

孙晓春, 黄文静, 李会容. 白蜡树 *FcF6'H1* 基因的密码子偏好性分析[J]. 山西农业科学, 2025, 53(1): 35-47.

SUN X C, HUANG W J, LI H R. Codon bias analysis of *FcF6'H1* gene in *Fraxinus chinensis*[J]. Journal of Shanxi Agricultural Sciences, 2025, 53(1): 35-47.

doi:10.3969/j.issn.1002-2481.2025.01.04

白蜡树 *FcF6'H1* 基因的密码子偏好性分析

孙晓春¹, 黄文静¹, 李会容²

(1. 陕西中医药大学 陕西中药资源产业化省部共建协同创新中心/秦药特色资源研究开发国家重点实验室(培育), 陕西 咸阳 712083; 2. 玉溪师范学院 化学生物与环境学院, 云南 玉溪 653100)

摘要: 利用 CodonW、EMBOSS 在线程序分析白蜡树 *FcF6'H1* 基因的密码子偏好性, 对芸香科、豆科、伞形科、木犀科以及模式植物拟南芥、烟草和番茄中 *F6'H1* 的基因密码子进行聚类分析、中性绘图、ENC-plot 分析和 PR2-plot 偏倚分析, 研究影响白蜡树 *FcF6'H1* 基因密码子偏好性形成的因素, 通过 *FcF6'H1* 与模式生物的密码子使用频率比较获得最佳受体。结果表明, 白蜡树 *FcF6'H1* 的 GC、GC1、GC2、GC3 和 GC12 含量分别为 0.431 2、0.512 5、0.349 0、0.432 1 和 0.430 8, CAI 值为 0.216, ENC 值为 57.84, 表明白蜡树 *FcF6'H1* 基因的偏好性较弱; 中性绘图、ENC-plot 和 PR2-plot 偏倚分析结果表明, 碱基突变和自然选择均会影响 *FcF6'H1* 密码子的偏好性; CDS 进化树和 RSCU 值聚类分析结果不完全一致, 但是均表明白蜡树 *FcF6'H1* 与木樨榄 XM_023028611.1、XM_023036788.1、XM_023034893.1 聚为一类, 确定了 *F6'H1* 基因的 6 个最优密码子, 分别是 CUC、AUC、AAG、GAG、UCG 和 ACA; 密码子使用频率分析表明, 模式生物中番茄和烟草适合作为遗传转化受体, 大肠杆菌表达系统适合作为 *FcF6'H1* 的异源表达载体。文章初步阐明了白蜡树 *FcF6'H1* 基因密码子的使用规律。

关键词: 白蜡树; *FcF6'H1*; 密码子偏好性; ENC-plot; PR2-plot; 异源表达

中图分类号: S792.41

文献标识码: A

文章编号: 1002-2481(2025)01-0035-13

Codon Bias Analysis of *FcF6'H1* Gene in *Fraxinus chinensis*

SUN Xiaochun¹, HUANG Wenjing¹, LI Huirong²

(1. Co-construction Collaborative Innovation Center for Chinese Medicine Resources Industrialization by Shaanxi & Education Ministry/State Key Laboratory of Research & Development of Characteristic Qin Medicine Resources(Cultivation), Shaanxi University of Chinese Medicine, Xianyang 712083, China;
2. School of Chemistry, Biology and Environment, Yuxi Normal University, Yuxi 653100, China)

Abstract: In this study, the codon bias analysis of *FcF6'H1* gene in *Fraxinus chinensis* was analyzed using CodonW software and EMBOSS online program, cluster analysis, neutral plot, ENC-plot and PR2-plot of the codon of *FcF6'H1* gene in Rutaceae, Leguminosae, Umbelliferae, Oleaceae, *Arabidopsis*, tobacco and tomato was conducted, the factors affecting information of the codon bias of the *FcF6'H1* gene in *Fraxinus chinensis* were studied, the best receptors were obtained by comparing the codon usage frequencies of *FcF6'H1* and model organisms. The results showed that the content of GC, GC1, GC2, GC3, and GC12 of *FcF6'H1* in *Fraxinus chinensis* was 0.431 2, 0.512 5, 0.349 0, 0.432 1, and 0.430 8, respectively, CAI was 0.216, and ENC was 57.84, indicating weak codon bias of *FcF6'H1* in *Fraxinus chinensis*. The results of neutral analysis, ENC-plot analysis, and PR2-plot analysis indicated that codon bias of *FcF6'H1* was affected by base mutation and natural selection. The results of CDS evolutionary tree and RSCU cluster analysis were not completely consistent, but both indicated that *FcF6'H1* in *Fraxinus chinensis* and XM_023028611.1, XM_023036788.1, and XM_023034893.1 of Olea

收稿日期: 2024-05-06

基金项目: 陕西省教育厅青年创新团队建设科研计划项目(21JP031); 陕西高校青年创新团队项目(陕教[2019]90号); 陕西中医药大学秦药特色资源研究开发重点实验室开放课题(KF202326); 陕西中医药大学校级科研课题(2023GP32); 云南省科技厅基础研究专项(202101AU070005)

作者简介: 孙晓春, 副教授, 主要从事中药资源研究, E-mail: sunxiaochun08@163.com

通信作者: 李会容, 副教授, 主要从事分子生药学研究, E-mail: 10942@yxnu.edu.cn

europaea var. sylvestris were aggregated into one category. Six optimal codons of *F6'H1* gene were identified, they were CUC, AUC, AAG, GAG, UCG, and ACA. The codon usage frequency analysis showed that in the model plants, tomato tobacco were suitable for the genetic transformation receptors, and the Escherichia coli expression system was suitable for the heterologous expression vector of *FcF6'H1*. This study preliminarily clarified the codon usage pattern of *FcF6'H1* gene in *Fraxinus chinensis*.

Keywords: *Fraxinus chinensis*; *FcF6'H1*; codon bias; ENC-plot; PR2-plot; heterologous expression

遗传密码具有通用性和简并性的特点^[1]。某种同义密码子的使用频率高于其他同义密码子的现象称为密码子使用偏好性(Codon usage bias, CUB)^[2-3]。密码子偏好性可影响蛋白质的翻译效率、翻译的准确性、共翻译折叠及转录等^[4-5]。此外,密码子偏好性在不同物种间差异较大^[6-7]。因此,研究密码子的偏好性有助于深入了解物种特定或相关基因的表达模式及机制,为基因功能验证、异源表达、宿主或供体基因选择提供理论基础^[5,8]。

芸香科、豆科、伞形科、木犀科等 74 科植物中均具有香豆素类化合物^[9]。香豆素具有抗病毒、抗炎、诱导癌细胞凋亡等药理作用^[10]。此外,在植物的生长发育、逆境应答、调节根系微生物群落结构等方面均具有重要作用^[11]。秦皮为陕西道地药材之一,具有清热燥湿、收涩止痢、止带、明目的功效^[12]。白蜡树(*Fraxinus chinensis* Roxb.)为秦皮的基原植物之一,秦皮的主要成分为秦皮甲素、秦皮乙素等香豆素类物质^[13]。在香豆素生物合成途径中,中间代谢产物阿魏酰 CoA,在阿魏酰 CoA6'-羟化酶(feruloyl-CoA6'-hydroxylase1, *F6'H1*)的作用下发生邻羟基化生成 6-羟基阿魏酰 CoA,然后经香豆素合酶作用形成东莨菪内酯^[14]。*F6'H*是东莨菪内酯生物合成的控制酶,在东莨菪内酯 8-羟化酶作用下,东莨菪内酯转化成秦皮素^[15]。*F6'H*属于 2-氧戊二酸依赖双加氧酶(2-oxoglutarate-dependent dioxygenases, 2OGDs)超家族^[16]。王蕾等^[17]对大豆(*Glycine max*(L.) Merr.)*GmF6'H1*基因进行了克隆与功能验证,采用农杆菌介导法将大豆 *GmF6'H1* 转入拟南芥(*Arabidopsis thaliana*)*Atf6'h1*突变体,转基因拟南芥中香豆素的含量高于拟南芥 *Atf6'h1* 突变体,与野生型拟南芥中香豆素的含量接近。

目前,对白蜡树中香豆素类物质的生物合成途径尚未有系统研究。前期已对白蜡树叶片和枝皮进行了全长转录组测序^[18],以获得的全长转录本为参考,通过分析得到 1 个白蜡树 *FcF6'H1* 基因。本研究通过将白蜡树 *FcF6'H1* 基因与芸香科、豆科、伞形科、木犀科以及模式植物烟草(*Nicotiana*

tabacum)、拟南芥、番茄(*Solanum lycopersicum*)中香豆素合成相关的 *F6'H1* 基因进行比较分析,明确 *FcF6'H1* 基因的密码子偏好性与进化关系,通过比较白蜡树 *FcF6'H1* 和模式生物密码子频率,确定合适的异源表达宿主,旨在为后续探究白蜡树 *FcF6'H1* 基因的功能特别是异源表达提供理论依据。

1 材料和方法

1.1 序列数据的获得

白蜡树 *FcF6'H1* (Isoform0057303) 的 CDS 序列来源于陕西中医药大学前期对白蜡树全长转录组测序结果^[18],芸香科、豆科、伞形科、木犀科以及模式植物拟南芥、烟草和番茄中 58 条序列来源于 GenBank,详细信息如表 1 所示。拟南芥、烟草、番茄、大肠杆菌(*Escherichia coli*)、酵母菌(*Saccharomyces cerevisiae*)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/>)^[8]。

1.2 密码子偏好性参数分析

密码子组成分析利用软件 codonW 和 EMBOSS (<https://www.bioinformatics.nl/emboss-explorer/>) 的 cusp^[19] 分析,主要包括密码子第 3 位各个碱基(A3s、C3s、G3s、T3s)的频率、密码子适应指数(Codon adaptation index, CAI)、第 1~3 位碱基的 G+C 含量(GC1、GC2 和 GC3)、第 1 和第 2 位碱基 G+C 含量(GC12)、同义密码子相对使用度(Relative synonymous codon usage value, RSCU)、有效密码子数(Effective number of codons, ENC)。参数的相关性利用 IBM SPSS Statistics 22 软件分析,利用联川生物云平台高级热图在线工具(<https://www.omicstudio.cn/tool/4>)进行 RSCU 的热图绘制。

1.3 中性绘图分析、ENC-plot 分析和 PR2-plot 偏倚分析

以 GC12 为纵坐标、GC3 为横坐标,利用 Excel 软件进行中性绘图,分析影响密码子偏好性的因素^[20]。若 GC12 与 GC3 无显著差异,散点沿对角线分布,表明突变是影响其密码子偏好性的主要因

素;若散点偏离对角线,则表明自然选择是主要的影响因素^[20]。

利用 ENC-plot 分析展示 ENC 和 GC₃ 的函数关系,标准曲线 $ENC=2+GC_3+29/(2\times GC_3^2-2\times GC_3+1)$ ^[20-21]。其中,若散点位于曲线上,表示密码子偏好性由碱基突变产生;若散点远离标准曲线,则表示自然选择是其主要影响因素^[22]。

以 A3/(A3+T3) 值为纵坐标、G3/(G3+C3) 值为横坐标绘图,图中中心 A=T 且 C=G 横纵坐标均为 0.5,从该中心出发的矢量表示奇偶性的方向与程度^[20,23]。

1.4 最优密码子分析

将各物种的 *F6'H1* 序列根据 ENC 值排序,从两端各选 10% 的基因建立高低偏性库,计算 $\Delta RSCU$ 值^[20]。最优密码子确定:高表达基因库中 $RSCU > 1$,低表达基因库中 $RSCU < 1$ 且 $\Delta RSCU \geq 0.2$ ^[24]。

1.5 分子进化树构建

使用 MEGAX 软件邻接法(Neighbor-joining)构建不同物种 *F6'H1* 基因的 CDS 系统进化树,参数重复值设置为 1 000^[8]。

1.6 *FcF6'H1* 基因受体的选择

选择 *FcF6'H1* 基因的合适表达宿主,计算其与多个模式物种基因组密码子使用频率之间的比值,确定最合适的异源表达受体^[25]。

2 结果与分析

2.1 *F6'H1* 基因密码子偏好性分析

白蜡树 *FcF6'H1* 的 A3s、C3s、G3s、T3s 频率分别为 0.311 1、0.236 7、0.318 4 和 0.424 0,GC、GC1、GC2、GC3 和 GC12 含量分别为 0.431 2、0.512 5、0.349 0、0.432 1 和 0.430 8,G+C 含量占比较低,A/T 含量 > G/C 含量,表明白蜡树 *FcF6'H1* 基因偏好以 A/T 结尾。CAI 值为 0.216,表明 *FcF6'H1* 基因偏好性较弱。由表 1 可知,20 个物种 59 条 *F6'H1* 序列的 GC 含量为 0.333 3~0.465 5,平均值 0.412 6;GC1 含量为 0.443 2~0.550 7,平均值 0.505 6;GC2 含量为 0.310 9~0.367 8,平均值 0.341 3;GC3 含量为 0.227 1~0.532 6,平均值 0.390 8;GC12 含量为 0.386 1~0.449 3,平均值 0.423 5;ENC 含量为 42.28~60.34,平均值 50.82;CAI 值 0.166~0.254,平均值 0.197,CBI 值 -0.201~0.075,平均值 -0.1,表明密码子的偏好性较弱。在木犀科、豆科、伞形科和芸香科中,伞形科 GC 含量最高(0.443 9),木犀科和豆科平均 GC 含量均较低,分别为 0.413 5 和 0.417 4。在 4 个科中,不同物种 *F6'H1* 的 ENC 值存在差异,表明密码子偏好性不同,其中伞形科的平均 ENC 值最高(56.15),表明密码子偏好性较弱;豆科平均 ENC 值最低(49.37),表明密码子偏好性相对较强。

表 1 不同物种 *F6'H1* 基因密码子偏好性分析
Tab.1 Condon bias of *F6'H1* gene in different species

| 分类 Sort | 物种 Species | 基因登录号 Gene accession number | T3s | C3s | A3s | G3s | CAI | CBI | ENC | GC3s | GC | GC1 | GC2 | GC3 | GC12 |
|------------------------|----------------|-----------------------------------|---------|---------|---------|---------|--------|--------|-------|---------|---------|---------|---------|---------|---------|
| 木犀科 Oleaceae | 白蜡树 | Isoform0057303 | 0.424 0 | 0.236 7 | 0.311 1 | 0.318 4 | 0.216 | -0.046 | 57.84 | 0.415 | 0.431 2 | 0.512 5 | 0.349 0 | 0.432 1 | 0.430 8 |
| | 木樨榄 | XM_023028611.1 | 0.452 0 | 0.224 2 | 0.338 3 | 0.286 9 | 0.191 | -0.114 | 54.21 | 0.375 | 0.410 0 | 0.498 6 | 0.338 0 | 0.393 4 | 0.418 3 |
| | 木樨榄 | XM_023034893.1 | 0.438 6 | 0.214 0 | 0.359 3 | 0.279 2 | 0.197 | -0.095 | 53.63 | 0.366 | 0.408 1 | 0.498 6 | 0.343 5 | 0.382 3 | 0.421 1 |
| | 木樨榄 | XM_023036788.1 | 0.449 0 | 0.207 5 | 0.395 6 | 0.239 7 | 0.177 | -0.160 | 51.32 | 0.331 | 0.400 7 | 0.498 7 | 0.353 1 | 0.350 4 | 0.425 9 |
| | 木樨榄 | XM_023043182.1 | 0.422 4 | 0.270 8 | 0.360 6 | 0.236 5 | 0.214 | -0.082 | 53.47 | 0.382 | 0.417 4 | 0.543 4 | 0.310 9 | 0.397 8 | 0.427 2 |
| 豆科 Legumi- nosae | 大豆 | NM_001253251.3 | 0.435 7 | 0.260 7 | 0.325 6 | 0.262 9 | 0.205 | -0.088 | 49.03 | 0.394 | 0.425 2 | 0.525 6 | 0.338 1 | 0.411 9 | 0.431 9 |
| | 大豆 | XM_003520322.5 | 0.393 6 | 0.266 0 | 0.326 8 | 0.294 4 | 0.167 | -0.201 | 48.71 | 0.423 | 0.436 9 | 0.502 8 | 0.361 6 | 0.446 3 | 0.432 2 |
| | 大豆 | XM_003530041.3 | 0.395 8 | 0.243 8 | 0.342 1 | 0.285 7 | 0.177 | -0.190 | 50.63 | 0.406 | 0.437 3 | 0.522 5 | 0.362 4 | 0.427 0 | 0.442 5 |
| | 菜豆 | XM_007134245.1 | 0.406 5 | 0.251 8 | 0.326 9 | 0.297 9 | 0.204 | -0.111 | 50.54 | 0.414 | 0.437 6 | 0.518 3 | 0.354 9 | 0.439 4 | 0.436 6 |
| | 菜豆 | XM_007146169.1 | 0.435 3 | 0.266 2 | 0.347 3 | 0.235 0 | 0.203 | -0.068 | 46.44 | 0.378 | 0.414 5 | 0.515 6 | 0.331 4 | 0.396 6 | 0.423 5 |
| | 菜豆 | XM_007146170.1 | 0.434 9 | 0.256 5 | 0.345 2 | 0.243 4 | 0.193 | -0.088 | 47.14 | 0.378 | 0.416 9 | 0.519 2 | 0.336 3 | 0.395 3 | 0.427 8 |
| | 绿豆 | XM_014660601.2 | 0.420 9 | 0.233 8 | 0.293 4 | 0.339 0 | 0.198 | -0.096 | 56.30 | 0.429 | 0.439 7 | 0.508 5 | 0.355 9 | 0.454 8 | 0.432 2 |
| | 蔓花生 | XM_016092410.3 | 0.439 3 | 0.221 4 | 0.386 4 | 0.241 7 | 0.181 | -0.158 | 49.58 | 0.348 | 0.408 2 | 0.493 2 | 0.350 7 | 0.380 8 | 0.422 0 |
| | 落花生 | XM_016318780.2 | 0.435 7 | 0.221 4 | 0.406 8 | 0.226 9 | 0.183 | -0.143 | 48.70 | 0.336 | 0.400 9 | 0.484 9 | 0.347 9 | 0.369 9 | 0.416 4 |
| 蒺藜苜蓿 | XM_003599672.4 | 0.445 2 | 0.229 7 | 0.397 7 | 0.213 0 | 0.198 | -0.086 | 46.98 | 0.332 | 0.395 5 | 0.514 1 | 0.322 0 | 0.350 3 | 0.418 1 | |
| 蒺藜苜蓿 | XM_039832318.1 | 0.455 8 | 0.222 6 | 0.393 0 | 0.222 7 | 0.199 | -0.092 | 45.00 | 0.331 | 0.392 5 | 0.509 9 | 0.318 3 | 0.349 3 | 0.414 1 | |

续表 1 不同物种 *F6'H1* 基因密码子偏好性分析Tab.1(Continued) Preference related parameters of *F6'H1* gene codons in different species

| 分类 Sort | 物种 Species | 基因登录号 Gene accession number | T3s | C3s | A3s | G3s | CAI | CBI | ENC | GC3s | GC | GC1 | GC2 | GC3 | GC12 |
|--------------------------|---------------|-----------------------------------|---------|---------|---------|---------|-------|--------|-------|-------|---------|---------|---------|---------|---------|
| | 赤豆 | XM_017586437.2 | 0.410 7 | 0.239 3 | 0.301 2 | 0.336 2 | 0.199 | -0.128 | 55.62 | 0.431 | 0.437 9 | 0.505 6 | 0.355 9 | 0.452 0 | 0.430 8 |
| | 狭叶羽扇豆 | XM_019609848.1 | 0.407 9 | 0.227 4 | 0.397 7 | 0.253 2 | 0.166 | -0.17 | 52.21 | 0.361 | 0.408 8 | 0.502 8 | 0.335 2 | 0.388 3 | 0.419 0 |
| | 木豆 | XM_020369703.2 | 0.422 5 | 0.235 9 | 0.333 3 | 0.283 3 | 0.172 | -0.18 | 48.09 | 0.392 | 42.91 | 0.509 9 | 0.360 6 | 0.416 9 | 0.435 3 |
| | 木豆 | XM_020384523.2 | 0.439 7 | 0.248 2 | 0.323 2 | 0.268 9 | 0.199 | -0.097 | 47.37 | 0.391 | 41.93 | 0.521 2 | 0.331 4 | 0.405 1 | 0.426 3 |
| | 落花生 | XM_021111754.1 | 0.417 4 | 0.239 1 | 0.421 1 | 0.201 1 | 0.177 | -0.112 | 50.66 | 0.333 | 39.55 | 0.481 4 | 0.335 6 | 0.369 5 | 0.408 5 |
| | 落花生 | XM_025769767.2 | 0.428 6 | 0.221 4 | 0.392 5 | 0.245 8 | 0.175 | -0.166 | 49.93 | 0.351 | 40.91 | 0.493 2 | 0.353 4 | 0.380 8 | 0.423 3 |
| | 落花生 | XM_025818674.1 | 0.435 7 | 0.221 4 | 0.406 8 | 0.226 9 | 0.183 | -0.143 | 48.70 | 0.336 | 40.09 | 0.484 9 | 0.347 9 | 0.369 9 | 0.416 4 |
| | 落花生 | XM_025822317.1 | 0.432 6 | 0.227 0 | 0.399 2 | 0.230 1 | 0.185 | -0.139 | 49.98 | 0.344 | 40.53 | 0.489 1 | 0.349 7 | 0.377 0 | 0.419 4 |
| | 相思子 | XM_027486381.1 | 0.435 4 | 0.241 5 | 0.319 0 | 0.278 9 | 0.186 | -0.105 | 48.66 | 0.394 | 42.41 | 0.522 9 | 0.336 9 | 0.412 4 | 0.429 9 |
| | 相思子 | XM_027486998.1 | 0.437 1 | 0.258 3 | 0.332 1 | 0.249 0 | 0.207 | -0.065 | 48.65 | 0.384 | 41.96 | 0.517 3 | 0.338 7 | 0.402 7 | 0.428 0 |
| | 相思子 | XM_027488094.1 | 0.461 3 | 0.235 9 | 0.324 1 | 0.265 5 | 0.205 | -0.111 | 49.33 | 0.374 | 41.50 | 0.510 0 | 0.344 7 | 0.390 3 | 0.427 4 |
| | 豇豆 | XM_028052360.1 | 0.417 9 | 0.217 9 | 0.296 6 | 0.351 5 | 0.199 | -0.114 | 55.22 | 0.426 | 44.26 | 0.515 4 | 0.358 5 | 0.453 8 | 0.437 0 |
| | 豇豆 | XM_028074878.1 | 0.438 8 | 0.244 6 | 0.351 4 | 0.257 6 | 0.208 | -0.063 | 45.50 | 0.374 | 40.89 | 0.509 9 | 0.322 9 | 0.393 8 | 0.416 4 |
| | 野大豆 | XM_028359722.1 | 0.435 7 | 0.260 7 | 0.325 6 | 0.262 9 | 0.205 | -0.088 | 49.03 | 0.394 | 42.52 | 0.525 6 | 0.338 1 | 0.411 9 | 0.431 9 |
| | 野大豆 | XM_028369093.1 | 0.393 6 | 0.266 0 | 0.326 8 | 0.294 4 | 0.167 | -0.201 | 48.71 | 0.423 | 43.69 | 0.502 8 | 0.361 6 | 0.446 3 | 0.432 2 |
| | 野大豆 | XM_028384402.1 | 0.395 8 | 0.243 8 | 0.342 1 | 0.285 7 | 0.177 | -0.188 | 50.46 | 0.406 | 43.63 | 0.522 5 | 0.359 6 | 0.427 0 | 0.441 1 |
| | 红车轴草 | XM_045956028.1 | 0.489 4 | 0.180 9 | 0.431 9 | 0.182 6 | 0.185 | -0.115 | 43.59 | 0.272 | 37.37 | 0.495 8 | 0.329 6 | 0.295 8 | 0.412 7 |
| | 红车轴草 | XM_045961390.1 | 0.489 4 | 0.180 9 | 0.434 1 | 0.177 5 | 0.188 | -0.105 | 43.53 | 0.269 | 37.46 | 0.498 6 | 0.332 4 | 0.293 0 | 0.415 5 |
| | 牧豆树 | XM_054926754.1 | 0.371 7 | 0.253 9 | 0.327 2 | 0.332 4 | 0.192 | -0.089 | 52.66 | 0.446 | 44.60 | 0.499 0 | 0.367 8 | 0.471 2 | 0.433 4 |
| | 牧豆树 | XM_054942588.1 | 0.430 8 | 0.253 8 | 0.342 6 | 0.274 8 | 0.214 | -0.045 | 53.42 | 0.391 | 42.41 | 0.535 5 | 0.322 5 | 0.414 2 | 0.429 0 |
| 伞形科 Umbel- liferae | 野胡萝卜 | XM_017366059.1 | 0.373 6 | 0.315 0 | 0.328 1 | 0.262 9 | 0.217 | 0.007 | 60.34 | 0.443 | 45.31 | 0.550 7 | 0.347 8 | 0.460 9 | 0.449 3 |
| | 野胡萝卜 | XM_017366063.1 | 0.378 6 | 0.239 3 | 0.341 3 | 0.342 0 | 0.176 | -0.101 | 54.42 | 0.432 | 43.37 | 0.517 0 | 0.332 4 | 0.451 7 | 0.424 7 |
| | 野胡萝卜 | XM_017367700.1 | 0.397 8 | 0.259 1 | 0.334 6 | 0.309 3 | 0.187 | -0.099 | 54.18 | 0.424 | 42.33 | 0.500 0 | 0.329 5 | 0.440 3 | 0.414 8 |
| | 野胡萝卜 | XM_017401370.1 | 0.340 7 | 0.337 0 | 0.280 2 | 0.353 4 | 0.201 | -0.059 | 55.64 | 0.513 | 46.55 | 0.526 9 | 0.337 1 | 0.532 6 | 0.432 0 |
| 芸香科 Rutaceae | 血橙 | XM_006445579.2 | 0.437 1 | 0.251 7 | 0.307 4 | 0.277 1 | 0.211 | -0.083 | 53.62 | 0.404 | 43.28 | 0.513 8 | 0.361 9 | 0.422 7 | 0.437 9 |
| | 血橙 | XM_006488844.4 | 0.437 1 | 0.251 7 | 0.306 3 | 0.276 0 | 0.209 | -0.086 | 53.77 | 0.404 | 43.28 | 0.516 6 | 0.359 1 | 0.422 7 | 0.437 9 |
| 茄科 Solana- ceae | 番茄 | XM_004228918.2 | 0.487 8 | 0.177 7 | 0.404 6 | 0.233 2 | 0.188 | -0.137 | 49.06 | 0.295 | 36.30 | 0.450 0 | 0.322 2 | 0.316 7 | 0.386 1 |
| | 番茄 | XM_004251190.4 | 0.476 9 | 0.210 0 | 0.349 8 | 0.260 5 | 0.224 | -0.098 | 53.14 | 0.349 | 40.83 | 0.541 7 | 0.313 9 | 0.369 4 | 0.427 8 |
| | 番茄 | XM_004251191.3 | 0.439 3 | 0.221 4 | 0.462 7 | 0.188 3 | 0.207 | -0.073 | 52.91 | 0.301 | 38.45 | 0.486 0 | 0.343 6 | 0.324 0 | 0.414 8 |
| | 番茄 | XM_004251192.4 | 0.446 1 | 0.220 6 | 0.357 1 | 0.266 7 | 0.200 | -0.019 | 54.64 | 0.363 | 40.70 | 0.469 0 | 0.360 5 | 0.391 5 | 0.414 8 |
| | 番茄 | XM_004251392.4 | 0.424 5 | 0.248 2 | 0.448 3 | 0.176 0 | 0.205 | -0.042 | 49.03 | 0.319 | 38.83 | 0.483 2 | 0.340 8 | 0.340 8 | 0.412 0 |
| | 番茄 | XM_010328989.3 | 0.503 6 | 0.185 7 | 0.390 6 | 0.227 3 | 0.197 | -0.125 | 49.72 | 0.297 | 36.43 | 0.450 7 | 0.323 9 | 0.318 3 | 0.387 3 |
| | 烟草 | XM_016581648.1 | 0.519 2 | 0.174 2 | 0.369 8 | 0.221 3 | 0.203 | -0.083 | 49.36 | 0.296 | 39.29 | 0.516 5 | 0.343 4 | 0.318 7 | 0.430 0 |
| | 烟草 | XM_016590293.1 | 0.449 6 | 0.230 2 | 0.318 4 | 0.294 6 | 0.217 | -0.085 | 53.96 | 0.391 | 42.27 | 0.539 1 | 0.318 4 | 0.410 6 | 0.428 8 |
| | 烟草 | XM_016605403.1 | 0.464 3 | 0.235 7 | 0.400 0 | 0.204 3 | 0.194 | -0.088 | 54.32 | 0.326 | 39.70 | 0.493 1 | 0.349 0 | 0.349 0 | 0.421 1 |
| | 烟草 | XM_016606995.1 | 0.573 9 | 0.119 7 | 0.442 3 | 0.160 9 | 0.180 | -0.192 | 42.28 | 0.203 | 33.33 | 0.443 2 | 0.329 6 | 0.227 1 | 0.386 4 |
| | 烟草 | XM_016609607.1 | 0.465 9 | 0.222 2 | 0.400 8 | 0.220 8 | 0.204 | -0.072 | 54.15 | 0.325 | 39.52 | 0.501 4 | 0.338 0 | 34.63 | 41.97 |
| | 烟草 | XM_016610138.1 | 0.508 8 | 0.208 5 | 0.334 6 | 0.248 9 | 0.204 | -0.11 | 49.37 | 0.333 | 39.17 | 0.483 3 | 0.336 1 | 0.355 6 | 0.409 7 |
| | 烟草 | XM_016619361.1 | 0.519 7 | 0.164 9 | 0.404 5 | 0.195 8 | 0.209 | -0.088 | 46.63 | 0.269 | 37.51 | 0.501 4 | 0.328 7 | 0.295 3 | 0.415 1 |
| | 烟草 | XM_016627673.1 | 0.508 7 | 0.181 2 | 0.358 8 | 0.242 7 | 0.192 | -0.115 | 50.53 | 0.314 | 39.56 | 0.511 0 | 0.337 9 | 0.337 9 | 0.424 5 |
| | 烟草 | XM_016635503.1 | 0.516 0 | 0.202 8 | 0.334 6 | 0.237 9 | 0.194 | -0.098 | 50.04 | 0.324 | 38.84 | 0.484 6 | 0.336 1 | 0.344 5 | 0.410 4 |
| 十字花科 Crucif- erae | 拟南芥 | NM_104404.4 | 0.371 9 | 0.343 9 | 0.334 6 | 0.226 7 | 0.254 | 0.075 | 51.28 | 0.440 | 0.436 5 | 0.505 5 | 0.348 1 | 0.455 8 | 0.426 8 |
| | 拟南芥 | NM_112207.4 | 0.392 4 | 0.322 9 | 0.284 6 | 0.280 5 | 0.239 | 0.053 | 55.41 | 0.462 | 0.451 2 | 0.535 9 | 0.342 5 | 0.475 1 | 0.439 2 |

利用 IBM SPSS Statistics 22 软件对密码子 GC、GC1、GC2、GC3、GC12 含量和 ENC 值进行相关性分析,结果表明(表 2),20 个物种 *F6'H1* 基因 GC 与 GC1、GC2、GC3、GC12 呈极显著正相关 ($P < 0.01$),表明密码子 3 个碱基组成较相似;GC、GC1、GC3、GC12 与 ENC 呈极显著正相关 ($P < 0.01$),GC2 与 ENC 呈显著正相关 ($P < 0.05$),表明 GC 含量越高,ENC 值越高,密码子偏好性越弱。

表 2 不同物种 *F6'H1* 基因密码子参数相关性分析
Tab.2 Correlation analysis of codon parameters of *F6'H1* gene in different species

| 参数 Parameter | GC | GC1 | GC2 | GC3 | GC12 |
|-----------------|---------|---------|---------|---------|---------|
| GC1 | 0.704** | | | | |
| GC2 | 0.470** | -0.063 | | | |
| GC3 | 0.975** | 0.587** | 0.418** | | |
| GC12 | 0.872** | 0.833** | 0.500** | 0.741** | |
| ENC | 0.589** | 0.344** | 0.280* | 0.601** | 0.454** |

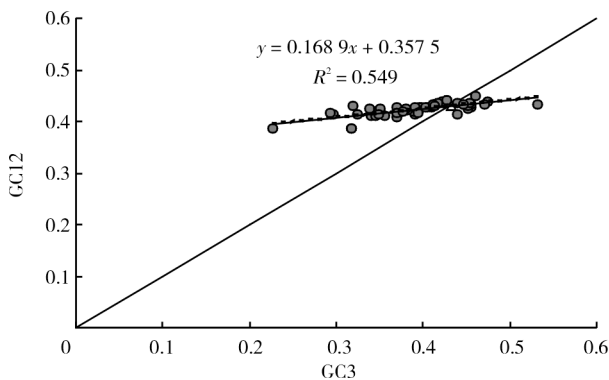
注:**表示极显著相关($P < 0.01$),*表示显著相关($P < 0.05$)。

Note: ** indicated extremely significant correlation($P < 0.01$),

* indicated significant correlation($P < 0.05$).

2.2 中性绘图分析、ENC-plot 分析和 PR2-plot 偏倚分析结果

对不同物种 *F6'H1* 基因进行中性绘图分析,结果显示如图 1 所示。



△表示白蜡树 *FcF6'H1* 基因。图 2、3 同

△indicated *Fraxinus chinensis FcF6'H1* gene. The same as Fig.2, 3

图 1 *F6'H1* 基因中性绘图分析

Fig.1 Neutrality plot analysis of *F6'H1* gene

由图 1 可知,GC12 与 GC3 的相关系数为 0.741,回归系数为 0.168 9,表明 20 个物种 *F6'H1* 基因中 3 个位置上碱基组成不存在明显差异。*FcF6'H1* 基因位于对角线上,部分基因分布在对角

线较近的位置,表明碱基突变是影响其密码子偏好性的主要因素。而其他物种的 *F6'H1* 基因分布在对角线较远的位置上,表明影响其密码子偏好性的主要因素是自然选择。

ENC-plot 分析结果如图 2 所示,少部分基因沿标准曲线分布,表明碱基突变对其密码子偏好性影响较大;而 *FcF6'H1* 和其他大部分基因位于标准曲线下方,表明自然选择是影响其密码子偏好性的主要因素。

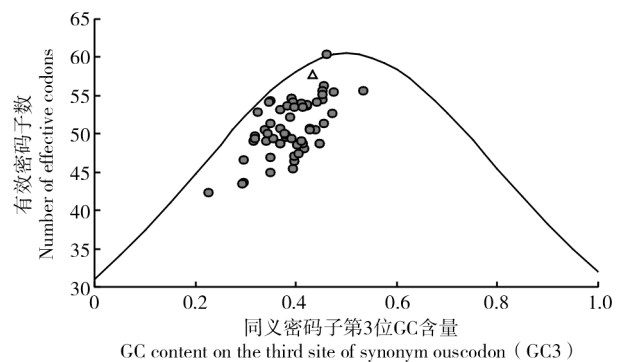


图 2 *F6'H1* 基因 ENC-plot 绘图分析

Fig.2 ENC-plot analysis of *F6'H1* gene

奇偶偏好性结果显示(图 3),*F6'H1* 基因在 4 个平面内分布不均衡,大多数 *F6'H1* 基因的坐标值偏离了 0.5,从横坐标来看,大多数基因小于 0.5,表明第 3 位上 U 的比例大于 A,密码子在 A 和 U 之间具有一定的偏好性;从纵坐标来看,大多数大于 0.5,表明第 3 位上密码子偏好 G 高于 C。*FcF6'H1* 基因位于左上方,A 小于 T,G 大于 C。若 *F6'H1* 基因完全受到突变压力的作用,A/T 和 G/C 的使用频率应相等^[24]。奇偶偏好性结果表明,20 个物种 *F6'H1* 基因进化过程中除碱基突变外,自然选择亦会影响密码子偏好性。

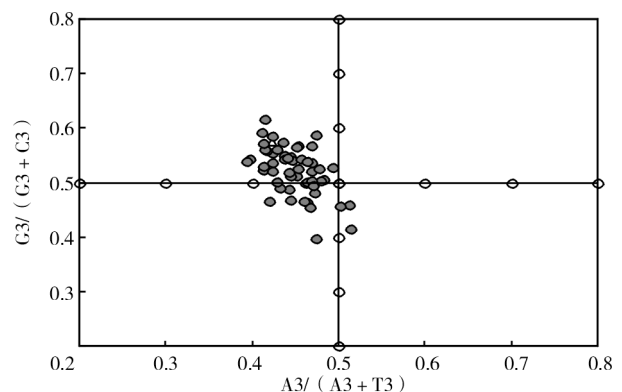


图 3 *F6'H1* 基因 PR2-plot 偏倚分析

Fig.3 PR2-plot analysis of *F6'H1* gene

2.3 *F6'H1* 基因密码子的相对使用度(RSCU)分析

当 RSCU=1.0 时,表示密码子不具有偏好性;当 RSCU>1.0 时,表示密码子偏好性较强;当 RSCU<1.0 时,表示密码子偏好性较弱;RSCU=0 时,表示密码子不参与翻译^[26]。白蜡树 *FcF6'H1* 基因密码子 RSCU 分析结果中有 24 个密码子的 RSCU 值>1,其中 16 个以 A/U 结尾,8 个以 G/C 结尾,表明 *FcF6'H1* 基因密码子偏好以 A/U 结尾密码子。密码子 UAA、AGG 和 GGU 的 RSCU 值分别为 3.0、2.5 和 2.0,表

明具有优先选择性;AGU、UAG 和 UGA 的 RSCU 值均为 0,表明 *FcF6'H1* 基因缺失这些密码子。

20 个物种 *F6'H1* 基因 RSCU 分析结果如表 3 所示,20 个物种 *F6'H1* 基因密码子的相对使用度(RSCU)分析结果中,有 28 个密码子的 RSCU 值>1,其中 22 个以 A/U 结尾,6 个以 G/C 结尾,表明 20 个物种 *F6'H1* 基因密码子偏好以 A/U 结尾密码子。RSCU 较高的 3 个密码子是 AGA(2.17)、AGG(2.04)和 UGA(1.9),表明其密码子偏好性较强。

表 3 20 个物种 *F6'H1* 基因编码序列同义密码子相对使用度

Tab.3 Relative synonymous codon usage(RSCU) analysis of *F6'H1* gene coding sequence of 20 species

| 氨基酸 | Amino acid | 密码子 | Codon | RSCU | 氨基酸 | Amino acid | 密码子 | Codon | RSCU |
|-----|------------|-----|-------|------|-----|------------|-----|-------|------|
| Phe | | UUU | | 1.03 | Trp | | UGG | | 1 |
| | | UUC | | 0.97 | TER | | UAA | | 0.7 |
| Leu | | UUA | | 0.79 | | | UAG | | 0.4 |
| | | UUG | | 1.28 | | | UGA | | 1.9 |
| | | CUU | | 1.44 | Pro | | CCU | | 1.82 |
| | | CUC | | 0.97 | | | CCC | | 0.67 |
| | | CUA | | 0.99 | | | CCA | | 1.23 |
| | | CUG | | 0.53 | | | CCG | | 0.28 |
| Ile | | AUU | | 1.37 | Thr | | ACU | | 1.42 |
| | | AUC | | 0.91 | | | ACC | | 0.95 |
| | | AUA | | 0.72 | | | ACA | | 1.38 |
| Met | | AUG | | 1 | | | ACG | | 0.25 |
| Val | | GUU | | 1.38 | Ala | | GCU | | 1.74 |
| | | GUC | | 0.63 | | | GCC | | 0.46 |
| | | GUA | | 0.62 | | | GCA | | 1.56 |
| | | GUG | | 1.37 | | | GCG | | 0.24 |
| Tyr | | UAU | | 1.23 | Cys | | UGU | | 0.94 |
| | | UAC | | 0.77 | | | UGC | | 1.06 |
| His | | CAU | | 1.49 | Ser | | AGU | | 0.91 |
| | | CAC | | 0.51 | | | AGC | | 1.16 |
| Gln | | CAA | | 1.55 | | | UCU | | 1.26 |
| | | CAG | | 0.45 | | | UCC | | 0.66 |
| Asn | | AAU | | 1.33 | | | UCA | | 1.55 |
| | | AAC | | 0.67 | | | UCG | | 0.46 |
| Lys | | AAA | | 0.98 | Arg | | CGU | | 0.53 |
| | | AAG | | 1.02 | | | CGC | | 0.25 |
| Asp | | GAU | | 1.51 | | | CGA | | 0.7 |
| | | GAC | | 0.49 | | | CGG | | 0.31 |
| Gly | | GGU | | 1.39 | | | AGA | | 2.17 |
| | | GGC | | 0.58 | | | AGG | | 2.04 |
| | | GGA | | 1.29 | Glu | | GAA | | 1.11 |
| | | GGG | | 0.74 | | | GAG | | 0.89 |

聚类热图结果显示(图 4),分为 8 个小支,烟草 XM_016619361.1、XM_016610138.1 和番茄 XM_004228918.2、XM_010328989.3 聚为一支;豆

科植物牧豆树 XM_054926754.1、XM_054942588.1,绿豆 XM_014660601.2 和豇豆 XM_028052360.1 聚为一支;番茄 XM_004251192.4,伞形科野胡萝卜

XM_017367700.1、芸香科血橙 XM_006445579.2、XM_006488844.4 和烟草 XM_016606995.1、XM_016635503.1 聚为一支。木犀科植物白蜡树 *FcF6'H1* 基因,木樨榄 XM_023036788.1、XM_023028611.1、XM_023034893.1 和豆科植物相思子 XM_027486381.1,大豆 XM_003530041.3,野大豆 XM_028384402.1 聚为一支;豆科植物落花生 XM_021111754.1、XM_025822317.1、XM_016318780.2、XM_025818674.1、XM_025769767.2 和蔓花生 XM_016092410.3 聚为一支;豆科植物相思子 XM_027486998.1、XM_027488094.1,木豆 XM_020384523.2,菜豆 XM_007146169.1、XM_007146170.1,豇豆 XM_028074878.1,大豆 NM_

001253251.3,野大豆 XM_028359722.1,红车轴草 XM_045956028.1、XM_045961390.1,蒺藜苜蓿 XM_003599672.4、XM_039832318.1 聚为一支;烟草 XM_016581648.1、XM_016627673.1、XM_016605403.1、XM_016609607.1、XM_016590293.1,番茄 XM_004251191.3、XM_004251392.4、XM_004251190.4,狭叶羽扇豆 XM_019609848.1,木豆 XM_020369703.2,大豆 XM_003520322.5,野大豆 XM_028369093.1,木樨榄 XM_023043182.1,菜豆 XM_007134245.1,赤豆 XM_017586437.2 聚为一支;伞形科野胡萝卜 XM_017366063.1、XM_017401370.1、XM_017366059.1 和拟南芥 NM_104404.4、NM_112207.4 聚为一支。

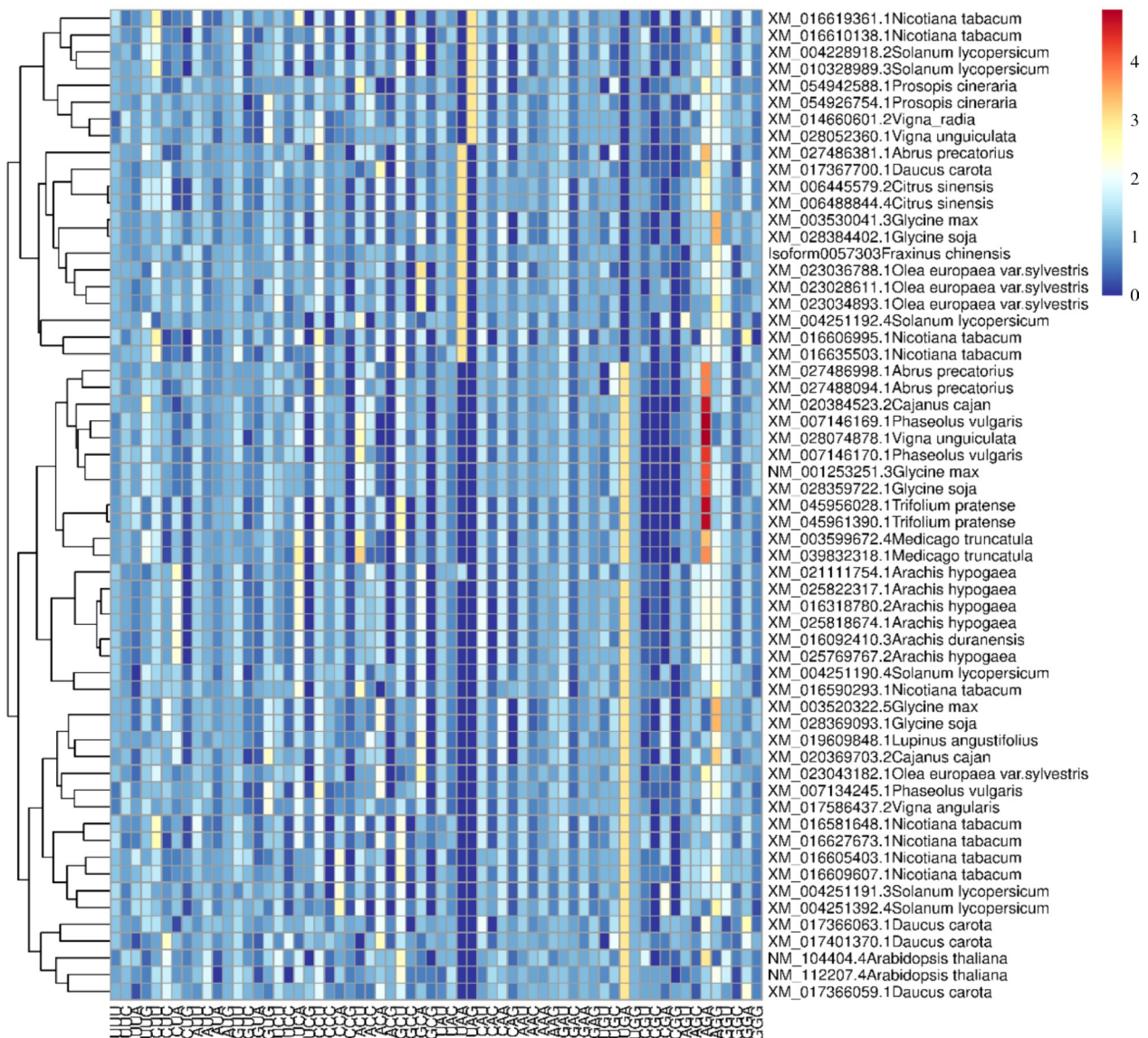


图4 *F6'H1* 基因密码子的RSCU热图分析

Fig.4 Heatmap of relative synonymous codon usage (RSCU) for *F6'H1* genes

2.4 最优密码子分析

表 4 结果确定了 *F6'H1* 基因的 6 个最优密码子, 分别是 CUC(Δ RSCU 0.69)、AUC(Δ RSCU

0.44)、AAG(Δ RSCU 0.23)、GAG(Δ RSCU 0.3)、UCG(Δ RSCU 0.97)、ACA(Δ RSCU 0.45)。

表 4 *F6'H1* 基因最优密码子
Tab.4 The optimal codons of *F6'H1* gene

| 氨基酸 Amino acid | 密码子 Codon | 高表达基因 Highly expressed gene | | 低表达基因 Low expressed gene | | Δ RSCU | 氨基酸 Amino acid | 密码子 Codon | 高表达基因 Highly expressed gene | | 低表达基因 Low expressed gene | | Δ RSCU |
|-------------------|--------------|--------------------------------|------|-----------------------------|------|---------------|-------------------|--------------|--------------------------------|------|-----------------------------|------|---------------|
| | | 数量 Number | RSCU | 数量 Number | RSCU | | | | 数量 Number | RSCU | | | |
| Phe | UUU | 64 | 0.93 | 71 | 0.92 | 0.01 | Ser | UCU | 53 | 1.22 | 43 | 1.08 | 0.14 |
| | UUC | 73 | 1.07 | 83 | 1.08 | -0 | | UCC | 30 | 0.69 | 33 | 0.83 | -0.1 |
| Leu | UUA | 39 | 0.8 | 55 | 1.03 | -0.2 | Uca | UCA | 56 | 1.29 | 83 | 2.09 | -0.8 |
| | UUG | 57 | 1.16 | 82 | 1.54 | -0.4 | | UCG | 44 | 1.02 | 2 | 0.05 | 0.97 |
| CUC | CUU | 56 | 1.14 | 89 | 1.67 | -0.5 | Pro | CCU | 88 | 1.74 | 119 | 2.04 | -0.3 |
| | CUC | 62 | 1.27 | 31 | 0.58 | 0.69 | | CCC | 25 | 0.5 | 39 | 0.67 | -0.2 |
| CUA | CUA | 37 | 0.76 | 43 | 0.81 | -0.1 | CCA | CCA | 56 | 1.11 | 72 | 1.24 | -0.1 |
| | CUG | 43 | 0.88 | 20 | 0.38 | 0.5 | | CCG | 33 | 0.65 | 3 | 0.05 | 0.6 |
| Ile | AUU | 93 | 1.21 | 139 | 1.51 | -0.3 | Thr | ACU | 28 | 1 | 66 | 2.13 | -1.1 |
| | AUC | 87 | 1.13 | 64 | 0.69 | 0.44 | | ACC | 27 | 0.96 | 28 | 0.9 | 0.06 |
| AUA | AUA | 51 | 0.66 | 74 | 0.8 | -0.1 | ACA | ACA | 39 | 1.39 | 29 | 0.94 | 0.45 |
| | AUG | 79 | 1 | 63 | 1 | 0 | | ACG | 18 | 0.64 | 1 | 0.03 | 0.61 |
| Met | AUG | 79 | 1 | 63 | 1 | 0 | Ala | GCU | 87 | 1.78 | 109 | 2.2 | -0.4 |
| | GUU | 100 | 1.37 | 105 | 1.39 | -0 | | GCC | 25 | 0.51 | 20 | 0.4 | 0.11 |
| Val | GUC | 60 | 0.82 | 41 | 0.54 | 0.28 | GCA | GCA | 53 | 1.09 | 64 | 1.29 | -0.2 |
| | GUA | 36 | 0.49 | 57 | 0.75 | -0.3 | | GCG | 30 | 0.62 | 5 | 0.1 | 0.52 |
| GUG | GUG | 97 | 1.32 | 100 | 1.32 | 0 | Cys | UGU | 18 | 0.84 | 14 | 0.9 | -0.1 |
| | UAU | 69 | 1.13 | 82 | 1.3 | -0.2 | | UGC | 25 | 1.16 | 17 | 1.1 | 0.06 |
| Tyr | UAC | 53 | 0.87 | 44 | 0.7 | 0.17 | TER | UGA | 6 | 1.8 | 8 | 2.4 | -0.6 |
| | UAA | 2 | 0.6 | 1 | 0.3 | 0.3 | | Trp | UGG | 52 | 1 | 46 | 1 |
| UAG | UAG | 2 | 0.6 | 1 | 0.3 | 0.3 | Arg | CGU | 16 | 0.71 | 2 | 0.12 | 0.59 |
| | CAU | 67 | 1.44 | 76 | 1.58 | -0.1 | | CGC | 8 | 0.35 | 0 | 0 | 0.35 |
| His | CAC | 26 | 0.56 | 20 | 0.42 | 0.14 | CGA | CGA | 20 | 0.88 | 5 | 0.29 | 0.59 |
| | CAA | 67 | 1.35 | 91 | 1.53 | -0.2 | | CGG | 11 | 0.49 | 3 | 0.18 | 0.31 |
| Gln | CAG | 32 | 0.65 | 28 | 0.47 | 0.18 | Arg | AGA | 38 | 1.68 | 67 | 3.94 | -2.3 |
| | AAU | 122 | 1.27 | 138 | 1.41 | -0.1 | | AGG | 43 | 1.9 | 25 | 1.47 | 0.43 |
| Asn | AAC | 70 | 0.73 | 58 | 0.59 | 0.14 | Gly | GGU | 65 | 1.31 | 63 | 1.39 | -0.1 |
| | AAA | 123 | 0.94 | 155 | 1.17 | -0.2 | | GGC | 30 | 0.6 | 22 | 0.49 | 0.11 |
| Lys | AAG | 140 | 1.06 | 110 | 0.83 | 0.23 | GGA | GGA | 74 | 1.49 | 63 | 1.39 | 0.1 |
| | GAU | 133 | 1.29 | 167 | 1.51 | -0.2 | | GGG | 30 | 0.6 | 33 | 0.73 | -0.1 |
| Asp | GAC | 74 | 0.71 | 54 | 0.49 | 0.22 | Ser | AGU | 44 | 1.02 | 40 | 1.01 | 0.01 |
| | GAA | 113 | 0.95 | 146 | 1.25 | -0.3 | | AGC | 33 | 0.76 | 37 | 0.93 | -0.2 |
| Glu | GAG | 125 | 1.05 | 88 | 0.75 | 0.3 | | | | | | | |

2.5 不同物种 *F6'H1* 基因的系统进化树

对 20 个物种 59 条序列构建系统进化树分析, 结果显示(图 5), 聚为 6 小枝, 白蜡树 *FcF6'H1* 与木樨榄 XM_023028611.1、XM_023036788.1、XM_023034893.1, 烟草 XM_016627673.1、XM_

016581648.1、XM_016619361.1、XM_016606995.1 聚为一类。伞形科野胡萝卜 XM_017366059.1、XM_017366063.1、XM_017367700.1、XM_017401370.1 与拟南芥 NM_112207.4、NM_104404.4 聚为一类。芸香科血橙 XM_006445579.2、XM_006488844.4

与部分豆科植物聚为一类,分别是牧豆树 XM_054926754.1, 落花生 XM_016318780.2、XM_021111754.1、XM_025769767.2、XM_025818674.1、XM_025822317.1、蔓花生 XM_016092410.3, 狭叶羽扇豆 XM_019609848.1, 赤豆 XM_017586437.2,

绿豆 XM_014660601.2, 菜豆 XM_007134245.1, 木豆 XM_020369703.2, 大豆 XM_003520322.5、XM_003530041.3, 野大豆 XM_028369093.1、XM_028384402.1 聚为一类。其他豆科植物聚为一类。

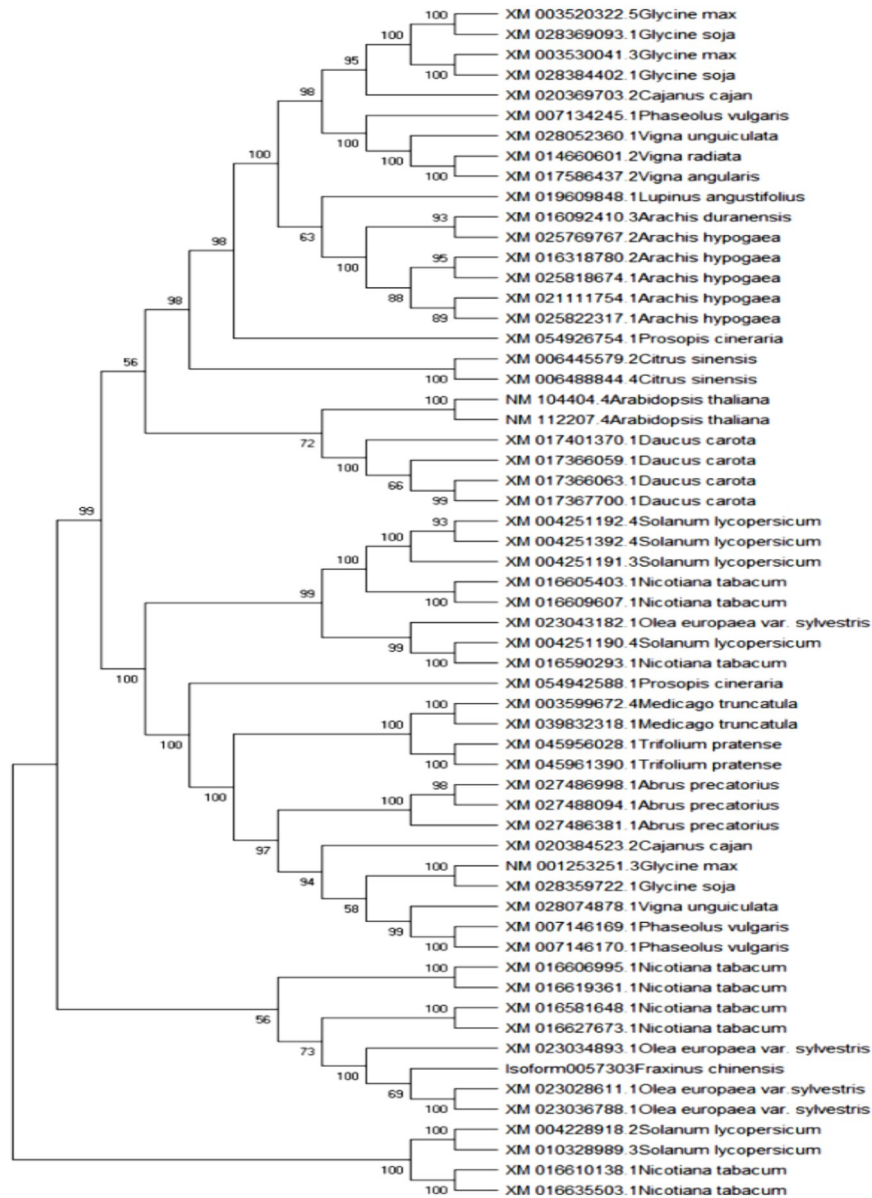


图5 *F6'H1* 基因分子进化树

Fig.5 Phylogenetic tree of *F6'H1* genes

2.6 *FcF6'H1* 基因的异源受体选择

白蜡树遗传转化体系尚未见报道,因此,其基因功能验证需借助模式生物进行异源表达。将 *FcF6'H1* 基因密码子使用频率与模式生物基因组密码子使用频率比较,比值在 0.5~2.0 范围内表示物种间密码子偏好性较一致。从表 5 可以看出,

FcF6'H1 基因与拟南芥、番茄、烟草的基因组密码子使用频率差异较大的个数分别是 11、8、8 个,表明番茄和烟草可作为 *FcF6'H1* 基因的遗传转化受体;与酵母菌、大肠杆菌的基因组密码子使用频率差异较大的个数分别是 15、9 个,表明大肠杆菌更适合 *FcF6'H1* 基因的异源表达。

表 5 白蜡树 *FcF6'H1* 基因与 5 种模式生物基因组密码子使用偏好性比较Tab.5 Comparison of codon usage bias of *FcF6'H1* gene in *Fraxinus chinensis* and 5 pattern organism genomes

| 密码子 | Codon | 氨基酸 | Amino acid | <i>FcF6'H1</i> | <i>FcF6'H1/At</i> | <i>FcF6'H1/Sl</i> | <i>FcF6'H1/Nt</i> | <i>FcF6'H1/Sc</i> | <i>FcF6'H1/Ec</i> |
|-----|-------|-----|------------|----------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| UUU | | F | | 22.16 | 1.02 | 0.85 | 0.88 | 0.85 | 0.91 |
| UUC | | F | | 19.39 | 0.94 | 1.11 | 1.08 | 1.05 | 1.40 |
| UUA | | L | | 11.08 | 0.87 | 0.77 | 0.83 | <u>0.42</u> | 0.64 |
| UUG | | L | | 13.85 | 0.66 | 0.57 | 0.62 | 0.51 | 1.07 |
| CUU | | L | | 19.39 | 0.80 | 0.78 | 0.81 | 1.58 | 1.34 |
| CUC | | L | | 16.62 | 1.03 | 1.48 | 1.35 | <u>3.08</u> | 1.75 |
| CUA | | L | | 8.31 | 0.84 | 0.83 | 0.88 | 0.62 | 1.48 |
| CUG | | L | | 22.16 | <u>2.26</u> | <u>2.11</u> | <u>2.17</u> | <u>2.11</u> | 0.59 |
| AUU | | I | | 36.01 | 1.67 | 1.28 | 1.30 | 1.20 | 1.22 |
| AUC | | I | | 11.08 | 0.60 | 0.79 | 0.80 | 0.64 | 0.57 |
| AUA | | I | | 22.16 | 1.76 | 1.58 | 1.58 | 1.25 | 1.67 |
| AUG | | M | | 16.62 | 0.68 | 0.67 | 0.66 | 0.80 | 0.70 |
| GUU | | V | | 24.93 | 0.92 | 0.89 | 0.93 | 1.13 | 1.15 |
| GUC | | V | | 16.62 | 1.30 | 1.65 | 1.50 | 1.41 | 1.27 |
| GUA | | V | | 16.62 | 1.68 | 1.48 | 1.46 | 1.41 | 1.27 |
| GUG | | V | | 22.16 | 1.27 | 1.39 | 1.33 | <u>2.05</u> | 1.11 |
| UCU | | S | | 22.16 | 0.88 | 1.05 | 1.11 | 0.94 | 1.69 |
| UCC | | S | | 5.54 | <u>0.49</u> | 0.56 | 0.54 | <u>0.39</u> | 0.57 |
| UCA | | S | | 19.39 | 1.06 | 0.94 | 1.10 | 1.04 | 1.48 |
| UCG | | S | | 11.08 | 1.19 | 1.98 | <u>2.09</u> | 1.29 | 1.35 |
| CCU | | P | | 30.47 | 1.63 | 1.59 | 1.63 | <u>2.26</u> | <u>3.21</u> |
| CCC | | P | | 11.08 | <u>2.09</u> | 1.94 | 1.68 | 1.63 | 1.79 |
| CCA | | P | | 13.85 | 0.86 | 0.72 | 0.70 | 0.76 | 1.52 |
| CCG | | P | | 8.31 | 0.97 | 1.81 | 1.66 | 1.57 | 0.57 |
| ACU | | T | | 11.08 | 0.63 | 0.56 | 0.55 | 0.55 | 0.85 |
| ACC | | T | | 13.85 | 1.34 | 1.61 | 1.43 | 1.09 | 0.73 |
| ACA | | T | | 13.85 | 0.88 | 0.77 | 0.80 | 0.78 | 0.92 |
| ACG | | T | | 8.31 | 1.08 | 1.81 | 1.85 | 1.04 | 0.61 |
| GCU | | A | | 22.16 | 0.78 | 0.72 | 0.71 | 1.05 | 1.17 |
| GCC | | A | | 2.77 | <u>0.27</u> | <u>0.27</u> | <u>0.22</u> | <u>0.22</u> | <u>0.13</u> |
| GCA | | A | | 13.85 | 0.79 | 0.62 | 0.60 | 0.85 | 0.60 |
| GCG | | A | | 11.08 | 1.23 | <u>2.13</u> | 1.91 | 1.79 | 0.53 |
| UAU | | Y | | 19.39 | 1.33 | 1.04 | 1.09 | 1.03 | 0.90 |
| UAC | | Y | | 13.85 | 1.01 | 1.12 | 1.03 | 0.94 | 1.18 |
| UAA | | * | | 2.77 | <u>3.08</u> | <u>3.08</u> | <u>2.52</u> | <u>2.52</u> | 1.39 |
| UAG | | * | | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| CAU | | H | | 19.39 | 1.41 | 1.25 | 1.45 | 1.43 | 1.56 |
| CAC | | H | | 8.31 | 0.96 | 1.07 | 0.96 | 1.07 | 1.14 |
| CAA | | Q | | 19.39 | 1.00 | 0.92 | 0.94 | 0.71 | 1.35 |
| CAG | | Q | | 8.31 | 0.55 | 0.59 | 0.55 | 0.69 | <u>0.31</u> |
| AAU | | N | | 22.16 | 0.99 | 0.73 | 0.79 | 0.62 | 0.76 |
| AAC | | N | | 27.70 | 1.33 | 1.60 | 1.55 | 1.12 | 1.36 |
| AAA | | K | | 41.55 | 1.35 | 1.34 | 1.27 | 0.99 | 1.12 |
| AAG | | K | | 41.55 | 1.27 | 1.34 | 1.24 | 1.35 | <u>2.72</u> |

续表5 白蜡树 *FcF6'H1* 基因与5种模式生物基因组密码子使用偏好性比较
Tab.5(Continued) Comparson of codon usage bias of *FcF6'H1* gene in *Fraxinus chinensis* and 5 pattern organism genomes

| 密码子 | Codon | 氨基酸 | Amino acid | <i>FcF6'H1</i> | <i>FcF6'H1/At</i> | <i>FcF6'H1/Sl</i> | <i>FcF6'H1/Nt</i> | <i>FcF6'H1/Sc</i> | <i>FcF6'H1/Ec</i> |
|-----|-------|-----|------------|----------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| GAU | | D | | 47.09 | 1.29 | 1.20 | 1.28 | 1.25 | 1.40 |
| GAC | | D | | 8.31 | <u>0.48</u> | 0.55 | <u>0.49</u> | <u>0.41</u> | <u>0.46</u> |
| GAA | | E | | 33.24 | 0.97 | 0.96 | 0.92 | 0.73 | 0.95 |
| GAG | | E | | 38.78 | 1.20 | 1.46 | 1.32 | <u>2.02</u> | 2.00 |
| UGU | | C | | 5.54 | 0.53 | 0.51 | 0.57 | 0.68 | 0.94 |
| UGC | | C | | 5.54 | 0.77 | 0.83 | 0.77 | 1.15 | 1.01 |
| UGA | | * | | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| UGG | | W | | 13.85 | 1.11 | 1.03 | 1.14 | 1.33 | 1.03 |
| CGU | | R | | <u>2.77</u> | <u>0.31</u> | <u>0.40</u> | <u>0.37</u> | <u>0.43</u> | <u>0.17</u> |
| CGC | | R | | 5.54 | 1.46 | 1.79 | 1.42 | <u>2.13</u> | 0.40 |
| CGA | | R | | <u>2.77</u> | <u>0.44</u> | 0.51 | 0.52 | 0.92 | 0.58 |
| CGG | | R | | <u>2.77</u> | 0.57 | 0.89 | 0.75 | 1.63 | <u>0.35</u> |
| AGU | | S | | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| AGC | | S | | 16.62 | 1.47 | 1.79 | 1.66 | 1.70 | 1.16 |
| AGA | | R | | 5.54 | <u>0.29</u> | <u>0.34</u> | <u>0.35</u> | <u>0.26</u> | 0.78 |
| AGG | | R | | 13.85 | 1.26 | 1.16 | 1.14 | 1.51 | <u>3.46</u> |
| GGU | | G | | 27.70 | 1.25 | 1.16 | 1.24 | 1.16 | 1.17 |
| GGC | | G | | <u>2.77</u> | <u>0.30</u> | <u>0.29</u> | <u>0.25</u> | <u>0.28</u> | <u>0.13</u> |
| GGA | | G | | 11.08 | <u>0.46</u> | <u>0.43</u> | 0.48 | 1.02 | 0.81 |
| GGG | | G | | 13.85 | 1.36 | 1.28 | 1.32 | <u>2.31</u> | 1.13 |

注: *At*为拟南芥; *Nt*为烟草; *Sl*为番茄; *Ec*为大肠杆菌; *Sc*为酵母菌。*表示终止密码子。下划线表示密码子使用频率差异较大。

Note: *At. Arabidopsis thaliana*; *Nt. Nicotiana tabacum*; *Sl. Solanum lycopersicum*; *Ec. Escherichia coli*; *Sc. Saccharomyces cerevisiae*. * Indicated termination codon. The underline data indicated that the codon usage frequency was quite different.

3 结论与讨论

本研究表明,白蜡树 *FcF6'H1* 的 GC、GC1、GC2、GC3 和 GC12 含量分别为 0.431 2、0.512 5、0.349 0、0.432 1 和 0.430 8, CAI 值为 0.216, ENC 值为 57.84, G+C 含量占比较低,表明白蜡树 *FcF6'H1* 基因偏好以 A/T 结尾,密码子偏好性较弱。这与西伯利亚杏 *PsSOC1-like* 基因密码子偏好性^[8]、罗布麻 *AvFLS* 基因密码子偏好性^[25] 分析结果类似。白蜡树 *FcF6'H1* 基因密码子中 UAA、AGG 和 GGU 具有优先选择性,缺失 AGU、UAG 和 UGA 这 3 个密码子,表明在进化过程中逐渐淘汰了以上 3 个密码子,而具体的进化机制有待于深入研究。

长期的自然进化使各近缘物种或基因表现出相似的密码子偏好模式,因此,密码子使用法则也常作为物种或基因分类、起源、进化和亲缘关

系研究的重要参考依据^[27]。*F6'H1* 基因的 CDS 进化树和 RSCU 值聚类分析结果并不完全一致。例如,CDS 进化树分析结果中,白蜡树 *FcF6'H1* 与木樨榄 XM_023028611.1、XM_023036788.1、XM_023034893.1,烟草 XM_016627673.1、XM_016581648.1、XM_016619361.1、XM_016606995.1 聚为一类。而 RSCU 值聚类分析结果中,木犀科植物白蜡树 *FcF6'H1*,木樨榄 XM_023036788.1、XM_023028611.1、XM_023034893.1 和豆科植物相思子 XM_027486381.1,大豆 XM_003530041.3,野大豆 XM_028384402.1 聚为一支。CDS 进化树和 RSCU 值聚类分析结果均表明,白蜡树 *FcF6'H1* 与木樨榄 XM_023028611.1、XM_023036788.1、XM_023034893.1 聚为一类。文献报道,基于 RSCU 值聚类分析,可将单子叶植物和双子叶植物查尔酮异构酶 CHI 归为 2 类,而单双子叶内部亚类群结构中的 RSCU 分类与 CDS 聚类结果仍存在较

大差异^[27];黄酮醇合成酶 *FLS* 基因的 CDS 序列进化树和 RSCU 值聚类分析的比较结果表明, *FLS* 基因 CDS 的差异大小更符合亲缘关系的分类规律,而密码子偏好性的聚类分析结果可能很好地反映了 *FLS* 基因的某些特异性进化规律^[25]。因此,为了准确地反映不同物种的进化历史,RSCU 值聚类分析结合 CDS 序列进化分析是最佳选择。

突变压力和自然选择效应是生物体密码子偏好性形成的关键因素,这 2 种作用力在不同物种中并非均等存在^[27-28]。蒺藜苜蓿基因组密码子偏好性主要受碱基突变影响^[29],而拟南芥密码子偏好性受碱基突变和自然选择的双重影响^[30],中华猕猴桃 *GRAS* 家族密码子偏好性主要受自然选择压力的影响^[31]。中性绘图分析结果中, *FcF6'H1* 基因位于对角线上;ENC-plot 分析结果中, *FcF6'H1* 基因位于标准曲线下方。表明碱基突变和自然选择均会影响 *FcF6'H1* 基因密码子的偏好性。奇偶偏好性结果中,大多数 *F6'H1* 基因的 A/T 和 G/C 的使用频率不相等,表明 20 个物种 *F6'H1* 基因进化过程中除碱基突变外,自然选择亦会影响密码子的偏好性。

香豆素类化合物是白蜡树的主要药效成分, *F6'H1* 是香豆素类化合物生物合成的关键基因,但是白蜡树的遗传背景和相关基因调控机制尚未阐明,严重制约了香豆素类化合物的理论研究和开发利用。采用模式生物进行基因外源表达是目前基因功能研究的有效手段之一。本研究通过白蜡树 *FcF6'H1* 基因与模式生物基因组密码子偏好性比较,表明番茄、烟草、大肠杆菌可作为 *FcF6'H1* 基因的遗传转化受体。但目标基因能否高效表达,主要受到表达载体、基因表达量、蛋白质合成、转化体系、培养基等因素的共同影响^[8,25]。因此,如何使 *FcF6'H1* 基因高效遗传转化还需更深入研究。

本研究对密码子偏好性参数进行分析,结果表明, *FcF6'H1* 基因密码子偏好以 A/U 结尾,且密码子偏好性较弱。RSCU 值和 CDS 序列聚类分析结果并不完全一致,但 CDS 序列进化分析更符合物种亲缘关系的远近程度。中性分析、ENC-plot 分析及 PR2-plot 分析结果表明, *FcF6'H1* 基因偏好性主要受碱基突变的影响,但是自然选择仍具有一定影响。模式生物中番茄和烟草适合作为遗传转化受体,大肠杆菌表达系统适合作为 *FcF6'H1*

的异源表达载体。

参考文献:

- [1] 张得芳,夏涛. 基于转录组的唐古特白刺密码子偏好性分析[J]. 干旱区资源与环境,2021,35(12):163-170.
ZHANG D F, XIA T. Analysis of *Codon* bias based on transcriptome in *Nitraria tangutorum*[J]. Journal of Arid Land Resources and Environment,2021,35(12):163-170.
- [2] 张扬,孙曙光,李晴,等. 莴苣线粒体基因组密码子使用偏好性分析[J]. 河南农业科学,2022,51(10):114-124.
ZHANG Y, SUN S G, LI Q, et al. *Codon* usage bias analysis of cultivated and wild lettuce mitochondrial genomes[J]. Journal of Henan Agricultural Sciences,2022,51(10):114-124.
- [3] ANDARGIE M, ZHU C Y. Genome-wide analysis of *Codon* usage in sesame (*Sesamum indicum* L.) [J]. Heliyon, 2022, 8(1):e08687.
- [4] YANNAI A D, KATZ S, HERSHBERG R. The *Codon* usage of lowly expressed genes is subject to natural selection[J]. Genome Biology and Evolution,2018,10(5):1237-1246.
- [5] 卢珍红,原晓龙,李绅崇,等. 基于转录组序列的非洲菊密码子偏好性分析[J]. 西南林业大学学报(自然科学),2023,43(5):77-86.
LU Z H, YUAN X L, LI S C, et al. *Codon* usage bias analysis of *Gerbera jamesonii* based on transcriptome sequences[J]. Journal of Southwest Forestry University,2023,43(5):77-86.
- [6] ZHANG W J, ZHOU J, LI Z F, et al. 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.
- [7] 高梦琦,邹建珍,霍小位,等. 基于转录组数据分析药用大黄的密码子使用偏好性[J]. 中草药,2021,52(20):6344-6349.
GAO M Q, ZOU J Z, HUO X W, et al. Analysis of *Codon* usage patterns in *Rheum officinale* based on transcriptome data [J]. Chinese Traditional and Herbal Drugs,2021,52(20):6344-6349.
- [8] 蔺悦,陈俊兴,红梅,等. 西伯利亚杏 PsSOC1-like 基因密码子偏好性与进化分析[J]. 内蒙古农业大学学报(自然科学版),2023,44(1):13-25.
LIN Y, CHEN J X, HONG M, et al. *Codon* bias and evolutionary analysis of PsSOC1-like gene in *Prunus sibirica*[J]. Journal of Inner Mongolia Agricultural University(Natural Science Edition),2023,44(1):13-25.
- [9] MATSUMOTO S, MIZUTANI M, SAKATA K, et al. Molecular cloning and functional analysis of the ortho-hydroxylases of p-coumaroyl coenzyme A/feruloyl coenzyme A involved in formation of umbelliferone and scopoletin in sweet potato, *Ipomoea batatas*(L.) Lam.[J]. Phytochemistry,2012,74:49-57.
- [10] 王荣香,宋佳,孙博,等. 香豆素类化合物功能及生物合成研究进展[J]. 中国生物工程杂志,2022,42(12):79-90.
WANG R X, SONG J, SUN B, et al. Research progress of function and biosynthesis of coumarins[J]. China Biotechnology,2022,42(12):79-90.
- [11] 段珍,吴凡,闫启,等. 植物香豆素生物合成途径及关键酶基因研究进展[J]. 草业学报,2022,31(1):217-228.

- DUAN Z, WU F, YAN Q, et al. Research progress on plant coumarin biosynthesis pathway and the genes encoding the key enzymes[J]. *Acta Prataculturae Sinica*, 2022, 31(1): 217-228.
- [12] 国家药典委员会. 中华人民共和国药典:2020年版一部[S]. 北京:中国医药科技出版社,2020:282-283.
National Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China 2020 Edition[S]. Beijing: China Medical Science Press, 2020:282-283.
- [13] 钱正明, 吴梦奇, 谭国英, 等. 高效液相色谱紫外等吸收波长法快速测定秦皮中秦皮甲素和秦皮乙素[J]. *色谱*, 2023, 41(8):690-697.
QIAN Z M, WU M Q, TAN G Y, et al. Rapid determination of aesculin and aesculetin in Fraxini Cortex by high performance liquid chromatography-ultraviolet at equal absorption wavelength[J]. *Chinese Journal of Chromatography*, 2023, 41(8):690-697.
- [14] HE B T, LIU Z H, LI B Z, et al. Advances in biosynthesis of scopoletin[J]. *Microbial Cell Factories*, 2022, 21(1):152.
- [15] SIWINSKA J, SIATKOWSKA K, OLRZY A, et al. Scopoletin 8-hydroxylase: a novel enzyme involved in coumarin biosynthesis and iron-deficiency responses in *Arabidopsis*[J]. *Journal of Experimental Botany*, 2018, 69(7):1735-1748.
- [16] SHIMIZU B I. 2-Oxoglutarate-dependent dioxygenases in the biosynthesis of simple coumarins[J]. *Frontiers in Plant Science*, 2014, 5:549.
- [17] 王蕾, 张娟, 魏丽娟, 等. 大豆 *GmF6'H1* 基因的克隆及功能验证[J]. *西北植物学报*, 2015, 35(2):213-219.
WANG L, ZHANG J, WEI L J, et al. Molecular cloning and functional characterization of *GmF6'H1* in soybean[J]. *Acta Botanica Boreali-Occidentalia Sinica*, 2015, 35(2):213-219.
- [18] SUN X C, LI H R. Full-length transcriptome combined with RNA sequence analysis of *Fraxinus chinensis*[J]. *Genes & Genomics*, 2023, 45(5):553-567.
- [19] 郑乔木, 黄思沛, 李秋娥, 等. 16种豆科植物 DXS 基因密码子偏好性分析[J]. *生物学杂志*, 2022, 39(4):50-54.
ZHENG Q M, HUANG S P, LI Q E, et al. Analysis of *Codon* bias of DXS gene in 16 Leguminous plants[J]. *Journal of Biology*, 2022, 39(4):50-54.
- [20] 龚婷婷, 李维仪, 陶媛, 等. 兰科 ABI₃ 基因的密码子偏性分析[J]. *热带作物学报*, 2023, 44(2):264-272.
GONG T T, LI W Y, TAO Y, et al. Analysis of *Codon* usage bias of ABI₃ gene in Orchidaceae[J]. *Chinese Journal of Tropical Crops*, 2023, 44(2):264-272.
- [21] WRIGHT F. The 'effective number of codons' used in a gene[J]. *Gene*, 1990, 87(1):23-29.
- [22] LÜ H, ZHAO W M, ZHENG Y, et al. Analysis of synonymous *Codon* usage bias in *Chlamydia*[J]. *Acta Biochimica et Biophysica Sinica*, 2005, 37(1):1-10.
- [23] 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.
- [24] NIE X J, DENG P C, FENG K W, et al. Comparative analysis of *Codon* usage patterns in chloroplast genomes of the Asteraceae family[J]. *Plant Molecular Biology Reporter*, 2014, 32(4):828-840.
- [25] 蔡德育, 郭晓农, 朱月滢, 等. 罗布麻 AvFLS 密码子偏好性与进化[J]. *应用与环境生物学报*, 2022, 28(2):423-431.
CAI D Y, GUO X N, ZHU Y Y, et al. *Codon* bias and evolutionary analysis of Av FLS in *Apocynum venetum* L[J]. *Chinese Journal of Applied and Environmental Biology*, 2022, 28(2):423-431.
- [26] 原晓龙, 王毅, 罗婷, 等. 思茅松及云南松转录组密码子偏好性比较分析[J]. *中南林业科技大学学报*, 2023, 43(9): 153-163.
YUAN X L, WANG Y, LUO T, et al. Comparison between the *Codon* preferences of the transcriptomes of *Pinus kesiya* and *Pinus yunnanensis*[J]. *Journal of Central South University of Forestry & Technology*, 2023, 43(9):153-163.
- [27] 赖瑞联, 冯新, 陈瑾, 等. 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):945-951.
LAI R L, FENG X, CHEN J, et al. *Codon* usage pattern of *Chalcone* isomerase gene (CHI) in *Canarium album* (Lour.) Raeusch[J]. *Chinese Journal of Applied and Environmental Biology*, 2017, 23(5):945-951.
- [28] KIMURA M. Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1981, 78(9):5773-5777.
- [29] 杨国锋, 苏昆龙, 赵怡然, 等. 蕹菜苜蓿叶绿体密码子偏好性分析[J]. *草业学报*, 2015, 24(12):171-179.
YANG G F, SU K L, ZHAO Y R, et al. Analysis of *Codon* usage in the chloroplast genome of *Medicago truncatula*[J]. *Acta Prataculturae Sinica*, 2015, 24(12):171-179.
- [30] QIU S, ZENG K, SLOTTE T, et al. Reduced efficacy of natural selection on *Codon* usage bias in selfing *Arabidopsis* and *Capsella* species[J]. *Genome Biology and Evolution*, 2011, 310:868-880.
- [31] 毛可欣, 王海荣, 安森, 等. 中华猕猴桃 GRAS 基因家族鉴定及低温胁迫表达分析[J]. *生物技术通报*, 2023, 39(11): 297-307.
MAO K X, WANG H R, AN M, et al. Identification of GRAS gene family and expression analysis under low temperature stress in *Actinidia chinensis*[J]. *Biotechnology Bulletin*, 2023, 39(11):297-307.