒿 <i>Artemisia annua</i>	GQ420346
长春花 <i>Catharanthus roseus</i>	HQ316638
积雪草 <i>Centella asiatica</i>	AY787627
向日葵 <i>Helianthus annuus</i>	AF0198921
拟南芥 <i>Arabidopsis thaliana</i>	NM-117823
檀香 <i>Santalum album</i>	KF011939

1.2 FPS 基因同义密码子偏好性分析

运用 EMBOSS 在线软件 (<http://vm-bioinfo.toulouse.inra.fr/emboss>) 中的密码子适应指数 (codon adaptation index, CAI)、CHIPS (codon heterozygosity in a protein coding sequence) 及 CUSP (create a codon usage table) 分别预测各物种 FPS 基因的 CAI, 有效密码子数 (Enc) 和 (G+C) 量, 并采用 Codon W 在线软件分析 FPS 基因的相对同义密码子使用度 (RSCU)。本研究中采用的大肠杆菌基因组密码子偏好性数据来源于 Codon Usage Database (<http://www.kazusa.or.jp/codon>)。

1.3 物种间 FPS 基因同义密码子偏好性比较

将膜荚黄芪 FPS 基因与来源于 GenBank 中的玉米、青蒿、长春花及积雪草等不同物种植物进行比较, 分析 FPS 基因在物种进化过程中其同义密码子的使用偏好性是否相似, 并通过密码子使用参数的比较预测 FPS 基因在不同植物中的表达水平。

1.4 基于密码子使用偏好性聚类分析

利用 SPSS 19.0 软件将膜荚黄芪与其他 7 个物种植物基于密码子使用偏好性进行聚类分析。在聚类分析过程中, 将被分析的物种作为对象, 59 个同义密码子 (不包括起始密码子 ATG 及 3 个终止密码子 TAG、TAA、TGA) 使用频率作为变量, 利用密

码子 RSCU 值计算密码子间的欧式平方距离系数, 并基于欧式平方距离系数进行两两间比对测定物种间的距离, 根据物种距离构建树状图。

2 结果与分析

2.1 膜荚黄芪 FPS 基因密码子偏好性分析

RSCU 是指对于某一特定的密码子在编码对应氨基酸的同义密码子中的相对概率。用 RSCU 值表示能直观地反映出密码子使用偏好程度^[7]。如果密码子的 RSCU 值等于 1, 表明基因对此密码子并没有明显的偏好; RSCU 值大于 1 时, 表明此密码子的使用频率相对较高。运用 CUSP 及 Condon W 在线软件分析膜荚黄芪 FPS 基因密码子使用偏好性的结果表明 (表 2), 有 26 个密码子的 RSCU 值大于 1, 其中, GCA、GCT、GAT、TTT、GGA、GGT、CAT、ATT、CTT、AAT、CCA、CCT、CAA、CGT、TCA、TCT、ACA、ACT、GTT、TAT 共 20 个密码子是以 A 或 T 碱基结尾, Fraction 值也较高, 是 FPS 基因偏好性密码子。大多数以 G 或 C 碱基结尾的密码子的 RSCU 值小于 1, Fraction 值也较低。整个编码区域内 (A+T) 量 > (G+C) 量, 表明膜荚黄芪 FPS 基因比较偏爱以 A 或 T 碱基结尾的密码子。

2.2 与其他植物 FPS 基因密码子偏好性比较

2.2.1 FPS 基因密码子使用参数比较 CAI 常用于衡量某一基因与高表达基因所用密码子的接近程度, 也是衡量某一基因的表达水平^[8]。此值范围在 0~1, 基因表达量越多, 密码子适应指数越接近 1, 反之亦然。Enc 是衡量特定基因对同义密码子使用不均衡情况, 其数值一般在 20~61, 基因表达量越高, 密码子使用偏好性越强, 其数值越接近 20。GC3s 则表示密码子的第 3 位碱基中 (G+C) 量在第 3 位碱基总量中所占的比率。运用 EMBOSS 中的 CAI、CHIPS 及 CUSP 在线软件分析了不同种属植物的 FPS 基因密码子使用参数并进行比较。由表 3 可见, 膜荚黄芪 FPS 基因的 CAI、Enc、(G+C) 量和 GC3s 分别为 0.674、54.460、41.79%、和 39.65%, 可见 CAI 和 Enc 值偏大, 表明膜荚黄芪 FPS 基因表达水平较高并且各密码子在编码相应的氨基酸时被使用的频率较为一致, (G+C) 量及 GC3s 较低, 说明膜荚黄芪 FPS 基因的编码区比较偏爱以 A 或 T 碱基结尾的密码子。与其他植物相比, FPS 基因在玉米中表达量较高, 并且基因编码区比较偏爱以 G 或 C 结尾的密码子。膜荚黄芪与青蒿、积雪草、

表 2 CUSP 和 Codon W 软件分析膜荚黄芪 FPS 基因密码子偏好性
Table 2 Codon bias of FPS gene analyzed by CUSP and Codon W program

氨基酸	密码子	比例	频率	个数	RSCU
Ala	GCA	0.476	29.155	10	1.90
	GCC	0.143	8.746	3	0.57
	GCG	0.048	2.915	1	0.19
	GCT	0.333	20.484	7	1.33
Cys	TGC	0.600	8.746	3	1.20
	TGT	0.400	5.831	2	0.80
Asp	GAC	0.296	23.324	8	0.59
	GAT	0.704	55.394	19	1.41
Glu	GAA	0.435	29.155	10	0.87
	GAG	0.565	37.901	13	1.13
Phe	TTC	0.385	14.577	5	0.77
	TTT	0.615	23.324	8	1.23
Gly	GGA	0.556	29.155	10	2.22
	GGC	0.056	2.915	1	0.22
	GGG	0.111	5.831	2	0.44
	GGT	0.278	14.577	5	1.11
His	CAC	0.429	8.746	3	0.86
	CAT	0.571	11.662	4	1.14
Ile	ATA	0.125	5.831	2	0.38
	ATC	0.125	5.831	2	0.38
Lys	ATT	0.750	34.985	12	2.25
	AAA	0.385	29.155	10	0.77
	AAG	0.615	46.647	16	1.23
Leu	CTA	0.071	8.746	3	0.43
	CTC	0.119	14.577	5	0.71
	CTG	0.214	26.239	9	1.29
	CTT	0.357	43.732	15	2.14
	TTA	0.095	11.662	4	0.57
	TTG	0.143	17.493	6	0.86
Met	ATG	1.000	20.408	7	1.00
Asn	AAC	0.400	17.493	6	0.80
	AAT	0.600	26.239	9	1.20
Pro	CCA	0.364	11.662	4	1.45
	CCC	0.091	5.915	1	0.36
	CCG	0.182	5.831	2	0.73
	CCT	0.364	11.662	4	1.45
Gln	CAA	0.538	20.408	7	1.08
	CAG	0.462	17.493	6	0.92
Arg	AGA	0.154	5.831	2	0.92
	AGG	0.154	5.831	2	0.92
	CGA	0.154	5.831	2	0.92
	CGC	0.077	2.915	1	0.46
Ser	CGG	0.077	2.915	1	0.46
	CGT	0.385	14.577	5	2.31
	AGC	0.190	11.622	4	1.14
	AGT	0.095	5.831	2	0.57
	TCA	0.238	14.577	5	1.43
	TCC	0.238	14.577	5	1.43
	TCG	0.000	0.000	0	0.00
	TCT	0.238	14.577	5	1.43
Thr	ACA	0.545	17.493	6	2.18
	ACC	0.091	2.915	1	0.36
	ACG	0.091	2.915	1	0.36
	ACT	0.273	8.746	3	1.09
Val	GTA	0.185	14.577	5	0.74
	GTC	0.074	5.831	2	0.30
	GTG	0.185	14.577	5	0.74
	GTT	0.556	43.732	15	2.22
Trp	TGG	1.000	17.493	6	1.00
Tyr	TAC	0.400	23.324	8	0.80
	TAT	0.600	34.985	12	1.20
未命名	TAG	1.000	2.915	1	3.00
未命名	TAA	0.000	0.000	0	0.00
未命名	TGA	0.000	0.000	0	0.00

表 3 不同物种密码子 CAI、Enc 和 (G+C) 量分析
Table 3 Analysis on CAI, Enc, and (G+C) contents of different species

物种	CAI 值	Enc 值	(G+C) /%	GC3s/%
膜荚黄芪	0.674	54.460	41.79	39.65
玉米	0.737	54.362	49.67	58.40
青蒿	0.685	51.556	41.96	43.31
长春花	0.690	56.311	42.39	41.91
积雪草	0.678	57.456	41.79	44.31
向日葵	0.658	56.248	41.72	40.94
拟南芥	0.656	53.638	41.24	47.21
檀香	0.685	55.209	41.50	42.16

向日葵、拟南芥等植物中 FPS 基因的表达水平及编码区域内 GC 量较为相似,表明 FPS 基因在物种进化中变异较小,可能在植物系统中以看家基因的形式存在。

2.2.2 基于密码子使用偏好性的系统聚类 为了进一步了解不同物种间 FPS 基因密码子使用偏好性差

异的大小,根据表 4 中各植物 RSCU 计算密码子间欧式平方距离系数(表 5),并基于密码子欧式平方距离系数进行聚类分析,构建树状图。各物种间的欧式平方距离系数越大,则表示它们在密码子使用上差异越大^[9]。由表 5 可知,膜荚黄芪与长春花和拟南芥间的欧式平方距离系数较小,与玉米间的欧式平方距离系数却较大,表明膜荚黄芪与长春花和拟南芥的关系较近,与玉米的关系较远。

亲缘关系较近的物种常表现出相似的密码子使用频率^[10],因此,基于密码子使用频率的聚类树状图常可用于推测不同物种、不同基因间密码子使用的相似度,也可推测不同物种间的进化关系。本研究的聚类树状图(图 1)将 8 个不同物种分为 4 大类,青蒿、向日葵较为相似聚为一类,膜荚黄芪和拟南芥 FPS 基因同源性较高聚为一类,长春花、积雪草及檀香聚为一类,而玉米则单为一类,说明在系统进化中,豆科与十字花科 FPS 基因的密码子使用偏好性较为相似。

表 4 物种 FPS 基因 RSCU

Table 4 Relatively synonymous codon usage (RSCU) of FPS gene from different species

氨基酸	密码子	RSCU/%							
		膜荚黄芪	玉米	青蒿	长春花	积雪草	向日葵	拟南芥	檀香
Ala	GCA	1.90	0.62	0.89	1.80	1.25	1.33	0.84	2.12
	GCC	0.57	1.23	1.33	1.00	0.50	1.33	0.42	0.47
	GCG	0.19	0.92	0.67	0.00	0.25	0.22	0.63	0.47
	GCT	1.33	1.23	1.11	1.20	2.00	1.11	2.11	0.94
Cys	TGC	1.20	1.00	0.40	1.00	0.50	0.50	1.25	0.80
	TGT	0.80	1.00	1.60	1.00	1.50	1.50	0.75	1.20
Asp	GAC	0.59	0.96	0.81	0.52	0.52	0.88	0.72	0.50
	GAT	1.41	1.04	1.19	1.48	1.48	1.12	1.28	1.50
Glu	GAA	0.87	0.67	1.08	1.04	0.88	1.43	1.19	0.92
	GAG	1.13	1.33	0.91	0.86	1.12	0.57	0.81	1.08
Phe	TTC	0.77	1.12	0.88	0.86	0.93	1.00	1.00	0.50
	TTT	1.23	0.88	1.12	1.14	1.07	1.00	1.00	1.50
Gly	GGA	2.22	0.64	2.32	1.47	1.56	2.10	1.25	1.33
	GGC	0.22	2.08	0.00	0.84	0.89	0.38	0.75	0.44
	GGG	0.44	0.64	0.42	0.42	0.89	0.19	0.75	0.89
	GGT	1.11	0.64	1.26	1.26	1.11	1.33	1.25	1.33
His	CAC	0.86	1.33	0.86	0.57	0.29	0.80	1.14	1.14
	CAT	1.14	0.67	1.14	1.43	1.71	1.20	0.86	0.86
Ile	ATA	0.38	0.14	0.30	0.47	0.91	0.18	0.50	0.90
	ATC	0.38	1.57	0.60	0.47	0.65	0.71	1.00	0.45
	ATT	2.25	1.29	2.10	2.05	1.43	2.12	1.50	1.65

续表 4

氨基酸	密码子	RSCU/%							
		膜荚黄芪	玉米	青蒿	长春花	积雪草	向日葵	拟南芥	檀香
Lys	AAA	0.77	0.54	0.81	0.74	0.69	0.90	0.73	1.00
	AAG	1.23	1.46	1.19	1.26	1.31	1.10	1.27	1.00
Leu	CTA	0.43	0.62	0.29	0.73	0.86	1.14	0.57	0.16
	CTC	0.71	1.38	0.88	1.02	0.57	1.14	1.29	0.65
	CTG	1.29	1.69	0.73	1.02	0.86	0.29	0.71	1.46
	CTT	2.14	1.69	2.93	1.76	2.00	2.00	1.71	2.11
	TTA	0.57	0.15	0.59	0.73	0.43	0.43	0.00	0.49
	TTG	0.86	0.46	0.59	0.73	1.29	1.00	1.71	1.14
	TTT	0.86	0.46	0.59	0.73	1.29	1.00	1.71	1.14
Asn	AAC	0.80	0.77	1.45	0.77	1.00	1.00	1.20	0.60
	AAT	1.20	1.23	0.55	1.23	1.00	1.00	0.80	1.40
Pro	CCA	1.45	1.00	1.33	1.45	1.09	1.20	1.14	0.89
	CCC	0.36	0.00	1.78	1.09	0.73	2.00	0.00	0.00
	CCG	0.73	1.00	0.00	0.00	0.36	0.40	0.00	0.44
	CCT	1.45	2.00	0.89	1.45	1.82	0.40	2.86	2.67
Gln	CAA	1.08	0.57	1.38	1.14	1.14	1.29	1.33	0.67
	CAG	0.92	1.43	0.62	0.86	0.86	0.71	0.67	1.33
Arg	AGA	0.92	0.00	1.50	2.12	0.92	1.64	0.43	0.75
	AGG	0.92	1.12	0.50	1.06	1.38	0.55	1.29	0.75
	CGA	0.92	0.38	0.50	0.71	0.92	0.55	0.00	1.50
	CGC	0.46	2.25	2.00	0.71	1.38	1.64	1.71	2.25
	CGG	0.46	1.12	0.50	0.71	0.92	1.09	0.86	0.75
	CGT	2.31	1.12	1.00	0.71	0.46	0.55	1.71	0.00
	CGU	2.31	1.12	1.00	0.71	0.46	0.55	1.71	0.00
Ser	AGC	1.14	1.85	0.55	0.95	1.26	0.82	0.67	1.67
	AGT	0.57	0.00	0.82	0.63	0.63	0.55	0.67	0.33
	TCA	1.43	2.31	1.09	1.89	1.58	1.09	1.33	2.00
	TCC	1.43	0.46	0.82	0.32	0.32	0.55	1.00	1.00
	TCG	0.00	0.00	0.55	0.32	0.32	0.55	0.00	0.00
	TCT	1.43	1.38	2.18	1.89	1.89	2.45	2.33	1.00
Thr	ACA	2.18	1.33	2.55	0.36	1.60	1.67	1.60	1.20
	ACC	0.36	1.33	0.36	1.45	2.00	0.67	1.07	0.80
	ACG	0.36	0.80	0.00	0.36	0.00	0.33	0.27	0.40
	ACT	1.09	0.53	1.09	1.82	0.40	1.33	1.07	1.60
Val	GTA	0.74	0.57	0.32	0.50	1.45	1.19	0.73	0.95
	GTC	0.30	0.38	0.80	0.33	0.55	0.30	0.18	0.38
	GTG	0.74	1.33	1.12	1.17	1.27	0.89	0.91	1.33
	GTT	2.22	1.71	1.76	2.00	0.73	1.63	2.18	1.33
Tyr	TAC	0.80	1.20	1.22	1.22	1.22	1.00	1.16	0.67
	TAT	1.20	0.80	0.78	0.78	0.78	1.00	0.84	1.33

表 5 FPS 基因间密码子使用偏好性的欧式平方距离系数

Table 5 Coefficients of absolute squared Euclidean distance of codon usage bias among FPS gene from different species

物种	膜荚黄芪	玉米	青蒿	长春花	积雪草	向日葵	拟南芥	檀香
膜荚黄芪	—							
玉米	25.955	—						
青蒿	16.437	33.593	—					
长春花	14.812	24.924	18.478	—				
积雪草	19.480	20.656	20.069	13.174	—			
向日葵	18.153	31.260	7.129	11.853	16.363	—		
拟南芥	14.855	18.238	22.266	18.128	13.989	22.555	—	
檀香	18.080	18.553	26.799	16.781	15.148	25.645	17.688	—

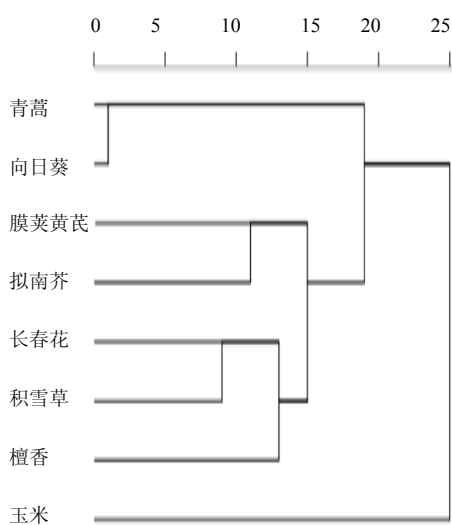


图 1 基于 FPS 基因 RSCU 的聚类树状图

Fig. 1 Cluster analysis dendrogram of RSCU values of different FPS genes

2.3 膜荚黄芪 FPS 基因与大肠杆菌密码子偏好性比较

大肠杆菌作为分子生物学中主要的表达载体，由于其具有转化效率高、生长繁殖快、成本低廉等特点，在基因表达研究中被广泛应用，因此大肠杆菌是目前应用最广泛的蛋白质表达系统^[11]。为了验证 FPS 基因能否在大肠杆菌表达系统中高效表达，将膜荚黄芪 FPS 基因与大肠杆菌进行密码子偏好性比较，密码子使用频率比值是衡量物种间密码子使用偏好差异的指标之一。表 6 为膜荚黄芪 FPS 基因和大肠杆菌基因组不同密码子出现频率以及它们的比值，比值在 0.5~2.0 表明两者的密码子偏好性比较接近，反之亦然。结果表明，膜荚黄芪 FPS 基因与大肠杆菌基因组密码子使用频率差异较大的有 22 个，若要使膜荚黄芪 FPS 基因在大肠杆菌中高效表达尚需对密码子进行优化。

表 6 膜荚黄芪 FPS 基因与大肠杆菌基因组的密码子出现频率及其比值

Table 6 Codon frequency and ratio of *A. membranaceus* FPS gene and *E. coli* genome

氨基酸	密码子	出现频率/%		二者密码子出现频率比值
		膜荚黄芪 FPS 基因	大肠杆菌基因	
Ala	GCA	29.155	20.7	1.41
	GCC	8.746	25.2	0.35
	GCG	2.915	32.3	0.09
	GCT	20.484	15.4	1.33
Cys	TGC	8.746	6.4	1.37
	TGT	5.831	5.2	1.12
Asp	GAC	23.324	19.2	1.21
	GAT	55.394	32.8	1.69
Glu	GAA	29.155	39.3	0.74
	GAG	37.901	18.7	2.03
Phe	TTC	14.577	15.9	0.92
	TTT	23.324	22.2	1.05

续表 6

氨基酸	密码子	出现频率/%		二者密码子出现频率比值
		膜荚黄芪 FPS 基因	大肠杆菌基因	
Gly	GGA	29.155	8.9	3.28
	GGC	2.915	28.1	0.10
	GGG	5.831	11.8	0.49
	GGT	14.577	24.2	0.60
His	CAC	8.746	9.4	0.93
	CAT	11.662	12.8	0.91
Ile	ATA	5.831	5.5	1.06
	ATC	5.831	23.9	0.24
	ATT	34.985	29.7	1.18
Lys	AAA	29.155	34.0	0.86
	AAG	46.647	11.0	4.24
Leu	CTA	8.746	3.9	2.24
	CTC	14.577	10.5	1.48
	CTG	26.239	51.1	0.51
	CTT	43.732	11.4	3.84
	TTA	11.662	13.8	0.85
	TTG	17.493	13.0	1.35
Met	ATG	20.408	27.2	0.75
Asn	AAC	17.493	21.7	0.81
	AAT	26.239	19.2	1.37
Pro	CCA	11.662	8.4	1.39
	CCC	5.915	5.6	1.06
	CCG	5.831	22.4	0.26
	CCT	11.662	7.2	1.62
Gln	CAA	20.408	14.7	1.39
	CAG	17.493	29.4	0.60
Arg	AGA	5.831	2.9	2.01
	AGG	5.831	1.8	3.24
	CGA	5.831	3.8	1.53
	CGC	2.915	20.8	0.14
	CGG	2.915	6.2	0.47
	CGT	14.577	20.2	0.72
Ser	AGC	11.622	16.0	0.73
	AGT	5.831	9.4	0.62
	TCA	14.577	8.1	1.80
	TCC	14.577	8.9	1.64
	TCG	0.000	8.8	0.00
	TCT	14.577	8.7	1.68
Thr	ACA	17.493	8.1	2.16
	ACC	2.915	22.8	0.13
	ACG	2.915	15.0	0.19
	ACT	8.746	9.1	0.96
Val	GTA	14.577	10.9	1.34
	GTC	5.831	14.8	0.39
	GTG	14.577	26.2	0.56
	GTT	43.732	18.1	2.42
Trp	TGG	17.493	15.3	1.14
Tyr	TAC	23.324	12.3	1.90
	TAT	34.985	16.5	2.12
未命名	TAA	2.915	2.0	1.46
未命名	TAG	0.000	0.3	0.00
未命名	TGA	0.000	1.1	0.00

3 讨论

膜荚黄芪 FPS 基因密码子偏爱性分析结果表明,膜荚黄芪 FPS 基因整个编码区域内(A+T)量>(G+C)量,密码子第3位置的碱基偏好于以A或T碱基结尾,共有26个偏好性密码子,并与拟南芥的亲缘关系较近。FPS基因在物种进化中,密码子使用偏好差异较小。

黄芪甲苷具有多种药理活性而被广泛研究,但在黄芪中含量较低,如何提高黄芪甲苷含量亟待解决。目前,通过密码子优化方法提高黄芪甲苷在宿主中的表达量的研究还鲜有报道。本研究分析了膜荚黄芪 FPS 基因同义密码子使用特点,并将膜荚黄芪 FPS 基因与大肠杆菌基因组密码子使用方法进行比较,但两者密码子使用偏好性差异较大,若要实现膜荚黄芪 FPS 基因在大肠杆菌中高效表达,尚需对部分密码子进行优化。

参考文献

- [1] 蒋 玮,吕贝贝,何建华,等. 草菇密码子偏好性分析[J]. 生物工程学报, 2014, 30(9): 1424-1435.
- [2] 吴松权,祖元刚,管清洁,等. 膜荚黄芪苯丙氨酸解氨酶基因的克隆与序列分析[J]. 中草药, 2012, 43(3): 456-460.
- [3] 段立军,孙博航. 黄芪甲苷的研究进展[J]. 沈阳药科大学学报, 2011, 28(5): 410-416.
- [4] 王 峰,吴秋红,高 辉,等. 红树林植物杯萼海桑法呢基焦磷酸合酶的基因克隆及序列分析[J]. 中草药, 2013, 44(16): 2294-2299.
- [5] 李永波,樊庆奇,王宝莲. 植物法呢基焦磷酸合酶基因(FPPS)研究进展[J]. 农业生物技术学报, 2012, 20(3): 321-330.
- [6] 蒙姣荣,陈本勇,黎起秦,等. 罗汉果法呢基焦磷酸合成酶基因的克隆及其序列分析[J]. 中草药, 2011, 42(12): 2512-2517.
- [7] 刘万达,张丙秀,魏媛媛. 草莓 DHAR 基因密码子偏好性分析[J]. 北方园艺, 2014(14): 92-97.
- [8] 王 侃,刘次全,曹 槐,等. 大肠杆菌 mRNA 编码长度、形成二级结构倾向与密码子偏好性关系[J]. 微生物学报, 2006, 46(6): 895-899.
- [9] 李 平,白云凤,冯瑞云,等. 粒子苺苹果酸酶(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.
- [10] Sharp P M, Cowe E, Higgins D G, et al. Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*: A review of the considerable within species diversity[J]. *Nucl Acids Res*, 1988, 16(17): 8207-8211.
- [11] 解庭波. 大肠杆菌表达系统的研究进展[J]. 长江大学学报: 自然科学版, 2008, 5(3): 77-82.