

惠小涵, 程婷婷, 柯卫东, 等. 莲藕 *PPO* 基因密码子偏好性特征分析[J]. 江苏农业学报, 2020, 36(2): 438-446.
doi:10.3969/j.issn.1000-4440.2020.02.026

莲藕 *PPO* 基因密码子偏好性特征分析

惠小涵¹, 程婷婷¹, 柯卫东², 郭宏波¹

(1.西北农林科技大学化学与药学院, 陕西 杨凌 712100; 2.武汉市农业科学院蔬菜研究所, 湖北 武汉 4300065)

摘要: 为了解莲藕多酚氧化酶(*PPO*)基因密码子使用偏好性、不同物种中 *PPO* 基因同源关系及异源表达受体,本研究运用 CondonW 和 EMBOSS 程序对莲藕 *PPO* 基因密码子进行分析。结果表明,莲藕 *PPO* 基因偏好性较弱但偏好于 G/C 结尾的密码子,其中高频密码子有 24 个(同义密码子相对使用度>1),偏好性最强的有 3 个(同义密码子相对使用度>2)。用 SPSS 和 MEGA 软件对 34 种单子叶和双子叶植物的 *PPO* 密码子进行偏性比较和聚类分析,发现不同物种 *PPO* 基因存在不同密码子偏好模式,多数单子叶植物 *PPO* 基因密码子偏好性比双子叶植物强。*PPO* 密码子相关性分析结果表明,莲藕 *PPO* 密码子使用偏好主要受突变压力的影响。在异源转化受体选择中,大肠杆菌真核表达系统适合莲藕 *PPO* 的异源表达,拟南芥则是莲藕 *PPO* 转基因的理想受体。莲藕 *PPO* 密码子偏好模式分析为后续的 *PPO* 基因遗传转化研究提供了理论依据。

关键词: 莲藕; *PPO*; 密码子; 偏好性

中图分类号: S645.1 **文献标识码:** A **文章编号:** 1000-4440(2020)02-0438-09

Analysis on codon preference of *PPO* gene in lotus root

HUI Xiao-han¹, CHENG Ting-ting¹, KE Wei-dong², GUO Hong-bo¹

(1.College of Chemistry and Pharmacy, Northwest A&F University, Yangling 712100, China; 2.Institute of Vegetable, Wuhan Academy of Agricultural Sciences, Wuhan 430065, China)

Abstract: To explore the codon usage preference of polyphenol oxidase(*PPO*) gene in lotus root, the homology of *PPO* gene in different species and heterologous expression receptors, the codon of *PPO* gene was analyzed by CondonW and EMBOSS programs. The results showed that the *PPO* gene favored codons ending in G or C. There were 24 high-frequency codons (relative synonymous codon usage>1), and three codons (relative synonymous codon usage>2) had the strongest preference. The *PPO* codon bias of 34 monocotyledonous and dicotyledonous plants was analyzed by SPSS and MEGA software. It was found that *PPO* genes from different species had different codon preference patterns, and most monocotyledonous *PPO* genes had stronger codon preference than dicotyledons. Correlation analysis of *PPO* codon indicated that the *PPO* codon usage preference in lotus root was mainly affected by mutation. In the selection of heterologous transforming receptors, eukaryotic expression system of *Escherichia coli* is suitable for heterologous expression of *PPO* in lotus root, and *Arabidopsis thaliana* is an ideal receptor for *PPO* transgene. The *PPO* codon preference pattern analysis of lotus root provides a theoretical basis for the subsequent study of *PPO* gene transformation.

Key words: lotus root; polyphenol oxidase (*PPO*); codon; preference

收稿日期:2019-10-29

基金项目:国家重点研发计划项目(2016YFD0100204-29);国家杨凌农业高新技术产业示范区科技攻关计划项目(201SF-08)

作者简介:惠小涵(1996-),女,陕西渭南人,硕士研究生,主要从事多酚氧化酶基因家族在莲藕等植物中的功能研究。(E-mail) huixh@nwsuaf.edu.cn

通讯作者:郭宏波, (E-mail) hbguo@nwsuaf.edu.cn

密码子是在信使 RNA 上相邻的 3 个核苷酸残基组成的基本单位,在合成蛋白质过程中,信使 RNA 上每个三联体核苷酸会编码对应某一个氨基酸,而一种氨基酸会有一种或多种密码子,称为密码子简并性^[1],这是氨基酸合成的规律。其中对应同一种氨基酸的密码子称为同义密码子^[2]。在生物长期进化过

程中,不同物种的密码子使用频率存在着较大差异,各种生物偏好使用同义密码子,称为密码子偏好性^[3]。在转基因研究中,在外源基因^[4]的表达和翻译水平等方面会受到不同物种密码子偏好性差异的影响。研究密码子偏好性,便于选择合适的宿主表达系统^[5],以达到提高基因表达水平的目的^[6]。莲藕(*Nelumbo nucifera Gaertn*)是莲的地下茎,属木兰亚纲,盛产于湖北等地,有较高的营养价值。莲藕味甘多液可清热凉血,含有黏液蛋白和膳食纤维,可通便止泻、健脾开胃。同时含有铁、钙等多种微量元素,可补气益血、增强人体免疫力,并富含大量单宁酸,可以止血散瘀^[7]。然而莲藕不耐储存,极易发生褐变现象,这与多酚氧化酶(PPO)的含量和活性密切相关。多酚氧化酶在自然界中分布极其广泛,在植物^[8]、动物^[9]和真菌^[10]中均可检测到且发挥重要作用,如增强抗性等。多酚氧化酶主要特点是存在分子氧时能催化酚类物质生成有色醌进而形成黑色素。密码子偏好性在芍药^[11]、小麦^[12]、杉木^[13]、橄榄^[14]等植物中均有研究,目前莲藕 PPO 基因密码子使用偏好性还未见

相关研究。通过对莲藕 PPO 基因密码子使用偏好性的分析,讨论其形成的主要因素,再与拟南芥、烟草、酵母和大肠杆菌基因组密码子进行比较分析,为莲藕 PPO 后续的基因研究中选择最佳受体提供依据,也为深入研究莲藕 PPO 的转基因受体、功能验证、蛋白质结构等提供基础数据。

1 材料与方法

1.1 数据来源

莲藕 PPO 基因序列全长 2 069 bp (GenBank 登录号为 FJ999635.1),其开放阅读框为 1 794 bp,编码 597 个氨基酸。凤梨、油棕、小麦、玉米、木薯、川桑、可可等 34 种植物的 PPO 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*)和烟草(*Nicotiana tabacum*)等生物基因的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/>)。

表 1 PPO 基因的编码区来源

Table 1 Sources of coding region of polyphenol oxidase (PPO) gene

物种	登录号	物种	登录号
凤梨 (<i>Ananas comosus</i>)	AY149882.1	莲 (<i>Nelumbo nucifera</i>)	FJ999635.1
二穗短柄草 (<i>Brachypodium distachyon</i>)	XM_010228972.3	苹果 (<i>Malus domestica</i>)	BAA21677.1
小麦 (<i>Triticum aestivum</i>)	HQ228152.1	木薯 (<i>Manihot esculenta</i>)	XM_021744153.1
玉米 (<i>Zea mays</i>)	XM_008665268.3	川桑 (<i>Morus notabilis</i>)	XP_010100570.1
谷子 (<i>Setaria italica</i>)	LC008435.1	毛果杨 (<i>Populus trichocarpa</i>)	AEH41424.1
高粱 (<i>Sorghum bicolor</i>)	KJ425575.1	可可 (<i>Theobroma cacao</i>)	EX097990.1
日本水稻 (<i>Oryza sativa Japonica Group</i>)	ABC23056.1	白梨 (<i>Pyrus bretschneideri</i>)	AEG78292.1
节节麦 (<i>Aegilops tauschii</i> subsp. <i>Tauschii</i>)	XP_020176750.1	沙梨 (<i>Pyrus pyrifolia</i>)	BAB64530.1
大麦 (<i>Hordeum vulgare</i> subsp. <i>Vulgare</i>)	BAJ10871.1	榲桲 (<i>Cydonia oblonga</i>)	ABY84850.1
二型花属 (<i>Dichanthelium oligosanthes</i>)	OEL17564.1	橡胶树 (<i>Hevea brasiliensis</i>)	XP_021649193.1
大叶藻 (<i>Zostera marina</i>)	KMZ58036.1	鲁桑 (<i>Morus alba</i> var. <i>multicaulis</i>)	ALF12286.1
油棕 (<i>Elaeis guineensis</i>)	XP_010910928.1	毛叶番荔枝 (<i>Annona cherimolia</i>)	AAO16865.1
枣 (<i>Ziziphus jujuba</i>)	ADR78836.1	蓝花楼斗菜 (<i>Aquilegia coerulea</i>)	PIA30668.1
金花茶 (<i>Camellia nitidissima</i>)	ACM43505.1	蜀罂粟 (<i>Argemone mexicana</i>)	ACJ76786.1
猕猴桃 (<i>Actinidia chinensis</i>)	PSS24660.1	西洋蒲公英 (<i>Taraxacum officinale</i>)	CAQ76694.1
橄榄 (<i>Canarium album</i>)	AEY78528.1	胡桃 (<i>Juglans regia</i>)	ACN86310.1
甜樱桃 (<i>Prunus avium</i>)	XP_021830806.1	葡萄 (<i>Vitis vinifera</i>)	AAB41022.1

1.2 数据统计分析

采用 CondonW 软件和 CHIPS 在线程序对 PPO 基因的编码序列 (CDS) 的有效密码子数 (ENC)、同义密码子相对使用度 ($RSCU$)、密码子适应指数

(CAI)、第 3 位碱基 G+C 含量 ($GC3s$) 和总 G+C 含量 (GC) 等密码子偏好性参数进行统计。将 34 种植物的 PPO 基因的密码子使用偏性用 SPSS 软件进行关联分析,进一步对 PPO 密码子进行中性绘图分

析、ENc-plot 分析以及 PR2-plot 分析。利用 MEGA 软件并结合组间邻接法 (Neighbour joining) 构建 *PPO* 基因 CDS 系统发育树。

2 结果与分析

2.1 莲藕 *PPO* 密码子的偏好性分析

2.1.1 莲藕 *PPO* 的 *ENc*、*GC*、*GC3s* 和 *CAI* 分析 由 CondonW 程序分析莲藕 *PPO* 基因密码子的 *GC* 和有效密码子数^[15]。结果显示莲藕 *PPO* 基因中密码子的 *GC* 和 *GC3s* 分别为 51.8% 和 66.3%, 说明莲藕 *PPO* 基因中的密码子比较偏好以 G/C 结尾。*ENc* 值为 52.12, 大于 35^[16], 接近 61, 说明莲藕 *PPO* 密码子偏好性较弱, 基因表达水平偏低。莲藕 *PPO* 的 *CAI* 为 0.204, 远小于 1, 进一步证明莲藕 *PPO* 对密码子的偏好性弱。

2.1.2 莲藕 *PPO* 同义密码子相对使用度 利用 CondonW、EMBOSS 在线程序分析各基因的 *RSCU* 值^[17]。同义密码子的相对使用度 (*RSCU*) 是指在不影响基因编码的氨基酸组成情况下, 密码子的实际频率与期望密码子频率的比值, 它能够直接反应密码子的使用偏好性^[18]。若 *RSCU* 大于 1 则表示密码子的偏好性较强, *RSCU* 小于 1 则表示密码子偏好性较弱, *RSCU* 值接近于 1 则表示密码子没有偏好性^[19]。计算结果 (表 2) 表明, 在莲 *PPO* 基因密码子中, *RSCU*<1 的有 35 个密码子, *RSCU*=1 的有 3 个密码子, *RSCU*>1 的有 24 个密码子, 其中 CAG、UCG、AGG 3 个密码子的 *RSCU* 值均大于 2, 说明莲藕 *PPO* 密码子中共有 24 个具有偏好性, 其中有 3 个偏好性极强。

2.2 不同物种间 *PPO* 密码子偏好性分析

2.2.1 不同物种间 *ENc*、*RSCU* 及 *GC* 分析 不同物种间 *PPO* 基因密码子偏好性参数如表 3 所示, 发现单子叶、双子叶植物 *PPO* 基因的 *CAI* 值相近。且范围为 0.165~0.318, 远小于 1, 表明这 34 个物种 *PPO* 密码子使用偏好性都较弱。单子叶植物凤梨、二穗短柄草、小麦等 *PPO* 基因的 *ENc* 值为 31.00~38.00, 平均值为 36.85, 双子叶植物莲、苹果、木薯等 *PPO* 基因的 *ENc* 值为 49.00~56.00, 平均值为 53.08, 表明这些物种的 *PPO* 密码子偏好性及基因表达水平有差异, 单子叶植物的 *PPO* 基因表达高于双子叶植物, 与基于 *CAI* 的分析结果相同。34 个物种的 *GC3s* 为 0.174~0.947, 均值为 0.682, *GC* 为 0.342~0.704, 均值为 0.569。表明不同物种 *PPO* 更偏向于 G/C 结尾的

密码子。其中有 10 种单子叶植物的 *GC3s* 含量均大于 0.8, 偏向于 G/C 结尾的密码子, 谷子 (*Setaria italica*) 对 G/C 结尾的密码子的偏性最强。双子叶植物中的 *GC* 含量均分布在 0.5 左右, 其中蓟罂粟 (*Argemone mexicana*) 的 *GC3s* 为 0.342, 偏向于 A/U 结尾的密码子。34 个物种 *PPO* 密码子偏好性的 *RSCU* 值^[20] 分析结果表明, *RSCU* 值>1.0 的密码子数最多有 31 个 (大叶藻、橡胶树和木薯); 最少的有 24 个 (莲藕)。 *RSCU* 值>1.5 的密码子数最多有 25 个 (大麦); 最少有 6 个 (可可、金茶花和葡萄)。

表 2 莲藕 *PPO* 的密码子偏性分析

Table 2 Codon bias analysis of *PPO* gene in lotus root

氨基酸	密码子	数目	相对同一密码子使用度	氨基酸	密码子	数目	相对同一密码子使用度
Phe	UUU	5	0.43	Ser	UCU	5	0.81
	UUC	18	1.57		UCC	6	0.97
Leu	UUA	2	0.23	Pro	UCA	3	0.49
	UUG	13	1.47		UCG	13	2.11
	CUU	8	0.91	Thr	CCU	8	0.70
	CUC	12	1.36		CCC	17	1.48
Ile	CUA	3	0.34		CCA	12	1.04
	CUG	15	1.70		CCG	9	0.78
	AUU	7	0.91	Ala	ACU	3	0.46
	AUC	9	1.17		ACC	12	1.85
Met	AUA	7	0.91		ACA	2	0.31
	AUG	18	1.00		ACG	9	1.38
Val	GUU	8	0.86	Cys	GCU	14	1.30
	GUC	8	0.86		GCC	20	1.86
	GUA	5	0.54		GCA	4	0.37
	GUG	16	1.73		GCG	5	0.47
Tyr	UAU	5	0.45	Trp	UGU	2	0.44
	UAC	17	1.55		UGC	7	1.56
	UAG	0	0	Arg	UGG	8	1.00
His	CAU	2	0.33		CGU	6	1.00
	CAC	10	1.67		CGC	4	0.67
Gln	CAA	7	0.82		CGA	3	0.50
	CAG	10	2.28	Ser	CGA	4	0.67
Asn	AAU	11	0.81		AGU	5	0.81
	AAC	16	1.19	Arg	AGC	5	0.81
Lys	AAA	11	0.48		AGA	5	0.81
	AAG	35	1.52	Gly	AGG	14	2.33
Asp	GAU	11	0.51		GGU	9	0.95
	GAC	32	1.49		GGC	9	0.95
Glu	GAA	12	0.73		GGA	11	1.16
	GAG	21	1.27		GGG	9	0.95

Phe: 苯丙氨酸; Leu: 亮氨酸; Ile: 异亮氨酸; Met: 蛋氨酸; Val: 缬氨酸; Tyr: 酪氨酸; His: 组氨酸; Glu: 谷氨酸; Asu: 天冬酰胺; Lys: 赖氨酸; Asp: 天冬氨酸; Glu: 谷氨酸; Ser: 丝氨酸; Pro: 脯氨酸; Thr: 苏氨酸; Ala: 丙氨酸; Cys: 半胱氨酸; Trp: 色氨酸; Arg: 精氨酸; Gly: 甘氨酸。

表 3 不同物种间 PPO 的偏好性参数分析

Table 3 Preference parameter analysis of PPO in different species

物种	ENc	GC	GC3s	CAI	RSCU>1.0	RSCU>1.5
凤梨 (<i>Ananas comosus</i>)	38.88	0.638	0.846	0.268	25	23
二穗短柄草 (<i>Brachypodium distachyon</i>)	36.29	0.666	0.885	0.293	27	20
小麦 (<i>Triticum aestivum</i>)	31.30	0.686	0.947	0.293	27	24
玉米 (<i>Zea mays</i>)	35.83	0.669	0.889	0.296	28	24
谷子 (<i>Setaria italica</i>)	31.29	0.702	0.953	0.300	27	23
高粱 (<i>Sorghum bicolor</i>)	34.40	0.672	0.895	0.318	27	23
日本水稻 (<i>Oryza sativa Japonica Group</i>)	31.47	0.692	0.943	0.295	26	24
节节麦 (<i>Aegilops tauschii</i> subsp. <i>Tauschii</i>)	31.26	0.686	0.947	0.298	27	24
大麦 (<i>Hordeum vulgare</i> subsp. <i>Vulgare</i>)	33.47	0.685	0.917	0.279	25	25
二型花属 (<i>Dichanthelium oligosanthes</i>)	31.55	0.704	0.946	0.293	27	24
大叶藻 (<i>Zostera marina</i>)	54.39	0.499	0.489	0.210	31	11
油棕 (<i>Elaeis guineensis</i>)	52.08	0.563	0.677	0.223	28	10
莲 (<i>Nelumbo nucifera</i>)	52.98	0.540	0.657	0.203	24	12
苹果 (<i>Malus domestica</i>)	56.29	0.505	0.554	0.197	27	8
木薯 (<i>Manihot esculenta</i>)	49.87	0.425	0.360	0.206	31	10
川桑 (<i>Morus notabilis</i>)	56.29	0.505	0.554	0.197	27	8
毛果杨 (<i>Populus trichocarpa</i>)	54.89	0.448	0.388	0.201	28	8
可可 (<i>Theobroma cacao</i>)	56.25	0.453	0.441	0.208	27	6
白梨 (<i>Pyrus bretschneideri</i>)	52.17	0.526	0.620	0.227	26	8
沙梨 (<i>Pyrus pyrifolia</i>)	52.42	0.524	0.621	0.230	26	8
榲桲 (<i>Cydonia oblonga</i>)	52.62	0.524	0.635	0.232	26	9
橡胶树 (<i>Hevea brasiliensis</i>)	52.73	0.430	0.372	0.205	31	13
鲁桑 (<i>Morus alba</i> var. <i>multicaulis</i>)	56.71	0.508	0.561	0.204	29	8
毛叶番荔枝 (<i>Annona cherimolia</i>)	44.26	0.594	0.781	0.240	30	17
蓝花耧斗菜 (<i>Aquilegia coerulea</i>)	48.77	0.421	0.317	0.165	25	15
薊罂粟 (<i>Argemone mexicana</i>)	38.67	0.342	0.174	0.184	28	21
西洋蒲公英 (<i>Taraxacum officinale</i>)	56.16	0.462	0.441	0.227	27	10
胡桃 (<i>Juglans regia</i>)	56.09	0.486	0.505	0.238	27	8
葡萄 (<i>Vitis vinifera</i>)	57.92	0.488	0.509	0.213	29	6
枣 (<i>Ziziphus jujuba</i>)	58.23	0.461	0.438	0.210	27	7
金花茶 (<i>Camellia nitidissima</i>)	57.41	0.487	0.490	0.189	29	6
猕猴桃 (<i>Actinidia chinensis</i>)	51.75	0.561	0.702	0.201	26	13
橄榄 (<i>Canarium album</i>)	51.64	0.432	0.373	0.216	28	9
甜樱桃 (<i>Prunus avium</i>)	53.58	0.510	0.578	0.182	28	8

ENc:有效密码子数;GC:G+C 含量;GC3s:第 3 位 G+C 含量;CAI:密码子适应指数;RSCU:同义密码子相对使用度。

2.2.2 PPO 碱基组成分析 用 SPSS 软件对 PPO 密码子成分进行相关性分析^[21],对 GC、GC3s、ENc、CAI、密码子偏爱指数(CBI)、最优密码子使用频率(Fop)、第 3 位碱基 U 含量(U3s)、第 3 位碱基 A 含量(A3s)之间的两两相关性进行分析^[22]。表 4 中 ENc 与 U3s、A3s 这两个参数呈正相关,ENc 和 GC、GC3s、CAI、CBI、Fop 这 5 个参数呈负相关,说明 PPO 密码子偏好性越强,其碱基中 GC 越高。如图 1A 所

示,PPO 分布在中性分析图回归线的两侧,第 1、2 位碱基 G+C 含量(GC12)和 GC3s 的线性回归系数(斜率)为 0.096 8,且在相关性分析中存在极显著相关($P<0.01$)。由此推断,密码子的偏好性主要受突变影响。如图 1B 所示,在 ENc-plot 分析中,PPO 基因密码子分布在标准曲线附近,再次验证其密码子偏好性仅受突变影响。此外,图 1C 显示,在 PR2-plot 分析中,基因较多集中在图的中间位置,即密码

子 C+G 和 A+U 频率基本相同,同时有较少的基因出现在平面图的其他位置,表明多数物种 *PPO* 密码

子偏好模式主要受突变压力作用。

表 4 *PPO* 密码子成分相关分析

Table 4 Correlation analysis of *PPO* related parameters

参数	GC	GC3s	ENc	CAI	CBI	Fop	U3s	A3s
GC	1.000							
GC3s	0.994 **	1.000						
ENc	-0.781	-0.749	1.000					
CAI	0.895 **	0.876 **	-0.836 **	1.000				
CBI	0.963 **	0.947 **	-0.833 **	0.962 **	1.000			
Fop	0.964 **	0.949 **	-0.830 **	0.967 **	0.999 **	1.000		
U3s	-0.991 **	-0.998 **	0.741 **	-0.872 **	-0.945 **	-0.948 **	1.000	
A3s	-0.992 **	-0.995 **	0.752 **	-0.869 **	-0.941 **	-0.944 **	0.988 **	1.000

GC、GC3s、ENc、CAI 见表 3 注;CBI:密码子偏爱指数;Fop:最优密码子使用频率;U3s:第 3 位碱基 U 含量;A3s:第 3 位碱基 A 含量。**表示在 0.01 水平上显著相关。

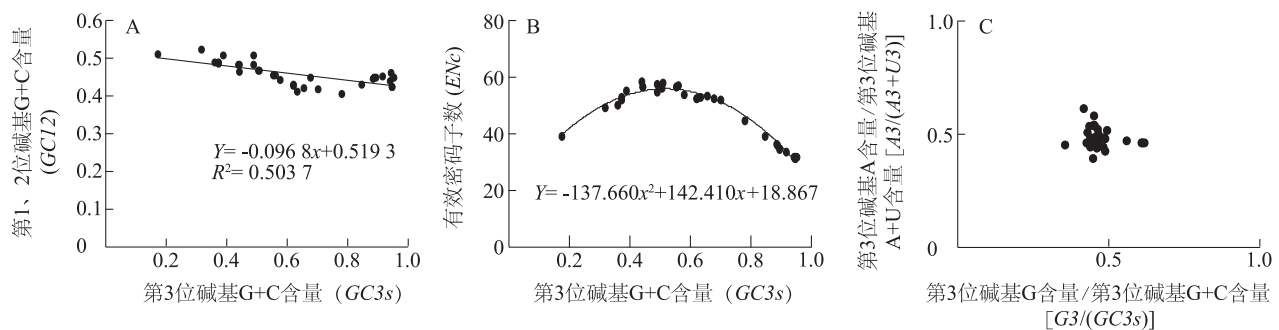


图 1 *PPO* 中性分析图 (A)、ENc-plot 分析图 (B)、PR2-plot 分析 (C)

Fig.1 Neutrality plot (A), ENc-plot (B) and PR2-plot (C) of *PPO*

2.2.3 密码子选用偏好性和 CDS 序列的系统聚类

34 个物种 *PPO* 基因的密码子通过邻接法^[23]进行系统聚类分析^[24],如图 2 所示。在系统发育树中,34 个物种大致按照原本的单子叶和双子叶植物进行分类,说明单子叶植物和双子叶植物的 *PPO* 密码子偏好模式不同。单子叶植物玉米、高粱、大麦、节节麦归为一小类,且同属禾本科,表明亲缘关系相近的物种表现出相近的 *PPO* 密码子使用偏好性。但是同属蔷薇科的苹果和白梨、沙梨没有优先聚类,且遗传距离大于不同属苹果和川桑,同属禾本科的小麦和日本水稻也遗传距离较远,表明不同物种 *PPO* 密码子偏好性的差异与物种进化没有必然联系。整体来看,莲藕 *PPO* 与其他植物 *PPO* 有较高同源性。其中莲藕 *PPO* 与葡萄 *PPO* 的亲缘关系最近,与猕猴桃 *PPO* 的亲缘关系相对于其他植物较近。

2.3 莲藕 *PPO* 基因受体系统的选择

探究相关基因的功能特征常用的研究植物为拟

南芥^[25]和烟草^[26],基因原核表达和真核表达系统常用大肠杆菌^[27-30]和酵母菌^[31]为受体。目前,莲藕的遗传转换没有成熟体系,*PPO* 相关基因的调控网络并不明确,莲藕 *PPO* 基因的研究通常需要异源表达^[32],借助如大肠杆菌、酵母菌等遗传转换体系成熟的受体。不同物种存在密码子使用模式差异,在异源表达过程中,应采用密码子使用模式较为相似的受体,提高转录、翻译效率,进而提高表达量。将莲藕 *PPO* 密码子分别与大肠杆菌、酵母菌、拟南芥、烟草基因组密码子使用频率进行对比分析。结果表明莲藕 *PPO* 基因与大肠杆菌有 20 组密码子有较大的差别,而与酵母有 27 组密码子存在较大差别(表 5),说明大肠杆菌真核异源表达系统更适用于莲藕 *PPO* 异源表达。莲藕 *PPO* 基因与拟南芥基因密码子存在较大差别的有 19 组,与烟草基因密码子存在较大差别的有 21 组,说明在莲藕 *PPO* 基因遗传转化试验中,拟南芥更适合作为异源受体。

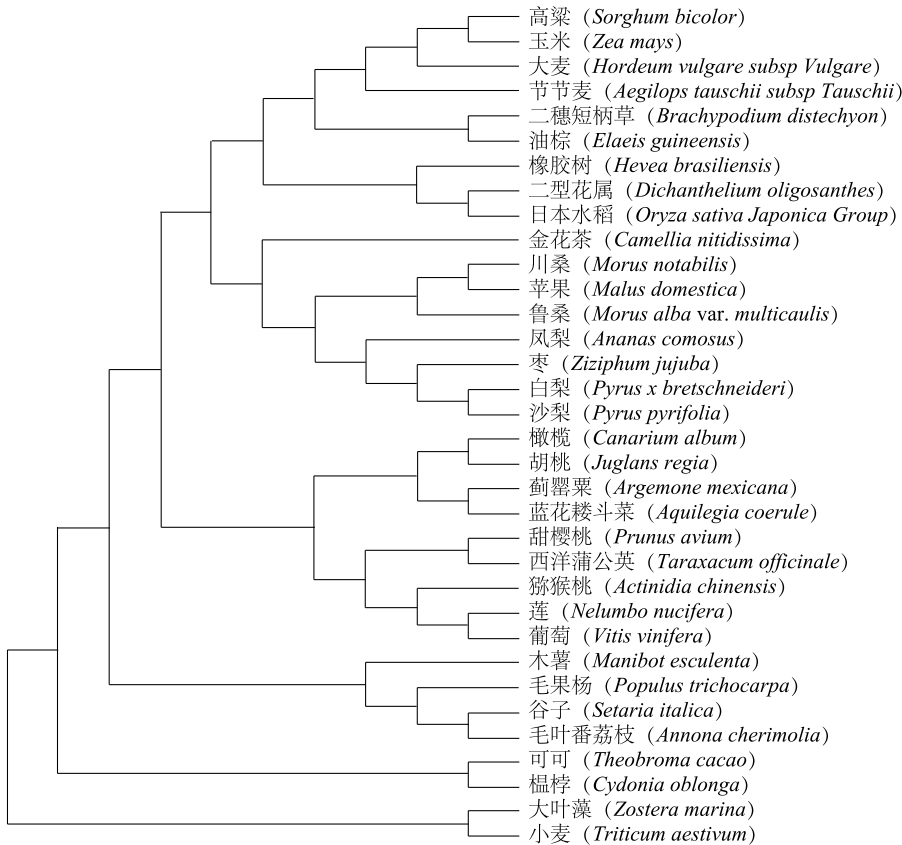


图 2 莲藕基于 PPO 密码子 CDS 序列的系统聚类
Fig.2 System cluster analysis of lotus root based on CDS sequence of PPO codon

表 5 莲藕 PPO 基因与部分模式生物基因组密码子使用偏性比较
Table 5 Comparison of codon usage preference between PPO in lotus root and other model organisms

密码子	氨基酸	密码子使用频率 (%)					密码子使用频率比值			
		PPO	拟南芥 基因组 (At)	烟草 基因组 (Nt)	大肠杆菌 基因组 (Ec)	酵母菌 基因组 (Sc)	PPO/At	PPO/Nt	PPO/Ec	PPO/Sc
GCA	Ala	0.067	0.175	0.231	0.206	0.161	0.38	0.29	0.33	0.42
GCC	Ala	0.334	0.103	0.125	0.255	0.125	3.24	2.67	1.31	2.67
GCG	Ala	0.084	0.090	0.058	0.317	0.061	0.93	1.45	0.26	1.38
GCU	Ala	0.234	0.283	0.312	0.156	0.211	0.83	0.75	1.50	1.11
UGC	Cys	0.117	0.072	0.072	0.069	0.047	1.63	1.63	1.70	2.49
UGU	Cys	0.033	0.105	0.098	0.055	0.080	0.31	0.34	0.60	0.41
GAC	Asp	0.535	0.172	0.169	0.186	0.202	3.11	3.17	2.88	2.65
GAU	Asp	0.184	0.367	0.369	0.321	0.378	0.50	0.50	0.57	0.49
GAA	Glt	0.201	0.343	0.360	0.382	0.485	0.59	0.56	0.53	0.41
GAG	Glt	0.351	0.323	0.294	0.177	0.191	1.09	1.19	1.98	1.84
UUC	Phe	0.301	0.207	0.180	0.169	0.182	1.45	1.67	1.78	1.65
UUU	Phe	0.084	0.218	0.251	0.232	0.261	0.39	0.33	0.36	0.32
GGA	Gly	0.184	0.242	0.232	0.090	0.109	0.76	0.79	2.04	1.69
GGC	Gly	0.151	0.092	0.112	0.279	0.097	1.64	1.35	0.54	1.56
GGG	Gly	0.151	0.102	0.105	0.113	0.060	1.48	1.44	1.34	2.52
GGU	Gly	0.151	0.222	0.223	0.244	0.240	0.68	0.68	0.62	0.63
CAC	His	0.167	0.087	0.087	0.098	0.077	1.92	1.92	1.70	2.17

续表5 Continued5

密码子	氨基酸	密码子使用频率 (%)					密码子使用频率比值			
		<i>PPO</i>	拟南芥 基因组 (<i>At</i>)	烟草 基因组 (<i>Nt</i>)	大肠杆菌 基因组 (<i>Ec</i>)	酵母菌 基因组 (<i>Sc</i>)	<i>PPO/At</i>	<i>PPO/Nt</i>	<i>PPO/Ec</i>	<i>PPO/Sc</i>
CAU	His	0.033	1.380	0.134	0.136	0.137	0.02	0.25	0.24	0.24
AUA	Ile	0.117	0.126	0.140	0.054	0.178	0.93	0.84	2.17	0.66
AUC	Ile	0.151	0.185	0.139	0.242	0.170	0.82	1.09	0.62	0.89
AUU	Ile	0.117	0.215	0.278	0.298	0.304	0.54	0.42	0.39	0.38
AAA	Lys	0.184	0.308	0.326	0.332	0.422	0.60	0.56	0.55	0.44
AAG	Lys	0.585	0.327	0.335	0.107	0.307	1.79	1.75	5.47	1.91
CUA	Leu	0.050	0.099	0.094	0.040	0.131	0.51	0.53	1.25	0.38
CUC	Leu	0.201	0.161	0.123	0.110	0.054	1.25	1.63	1.83	3.72
CUG	Leu	0.251	0.098	0.102	0.509	0.104	2.56	2.46	0.49	2.41
CUU	Leu	0.135	0.241	0.240	0.117	0.121	0.56	0.56	1.15	1.12
UUA	Leu	0.033	0.127	0.134	0.139	0.267	0.26	0.25	0.24	0.12
UUG	Leu	0.217	0.209	0.223	0.140	0.270	1.04	0.97	1.55	0.80
AUG	Met	0.301	0.245	0.250	0.270	0.209	1.23	1.20	1.11	1.44
AAC	Asn	0.268	0.209	0.179	0.214	0.249	1.28	1.50	1.25	1.08
AAU	Asn	0.184	0.223	0.280	0.186	0.363	0.83	0.66	0.99	0.51
CCA	Pro	0.201	0.162	0.198	0.085	0.182	1.24	1.02	2.36	1.10
CCC	Pro	0.284	0.053	0.066	0.058	0.068	5.36	4.30	4.90	4.18
CCG	Pro	0.151	0.086	0.050	0.218	0.053	1.76	3.02	0.69	2.85
CCU	Pro	0.134	0.187	0.187	0.073	0.136	0.72	0.72	1.84	0.99
CAA	Glu	0.117	0.195	0.207	0.150	0.275	0.60	0.57	0.78	0.43
CAG	Glu	0.167	0.152	0.150	0.295	0.121	1.10	1.11	0.57	1.38
AGA	Arg	0.084	0.190	0.160	0.029	0.213	0.44	0.53	2.90	0.39
AGG	Arg	0.234	0.110	0.122	0.019	0.092	2.13	1.92	12.32	2.54
CGA	Arg	0.050	0.063	0.053	0.039	0.030	0.79	0.94	1.28	1.67
CGC	Arg	0.067	0.038	0.039	0.210	0.026	1.76	1.72	0.32	2.58
CGG	Arg	0.067	0.049	0.037	0.063	0.017	1.37	1.81	1.06	3.94
CGU	Arg	0.100	0.090	0.075	0.203	0.065	1.11	1.33	0.49	1.54
AGC	Ser	0.084	0.113	0.100	0.160	0.097	0.74	0.84	0.53	0.87
AGU	Ser	0.084	0.140	0.133	0.095	0.142	0.60	0.63	0.88	0.59
UCA	Ser	0.050	0.183	0.176	0.078	0.188	0.27	0.28	0.64	0.27
UCC	Ser	0.100	0.112	0.102	0.089	0.142	0.89	0.98	1.12	0.70
UCG	Ser	0.217	0.093	0.053	0.087	0.085	2.33	4.09	2.49	2.55
UCU	Ser	0.084	0.252	0.200	0.087	0.235	0.33	0.42	0.97	0.36
ACA	Thr	0.033	0.157	0.174	0.082	0.178	0.21	0.19	0.40	0.19
ACC	Thr	0.201	0.103	0.097	0.228	0.126	1.95	2.07	0.88	1.60
ACG	Thr	0.151	0.077	0.045	0.148	0.079	1.96	3.36	1.02	1.91
ACU	Thr	0.050	0.175	0.203	0.091	0.203	0.29	0.25	0.55	0.25
GUA	Val	0.084	0.099	0.114	0.111	0.118	0.85	0.74	0.76	0.71
GUC	Val	0.134	0.128	0.111	0.151	0.116	1.05	1.21	0.89	1.16
GUG	Val	0.268	0.174	0.167	0.255	0.106	1.54	1.60	1.05	2.53
GUU	Val	0.134	0.272	0.268	0.185	0.220	0.49	0.50	0.72	0.61
UGG	Trp	0.134	0.125	0.122	0.152	0.103	1.07	1.10	0.88	1.30
UAC	tyr	0.284	0.137	0.135	0.121	0.146	2.07	2.10	2.35	1.95
UAU	tyr	0.084	0.146	0.178	0.165	0.189	0.58	0.47	0.51	0.44

各氨基酸见表 2 注。

3 讨论

用生物信息学方法对莲藕 *PPO* 密码子偏好模式进行了系统分析研究, ENc 为 52.12 以及 CAI 为 0.204 均表明莲藕 *PPO* 基因密码子偏好性较弱, 基因表达水平偏低。 $GC3s$ 为 66.3%, 表明 *PPO* 基因密码子较偏好以 G/C 结尾的密码子。这一结果不符合大部分双子叶植物密码子偏好使用 A/U 结尾的特征^[33], 这可能是在进化过程中受突变的影响, 密码子偏好由 A、U 变为 G、C, 密码子偏好性产生变化的具体机制和原因还有待深入研究。

中性绘图分析、 ENc -plot 分析和 PR2-plot 分析结果表明, 多数物种 *PPO* 密码子偏好模式主要受突变压力影响, 其中少数物种 *PPO* 受突变作用压力和其他因素的综合作用。34 个物种 *PPO* 基因的偏好性分析结果显示, 不同物种 *PPO* 基因的 ENc 值和 CAI 值均不相同, *PPO* 基因在各物种内密码子使用偏好性较弱, 单子叶植物 *PPO* 基因表达高于双子叶植物, 但基因表达量普遍不高。经过长期的自然进化, 在突变压力、自然选择和基因功能等各方面综合影响下^[34], 密码子的偏好模式会出现差异, 整体来说, 亲缘关系相近的物种的基因密码子偏好模式比较相似。分析不同物种基因的密码子偏好模式和聚类对物种间系统分类和亲缘关系研究有重要的参考意义。

表达翻译水平与密码子偏好性相关, 本研究结果对莲藕 *PPO* 基因进行简并引物设计、点突变引入等试验提供了参考价值。目前, 莲藕的遗传转换没有成熟体系, *PPO* 相关基因的调控网络并不明确, 莲藕 *PPO* 基因的研究通常需要异源表达, 借助如大肠杆菌、酵母菌等遗传转换体系成熟的受体。本研究通过比较莲藕与拟南芥、拟南芥、大肠杆菌、酵母菌的 *PPO* 基因密码子使用模式的差异, 发现大肠杆菌真核异源表达系统更适用于莲藕 *PPO* 异源表达试验, 在莲藕 *PPO* 基因遗传转化功能验证中, 拟南芥更适用于异源受体。

通过对莲藕多酚氧化酶密码子的分析发现, 该基因密码子偏好性较弱, 但较偏好以 G/C 结尾的密码子。这种使用偏好模式主要受突变压力作用, 其次受多种因素共同影响。通过比较密码子使用频率发现, 大肠杆菌和拟南芥可作为莲藕目标基因异源转化的受体。本试验结果初步揭示了莲藕 *PPO* 密

码子使用规律, 对以后进一步开展转基因功能验证模式作物的筛选有重要的指导作用, 同时也可研究莲藕分子进化提供一定的科学依据。

参考文献:

- [1] CASKEY C T, TOMPKINS R, SCOLNICK E, et al. Sequential translation of trinucleotide codons for the initiation and termination of protein synthesis[J]. *Science*, 1968, 162(3849): 135-138.
- [2] CHAN C, PHAM P, DEDON P C, et al. Lifestyle modifications: coordinating the tRNA epitranscriptome with codon bias to adapt translation during stress responses[J]. *Genome Biology*, 2018, 19(1): 228.
- [3] QUAX T F, CLAASSENS N, SÖLL D, et al. Codon bias as a means to fine-tune gene expression[J]. *Molecular Cell*, 2015, 59(2): 149-161.
- [4] GINGOLD H, DAHAN O, PILPEL Y. Dynamic changes in translational efficiency are deduced from codon usage of the transcriptome.[J]. *Nucleic Acids Research*, 2012, 40(20): 53-63.
- [5] PLOTKIN J B, KUDLA G. Synonymous but not the same: the causes and consequences of codon bias[J]. *Nature Reviews Genetics*, 2010, 12(1): 32-42.
- [6] GUSTAFSSON C, GOVINDARAJAN S, MINSHULL J. Codon bias and heterologous protein expression[J]. *Trends in Biotechnology*, 2004, 22(7): 346-353.
- [7] 刁英, 韩延闯, 何建军, 等. 莲藕研究进展[J]. *氨基酸和生物资源*, 2004, 26(1): 8-11.
- [8] MAYER A M, HAREL E. Polyphenol oxidases in plants[J]. *Phytochemistry*, 1979, 18(2): 193-215.
- [9] YORUK R, MARSHALL M R. Physicochemical properties and function of plant polyphenol oxidase: a review.[J]. *Journal of Food Biochemistry*, 2010, 27(5): 361-422.
- [10] LILLY V V, HAARD N F. Polyphenol oxidases and peroxidases in fruits and vegetables[J]. *C R C Critical Reviews in Food Technology*, 2009, 15(1): 49-127.
- [11] 吴彦庆, 赵大球, 陶俊, 等. 芍药花色调控基因的密码子使用模式及其影响因素分析[J]. *中国农业科学*, 2016, 49(12): 2368-2378.
- [12] 雷梦林, 冯瑞云, 郝雅萍, 等. 小麦抗逆相关转录因子 DREB 密码子偏好性特征分析[J]. *麦类作物学报*, 2019, 39(1): 5-13.
- [13] TORRE A R D L, LIN Y C, PEER Y V D, et al. Genome-wide analysis reveals diverged patterns of codon bias, gene expression, and rates of sequence evolution in *Picea* gene families [J]. *Genome Biol Evol*, 2015, 7(4): 1002-1015.
- [14] 赖瑞联, 冯新, 陈瑾, 等. 橄榄转录组密码子使用偏好性及其影响因素[J]. *核农学报*, 2019, 33(1): 31-38.
- [15] PEDEN J F. Analysis of codon usage [J]. *Univ Nott*, 2000, 90(1): 73-74.
- [16] SHARP P M, LI W H. An evolutionary perspective on synonymous codon usage in unicellular organisms [J]. *J Mol Evol*, 1986, 24

- (1): 28-38.
- [17] RICE P, LONGDEN I, BLEASBY A. EMBOSS: the european molecular biology open software suite [J]. Trends Genet, 2000, 16 (6): 276-277.
- [18] SHI X, WANG X, LI Z, et al. Nucleotide substitution pattern in rice paralogues: implication for negative correlation between the synonymous substitution rate and codon usage bias [J]. Gene, 2006, 376 (2): 199-206.
- [19] MURRAY E E, LOTZER J, EBERLE M. Codon usage in plant genes[J]. Nucleic Acids Research, 1989, 17(2): 477.
- [20] SUEOKA N. Directional mutation pressure and neutral molecular evolution.[J]. Proceedings of the National Academy of Sciences, 1988, 85(8):2653-2657.
- [21] WRIGHT F. The effective number of codons used in a gene[J]. Gene, 1990, 87(1):23-29.
- [22] SUEOKA N. Two aspects of DNA base composition: G+C content and translation-coupled deviation from intra-strand rule of A=T and G=C[J]. J Mol Evol, 1999, 49(1):49-62.
- [23] 程丽,李宜奎,李晓丹,等. 植物CPR基因密码子偏好性及聚类分析[J]. 分子植物育种, 2017, 15(5): 1672-1682.
- [24] 刘潮,韩利红,王海波,等. 谷子类甜蛋白基因家族的鉴定与密码子偏性分析[J]. 西北农业学报, 2018, 27(1): 52-61.
- [25] RIECHMANN J L, HEARD J, MARTIN G, et al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes[J]. Science, 2000, 290(5499):2105-2110.
- [26] LINSMAIER E M, FOLKE S. Organic growth factor requirements of tobacco tissue cultures[J]. Physiologia Plantarum, 2010, 18 (1):100-127.
- [27] 黄雪龙,单敏,陈烨,等. 鸭疫里默氏杆菌 *Tet(X)* 基因原核表达及其功能鉴定[J]. 南方农业学报, 2019, 50(4): 844-850.
- [28] HANAHAN D. Studies on transformation of *Escherichia coli* with plasmids[J]. Journal of Molecular Biology, 1983, 166(4):557-580.
- [29] 白雪,张慧莉,黄冲,等. 解淀粉芽孢杆菌草酸脱羧酶基因的克隆、原核表达与活力测定[J]. 江苏农业科学, 2019, 47 (12):66-70.
- [30] 刘政伟,李瑞芳,张瑞玲. 热带念珠菌 β -葡聚糖合成酶 *KRE9* 基因原核表达及其比活力测定[J]. 南方农业学报, 2018, 49 (4):628-634.
- [31] GASCH A P, SPELLMAN P T, KAO C M, et al. Genomic expression programs in the response of yeast cells to environmental changes.[J]. Molecular Biology of the Cell, 1998, 11(12):4241-4257.
- [32] KURLAND C G. Codon bias and gene expression[J]. Febs Letters, 1991, 285(2):165-169.
- [33] FEDOROV A, SAXONOV S, GILBERT W. Regularities of context-dependent codon bias in eukaryotic genes[J]. Nucleic Acids Research, 2002, 30(5):1192-1197.
- [34] HERSHBERG R, PETROV D A. Selection on codon bias[J]. Annual Review of Genetics, 2008, 42(42):287-299.

(责任编辑:陈海霞)