

## 基于转录组数据分析药用真菌猪苓密码子使用偏好性

刘蒙蒙<sup>1,2,3</sup>, 邢咏梅<sup>1\*</sup>, 郭顺星<sup>1\*</sup>

(1. 中国医学科学院、北京协和医学院药用植物研究所, 北京 100193; 2. 河北大学中医学院, 河北 保定 071002;  
3. 江苏理工学院电气信息工程学院, 生物信息与医药工程研究所, 江苏 常州 213001)

**摘要:** 本研究以猪苓转录组数据 (SRP058382) 为基础, 利用 BLASTx 和 ESTScan 对猪苓的 CDS 进行预测, 共筛选出 332 条 unigenes, 再利用 CodonW 分析这些基因密码子使用偏向性。结果显示, 猪苓基因密码子 GC 含量范围在 44.51%~64.6%, 平均含量为 53.57%, GC3 平均含量为 57.98%, 说明猪苓基因偏向使用以 G 或 C 结尾的密码子; ENC 值在 38.46 到 61 之间, 说明其密码子偏向性较弱; 中性绘图分析及 ENC-plot 曲线分析显示, 猪苓基因 GC12 与 GC3 的回归系数为 0.075, 说明 GC12 和 GC3 的相关性微弱, 并且大部分基因 ENC 值偏离了理论值。这些分析表明在猪苓基因的密码子使用偏向性的形成过程中, 突变和选择压力等诸多因素都发挥着重要作用。通过高表达优越密码子分析确定了猪苓基因中 22 个最优密码子, 多以 G 或 C 结尾。研究结果为猪苓的转基因研究中载体的选择及分子进化研究提供理论依据。

**关键词:** 猪苓; 转录组; 密码子偏向性; 最优密码子

中图分类号: R931 文献标识码: A 文章编号: 0513-4870(2020)05-1050-06

## Analysis of codon usage patterns in *Polyporus umbellatus* based on transcriptome data

LIU Meng-meng<sup>1,2,3</sup>, XING Yong-mei<sup>1\*</sup>, GUO Shun-xing<sup>1\*</sup>

(1. Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China; 2. College of Traditional Chinese Medicine, Hebei University, Baoding 071002, China;  
3. Institute of Bioinformatics and Medical Engineering, School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou 213001, China)

**Abstract:** We obtained 332 coding sequences from the *Polyporus umbellatus* transcriptome based on the BLASTx and ESTScan analyses. The codon usage patterns of *P. umbellatus* were calculated and statistically analyzed using CodonW. The results showed that the average GC content of genes was 53.57% and the average GC3 content was 57.98%, suggesting that genes favored codons ending with G or C. The effective number of codons (ENC) value range from 38.46 to 61, which indicates that these genes have low codon usage bias. The neutrality plot and ENC-plot analysis revealed that many factors such as mutation and selective pressure play an important role in shaping codon usage bias in *P. umbellatus* genes. Twenty-two optimal codons were identified as being biased toward codons ending with G or C using the high expression superior codon analysis method. This study will lay a foundation for future research on genetic engineering and molecular evolution in *P. umbellatus*.

**Key words:** *Polyporus umbellatus*; transcriptome; codon usage bias; preferred codon

收稿日期: 2019-11-14; 修回日期: 2020-02-02.

基金项目: 国家自然科学基金资助项目 (81773843, 81973425); 河北省教育厅基金资助项目 (QN2018128); 北京协和医学院学科建设项目 (201920100901).

\*通讯作者 Tel / Fax: 86-10-62829619, E-mail: sxguo1986@163.com;

Tel / Fax: 86-10-57833240, E-mail: meimary@163.com

DOI: 10.16438/j.0513-4870.2019-0900

密码子是遗传信息传递的基本单位,每一种氨基酸至少对应一种密码子(如甲硫氨酸和色氨酸都只有一个密码子,其他的氨基酸都对应2~6种密码子)。编码同一种氨基酸的密码子称为同义密码子。在早期的研究中显示,蛋白质翻译过程中,密码子的使用并不是随机发生的,并且发现特定物种甚至特定基因通常会偏向于使用一种或几种固定的密码子,这种现象被称为密码子使用偏好性<sup>[1,2]</sup>。在密码子偏好性的形成过程中受到多种因素的影响,例如GC含量、基因表达量、tRNA丰度及突变压力等<sup>[3,4]</sup>。不同物种之间存在着密码子使用偏好性的差异,即使在同种生物中,不同基因的密码子使用偏好性也不相同。通过分析物种的密码子使用偏好性,揭示物种密码子使用偏好性的特征,对于生物分子进化机制及基因工程等研究等都具有重要的意义。例如研究植物的叶绿体基因组的序列特征密码子碱基的组成规律,不仅叶绿体遗传转化效率的提高,也有利于目的基因的高效稳定表达<sup>[5]</sup>。

猪苓(*Polyporus umbellatus*)是我国传统珍贵药用真菌,其药用记载最早出现在《神农本草经》上,被列为中品<sup>[6]</sup>。猪苓性味归经及功效<sup>[7]</sup>与现代药理学证明,猪苓的主要活性成分是猪苓多糖、麦角甾醇等,对膀胱癌、急性肾炎和水肿都具有良好的疗效<sup>[8]</sup>。目前,猪苓的研究主要集中在其有效药用活性成分、药理作用及其栽培技术等方面<sup>[9,10]</sup>。课题组在长期研究猪苓栽培、资源等基础上,利用二代高通量测序技术研究猪苓转录组,揭示了猪苓防御蜜环菌入侵的分子机制<sup>[11]</sup>。为了进一步深入认识猪苓分子系统学及功能基因组学,本研究基于本课题组完成的猪苓转录组数据,对猪苓的密码子使用偏向性进行了系统研究,并分析影响猪苓基因密码子使用偏好性的因素,进一步对筛选出的猪苓基因密码子使用模式及最优密码子进行了分析,为下一步开展猪苓系统进化的分子机制、基于密码子改造的基因功能验证研究提供理论依据。

## 材料与方法

**猪苓转录组** 在NCBI序列片段存档(sequence read archive, SRA)数据库选取之前本课题组发表的猪苓转录组数据(SRP058382)<sup>[11]</sup>。利用Trinity 2.2.0软件拼接组装转录谱<sup>[12]</sup>,应用BLASTx和ESTScan软件分析猪苓转录组中的编码序列(coding sequences, CDS)。由于短序列无法正确计算有效密码子数,为了减少样本误差,本研究选择以ATG为起始密码子的非重复且长度大于300 bp的序列进行分析,最终得到332条用于之后密码子偏好性分析的猪苓CDS序列。

**密码子偏好性分析** 利用CodonW 1.4.2 ([\[codonw.sourceforge.net/\]\(http://codonw.sourceforge.net/\)\)软件计算密码子组成和使用偏好性分析。密码子组成的度量指标包括G+C含量\(鸟嘌呤和胞嘧啶含量\);A3s、T3s、G3s、C3s\(同义密码子在第3位上腺嘌呤、胸腺嘧啶、鸟嘌呤和胞嘧啶的出现频率\);GC3s\(同义密码子第3位的G+C含量\)等。密码子使用偏好性的度量指标包括:同义密码子相对使用度\(relative synonymous codon usage, RSCU\)、有效密码子数\(effective number of codons, ENC\)、密码子适应指数\(codon adaption index, CAI\)等。](http://</a></p></div><div data-bbox=)

ENC可用于检测单个基因CDS区密码子偏好程度,它的取值范围在20到61之间,这个值越接近20表示该基因所有氨基酸越偏向使用其同义密码子的某一密码子,这个值越接近61表示该基因所有氨基酸的各自同义密码子越被平均使用,所以ENC的值越低,表明该基因的密码子使用偏好性越强<sup>[13]</sup>。

CAI可以检测猪苓基因编码区同义密码子与密码子最佳使用相符合程度,范围在0到1之间。CAI值较高,基因具有较高的表达量,CAI值较低,基因具有较低的表达量<sup>[14]</sup>。

RSCU是对59个同义密码子(不包括3个终止密码子TAG、TGG、TGA和仅由一个密码子编码的甲硫氨酸ATG及色氨酸TGG密码子)的使用偏好性评估。RSCU值是用同义密码子的实际观测值和同义密码子平均使用期望值之间的比值获得的<sup>[15]</sup>。

**中性绘图分析** 在生物体编码氨基酸的过程中,除了编码精氨酸、亮氨酸和丝氨酸的密码子外,在密码子第三位的碱基经常会出现同义突变,中性绘图是以GC12(密码子第一、二位GC含量的平均值)为纵坐标,以GC3(密码子第三位GC含量)为横坐标,每个散点代表一个基因,分析密码子3个位点碱基组成之间的相关性,确定影响密码子偏好性的因素。

**ENC-plot绘图** 采用ENC绘图分析ENC值和GC3分布的关系,主要用于检测碱基组成对密码子偏好性的影响。本文利用GC3为横坐标,ENC为纵坐标,创建二维散点图,并在图中附加ENC值的标准曲线。如果基因的点分布在标准曲线附近,这表明密码子的使用偏好性主要受到了突变作用的影响,如果基因点主要分布在期望曲线下方,表明密码子使用偏好性主要受选择压力的影响。

**PR2 (parity rule 2) 奇偶偏好分析** 奇偶规则表明,在DNA的两条互补链之间如果不存在任何突变或者选择压力上的偏倚,那么碱基含量上应该存在A=T和G=C关系。这种方法主要针对4个密码子编码的氨基酸,包括Ser、Leu、Pro、Arg、Thr、Val、Ala和Gly。分别以基因中上述氨基酸对应密码子A3/(A3+T3)为纵

坐标, 以  $G3/(G3+C3)$  为横坐标作图, 以  $(0.5, 0.5)$  坐标为中心, 从这个中心出发的矢量表示了 PR2 偏倚的程度和方向<sup>[16]</sup>。

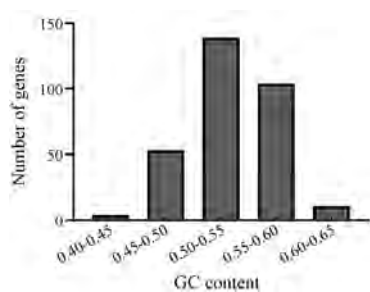
**最优密码子分析** 对于最优密码子的确定, 以 ENC 值进行排序然后作为偏性标准, 选取 ENC 最大值和 ENC 最小值各 10% 的基因, 分别构建高低偏性库, 之后以  $\Delta RSCU$  值为标准确定最优密码子, 取  $\Delta RSCU > 0.08$  的密码子作为高表达优越密码子。

**对应性分析** 对应性分析是指采用多元统计方法研究不同基因密码子变异趋势。利用 CodonW 软件的 RSCU 对应分析功能, 将所有基因分布到 59 维向量空间, 从而得到密码子使用偏好性的特征。通过分析基因位置探究基因向量及基因间值的变异程度, 推测导致密码子偏好发生的原因。

## 结果

### 1 密码子组成分析

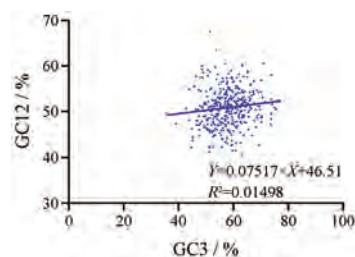
利用 CodonW 软件分析猪苓转录组中所有候选基因的编码区序列 GC 含量。将猪苓转录组候选基因的 CDS 序列作为一个整体进行分析, 结果如图 1 所示, GC 含量范围在 44.51%~64.6%, 主要分布在 50%~60% 之间, 平均含量为 53.57%, GC1、GC2、GC3 平均含量分别为 56.28%、45.48% 和 57.98%。总体来看, GC2 的值最小, GC1 和 GC3 的值相近, 说明猪苓的 CDS 的密码子更倾向于使用鸟嘌呤 (G) 和胞嘧啶 (C) 这两种碱基。



**Figure 1** Distribution of the GC content of coding sequences (CDS) in *Polyporus umbellatus*

### 2 中性绘图分析

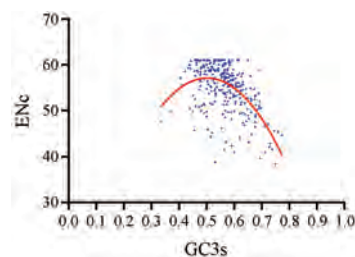
猪苓转录组 CDS 中性绘图分析结果 (图 2) 显示, GC12 的取值范围为 0.407~0.705, GC3 的范围为 0.356~0.771。GC12 与 GC3 的回归系数为 0.075, 说明 GC12 和 GC3 的相关性微弱。密码子第一、二位碱基和第三位碱基受突变的影响不同, 表明猪苓的 CDS 在进化过程中受到一定的突变压力, 影响了 CDS 本身的碱基组成。



**Figure 2** Neutrality plot (GC12 vs GC3)

### 3 ENC-plot 绘图分析

本研究显示猪苓 CDS 的 ENC 取值范围为 38.46~61, 因此认为猪苓 CDS 的密码子偏向性较弱。以 GC3 为横坐标, ENC 为纵坐标作图 (图 3), 图 3 中各点为基因分布情况。代表基因的大部分点都分布在远离期望曲线的地方, 只有小部分的基因点分布在期望曲线周围, 说明除了突变压力在猪苓密码子偏向性的形成中起到重要作用外, 其他因素比如遗传选择作用也在猪苓密码子偏向性的形成中起到重要作用。

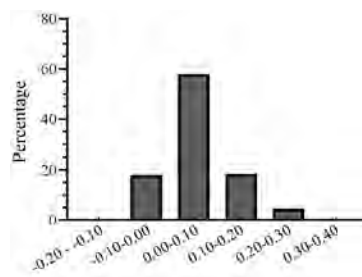


**Figure 3** Neutrality plot (ENC vs GC3). ENC: Effective number of codons

为了准确的区分期望 ENC 与观测 ENC 之间的区别, 本研究进行了  $(ENC_{exp} - ENC_{obs}) / ENC_{exp}$  计算<sup>[17]</sup>, 如图 4 所示,  $(ENC_{exp} - ENC_{obs}) / ENC_{exp}$  的比值主要分布在 0.00~0.10 之间, 说明期望 ENC 与观测 ENC 之间存在很小差异。因此, 这个结果说明突变不是影响猪苓密码子偏向性形成的唯一因素, 但确实是一个很重要的因素。

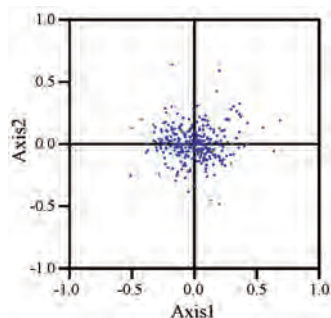
### 4 对应性分析

对应性分析结果表明, 第一向量轴的差异占 11.31%,



**Figure 4** Frequency distribution of the ENC ratio

第二向量轴占9.64%的差异,说明第一轴对猪苓转录组密码子偏向性造成的影响最大。为了进一步说明GC含量对猪苓密码子偏向性的影响,GC含量不同的基因被着上不同的颜色,GC含量高于60%的基因被标记为红色,GC含量在45%~60%之间的基因被标记为蓝色,GC含量低于45%的基因被标记为绿色。如图5所示,GC含量高于60%的基因在坐标系中分布的比较分散,而GC含量小于60%的基因则分布的比较集中。



**Figure 5** Correspondence analysis: genes with GC content higher than 60%, within 45%-60% and lower than 45% were plotted as red, blue and green dots, respectively

另外,本研究做了6个参数的相关性分析,包括GC12、GC3、GCall、ENC、CAI以及Axis 1。如表1所示,Axis 1与3个重要的参数存在显著相关性,包括GCall、ENC及Axis 1 ( $r = 0.679$ ,  $r = 0.466$ ,  $r = 0.437$ ,  $r = 0.218$ ;  $P < 0.01$ ),说明突变压力及遗传选择作用在猪苓密码子偏向性形成过程中起到重要作用。

**Table 1** Correlation analysis of *Polyporus umbellatus* gene-related parameters. \* $P < 0.05$ ; \*\* $P < 0.01$ . CAI: Codon adaption index

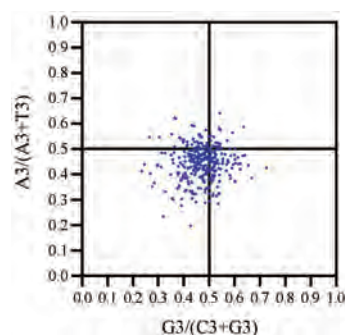
Parameters	GC3	GCall	ENC	CAI	Axis1
GC12	0.141*	0.086	-0.058	0.055**	0.091
GC3		0.041	-0.135*	0.100	0.018
GCall			-0.357**	0.106**	0.679**
ENC				-0.529**	0.466**
CAI					0.437**

## 5 PR2奇偶偏好分析

采用PR2-plot分析了猪苓转录组332条CDS嘌呤(A和G)与嘧啶(T和C)之间的关系,结果如图6所示,大部分基因位于y轴0.5以下,矢量向下及左右两侧偏倚,表明猪苓转录组基因的密码子第三位中T、G和C含量较高,在密码子第三位A、T、C和G使用频率并不相等,表明猪苓密码子偏向性不仅仅由突变引起,还有其他因素影响,比如说遗传选择因素。

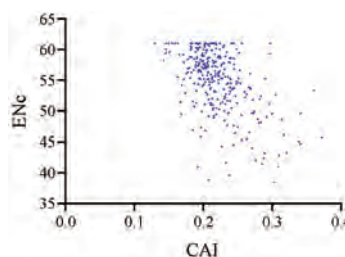
## 6 基因表达水平对密码子偏好性的影响

猪苓CDS的CAI取值范围为0.13~0.372,说明猪苓基因表达水平不高。同时,对CAI和其他几个重要



**Figure 6** Parity rule 2 (PR2)-bias plot [ $A_3/(A_3+T_3)$  vs  $G_3/(C_3+G_3)$ ]

参数(ENC、GC3、GC)进行了相关性分析(图7,表1),结果显示CAI值与ENC之间存在显著的负相关性( $r = -0.529$ ,  $P < 0.01$ ),CAI值与GC含量之间存在显著的正相关性( $r = 0.106$ ,  $P < 0.01$ )。因此,猪苓密码子偏好性的形成过程中受到基因表达水平的影响,GC含量高且表达水平高,基因的密码子偏好性的程度越高。



**Figure 7** Neutrality plot (ENC vs CAI)

## 7 最优密码子分析

以ENC值为偏性标准,建立偏性库,确定了猪苓CDS密码子最优密码子为23个(表2)。在这些最优密码子中,以G(7个)或C(14个)结尾的较多,表明G或C结尾的密码子为猪苓转录组CDS的偏好使用。

## 讨论

研究物种的密码子使用偏好性对于指导基因工程相关的研究,如重要基因在异源生物中的表达起到重要的作用,但是由于缺少多数高等生物的基因组数据,使得我们不能基于基因组数据对生物密码子使用偏好性进行研究。在2015年郭顺星等公布了药用真菌猪苓的转录组数据<sup>[11]</sup>,为猪苓的密码子使用偏好性奠定了基础。本研究为了了解猪苓转录组CDS密码子使用的偏好性特点,首先对预测猪苓CDS序列进行筛选,去除编码区小于300 bp的序列,最后得到332条CDS密码子使用偏好性分析。

密码子使用偏好性与选择压相关,通过影响RNA转录与翻译等多种过程以决定整体编译的合理性<sup>[18]</sup>。

**Table 2** Optimal codons of *Polyporus umbellatus* genes based on the RSCU values. RSCU: Relative synonymous codon usage; \*Optimal codons

Amino acid	Codon	High		Low		Amino acid	Codon	High		Low	
		RSCU	N	RSCU	N			RSCU	N	RSCU	N
Phe	UUU	0.53	43	1.07	71	Ser	GAG*	1.39	115	0.89	85
	UUC*	1.47	119	0.93	62		UCU	1.18	54	1.10	50
Leu	UUA	0.09	5	0.57	32	UCC*	1.71	78	1.21	55	
	UUG	0.72	39	1.33	74	UCA	0.55	25	1.37	62	
	CUU	0.94	51	1.47	82	UCG*	0.62	28	1.12	51	
	CUC*	3.03	165	0.90	50	AGU	0.28	13	0.79	36	
	CUA	0.26	14	0.63	35	AGC	1.16	53	0.90	41	
	CUG	0.97	53	1.10	61	Arg	AGA	0.23	8	1.36	29
Ile	AUU	0.91	54	1.36	76		AGG	0.44	15	1.36	29
	AUC*	1.91	114	0.88	49		CGU*	1.61	55	0.66	14
	AUA	0.18	11	0.77	43		CGC*	2.55	87	0.33	7
Met	AUG	1.00	60	1.00	91	CGA	0.47	16	1.45	31	
Val	GUU	0.59	40	1.05	69	CGG	0.70	24	0.84	18	
	GUC*	2.07	140	1.05	69	Pro	CCU	0.80	46	1.27	58
	GUA	0.43	29	0.74	49		CCC*	1.57	90	0.90	41
	GUG	0.90	61	1.17	77		CCA	0.52	30	1.31	60
Tyr	UAU	0.51	30	1.11	62	CCG*	1.10	63	0.52	24	
	UAC*	1.49	87	0.89	50	Thr	ACU	0.74	40	1.34	57
TER	UAA	0.75	4	1.50	8		ACC*	1.73	94	0.82	35
	UAG	0.75	4	1.13	6		ACA	0.50	27	1.20	51
	UGA	1.50	8	0.38	2		ACG*	1.03	56	0.64	27
His	CAU	0.57	29	1.34	51	Ala	GCU	0.94	85	1.12	69
	CAC*	1.43	72	0.66	25		GCC*	1.44	130	0.94	58
Gln	CAA	0.69	45	1.12	74		GCA	0.59	53	1.44	89
	CAG*	1.31	86	0.88	58	GCG*	1.02	92	0.50	31	
Asn	AAU	0.51	39	1.13	71	Cys	UGU	0.40	12	0.72	25
	AAC*	1.49	114	0.87	55		UGC*	1.60	48	1.28	44
Lys	AAA	0.45	40	0.79	71	Trp	UGG	1.00	43	1.00	61
	AAG*	1.55	139	1.21	109	Gly	GGU	1.22	74	1.26	77
Asp	GAU	0.62	60	1.19	63		GGC*	1.56	95	0.87	53
	GAC*	1.38	134	0.81	43		GGA	0.66	40	1.31	80
Glu	GAA	0.61	51	1.11	107	GGG	0.56	34	0.56	34	

编码序列 GC 含量、转录终止子、转录因子识别序列、隐藏的终止密码、重复序列、CpG 二核苷酸偏向性和隐藏的剪接信号等基因序列特征,都与密码子使用偏好性相关<sup>[19]</sup>。刘思思等<sup>[20]</sup>最新研究指出,GC 含量和突变压力的平衡调节在基因组的碱基组成中起到重要的作用,这种调节是影响密码子使用偏好性的最普遍的作用之一。ENC-plot 绘图分析结果表明,猪苓部分 CDS 分布于标准曲线附近,基因的实际 ENC 值与理论 ENC 值基本保持一致,表明这些 CDS 的密码子使用偏向性受到 GC 含量(突变)的影响比较大;而大多数 CDS 分布的离标准曲线较远,表明实际 ENC 值与理论 ENC 值存在较大的差异,受遗传选择的影响较大。因此,猪苓转录组 CDS 密码子使用偏好性受到突变的影响,更多的是受到了选择的作用。

本研究重点分析了猪苓转录组中 CDS 密码子使用特点,发现猪苓 CDS 的最优密码子有 23 个。在这些最优密码子中,以 G 或 C 结尾的较多,表明 G 或 C 结尾的密码子为猪苓密码子 CDS 的偏好使用。GC 含量被

认为是衡量基因组结构进化的重要因素,基因组不同简并密码子变化的方向和幅度与密码子自身的 GC 含量相关,GC 含量会影响基因的密码子使用偏向性<sup>[21,22]</sup>。从猪苓 CDS 中的 GC 含量可以看出,其偏好使用 G 或 C 碱基结尾的密码子。猪苓转录组的 CDS 在密码子使用上的这一特征与赤芝基因<sup>[23]</sup>、大豆<sup>[24]</sup>、石榴<sup>[25]</sup>基因密码子使用规律相同。由此推断猪苓 CDS 的密码子偏好性的产生与猪苓 CDS 进化过程中 GC 到 AT 突变压力要比 AT 到 GC 突变压力高有关。

本研究对猪苓转录组数据中的 CDS 进行了密码子偏好性及其偏好程度研究,结果显示猪苓的 CDS 密码子偏好性较弱,突变压力和遗传选择都对其密码子使用偏好性产生影响,这些结果可为猪苓的分子进化研究提供理论基础;另外,通过最优密码子分析共计得到 21 个偏好以 G 或 C 结尾的最优密码子,这个结果可为猪苓基因在进行转基因研究时宿主的选择提供依据。

## References

- [1] Grantham R, Gautier C, Gouy M. Codon frequencies in 119 individual genes confirm consistent choices of degenerate bases according to genome type [J]. *Nucleic Acids Res*, 1980, 8: 1893-1912.
- [2] 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: r43-74.
- [3] Holmquist GP, Filipiski J. Organization of mutations along the genome: a prime determinant of genome evolution [J]. *Trends Ecol Evol*, 1994, 9: 65-69.
- [4] Bernardi G. The isochore organization of the human genome and its evolutionary history—a review [J]. *Gene*, 1993, 135: 57-66.
- [5] Chao YE, Chang Y, Wang MF, et al. Codon usage bias and cluster analysis on chloroplastic genes from seven crop species [J]. *Acta Agric Boreali-Sin (华北农学报)*. 2012, 27: 60-64.
- [6] Xing XK, Ma XT, Hart MM, et al. Genetic diversity and evolution of Chinese traditional medicinal fungus *Polyporus umbellatus* (Polyporales, Basidiomycota) [J]. *PLoS One*, 2013, 8: e58807.
- [7] Xing YM, Chen J, Lv YL, et al. Determination of optimal carbon source and pH value for sclerotial formation of *Polyporus umbellatus* under artificial conditions [J]. *Mycol Prog*, 10: 121-125.
- [8] Xing YM, Zhang LC, Liang HQ, et al. Sclerotial formation of *Polyporus umbellatus* by low temperature treatment under artificial conditions [J]. *PLoS One*, 2013, 8: e56190.
- [9] Liu YY, Guo SX. Nutritional factors determining sclerotial formation of *Polyporus umbellatus* [J]. *Lett Appl Microbiol*, 2009, 49: 283-288.
- [10] Sun Y, Yasukawa K. New anti-inflammatory ergostane-type ecdysteroids from the sclerotium of *Polyporus umbellatus* [J]. *Bioorg Med Chem Lett*, 18: 3417-3420.
- [11] Liu MM, Xing YM, Zhang DW, et al. Transcriptome analysis of genes involved in defence response in *Polyporus umbellatus* with *Armillaria mellea* infection [J]. *Sci Rep*, 2015, 5: 16075.
- [12] Haas BJ, Papanicolaou A, Yassour M, et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis [J]. *Nat Protoc*, 2013, 8: 1494-1512.
- [13] Fuglsang A. Impact of bias discrepancy and amino acid usage on estimates of the effective number of codons used in a gene, and a test for selection on codon usage [J]. *Gene*, 2008, 410: 82-88.
- [14] Gupta SK, Bhattacharyya TK, Ghosh TC. Synonymous codon usage in *Lactococcus lactis*: mutational bias versus translational selection [J]. *J Biomol Struct Dyn*, 2004, 21: 527-536.
- [15] Zalucki YM, Power PM, Jennings MP. Selection for efficient translation initiation biases codon usage at second amino acid position in secretory proteins [J]. *Nucleic Acids Res*, 2007, 35: 5748-5754.
- [16] Sinclair G, Choy FY. Synonymous codon usage bias and the expression of human glucocerebrosidase in the methylotrophic yeast, *Pichia pastoris* [J]. *Protein Exp Purif*, 2002, 26: 96-105.
- [17] Wright F. The 'effective number of codons' used in a gene [J]. *Gene*, 1990, 87: 23-29.
- [18] Kumar N, Bera BC, Greenbaum BD, et al. Revelation of influencing factors in overall codon usage bias of equine influenza viruses [J]. *PLoS One*, 2016, 11: e0154376.
- [19] Quax TEF, Claassens NJ, Soell D, et al. Codon bias as a means to fine-tune gene expression [J]. *Mol Cell*, 2015, 59: 149-161.
- [20] Liu S, Qiao Z, Wang X, et al. Analysis of codon usage patterns in "Lonicerae Flos" (*Lonicera macranthoides* Hand. -Mazz.) based on transcriptome data [J]. *Gene*, 2019, 705: 127-132.
- [21] Shackelton LA, Parrish CR, Holmes EC. Evolutionary basis of codon usage and nucleotide composition bias in vertebrate DNA viruses [J]. *J Mol Evol*, 2006, 62: 551-563.
- [22] Van Hemert F, Berkhout B. Nucleotide composition of the Zika virus RNA genome and its codon usage [J]. *Virology*, 2016, 13: 95.
- [23] Zhu XX, Zhu YJ, Song JY, et al. Comparative studies on codon usage bias of *Ganoderma lucidum* based on analysis of genomic and transcriptomic data [J]. *Acta Pharm Sin (药学报)*, 2014, 49: 1340-1345.
- [24] Zhang L, Jin LG, Luo L, et al. Analysis of nuclear gene codon bias on soybean genome and transcriptome [J]. *Acta Agron Sin (作物学报)*, 2011, 37: 965-974.
- [25] Zhang TK, Qi GH, Ye HL, et al. Codon usage bias in pomegranate transcriptome [J]. *Acta Hort Sin (园艺学报)*, 2017, 44: 675-690.