

## 藻类 $\Delta 6$ 脂肪酸脱饱和酶基因密码子偏好性分析

赵薇, 陈必链, 王明兹, 黄建忠

(福建师范大学工业微生物教育部工程研究中心, 生命科学学院, 福建 福州 350108)

**摘要:**  $\Delta 6$  脂肪酸脱饱和酶是在长链脂肪酸 18C 位置上引入双键的关键酶。本文分析硅藻门、绿藻门和蓝藻门等藻类的  $\Delta 6$  脂肪酸脱饱和酶基因密码子偏好及碱基位置偏好, 同时比较藻类  $\Delta 6$  脂肪酸脱饱和酶基因与大肠杆菌、酵母及拟南芥基因组的密码子偏爱性。结果发现, 硅藻门 *T. pseudonana* 与 *P. tricornutum* 及绿藻门 *Ostreococcus tauri*  $\Delta 6$  脂肪酸脱饱和酶基因明显偏爱使用以 G 或 C 结尾的密码子, GC 含量高达 61.61%。蓝藻门三种藻  $\Delta 6$  脂肪酸脱饱和酶基因与之相反, 偏爱 A 或 U 结尾的密码子, GC 含量仅为 39.1%。所选 9 条不同藻类的  $\Delta 6$  脱饱和酶基因密码子偏好不明显, 说明藻类物种间差异对  $\Delta 6$  脱饱和酶基因密码子偏好性的影响较小, 在进化上较为保守。然而, 藻类  $\Delta 6$  脱饱和酶基因密码子用法与大肠杆菌、酵母和拟南芥的  $\Delta 6$  脱饱和酶基因密码子用法差异较大, 若要实现该基因在以上宿主中的高效表达则需对部分密码子进行改造。

**关键词:** 藻类;  $\Delta 6$  脱饱和酶基因; 密码子偏好性; 聚类分析

中图分类号: P735.4 文献标识码: A 文章编号: 1009-5470(2010)03-0126-09

## Analysis on codon usage bias of fatty acid delta-6 desaturase genes from different algae

ZHAO Wei, CHEN Bi-lian, WANG Ming-zi, HUANG Jian-zhong

(Engineering Research Center of Industrial Microbiology, Ministry of Education, College of Life Sciences, Fujian Normal University, Fuzhou 350108, China)

**Abstract:** Delta-6 desaturases are the key enzymes required for the synthesis of unsaturated fatty acids, which introduce double bond at the position of 18C in algae. In this study, the authors investigate codon usage bias and base composition variations in nine available complete desaturase gene sequences from Bacillariophyta, Chlorophyta and Cyanophyta. The results showed that  $\Delta 6$  desaturase genes of *T. pseudonana* and *P. tricornutum* (Bacillariophyta) as well as *Ostreococcus tauri* (Chlorophyta) prefer to use the codons ended with G or C, with GC content as high as 61.61%; while  $\Delta 6$  desaturase genes of three algae of Cyanophyta prefer to use the codons ended with A or U, with GC content 39.1%. Codon usage bias in alga  $\Delta 6$  desaturase gene was very small, which was mainly determined by base compositions on the third codon position and the value of effective number of codons (ENC). It was implied that  $\Delta 6$  desaturase genes were phylogenetically conservative. The results of comparing algae  $\Delta 6$  desaturase genes to genomic codon usage of *E.coli*, yeast and *Arabidopsis thaliana* preference revealed that codon usage pattern among them was significantly different. In order to significantly increase the efficiency of algae  $\Delta 6$  desaturase gene expression, partial codon in gene's open reading frame needs to be reconstituted. The findings presented in this study offer guidance for expressing algae  $\Delta 6$  desaturase gene in prokaryotic and eukaryotic expression systems.

**Key words:** alga; delta-6 desaturase gene; codon bias; cluster analysis

由于密码子的简并性, 大多数编码同一氨基酸的密码子不止 1 种。研究密码子使用特别是同义密

码子的使用不仅可以揭示基因组内部或基因组间的演化关系, 提供数据用于验证多种基因遗传算法的

收稿日期: 2009-07-31; 修订日期: 2009-11-08。蔡卓平编辑

基金项目: 福建省教育厅项目(JA09061); 福建省科技厅重点项目(2007I0039, 2008I0047); 教育部留学回国人员科研启动基金资助项目

作者简介: 赵薇(1984—), 女, 福建省福州市人, 硕士, 主要从事微生物生物技术研究, 现工作单位: 福建省微生物研究所

通信作者: 陈必链。E-mail: chenbil@fjnu.edu.cn

准确性并揭示其各自的特点,还可以检测未经鉴定基因组 DNA 的蛋白编码区。人们通过对不同物种的密码子使用偏好性进行研究,发现单细胞和多细胞有机体中没有统一的密码子使用模式<sup>[1-2]</sup>。在单细胞有机体(大肠杆菌、酿酒酵母)中发现其高表达基因有很强的密码子选择偏好,并且这些密码子对应着高浓度的 tRNA;然而低表达基因则表现出比较单一的密码子使用方式,常使用某些罕见密码子。在多细胞有机体中,因为同一基因在不同组织及细胞生长不同阶段的表达水平与 tRNA 丰度存在着变化,所以无法对一个整体的多细胞有机体基因的表达水平和 tRNA 丰度进行很好的定量分析。

从原核生物到真核生物,其基因组中同义密码子使用偏好性的现象广泛存在,造成密码子使用偏好的理论有:自然选择效应<sup>[3]</sup>、突变效应<sup>[3]</sup>、基因漂移<sup>[4]</sup>。与密码子使用偏性有关的因素有基因的长度、G+C 含量<sup>[5]</sup>、碱基组成限制、重组率、基因的表达水平<sup>[6]</sup>、翻译起始效率<sup>[7]</sup>、tRNA 的相对丰度、蛋白质的空间结构<sup>[8]</sup>以及密码子反密码子间结合能的大小等<sup>[9]</sup>。近年来,虽然有许多研究机构开展藻类脂肪酸脱饱和酶基因的分子克隆及功能鉴定方面的研究工作,但藻类脱饱和酶基因的研究不论从广度还是深度上都有限。NCBI(National Center for Biotechnology Information)数据库中黄藻门、甲藻门、隐甲藻门的脂肪酸脱饱和酶基因数据处于空白状态,红藻门仅含 1 条  $\Delta 9$  脱饱和酶全长基因与 1 条并不完整的  $\Delta 5$  脱饱和酶基因,而对于脂肪酸脱饱和酶基因密码子的分析研究更是未见报道。本文以 NCBI 数据库中已知的硅藻、绿藻、蓝藻的  $\Delta 6$  脂肪酸脱饱和酶基因为基础,应用分层聚类分析法,探讨 3 个不同门类藻种基因在密码子使用偏好性方面的相似性和差异性,并进行物种间和物种内基因的分类和功能研究;分析藻类  $\Delta 6$  脂肪酸脱饱和酶基因密码子偏好性,并与常用的原核、真核表达宿主(大肠杆菌、酵母、拟南芥)全基因组密码子使用偏好性进行比较,这不仅可以为该基因的表达选择合适的表达系统,还可以为改造密码子提高该基因的表达量提供依据。

## 1 实验材料与方法

### 1.1 数据来源

本文所用的硅藻门、绿藻门和蓝藻门藻类基因及资料来自于 NCBI (<http://www.ncbi.nlm.nih.gov>),所用的大肠杆菌和酵母基因组的密码子偏好性数据

来源于 Codon Usage Database(<http://www.kazusa.or.jp/codon/>)。

### 1.2 软件和程序

使用密码子分析软件 CodonW (<http://www.molbio1.ox.ac.uk>) 和序列比对软件 Clustalx (<http://bioweb.pasteur.fr/seqanal/phylogeny/phylyp-uk.html>);使用欧洲分子生物学开放软件系统(European Molecular Biology Open Software Suite, EMBOSS)的 CHIPS (codon heterozygosity in a protein coding sequence)和 CUSP (create a condon usage table)程序。

### 1.3 方法

#### 1.3.1 序列获得

从 NCBI 上取得藻类 19 条  $\Delta 6$  脱饱和酶基因 CDS(coding sequence)序列(登录号:CS486251、CS160911、AY817155、CS020143、EF494667、CS486239、CS160835、AX951573、CS020067、CQ891303、AY082393、CS486249、CS351573、CS161015、CS160903、CS020135、EU606018、EF222476、L11421)经过 ClusterX 比对除去核酸同源性 100%的序列,最终选择 9 条  $\Delta 6$  脱饱和酶基因 CDS 序列进行密码子分析(表 1)。

表 1 藻类  $\Delta 6$  脱饱和酶基因序列

Tab. 1 Gene sequences of algae  $\Delta 6$  desaturase

编号	登录号	藻种名
1	AY817155	<i>Thalassiosira pseudonana</i>
2	CS486251	<i>Thalassiosira pseudonana</i>
3	EF494667	<i>Phaeodactylum tricornutum</i>
4	AY082393	<i>Phaeodactylum tricornutum</i>
5	CS161015	<i>Ostreococcus tauri</i>
6	CS486249	<i>Ostreococcus tauri</i>
7	L11421	<i>Synechocystis</i> sp.
8	EF222476	<i>Arthrospira platensis</i> CG590
9	AG20590	<i>Arthrospira platensis</i> KCTC AG20590

#### 1.3.2 密码子分析方法

有效密码子数(effective number of codons, ENC)用于定量分析 CDS 序列的密码子使用偏好,能反映密码子家族中同义密码子非均衡使用的偏好程度,其取值范围在 20(每种氨基酸只偏好使用 1 种密码子)至 61(各种同义密码子的使用机会完全均等)之间,即基因的密码子偏爱程度越大,ENC 值越小。

GC3<sub>s</sub>(the frequency of G+C at the synonymous third position of sense codons)是氨基酸密码子第 3 位 GC 出现的频率,GC12<sub>s</sub>是密码子 1、2 位置上 GC 出

现的频率, GC12 的改变更受到功能性的约束, GC12 与 GC 用于研究密码子使用模式与碱基位置限制的关系<sup>[10]</sup>。

密码子适应指数 (codon adaptation index, CAI) 指实际编码该蛋白的所有密码子对于完全使用最优密码子编码该蛋白情况下的适应性指数。当基因中密码子偏好性达到最大程度, 即所有氨基酸都是由最优密码子编码, CAI 值为 1; CAI 值越小, 说明密码子偏好程度越低, 反之亦然; 该指标的大小和基因表达水平呈正相关。

### 1.3.3 基因密码子使用概率聚类分析

相对同义密码子使用频率(relative synonymous codon usage, RSCU)指对于某一特定的密码子在编码对应氨基酸同义密码子间的相对概率, 用于检测同义密码子使用模式, 能够排除 CDS 序列氨基酸位置偏好的影响; 如果密码子的使用没有偏好, 则该密码子的 RSCU 值等于 1。当某一密码子的 RSCU 值大于 1, 则表明该密码子的使用频率相对较高, 反之亦然<sup>[11]</sup>。根据物种名称分类的硅藻门、绿藻门、蓝藻门计算其 $\Delta 6$  脂肪酸脱饱和酶基因高偏倚和低偏倚的 RSCU 值, 通过相互比较, 预测不同门类藻之间的一些最优密码子。

为了研究藻类 $\Delta 6$  脂肪酸脱饱和酶基因同义密码子的进化模式, 以每 1 条基因的 RSCU 值为依据, 采用分层聚类法对基因密码子使用偏好性进行聚类分析, 除去以 ATG 为起始密码子的蛋氨酸和编码色氨酸 Trp 的密码子 TGG, 再除去 3 个不编码氨基酸的终止密码子, 取余下的 59 个密码子的 RSCU 值。

### 1.3.4 藻类脂肪酸 $\Delta 6$ 脱饱和酶基因与大肠杆菌、酵母、拟南芥的密码子偏好特征比较

藻类脂肪酸 $\Delta 6$  脱饱和酶基因与大肠杆菌、酵母、拟南芥的密码子偏好特征比较的具体方法参照

参考文献[12]。

## 2 结果与讨论

### 2.1 藻类脂肪酸 $\Delta 6$ 脱饱和酶基因同义密码子使用模式

为了研究不同藻类间 $\Delta 6$  脂肪酸脱饱和酶基因的密码子使用方式, 9 条所选基因片段的长度、ENC 值、GC 含量信息和 59 个密码子的 RSCU 值分别在表 2 和表 3 中列出。

由 GC 含量可以看出, 蓝藻门 3 种藻都偏好使用 A 或 U 结尾的密码子; 硅藻门与绿藻门有偏好使用 GC 末尾密码子的藻株, 也有偏好使用 AU 结尾的密码子的藻株; 即使是同属的藻, 藻名称相同, 其第 3 位密码子的使用偏好也会不同。相对来说, 藻类 $\Delta 6$  脂肪酸脱饱和酶基因一般在密码子 1、2 位置上偏好使用 A 或 U; 不同藻株 ENC 值变化从 38.2 至 54.23, 平均值为 47.99, 标准偏差为 5.32。88.9% 的 ENC 值大于 40, 反映了大部分已知藻类 $\Delta 6$  脂肪酸脱饱和酶基因均一性较好, 偏好性不强。同样地, GC3<sub>s</sub> 值也能用来进一步确认同义密码子使用方式的均一性, 藻类 $\Delta 6$  脂肪酸脱饱和酶基因 GC3<sub>s</sub> 变化范围值为 0.351—0.796, 平均值为 0.52, 标准偏差为 0.17。

一般认为基因的表达水平越高, 所受翻译效率的选择压力越大, 密码子的偏好性也就越高, 反之则偏好性较低, 因而现在常把基因的密码子偏好程度作为该基因表达水平的标志。具有低密码子 CAI 值<sup>[13]</sup>的基因倾向于低水平表达, 无论是依据 mRNA 表达水平或蛋白质水平的分析都是如此。当基因的密码子 CAI 值为 0.25 或更低时(如大多数基因), mRNA 水平与蛋白质水平的相关性很差。对大多数高表达的基因(密码子 CAI 值>0.5 的基因), mRNA 水

表 2  $\Delta 6$  脱饱和酶基因同义密码子的使用

Tab. 2 Usage pattern of  $\Delta 6$  desaturase gene synonymous codon

登录号	ENC	GC3 <sub>s</sub>	GC	GC12 <sub>s</sub>	CAI	长度/bp
AY817155	46.18	0.617	0.504	0.431	0.215	1 452
CS486251	46.36	0.615	0.503	0.431	0.215	1 452
EF494667	38.20	0.751	0.581	0.486	0.273	1 431
AY082393	51.35	0.399	0.554	0.620	0.146	1 653
CS161015	53.31	0.426	0.499	0.497	0.194	1 302
CS486249	42.07	0.796	0.599	0.492	0.269	1 368
L11421	54.23	0.401	0.466	0.475	0.169	1 782
EF222476	51.11	0.351	0.434	0.443	0.165	1 065
EU606018	49.12	0.355	0.419	0.431	0.256	1 104

表 3 不同藻株 $\Delta 6$ 脱饱和酶相对同义密码子使用频率Tab. 3 Relative synonymous codon usage in different stereotypes of  $\Delta 6$  desaturase genes

氨基酸	密码子	AY817155	CS486251	EF494667	AY082393	CS161015	CS486249	L11421	EF222476	AG20590
Ala	GCG	0.44	0.44	0.65	0.27	0.47	<u>2.46</u>	0.52	0.89	0.69
Ala	GCA	0.89	0.89	0.41	<u>2.00</u>	<u>1.88</u>	0.23	1.39	0.89	0.83
Ala	GCU	1.00	0.89	0.73	<u>1.33</u>	0.94	0.31	<u>1.91</u>	<u>2.22</u>	<u>1.24</u>
Ala	GCC	<u>1.71</u>	<u>1.78</u>	<u>2.40</u>	0.34	0.57	1.02	0.2	0.00	<u>1.38</u>
Cys	UGU	0.57	0.57	0.86	1.00	0.94	0.00	0.93	1.37	1.00
Cys	UGC	<u>1.43</u>	<u>1.43</u>	<u>1.14</u>	<u>1.00</u>	<u>1.06</u>	<u>2.00</u>	<u>1.07</u>	0.63	1.00
Asp	GAU	1.23	1.23	0.67	0.91	1.05	0.86	<u>1.08</u>	<u>1.27</u>	<u>1.43</u>
Asp	GAC	0.77	0.77	<u>1.33</u>	<u>1.09</u>	0.95	<u>1.14</u>	0.92	0.73	0.57
Glu	GAG	<u>1.75</u>	<u>1.75</u>	0.47	0.00	0.67	<u>1.55</u>	1.00	0.25	0.38
Glu	GAA	0.25	0.25	<u>1.53</u>	2.00	<u>1.33</u>	0.45	1.00	<u>1.75</u>	<u>1.63</u>
Phe	UUU	0.42	0.42	0.88	1.29	1.20	0.62	<u>1.44</u>	<u>1.45</u>	<u>1.62</u>
Phe	UUC	<u>1.58</u>	<u>1.58</u>	<u>1.12</u>	0.71	0.80	1.38	0.56	0.55	0.38
Gly	GGG	0.21	0.21	1.08	1.10	0.73	0.65	0.78	0.57	0.91
Gly	GGA	<u>2.63</u>	<u>2.63</u>	<u>1.85</u>	<u>1.57</u>	<u>1.39</u>	0.90	<u>1.04</u>	<u>1.33</u>	<u>1.27</u>
Gly	GGU	0.74	0.74	1.08	0.39	0.73	1.03	0.87	1.33	1.64
Gly	GGC	0.42	0.42	<u>1.79</u>	0.94	<u>1.14</u>	<u>1.42</u>	<u>1.30</u>	0.76	0.18
His	CAU	0.36	0.45	0.21	<u>1.32</u>	0.83	0.36	<u>1.16</u>	<u>1.57</u>	<u>1.52</u>
His	CAC	<u>1.64</u>	<u>1.55</u>	<u>2.38</u>	0.68	<u>1.17</u>	<u>1.64</u>	0.84	0.43	0.48
Ile	AUA	0.00	0.00	0.00	0.19	0.63	0.00	0.70	0.56	0.31
Ile	AUU	<u>1.06</u>	<u>1.06</u>	<u>0.62</u>	<u>2.06</u>	<u>1.89</u>	0.50	<u>1.20</u>	<u>1.50</u>	<u>1.66</u>
Ile	AUC	<u>1.94</u>	<u>1.94</u>	0.52	0.75	0.47	<u>2.50</u>	<u>1.10</u>	0.94	<u>1.03</u>
Lys	AAG	<u>1.71</u>	<u>1.66</u>	<u>1.48</u>	0.00	0.67	<u>1.77</u>	0.75	0.00	0.43
Lys	AAA	0.29	0.34	0.64	<u>2.00</u>	<u>1.33</u>	0.23	<u>1.25</u>	<u>2.00</u>	<u>1.57</u>
Leu	UUG	<u>2.61</u>	<u>2.61</u>	<u>2.14</u>	0.38	0.84	<u>1.59</u>	<u>1.36</u>	<u>1.26</u>	0.62
Leu	UUA	0.13	0.13	0.21	0.48	0.56	0.18	0.48	1.42	2.90
Leu	CUG	0.00	0.00	2.36	1.05	0.84	<u>1.24</u>	0.88	0.95	0.83
Leu	CUA	0.00	0.00	0.00	1.05	0.98	0.00	<u>1.52</u>	<u>1.11</u>	<u>1.24</u>
Leu	CUU	<u>1.57</u>	<u>1.57</u>	0.64	2.10	2.09	0.35	<u>1.20</u>	<u>1.11</u>	<u>0.21</u>
Leu	CUC	<u>1.7</u>	<u>1.70</u>	<u>1.00</u>	0.95	0.70	2.65	0.56	0.16	0.21
Asn	AAU	0.86	0.86	0.24	0.80	1.25	0.27	<u>1.58</u>	0.86	0.82
Asn	AAC	<u>1.14</u>	<u>1.14</u>	<u>2.89</u>	1.20	0.75	1.73	0.42	1.14	1.18
Pro	CCG	0.53	0.53	<u>1.11</u>	0.57	0.27	<u>1.88</u>	0.61	0.17	<u>1.33</u>
Pro	CCA	0.80	0.80	0.00	1.71	2.13	0.00	2.24	1.91	0.22
Pro	CCU	0.80	0.80	0.00	1.06	1.07	0.24	0.61	0.35	<u>1.11</u>
Pro	CCC	<u>1.87</u>	<u>1.87</u>	0.83	0.65	0.53	<u>1.88</u>	0.54	<u>1.57</u>	<u>1.33</u>
Gln	CAG	<u>1.08</u>	<u>1.08</u>	<u>1.17</u>	0.39	0.67	<u>1.09</u>	0.35	0.00	0.25
Gln	CAA	0.92	0.92	<u>4.00</u>	<u>1.61</u>	<u>1.33</u>	0.91	<u>1.65</u>	<u>2.00</u>	<u>1.75</u>
Arg	AGG	0.50	0.50	0.00	0.29	0.40	0.52	1.20	0.71	0.00
Arg	AGA	<u>1.50</u>	<u>1.50</u>	0.00	0.22	0.40	0.52	1.05	1.24	0.00
Arg	CGG	0.50	0.50	0.50	<u>1.17</u>	0.93	1.04	1.20	0.88	0.67
Arg	CGA	1.00	1.00	0.50	<u>1.39</u>	<u>1.33</u>	<u>1.04</u>	<u>1.05</u>	0.53	0.67
Arg	CGU	2.00	2.00	1.00	0.95	0.80	0.78	0.45	1.24	1.33
Arg	CGC	0.50	0.50	<u>1.93</u>	<u>1.98</u>	<u>2.13</u>	<u>2.09</u>	<u>1.05</u>	<u>1.41</u>	<u>3.33</u>
Ser	AGU	<u>1.60</u>	<u>1.60</u>	0.21	1.00	0.77	0.25	0.78	1.02	0.63
Ser	AGC	0.00	0.00	0.86	0.83	1.16	1.50	0.70	1.40	0.95
Ser	UCG	<u>1.60</u>	<u>1.60</u>	<u>2.36</u>	0.67	0.00	1.25	0.61	0.77	0.00
Ser	UCA	0.40	0.40	0.21	<u>1.17</u>	<u>2.13</u>	0.25	0.78	1.53	0.95
Ser	UCU	1.00	1.00	0.43	1.50	0.19	0.75	1.39	0.38	<u>2.21</u>
Ser	UCC	<u>1.40</u>	<u>1.40</u>	<u>2.29</u>	0.83	<u>1.74</u>	<u>2.00</u>	<u>1.74</u>	0.89	<u>1.26</u>
Thr	ACG	0.21	0.21	0.95	0.77	0.00	2.71	0.76	0.00	0.20
Thr	ACA	0.63	0.63	0.38	<u>1.69</u>	<u>2.18</u>	0.14	0.76	1.45	1.00
Thr	ACU	1.05	1.05	0.38	<u>1.08</u>	0.36	0.00	0.95	1.82	0.80
Thr	ACC	<u>2.11</u>	<u>2.11</u>	<u>2.44</u>	0.46	<u>1.45</u>	<u>1.14</u>	<u>1.52</u>	0.73	2.00
Val	GUG	<u>2.34</u>	<u>2.34</u>	0.89	1.05	0.92	<u>2.63</u>	0.89	0.00	0.91
Val	GUA	0.14	0.14	0.67	0.63	0.92	0.11	0.67	0.80	1.14
Val	GUU	0.28	0.28	0.00	<u>1.05</u>	<u>1.54</u>	0.34	<u>2.44</u>	<u>2.13</u>	<u>1.26</u>
Val	GUC	<u>1.24</u>	<u>1.24</u>	1.00	<u>1.26</u>	0.62	0.91	0.00	1.07	0.69
Tyr	UAU	0.50	0.50	0.00	0.67	0.93	0.32	<u>1.58</u>	<u>1.87</u>	<u>1.71</u>
Tyr	UAC	<u>1.50</u>	<u>1.50</u>	<u>3.00</u>	<u>1.33</u>	<u>1.07</u>	<u>1.68</u>	0.42	0.13	0.29

数值带下划线表示 RSCU 值大于 1 的密码子为相应藻株的偏好密码

平与蛋白质水平的相关性要高很多。由表 2 中数据得知, CAI 值变化范围 0.146—0.273, 平均值 0.211, 标准偏差 0.047, 藻类 $\Delta 6$  基因密码子 CAI 明显偏低, 所以其基因倾向于低水平表达。由于藻类 $\Delta 6$  脂肪酸脱饱和酶基因表达水平低, 所以酶量很低, 至今未能从藻中得到纯化后的 $\Delta 6$  脂肪酸脱饱和酶, 也无法研究其蛋白质结构、理化性质及稳定性, 尤其是无法直接测定酶活性, 只能检测通过其作用后所产生相应的不饱和脂肪酸的含量来间接反映酶活或基因的表达情况。所以, 根据 $\Delta 6$  脱饱和酶基因同义密码子偏性来选择与其密码子偏性较为接近的、比较合适的表达宿主进行密码子改造对于提高 $\Delta 6$  基因的表达量具有重要作用。

如表 3 所示, 硅藻门、绿藻门和蓝藻门 $\Delta 6$  脂肪酸脱饱和酶基因密码子 RSCU 值大于 1 且偏好一致的氨基酸所对应的密码子有: 编码 Gly 的 GGA, 编码 Ile 的 AUU 与 AUC, 编码 Leu 的 UUG, 编码 Arg 的 CGC, 编码 Ser 的 UCC, 编码 Thr 的 ACC。相比较而言, 硅藻门与绿藻门 $\Delta 6$  脂肪酸脱饱和酶基因密码子偏好性使用较为一致, 偏好于 C 或 G 结尾的最优密码子; 而蓝藻门与其相比差异较大。蓝藻门属于原核生物, 绿藻门、硅藻门属于真核生物, 这说明从原核藻进化至真核藻, 其密码子的使用偏性有明显不同, 最优密码子的出现可能与翻译水平相关, 也与弱的选择效应有关。

## 2.2 密码子使用模式的影响因素

造成密码子使用差异的因素有很多, 目前对密码子偏爱性的解说获得较为广泛认可的是“突变-选择平衡”假说<sup>[14]</sup>。该假说认为, 由于选择压力的存在, 生物体倾向于选用最优密码子来编码氨基酸, 但由于突变的发生, 仍会有非最优密码子的存在, 不同物种基因组的密码子偏好情况主要就是在这两大因素的动态平衡中形成<sup>[15]</sup>。根据表 2 和表 3 中的数据, 分析影响藻类密码子使用模式的两个主要因素。

### 2.2.1 ENC 与 GC3s 相关性分析

ENC 绘图分析是一种利用图像展示基因数据密码子使用偏好性状况的可视化手段。ENC 值和 GC3s 分布的关系图被用于评估基因序列 ENC 实际值与期望值之间的差异。为了分辨基因的偏好性是否受到除了组成突变压力外其他因素的影响, 图 1 中标准曲线表示不存在选择压力时, 密码子偏好性完全取决于突变压, 即序列本身核酸组成偏好性, 有:

$$N_c = 2 + s + \frac{29}{s^2 + (1-s)^2}$$
, 式中,  $N_c$  表示标准曲线中的

ENC 值,  $s$  表示 GC3s<sup>[16]</sup>。

图 1 显示所有的样本点实际观察值都低于标准曲线, 表明其值明显低于期望值, 也就是说, 除了 GC 含量这一由于中性突变而造成的序列背景差异外, 还存在其他因素, 如与选择效应相关联的基因表达水平, 以及组成突变效应一起影响着藻类 $\Delta 6$  基因密码子的偏性。由于目前已知藻类 $\Delta 6$  脂肪酸脱饱和酶基因序列都落在显著低于标准曲线的区域, 预示着此基因可能直接位于表达选择压力之下, 所以接下来对 GC12s 与 GC3s 含量进行中性绘图分析。

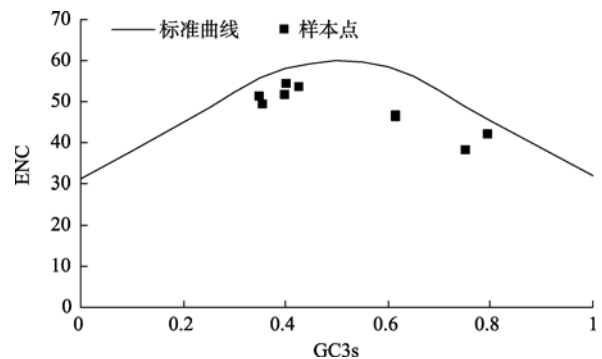


图 1 ENC 与 GC3s 之间的关系图

Fig. 1 Relationship between ENC and GC3s

### 2.2.2 GC12s 与 GC3s 中性绘图分析<sup>[10]</sup>

为了探讨密码子出现偏爱现象的主要原因, 采用 SPSS 对密码子第 1 位与第 2 位 GC 含量的平均值 GC12s 与 GC3s 值进行相关性分析, 得出相关系数  $R=0.192$ ,  $P=0.620$ ,  $P>0.5$  说明两变量间相关性不显著。对 GC12s 与 GC3s 含量进行直观比较, 结果如图 2。从理论上来说, 如果 GC12s 的变异和 GC3s 一样中性而无选择压力作用, 则代表基因序列的点会沿着对角线分布; 反之, GC12s 的变异与 GC3s 不同, 完全是中性的情况下, 那么图中的点将会排列

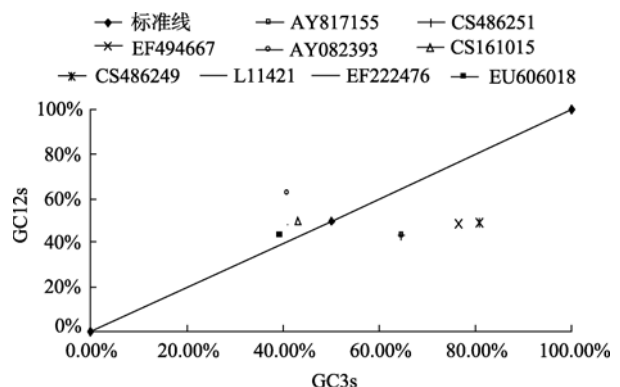


图 2 GC3s 与 GC12s 的相关性

Fig. 2 Correlation between GC3s and GC12s

成与横轴几乎平行的线。从图 2 中看出, 代表相应基因序列的点绝大多数处于 GC12<sub>s</sub> 含量 50% 水平线以下, 这表明密码子碱基组成模式的形成主要是由于选择效应, 因为自然选择对密码子不同位置上碱基的影响是均衡的, 从而造成 GC 含量的变异程度低。

2.3 密码子偏好性的物种特异性

表 1 所列 9 条基因序列所有编码区密码子的相对使用概率(RSCU)见表 4, 根据 59 个密码子的相对使用概率进行聚类分析<sup>[9]</sup>(图 3)。

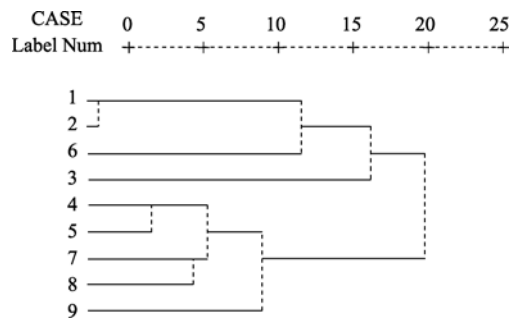


图 3 藻类  $\Delta 6$  脱饱和酶基因密码子偏好性聚类  
Fig. 3 Cluster tree based on the relative synonymous codon usage (RSCU) values of nine algae  $\Delta 6$  desaturase genes

表 4 硅藻门、绿藻门、蓝藻门  $\Delta 6$  脱饱和酶基因的聚类  
Tab. 4 Cluster results of Bacillariophyta, Chlorophyta, Cyanophyta desaturase genes

分类号	编号
I	1. AY817155; 2. CS486251; 6. CS486249; 3. EF494667
II	4. AY082393; 5. CS161015; 7. L11421; 8. EF222476; 9. AG20590

对目前已知藻类  $\Delta 6$  脱饱和酶基因的同义密码子使用偏性聚类分析发现, 主要分为两大类, 其中绿藻门的 1 株青绿藻 *Ostreococcus tauri*(CS486249) 与硅藻门的假矮海链藻 *Thalassiosira pseudonana* 及 *Phaeodactylum tricornutum* 其中 1 株(EF494667)的  $\Delta 6$  脱饱和酶同义密码子使用偏性相近; 而第 2 大类中包括了硅藻门的 *Phaeodactylum tricornutum*, 绿藻门的 *Ostreococcus tauri* 以及蓝藻门的 *Arthrospira platensis*, *Synechocystis* sp.。因此, 作者认为当基因功能相同时, 藻类物种间的差异对  $\Delta 6$  脱饱和酶基因密码子偏好性的影响较小。藻类是地球上分布最广的水生植物, 在进化地位上处于最底层, 为了生存, 需要适应各式各样的生态环境, 由于环境选择压力的存在, 生物体倾向于选用最优密码子(optimal codon)来编码氨基酸, 所以, 同一物种

对密码子使用偏好的不同可能是藻种在长期进化过程中生态适应和生态进化的结果。

2.4 藻类  $\Delta 6$  脱饱和酶基因与大肠杆菌、酵母、拟南芥的密码子偏好特征比较

来源于 Codon Usage Database 的大肠杆菌、酵母、拟南芥基因组 G+C 含量及 GC 在密码子中的位置情况如表 5 所示。

表 5 藻类及大肠杆菌、酿酒酵母、毕赤酵母、拟南芥基因组 GC 含量  
Tab. 5 GC contents of algae and *E.coli*, *Saccharomyces cerevisiae*, *Pichia pastoris* and *Arabidopsis thaliana*

种类	GC/%	GC <sub>1s</sub> /%	GC <sub>2s</sub> /%	GC <sub>3s</sub> /%
硅藻门	53.41	57.39	41.92	60.93
绿藻门	53.51	53.33	45.59	61.61
蓝藻门	43.30	44.61	46.20	39.10
藻类*	50.20	52.42	44.08	54.09
大肠杆菌	51.50	58.44	40.88	55.17
酿酒酵母	39.77	44.58	36.64	38.10
毕赤酵母	42.73	48.72	37.32	42.16
拟南芥	44.59	50.84	40.54	42.38

\*: 硅藻门、绿藻门、蓝藻门  $\Delta 6$  脱饱和酶基因序列的 GC 含量及密码子 1、2、3 位 GC 含量的综合统计结果

不同密码子的使用模式和基因的 GC 含量(尤其是第 3 位密码子的 GC 含量)有密切的关系。从表 5 来看, 酿酒酵母、毕赤酵母以及拟南芥的 GC 含量都比较低, 第 3 位密码子明显偏爱 A 或 U 结尾的密码子, 而且藻类  $\Delta 6$  脱饱和酶与大肠杆菌、酵母、拟南芥在密码子第 2 位 GC 含量都偏低。

一个物种的基因在长期进化过程中会逐渐适应宿主的基因组环境而采用符合宿主基因组的密码子用法, 同时, 对于同一物种, 不同蛋白质编码氨基酸在基因中出现的频率不同。虽然影响外源基因在宿主中表达的因素较多, 但是密码子的选用是影响表达的重要因素之一。如果外源基因含有大量宿主表达系统的罕见密码子, 特别是当罕见密码子成连续分布的时候, 就会造成表达量极低或翻译提前终止<sup>[11]</sup>。由于宿主密码子与藻类  $\Delta 6$  脱饱和酶基因密码子使用偏性的差异将影响到表达水平, 所以,  $\Delta 6$  脱饱和酶基因在异源宿主内表达时必须要考虑密码子偏爱性。将藻类  $\Delta 6$  脱饱和酶基因、大肠杆菌、酿酒酵母、毕赤酵母和拟南芥密码子使用频率进行比较, 结果如表 6 所示: A、E、S、P、A.th 分别表示其一种密码子在编码基因总密码子中出现的频率, 即在 1 000 个密码子中出现的次数, A/E、A/S、A/P、A/A.th 分别表示其密码子出现频率的比值。藻类与

表 6 藻类 $\Delta 6$ 脱饱和酶基因密码子、大肠杆菌、酿酒酵母、毕赤酵母、拟南芥基因组密码子的“/1 000”值以及藻类(A)与大肠杆菌(E)、酿酒酵母(S)、毕赤酵母(P)、拟南芥(A.th)之间的比值。

Tab. 6 “/1000” values of codons for every amino acid and comparision of codon preference among algae(A)  $\Delta 6$  desaturase gene, *E.coli*(E), *Saccharomyces cerevisia*(S), *Pichia pastoris*(P) and *Arabidopsis thaliana* (A.th)genome

密码子	A	E	S	P	A . th	A/E	A/S	A/P	A/A.th
GCG	14.60	32.32	6.18	3.86	14.60	<u>0.45</u>	<u>2.36</u>	<u>3.78</u>	1.62
GCA	20.16	20.69	16.21	15.10	20.16	0.97	1.24	1.34	1.15
GCU	14.37	15.42	21.17	28.92	14.37	0.93	0.68	0.50	0.51
GCC	15.99	25.18	12.60	16.58	15.99	0.64	1.27	0.96	1.55
UGU	15.06	5.25	8.10	7.70	15.06	<u>2.87</u>	1.86	1.96	1.43
UGC	12.98	6.38	4.76	4.38	12.98	<u>2.03</u>	<u>2.73</u>	<u>2.96</u>	1.81
GAU	19.00	32.84	37.59	35.66	19.00	0.58	0.51	0.53	0.52
GAC	22.02	19.21	20.21	25.87	22.02	1.15	1.09	0.85	1.28
GAG	15.06	18.71	19.24	29.03	15.06	0.81	0.78	0.52	<u>0.47</u>
GAA	14.37	39.30	45.60	37.43	14.37	<u>0.37</u>	<u>0.32</u>	<u>0.38</u>	<u>0.42</u>
UUU	30.13	22.24	26.12	24.14	30.13	1.35	1.15	1.25	1.38
UUC	29.90	15.86	18.44	20.60	29.90	1.88	1.62	1.45	1.45
GGG	30.59	11.76	6.02	5.76	30.59	<u>2.60</u>	<u>5.08</u>	<u>5.31</u>	<u>3.01</u>
GGA	19.00	8.87	10.90	19.06	19.00	<u>2.14</u>	1.74	1.00	0.79
GGU	12.98	24.19	23.89	25.52	12.98	0.54	0.54	0.51	0.59
GGC	15.99	28.06	9.78	8.06	15.99	0.57	1.64	1.98	1.75
CAU	25.96	12.80	13.62	11.81	25.96	<u>2.03</u>	1.91	<u>2.20</u>	1.88
CAC	19.24	9.38	7.77	9.07	19.24	<u>2.05</u>	<u>2.48</u>	<u>2.12</u>	<u>2.21</u>
AUA	4.17	5.48	17.79	11.14	4.17	0.76	<u>0.23</u>	<u>0.37</u>	<u>0.33</u>
AUU	16.92	29.69	30.13	31.14	16.92	0.57	0.56	0.54	0.79
AUC	18.31	23.89	17.17	19.43	18.31	0.77	1.07	0.94	0.99
AAG	15.76	11.00	30.82	33.80	15.76	1.43	0.51	<u>0.47</u>	<u>0.48</u>
AAA	28.04	33.96	41.87	29.93	28.04	0.83	0.67	0.94	0.91
UUG	11.59	12.97	27.17	31.51	11.59	0.89	<u>0.43</u>	<u>0.37</u>	0.56
UUA	19.93	13.85	26.15	15.56	19.93	1.44	0.76	1.28	1.57
CUG	12.98	51.07	10.48	14.94	12.98	<u>0.25</u>	1.24	0.87	1.32
CUA	20.63	3.88	13.41	10.74	20.63	<u>5.32</u>	1.54	1.92	<u>2.09</u>
CUU	9.73	11.40	12.25	15.85	9.73	0.85	0.79	0.61	<u>0.40</u>
CUC	22.48	10.47	5.44	7.63	22.48	<u>2.15</u>	<u>4.13</u>	<u>2.95</u>	1.40
AUG	22.71	27.21	20.94	18.66	22.71	0.83	1.08	1.22	0.93
AAU	17.85	19.15	35.68	25.07	17.85	0.93	0.50	0.71	0.80
AAC	12.98	21.66	24.82	26.67	12.98	0.60	0.52	<u>0.49</u>	0.62
CCG	18.54	22.45	5.29	3.94	18.54	0.83	<u>3.50</u>	<u>4.71</u>	<u>2.15</u>
CCA	15.06	8.38	18.31	18.94	15.06	1.80	0.82	0.80	0.93
CCU	9.50	7.23	13.51	15.77	9.50	1.31	0.70	0.60	0.51
CCC	9.27	5.56	6.78	6.80	9.27	1.67	1.37	1.36	1.75
CAG	23.64	29.37	12.11	16.27	23.64	0.80	1.95	1.45	1.55
CAA	14.37	14.71	27.28	25.45	14.37	0.98	0.53	0.56	0.74
AGG	6.49	1.81	9.23	6.63	6.49	<u>3.59</u>	0.70	0.98	0.59
AGA	5.33	2.89	21.28	20.10	5.33	1.84	<u>0.25</u>	<u>0.27</u>	<u>0.28</u>
CGG	11.36	6.18	1.74	1.94	11.36	1.84	<u>6.53</u>	<u>5.85</u>	<u>2.33</u>
CGA	19.00	3.83	2.99	4.18	19.00	<u>4.96</u>	<u>6.36</u>	<u>4.55</u>	<u>3.02</u>
CGU	10.20	20.22	6.40	6.94	10.20	0.50	1.59	1.47	1.13
CGC	10.20	20.82	2.60	2.15	10.20	<u>0.49</u>	<u>3.92</u>	<u>4.74</u>	<u>2.70</u>
AGU	9.97	9.38	14.15	12.55	9.97	1.06	0.70	0.79	0.71
AGC	10.89	16.01	9.75	7.64	10.89	0.68	1.12	1.43	0.96
UCG	11.12	8.82	8.56	7.36	11.12	1.26	1.30	1.51	1.19
UCA	18.31	8.14	18.67	15.18	18.31	<u>2.25</u>	0.98	1.21	1.00
UCU	11.36	8.70	23.50	24.39	11.36	1.31	<u>0.48</u>	<u>0.47</u>	<u>0.45</u>
UCC	11.82	8.95	14.22	16.53	11.82	1.32	0.83	0.72	1.06
ACG	9.04	14.97	7.96	6.04	9.04	0.60	1.14	1.50	1.17
ACA	15.30	8.09	17.76	13.75	15.30	1.89	0.86	1.11	0.98
ACU	8.34	9.08	20.28	22.39	8.34	0.92	<u>0.41</u>	<u>0.37</u>	<u>0.48</u>
ACC	7.88	22.79	12.73	14.45	7.88	<u>0.35</u>	0.62	0.55	0.76
GUG	7.88	26.22	10.76	12.28	7.88	<u>0.30</u>	0.73	0.64	0.45
GUA	16.92	10.90	11.77	9.89	16.92	1.55	1.44	1.71	1.71
GUU	20.63	18.12	22.07	26.91	20.63	1.14	0.93	0.77	0.76
GUC	13.21	14.81	11.78	14.88	13.21	0.89	1.12	0.89	1.03
UGG	29.90	15.30	10.37	10.26	29.90	1.95	<u>2.88</u>	<u>2.91</u>	<u>2.40</u>
UAU	19.93	16.54	18.78	15.99	19.93	1.20	1.06	1.25	1.36
UAC	17.15	12.27	14.78	18.12	17.15	1.40	1.16	0.95	1.25
UGA	10.89	1.09	0.68	0.33	10.89	<u>9.99</u>	<u>16.02</u>	<u>33.01</u>	<u>9.31</u>
UAG	4.87	0.27	0.51	0.49	4.87	<u>18.03</u>	<u>9.54</u>	<u>9.93</u>	<u>9.18</u>
UAA	10.20	1.96	1.06	0.85	10.20	<u>5.20</u>	<u>9.62</u>	<u>12.00</u>	<u>10.73</u>

下划线表示的频率比值在小于 0.5, 大于 2.0 之间

表 7 藻类 $\Delta 6$  脂肪酸脱饱和酶基因在不同表达宿主中表达时需要改造的密码子  
Tab. 7 Codons that require modification in different expression hosts for alga  $\Delta 6$  desaturase genes

以大肠杆菌为表达宿主		以酿酒酵母为表达宿主		以毕赤酵母为表达宿主		以拟南芥为表达宿主	
氨基酸	密码子	氨基酸	密码子	氨基酸	密码子	氨基酸	密码子
Ala	GCG	Ala	GCG	Ala	GCG	Ala	GAG
Cys	UGU	Cys	UGC	Cys	UGC	Glu	GAA
Cys	UGC	Glu	GAA	Glu	GAA	Gly	GGG
Glu	GAA	Gly	GGG	Gly	GGG	His	CAC
Gly	GGG	His	CAC	His	CAU	Ile	AUA
Gly	GGA	Ile	AUA	His	CAC	Lys	AAG
His	CAU	Leu	UUG	Ile	AUA	Leu	CUA
His	CAC	Leu	CUC	Lys	AAG	Leu	CUU
Leu	CUG	Pro	CCG	Leu	UUG	Pro	CCG
Leu	CUA	Arg	AGA	Leu	CUC	Arg	AGA
Leu	CUC	Arg	CGG	Asn	AAC	Arg	CGG
Arg	AGG	Arg	CGA	Pro	CCG	Arg	CGA
Arg	CGA	Arg	CGC	Arg	AGA	Arg	CGC
Arg	CGC	Ser	UCU	Arg	CGG	Ser	UCU
Ser	UCA	Thr	ACU	Arg	CGA	Thr	ACU
Thr	ACC	Trp	UGG	Arg	CGC	Trp	UGG
Val	GUG		UGA*	Ser	UCU		UGA*
	UGA*		UAG*	Thr	ACU		UAG*
	UAG*		UAA*	Trp	UGG		UAA*
	UAA*				UGA*		
					UAG*		
					UAA*		

\*: 终止密码子

大肠杆菌、酵母、拟南芥密码子使用频率的比值在 0.5—2 之间(表 6), 表示密码子的偏好性较接近, 反之则差别较大<sup>[9,11]</sup>。大肠杆菌有 20 个, 酿酒酵母有 19 个, 毕赤酵母有 22 个, 拟南芥有 19 个密码子与藻类 $\Delta 6$  脱饱和酶基因密码子使用偏爱性差异较大(表 7)。其中在酵母与拟南芥中有 15 个相同的密码子, 其使用方式与藻类  $\Delta 6$  脱饱和酶基因使用偏性差异大, 所以应尽量避免编码这些氨基酸的密码子在这些真核表达载体中使用。

3 结论

本文对已知藻类 $\Delta 6$  脂肪酸脱饱和酶基因密码子的使用模式进行研究, 揭示了其产生密码子偏爱性的主要原因是由于选择压力, 同时对藻类 $\Delta 6$  脂肪酸脱饱和酶基因密码子的进化过程进行研究, 证实了其密码子使用模式形成的因素; 另外, 还把分析结果与常用表达宿主密码子的使用偏性进行比较, 为选择表达宿主提供理论依据。

研究结果发现, 硅藻门 *T. pseudonana* 与

*P.tricornutum* 及绿藻门 *Ostreococcus tauri*  $\Delta 6$  脂肪酸脱饱和酶基因明显偏爱使用以 G 或 C 结尾的密码子, GC 含量高达 61.61%。蓝藻门 3 种藻 $\Delta 6$  脂肪酸脱饱和酶基因与之相反, 偏爱 A 或 T 结尾的密码子, GC 含量较低, 仅为 39.1%。3 种门类的藻偏向使用以 A 结尾的 UGA、UAA 终止密码子。硅藻的偏好密码子有 28 个, 绿藻的偏好密码子有 30 个, 蓝藻有 28 个。硅藻与绿藻所偏好的密码子较为一致, 与蓝藻偏好相异较大。从藻 $\Delta 6$  脂肪酸脱饱和酶基因密码子的 “/1 000 值” 与大肠杆菌、酵母、拟南芥基因组密码子的比值可以得出: 从藻类整个群体出发,  $\Delta 6$  脂肪酸脱饱和酶基因密码子偏性与大肠杆菌、酵母、拟南芥基因组密码子使用偏性相差较大, 直接将 $\Delta 6$  脂肪酸脱饱和酶基因导入这些表达宿主中会造成其表达水平与原宿主相比大大偏低, 若要实现该基因在以上宿主中的高效表达则须对部分密码子进行改造。该藻类脂肪酸脱饱和酶基因序列与其他表达宿主比较, 差异性较强的密码子(对照表 6 和表 7 得出), 如甘氨酸 Gly(GGG)、组氨酸 His(CAC)、



亮氨酸 Leu(CUC)、脯氨酸 Pro(CCG)、精氨酸(CGG、CGA、CGC)、苏氨酸 Thr(ACU)的改造方法可以使用定点突变法,根据其密码子使用偏性对 $\Delta 6$  脱饱和酶基因的 CDS 序列进行改造,转换成与表达宿主密码子使用方式差异性较小的决定同一种氨基酸的

同义密码子。对藻类 $\Delta 6$  脱饱和酶基因功能、表达的研究是揭示藻类多样化脂肪酸生物合成机理和实现其体外高表达的基础,不同生物物种的密码子偏爱性有一定差别,因此外源基因在宿主系统中的表达有强弱之分。

## 参考文献

- [1] COMERON J M, AGUADE M. An evaluation of measures of synonymous codon usage bias[J]. *Journal of Molecular Evolution*, 1998, 47(3): 268–274.
- [2] ZHOU T, GU W, MA J, et al. Analysis of synonymous codon usage in H5N1 virus and other influenza A viruses[J]. *Biosystems*, 2005, 81(1): 77–86.
- [3] SHARP P M, STENICO M, PEDEN J F, et al. Codon usage: mutational bias, translational selection, or both? [J]. *Biochemical Society Transactions*, 1993, 21(4): 835–841.
- [4] LIVINGSTONE F B. Genetic drift and polygenic inheritance[J]. *American Journal of Physical Anthropology*, 1972, 37(1): 117–126.
- [5] 石秀凡, 黄京飞, 柳树群, 等. 人类基因同义密码子偏好的特征以及与基因 GC 含量的关系[J]. *生物化学与生物物理进展*, 2002, 29 (3): 411–414.
- [6] DURET L, MOUCHIROUD D. Expression pattern and surprisingly gene length shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*[J]. *Proceedings of the National Academy of Sciences*, 1999, 96(8): 4 482–4 487.
- [7] SAKAI H, WASHIO T, SATIO R, et al. Correlation between sequence conservation of the 5' untranslated region and codon usage bias in *Mus musculus* genes[J]. *Gene*, 2001, 276(1/2): 101–105.
- [8] 顾万君, 马建民, 周童, 等. 不同结构的蛋白编码基因的密码子偏性研究[J]. *生物物理学报*, 2002, 18(1): 81–86.
- [9] 石秀凡, 黄京飞, 梁宠荣, 等. 人类基因中同义密码子的偏好与密码子-反密码子间的结合强度密切相关吗? [J]. *科学通报*, 2000, 45 (23): 2 520–2 525.
- [10] ZHAO S, ZHANG Q, LIU X, et al. Analysis of synonymous codon usage in 11 human bocavirus isolates [J]. *Biosystems*, 2008, 92(3): 207–214.
- [11] 赵耀, 刘汉梅, 顾勇, 等. 玉米 waxy 基因密码子偏好性分析[J]. *玉米科学*, 2008, 16(2): 16–21.
- [12] 吴宪明, 吴松锋, 任大明, 等. 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.
- [13] POWELL J P, SEZZI E, MORIYAMA E N, et al. Analysis of a shift in codon usage in *Drosophila*[J]. *Journal of Molecular Evolution*, 2003, 57: 214–225.
- [14] BULMER M. The selection-mutation-drift theory of synonymous codon usage[J]. *Genetics*, 1991, 129(3): 897–907.
- [15] 王侃, 刘次全, 曹槐, 等. 大肠杆菌 mRNA 编码区长度、形成二级结构倾向与密码子偏好性的关系[J]. *微生物学报*, 2006, 46 (6): 895–899.
- [16] 张文娟. 基于密码子水平的生物信息学分析及进化研究[D]. 上海: 复旦大学生命科学学院, 2006.