

DOI: 10.11931/guihaia.gxzw202003074

梁晓静, 朱昌叁, 李开祥, 等. 香樟转录组基因密码子偏好性分析 [J]. 广西植物, 2021, 41(12): 2077–2083.

LIANG XJ, ZHU CS, LI KX, et al. Codon bias of transcriptomic genes in *Cinnamomum camphora* [J]. *Guihaia*, 2021, 41(12): 2077–2083.



# 香樟转录组基因密码子偏好性分析

梁晓静<sup>1</sup>, 朱昌叁<sup>1</sup>, 李开祥<sup>1</sup>, 安家成<sup>1</sup>, 王鹏良<sup>2\*</sup>

(1. 广西林业科学研究院, 广西特色经济林培育与利用重点实验室, 国家林业局八角肉桂工程技术研究中心, 广西木本香料工程技术研究中心, 南宁 530002; 2. 北部湾大学, 广西 钦州 535011)

**摘要:** 为了解香樟基因密码子偏好性, 该文以 NCBI 网站中香樟转录组数据为材料, 利用生物信息学手段评价转录组数据质量, 选取高质量数据的转录组, 去除低质量序列, 组装转录组, 预测基因结构, 再利用自编 perl 脚本提取以 AUG 开头的基因序列 37 Mb 序列 34 931 个基因, 进一步利用 CodonW 分析基因密码子偏好性。结果表明: GC 含量的变化范围为 0.273~0.742, 均值为 0.452; ENC 的范围为 26.29~61.00, 均值为 52.76; CAI 的范围为 0.064~0.401, 均值为 0.199; RSCU 值大于 1 的密码子数目为 27 个, 其中以 U 或 A 结尾的有 22 个; 中性分析表明, 小部分基因在对角线上, 大多数基因偏离对角线; ENC-plot 分析表明小部分基因在标准曲线上, 大多数基因偏离标准曲线。上述研究结果表明, 香樟基因的密码子偏好性比较弱, 密码子常以 A/U 结尾; 突变和选择两者都在密码子偏好中起作用, 而选择作用更大; 最终确定了 GUU、CAG、GAA、UCU、GCU、GGU 为最优密码子, 通过对目标基因密码子的校正, 提高表达效率, 从而为利用基因工程技术改良香樟重要性状奠定了基础。

**关键词:** 香樟, 转录组, 基因, 密码子, 偏好性

**中图分类号:** Q945.4 **文献标识码:** A **文章编号:** 1000-3142(2021)12-2077-07

## Codon bias of transcriptomic genes in *Cinnamomum camphora*

LIANG Xiaojing<sup>1</sup>, ZHU Changsan<sup>1</sup>, LI Kaixiang<sup>1</sup>, AN Jiacheng<sup>1</sup>, WANG Pengliang<sup>2\*</sup>

(1. Guangxi Forestry Research Institute, Guangxi Key Laboratory for Cultivation and Utilization of Special Non-Timber Forest Crops, Engineering and Technology Research Center for Anise and Cinnamon of State Forestry Administration, Guangxi Engineering and Technology Research Center for Woody Spices, Nanning 530002, China; 2. Beibu Gulf University, Qinzhou 535011, Guangxi, China)

**Abstract:** In order to understand the genes codon bias of transcriptome in *Cinnamomum camphora*, all transcriptome data of *C. camphora* were downloaded, screened against their quality. The screened data were assembled and annotated in gene structure after low-quality reads were deleted. Then 37 Mb gene sequences initiated with AUG were extracted by perl script with 34 931 genes. The codon bias was analyzed using CodonW software. The results were as follows: GC content ranged from 0.273 to 0.742 with the average of 0.452; ENC ranged from 26.29 to 61.00 with the average of 52.76; CAI ranged from 0.064 to 0.401 with the average of 0.199; There were 27 genes whose RSCU were greater than 1, of which 22 genes ended with U or A; Neutral plot analysis indicated that a great many genes were not on the line and

收稿日期: 2020-07-22

基金项目: 自治区主席科技资金项目(1517-06); 第八批广西特聘专家专项; 广西林科院基本科研业务费项目(林科 201806 号); 广西重点实验室开放课题(19-B-04-01) [Supported by the Scientific Program of Chairman in Guangxi (1517-06); the Eighth Batch Guangxi Special Expert Program; Basic Research Fund of Guangxi Forestry Research Institute (LinKe 201806); Open Project of Guangxi Key Laboratory (19-B-04-01)].

作者简介: 梁晓静(1983-), 硕士, 高级工程师, 主要从事经济林栽培育种研究, (E-mail) liangxj2013@126.com。

\*通信作者: 王鹏良, 博士, 教授, 硕士研究生导师, 主要从事植物遗传育种研究, (Email) pengliang\_wang@163.com。

ENC-plot showed the similar results. The above results elucidate that the gene codon bias of transcriptome was weak, and most ended with U or A; Selection played more important role than mutation in codon bias; Finally, GUU, CAG, GAA, UCU, GCU and GGU were taken as the optimal codons. This will provide a solid foundation for improving traits of *C. camphora* via the codon revision of the targeted genes.

**Key words:** *Cinnamomum camphora*, transcriptome, genes, codon, bias

香樟 (*Cinnamomum camphora*) 是樟科樟属的一种常绿大乔木,为我国国家二级保护植物,广泛分布于我国南方及西南各省(中国科学院中国植物志编辑委员会,1982)。香樟的木材是良好的建筑用材,其根、枝叶、木材和种子均可提取樟脑和樟油,用途十分广泛,具有很高的应用价值。

香樟群体及个体间均存在较大遗传变异。根据樟油化学成分差异,香樟可分为本樟、芳樟和油樟(中国科学院中国植物志编辑委员会,1982);从外部表型上看,材用和油用樟树木材构造存在差异(王军锋等,2019),不同化学型油用樟树叶片解剖结构和抗旱性也存在差异(王坤等,2019),从内部遗传物质上看,等位基因也存在较大差异(刘爽,2019)。丰富的遗传变异为香樟的定向遗传改良提供了良好的资源。遗传改良通常利用杂交等传统的方法开展,然而这些方法存在周期长、成本高、盲目性大等不足。而分子育种能有效克服这些缺点成为现代育种的重要手段。

尽管香樟的分子研究起步较迟,但已有一些报道。Liu et al. (2015)利用同源克隆的方法克隆了香樟 *CcBBM*,并进行了转化验证;Chen et al. (2018)利用转录组对香樟萜类生物合成基因开展了鉴定分析;另外,研究人员对香樟不同化学成分合作途径的关键基因 *CcHMGRs*(郑汉等,2020a)、*CcPMK*(郑汉等,2020b)及 NBS-LRR 类抗病基因家族开展克隆和部分功能分析的工作(郑永杰等,2018);同时施雪萍(2009)和陈甘明(2014)先后建立了香樟体细胞再生体系并成功将 *Barnase*、*PaFT*、*CBF3* 和 *DREB1* 等基因转入香樟,为香樟不同成分的合成、抗逆机理探索和遗传改良奠定基础。基因工程育种是分子育种的重要方法之一,具有周期短、效率高、目标明确等优点(王关林和方宏筠,2014)。密码子指导氨基酸的合成(Krebs et al., 2018)。不同的密码子对基因的表达和功能产生较大影响(Hershberg & Petrov, 2008; Zhou et al., 2016)。不同物种密码子偏好性存在较大差异(陆育生等,2018),了解香樟基因的密码子偏好性,对利用基因工程技术培育香樟新品种具有十分重要的作用。因此,本文分析香樟转录组基因密码子偏好性,确定最优密码子,从而为利用基

因工程改良香樟奠定坚实基础。

## 1 材料与amp;方法

### 1.1 材料

从 NCBI 网站中下载 6 个 SRR6436155、SRR6436156、SRR8316438、SRR6436010、SRR6436011、SRR6374671 香樟转录组测序的结果。本文利用 FastQC 检测转录组的序列质量,筛选数据质量较高的转录组为材料,再利用 Trimmomatic 修剪去除低质量的序列和接头。在此基础上利用 Trinity 对转录组进行组装,利用 TransDecoder 预测基因结构。提取以 AUG 开头的基因用于本文的后续分析。

### 1.2 数据分析

1.2.1 密码子偏好参数估计 以所选的编码蛋白基因为研究对象,采用 CodonW 1.4.2 软件分析密码子偏好性的参数:有效密码子数目(ENC)、密码子适应指数(CAI)、密码子偏好性指数(CBI)、最优密码子使用频率(Fop)、基因表达的蛋白质疏水性(Gravy)和芳香族氨基酸的频率(Aromo)及密码子不同位置碱基的比例(GC1、GC2 和 GC3 分别代表第一位、第二位和第三位密码子的 GC 含量,GC3s 表示第三位同义密码子的 GC 含量)。

1.2.2 中性绘图分析 为了初步摸索对密码子偏好性的影响因素,中性绘图根据 GC1 和 GC2 均值记为 GC12,其作为纵坐标,GC3 作为横坐标,用基因散点图的方式进行作图。根据基因所在位置与对角线的关系从而判断影响基因密码子偏好性的因素。

1.2.3 ENC-plot 绘图 为进一步了解造成密码子偏好的原因,本文以 ENC 为纵坐标,以 GC3s 为横坐标建立坐标系,根据基因坐落的位置作散点图。再利用方程添加其理论标准曲线  $ENC_{exp} = 2 + GC3s + \frac{29}{GC3s^2 + (1 - GC3s)^2}$ , 并计算  $ENC_{Ratio} = \frac{ENC_{exp} - ENC_{obs}}{ENC_{exp}}$ 。根据散点图及 ENC 比值的结果推断造成密码子偏好的原因。

1.2.4 最佳密码子确定 为了确定最优密码子,本文以 CAI 密码子偏好性参数为标准对参试基因进行排序,分别从高低两端取总数的 1% 建立高表达

库和低表达库,并以高表达库的 RSCU 与低表达库的对应的 RSCU 的差值为  $\Delta$ RSCU,以  $\Delta$ RSCU 大于 0.08 且 RSCU 大于 1 的密码子为最优密码子。

## 2 结果与分析

### 2.1 香樟转录组选取

为了准确分析香樟密码子偏好性,本文从 NCBI 中下载了 6 个香樟转录组测序的结果。但由于许多转录组没有明确注明转录组的组织,或者对转录测序质量比较差的结果进行剔除,最后只剩下 SRR6436155 一个转录组用作本文的实验数据。该转录组数据量为 9.7 G。利用 Trinity 组装转录组后得到 305 Mb 序列,利用 cd-hit-est 命令以默认参数进一步延伸得到 229 Mb。在此基础上,利用 TransDecoder 将转录组序列与 Swissprot 和 pfam-A 数据库同源比对后得到 54 Mb 数据。利用自编 perl 脚本提取以 AUG 开头的 37 Mb 基因序列用于后续分析。

### 2.2 密码子偏好性参数

本文利用 CodonW 软件(<https://sourceforge.net/projects/codonw/>)计算所有参试基因密码子偏好性参数。由表 1 可知,密码子不同位置碱基的 GC 含量差异较大,第一位、第二位和第三位密码子 GC 含量的变化范围分别为 0.245~0.913、0.179~0.858 和 0.108~0.939,均值分别为 0.511、0.408 和 0.434;总 GC 含量的变化范围为 0.273~0.742,均值为 0.452;CAI 的范围为 0.064~0.401,均值为 0.199;CBI 的范围为 -0.513~0.364,均值为 -0.056;Fop 的范围为 0.153~0.634,均值为 0.384;Gravy 的范围为 -2.747~1.436,均值为 -0.259;其氨基酸长度变化范围 86~476,均值为 366;Aromo 的变化范围为 0.000~0.324,均值为 0.085。ENC 的范围为 26.29~61.00,均值为 52.76,在 34 931 基因中,有 80 个基因的 ENC 小于 35 (Jiang et al., 2008),其余基因的 ENC 均大于 35,这说明香樟转录组中只有极少数基因偏好性较强,绝大多数基因的偏好性较弱,甚至有些基因没有密码子偏好性。

密码子偏好性参数相关性分析结果(表 2)表明,GC1、GC2、GC3s 与其余参数都呈极显著相关;GC3 除与 Gravy 不相关外,与其他参数均显著相关。上述说明基因的碱基组成对密码子偏好性参数存在一定影响。基因长度(N)与基因不同位置的碱基含量极显著相关,GC1、GC2 及 GC3 的相关系数从 0.145 到 -0.103,说明基因越长,GC1 越高,

GC3 越低。基因长度与 CAI 无显著相关;尽管基因长度与密码子偏好性参数 CBI、Fop、ENC 为极显著相关,但其相关系数极低,说明基因长度对密码子偏好影响不大。

RSCU 分析结果(表 3)表明,RSCU 值大于 1 的密码子数目为 27 个,其中以 U 结尾的有 14 个,以 A 结尾的有 8 个,以 G 结尾的有 5 个;以 A 或 U 结尾的密码子占 81.48%。

### 2.3 中性绘图

香樟叶片转录组基因中性绘图结果(图 1)表明,GC12 的变化范围为 0.271~0.759,GC3 变化范围为 0.108~0.939。图 1 中代表基因的点有一些落在对角线上,有更多的点偏离对角线,GC12 与 GC3 的相关性不显著。上述结果综合表明香樟转录组基因密码子偏好性同时受到突变和选择的作用。

### 2.4 ENC-plot 绘图

以 ENC 作为 y 轴,以 GC3s 作为 x 轴,开展 ENC-plot 绘图。将参试基因都布置于该坐标系,GC3s 分布于 0.077~0.937 之间,代表基因的点有的着落在标准曲线的上方,有的着落在标准曲线的下方及另外一部分在标准曲线上(图 2)。为了进一步分析基因分布的具体情况,本文分析了 ENC 比值的频率分布(表 4),32% 的基因分布于 -0.05~0.05 之间,54.99% 的基因分布于 0.05~0.15 之间,1.66% 分布于 -0.15~-0.05 之间,9.62% 分布于 0.15~0.25 之间,1.67% 分布于 0.25~0.5 之间。说明突变在香樟叶片转录组基因偏好性形成过程中起重要作用,而选择的作用更大。

### 2.5 最优密码子

按照基因密码子偏好性参数 CAI 数值大小进行排序,从两端分别取 1% 的基因分别建立高表达库和低表达库,再分别计算两个库的 RSCU,进一步计算  $\Delta$ RSCU。64 个密码子中,有 25 个密码子的  $\Delta$ RSCU 大于 0.08,其中以 G 结尾的有 3 个,以 C 结尾的有 15 个,以 A 结尾的有 2 个,以 U 结尾的有 5 个。结合表 3 中 RSCU 大于 1 的高频密码子与表 5 中标星号的高表达密码子,从而确定最优密码子。最终确定香樟叶片转录组基因中有 6 个最优密码子分别为 GUU、CAG、GAA、UCU、GCU、GGU,其中 5 个密码子以 A 或 U 结尾,另外 1 个以 G 结尾。

## 3 讨论与结论

本文以测序质量较高转录组数据为材料研究香樟密码子偏好性。在转录组数据组装、延伸去

表 1 密码子偏好性参数变化范围  
Table 1 Range of codon bias parameters

参数 Parameter	最小值 Min.	最大值 Max.	均值 Average	参数 Parameter	最小值 Min.	最大值 Max.	均值 Average
GC1	0.245	0.913	0.511	ENC	26.29	61.00	52.76
GC2	0.179	0.858	0.408	GC3s	0.077	0.937	0.412
GC3	0.108	0.939	0.434	GC	0.273	0.742	0.452
CAI	0.064	0.401	0.199	L_sym	83	4 586	351.885
CBI	-0.513	0.364	-0.056	L_aa	86	4 765	366
Fop	0.153	0.634	0.384	Gravy	-2.747	1.436	-0.259
Aromo	0.000	0.324	0.085				

表 2 密码子偏好性参数间的相关性  
Table 2 Correlation of codon bias parameters

参数 Parameter	GC1	GC2	GC3	CAI	CBI	Fop	ENC	GC3s	GC	N	Gravy
GC2	0.172 **										
GC3	0.185 **	0.100 **									
CAI	0.176 **	-0.096 **	0.166 **								
CBI	0.150 **	0.173 **	0.508 **	0.605 **							
Fop	0.143 **	0.210 **	0.478 **	0.698 **	0.962 **						
ENC	0.060 **	-0.45 **	0.066 **	0.002	0.049 **	0.037 **					
GC3s	0.212 **	0.113 **	0.994 **	0.179 **	0.521 **	0.494 **	0.065 **				
GC	0.599 **	0.555 **	0.785 **	0.142 **	0.473 **	0.465 **	0.048 **	0.798 **			
N	0.145 **	0.023 **	-0.103 **	0	-0.037 **	-0.024 **	0.036 **	-0.094 **	-0.014 **		
Gravy	-0.222 **	-0.122 **	0	-0.094 **	0.079 **	-0.102 **	0.031 **	-0.025 **	-0.140 **	-0.088 **	
Aromo	-0.451 **	-0.201 **	0.016 **	-0.008	-0.013 *	-0.055 **	0.039 **	-0.023 **	-0.255 **	-0.159 **	0.442 **

注: \*\* 表示极显著水平; \* 表示显著水平。

Note: \*\* indicates extremely significant level; \* indicates significant level.

除冗余序列的基础上,利用 TrasnDecoder 结合 Swissprot 数据库和 pfam-A 数据库的同源搜索大大提高了捕捉功能开放阅读框(opening reading framework, ORF)的敏感性和准确性,同时确实也可能存在一些操纵子和嵌合体。然而在目前没有香樟基因组的情况下,转录组基因密码子分析仍然可以提供重要参考。

64 种密码子编码 20 种氨基酸,其中 3 种终止密码子,另有 AUG 和 UGG 分别只单独编码甲硫氨酸和色氨酸,其余 59 种密码子编码 18 种氨基酸(朱圣庚和徐长发,2016; Krebs et al., 2018)。18 种氨基酸中每种氨基酸至少有 2 种密码子编码,最多的有 6 种密码子编码(Nelson & Cox, 2017),由于不同物种或同一物种不同基因组利用不同密码子编码同一氨基酸(王鹏良等, 2018, 2019; 赖瑞联等, 2019),因此产生密码子使用的偏好性。

香樟叶绿体基因组密码子偏好性报道表明其密码子偏好性较弱(秦政等, 2018),本文中密码子偏好性参数 ENC 均值为 52.76,大于叶绿体基因组

的 ENC,表明香樟转录组基因密码子偏好性比叶绿体更弱,也就是说在同一密码子中选择任意密码子对翻译效率和蛋白功能影响较小。RSCU 分析表明,香樟转录组基因密码子仍然偏向以 A 或 U 结尾。这与橄榄转录组密码子偏好性相似(赖瑞联等, 2019)。

密码子偏好受碱基组成、选择、突变、tRNA 丰度、蛋白质长度、疏水性和脂肪族氨基酸含量的影响。ENC 几乎与所有的密码子参数极显著相关,相关系数均低于 0.1,与 GC2 呈极显著相关,相关系数为-0.45;而另一密码子偏好性参数 CAI 与不同位置 GC 含量相关,与氨基酸长度相关性较小,均说明密码子偏好性与碱基组成有关,氨基酸长度相关性小。CAI 与 CBI 及 Fop 的紧密相关表明基因选择高表达的密码子。中性分析和 ENC-plot 分析表明突变在密码子偏好形成中起作用,而选择在这个过程中起更大作用。本文筛选出 6 个香樟基因最优密码子分别为 GUU、CAG、GAA、UCU、GCU、GGU。最优密码子的确定为利用基因工程技术改良香樟奠定了坚实基础。

表 3 密码子相对使用度

Table 3 Relative synonymous codon usage analysis

氨基酸 Amino acid	密码子 Codon	数量 Number	RSCU	氨基酸 Amino acid	密码子 Codon	数量 Number	RSCU
苯丙氨酸 Phe	UUU	304 100	1.13	丝氨酸 Ser	UCU	311 649	1.61
	UUC	233 350	0.87		UCC	166 140	0.86
亮氨酸 Leu	UUA	147 757	0.71		UCA	257 999	1.33
	UUG	300 758	1.45	UCG	89 171	0.46	
	CUU	302 166	1.45	脯氨酸 Pro	CCU	233 779	1.48
	CUC	186 741	0.90		CCC	107 080	0.68
	CUA	121 249	0.58		CCA	224 344	1.42
异亮氨酸 Ile	CUG	188 771	0.91	CCG	67 821	0.43	
	AUU	312 103	1.38	苏氨酸 Thr	ACU	208 488	1.38
	AUC	198 381	0.88		ACC	126 441	0.84
	AUA	167 349	0.74		ACA	205 703	1.36
甲硫氨酸 Met	AUG	322 582	1.00	ACG	63 477	0.42	
	缬氨酸 Val	GUU	318 663	1.53	丙氨酸 Ala	GCU	325 497
GUC		153 611	0.74	GCC		160 934	0.74
GUA		132 957	0.64	GCA		301 653	1.38
GUG		228 953	1.10	GCG		86 780	0.40
酪氨酸 Tyr	UAU	210 442	1.21	半胱氨酸 Cys		UGU	125 865
	UAC	136 369	0.79		UGC	114 030	0.95
终止子 TER	UAA	9 611	0.92	终止子 TER	UGA	14 082	1.35
	UAG	7 489	0.72	色氨酸 Trp	UGG	162 726	1.00
组氨酸 His	CAU	208 545	1.32	精氨酸 Arg	CGU	90 400	0.77
	CAC	106 626	0.68		CGC	69 857	0.59
谷氨酰胺 Gln	CAA	234 428	0.99	丝氨酸 Ser	CGA	81 950	0.70
	CAG	241 418	1.01		CGG	78 056	0.66
天冬酰胺 Asn	AAU	331 926	1.26		AGU	185 042	0.95
	AAC	192 873	0.74	AGC	154 319	0.80	
赖氨酸 Lys	AAA	339 397	0.91	精氨酸 Arg	AGA	203 756	1.73
	AAG	409 146	1.09		AGG	180 692	1.54
天冬氨酸 Asp	GAU	475 997	1.40	甘氨酸 Gly	GGU	242 575	1.14
	GAC	205 874	0.60		GGC	157 721	0.74
谷氨酸 Glu	GAA	432 745	1.04		GGA	264 126	1.25
	GAG	399 622	0.96		GGG	184 049	0.87

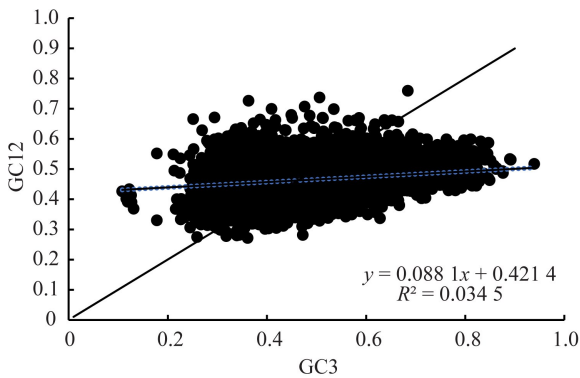


图 1 香樟转录组基因中性绘图

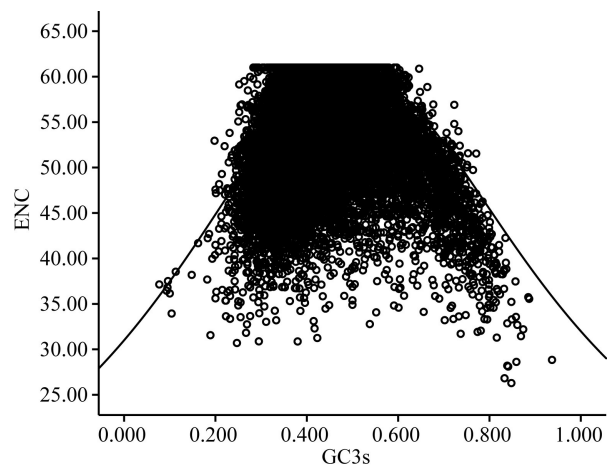
Fig. 1 Neutral plot of transcriptomic genes in *Cinnamomum camphora*

图 2 香樟转录组基因 ENC-plot 分析

Fig. 2 ENC-plot analysis on transcriptomic genes in *Cinnamomum camphora*

目前,香樟的转录组研究、重要基因分子克隆正在迅猛开展。香樟的遗传转化体系正逐步完善。施雪萍(2009)将 *GFP* 基因转入香樟体细胞胚中,通过绿色荧光检测、PCR 检测和 Southern blotting 检测证实了 *GFP* 基因成功导入香樟基因组中,建立了相

对完善的遗传转化体系;施雪萍(2009)将 *Barnase* 和 *PaFT* 基因分别转入香樟基因组,然而未对基因

表 4 香樟叶片转录组基因 ENC 比值频率分布

Table 4 ENC ratio of transcriptomic genes in *Cinnamomum camphora*

ENC 范围 Range of ENC	数目 Number	频率 Frequency (%)
-0.25 ~ -0.15	22	0.06
-0.15 ~ -0.05	580	1.66
-0.05 ~ 0.05	11 178	32.00
0.05 ~ 0.15	19 209	54.99
0.15 ~ 0.25	3 359	9.62
0.25 ~ 0.50	583	1.67

转化的表型效应进行鉴定;王长宪(2009)将沙冬青抗寒基因 *AmEBP1* 和 *AmGS* 转入香樟基因组,并通过抗寒能力测定表明,转基因植株的抗寒能力有所提高。然而施雪萍(2009)和王长宪(2009)都没有对目的基因的密码子按照物种最优密码子进行调整。Young & Purton (2016)报道了不适合的遗传密码子在一定程度上降低或阻止转化目的基因的表达。按照密码子偏好性结果对转化目的基因进行密码子的修饰能够大大促进基因表达,提高目的基因的应用效果。本研究的成果对利用基因工程技术改良香樟具有重要的指导意义。

表 5 香樟转录组基因最优密码子分析

Table 5 Analysis of the optimal codons of transcriptomic genes in *Cinnamomum camphora*

氨基酸 Amino acid	密码子 Codon	低表达库 Low expression genes		高表达库 High expression genes		$\Delta$ RSCU	氨基酸 Amino acid	密码子 Codon	低表达库 Low expression genes		高表达库 High expression genes		$\Delta$ RSCU
		数目 Number	RSCU	数目 Number	RSCU				数目 Number	RSCU	数目 Number	RSCU	
Phe	UUU	1 384	1.32	1 095	0.68	-0.64	Tyr	UAU	969	1.35	894	0.70	-0.65
	UUC***	716	0.68	2 134	1.32	0.64		UAC***	464	0.65	1 643	1.30	0.65
Leu	UUA	750	0.83	520	0.51	-0.32	TER	UAA**	90	0.87	118	1.19	0.32
	UUG	1 474	1.63	1 309	1.29	-0.34		UAG	88	0.85	67	0.67	-0.18
Ile	CUU	1 208	1.34	962	0.95	-0.39	His	CAU	730	1.32	909	0.93	-0.39
	CUC**	767	0.85	1 311	1.29	0.44		CAC**	380	0.68	1 054	1.07	0.39
	CUA	631	0.70	369	0.36	-0.34	Gln	CAA	967	1.11	1 277	0.84	-0.27
	CUG***	582	0.65	1 620	1.60	0.95		CAG*	770	0.89	1 766	1.16	0.27
Met	AUU	1 249	1.22	1 441	1.07	-0.15	Asn	AAU	1 391	1.35	1 407	0.79	-0.56
	AUC***	710	0.70	2 250	1.67	0.97		AAC***	664	0.65	2 140	1.21	0.56
Val	AUA	1 102	1.08	347	0.26	-0.82	Lys	AAA	1 237	0.88	3 001	0.88	0.00
	AUG	1 533	1.00	1 889	1.00	0.00		AAG	1 560	1.12	3 858	1.12	0.00
Ser	GUU*	903	1.22	2 088	1.35	0.13	Asp	GAU	1 771	1.42	3 183	1.13	-0.29
	GUC*	550	0.75	1 589	1.03	0.28		GAC*	728	0.58	2 455	0.87	0.29
	GUA	460	0.62	836	0.54	-0.08	Glu	GAA*	1 554	0.91	3 953	1.06	0.15
	GUG	1 037	1.41	1 683	1.09	-0.32		GAG	1 874	1.09	3 502	0.94	-0.15
Pro	UCU*	1 193	1.30	1 783	1.50	0.20	Cys	UGU	526	1.18	555	0.80	-0.38
	UCC***	697	0.76	1 681	1.41	0.65		UGC**	366	0.82	836	1.20	0.38
	UCA	1 109	1.20	869	0.73	-0.47	TER	UGA	132	1.28	113	1.14	-0.14
	UCG	609	0.66	788	0.66	0.00	Trp	UGG	882	1.00	895	1.00	0.00
Thr	CCU	1 071	1.38	1 278	1.15	-0.23	Arg	CGU***	500	0.48	718	1.40	0.92
	CCC	636	0.82	873	0.79	-0.03		CGC***	428	0.41	642	1.25	0.84
	CCA	1 089	1.40	1 318	1.19	-0.21		CGA	761	0.72	389	0.76	0.04
	CCG**	306	0.39	962	0.87	0.48		CGG	646	0.61	242	0.47	-0.14
Ala	ACU	643	1.16	1 370	1.11	-0.05	Ser	AGU	1 078	1.17	756	0.63	-0.54
	ACC***	420	0.76	1 897	1.54	0.78		AGC*	840	0.91	1 268	1.06	0.15
	ACA	833	1.50	870	0.71	-0.79	Arg	AGA	1 988	1.89	596	1.16	-0.73
Gly	ACG	325	0.59	786	0.64	0.05		AGG	1 988	1.89	490	0.96	-0.93
	GCU*	827	1.20	2 728	1.50	0.30		GGU***	954	0.85	2 226	1.49	0.64
	GCC**	548	0.80	2 072	1.14	0.34		GGC***	637	0.57	1 815	1.22	0.65
	GCA	950	1.38	1 467	0.81	-0.57		GGA	1 562	1.39	1 316	0.88	-0.51
	GCG	424	0.62	1 016	0.56	-0.06		GGG	1 349	1.20	614	0.41	-0.79

注: \*\*\* 表示  $\Delta$ RSCU>0.5; \*\* 表示  $\Delta$ RSCU>0.3; \* 表示  $\Delta$ RSCU>0.08。具体氨基酸缩写见表 3。

Note: \*\*\* indicates  $\Delta$ RSCU>0.5; \*\* indicates  $\Delta$ RSCU>0.3; \* indicates  $\Delta$ RSCU>0.08. See Table 3 for abbreviations of specific amino acids.

## 参考文献:

- CHEN CH, ZHENG YJ, ZHONG YD, et al., 2018. Transcriptome analysis and identification of genes related to terpenoid biosynthesis in *Cinnamomum camphora* [J]. BMC Genomics 19: 550–564.
- CHEN GM, 2014. Optimization of genetic transformation system and transformation of cold resistant genes in *Cinnamomum camphora* L. [D]. Wuhan: Huazhong Agricultural University [陈甘明, 2014. 樟树遗传转化体系的优化及抗寒基因的转化 [D]. 武汉: 华中农业大学.]
- Editorial Board Committee of Flora in China, 1982. Flora Reipublicae Popularis Sinicae [M]. Beijing: Science Press. [中国科学院中国植物志编辑委员会, 1982. 中国植物志 [M]. 北京: 科学出版社.]
- HERSHBERG R, PETROV DA, 2008. Selection on codon bias [J] Ann Rev Genet, 42: 287–299.
- JIANG Y, DENG F, WANG HL, et al., 2008. An extensive analysis on the global codon usage pattern of baculoviruses [J] Arch Virol, 153(12): 2273–2282.
- KREBS JE, GOLDSTEIN ES, KILPATRICK ST, 2018. GENE XII [M]. USA Massachusetts Burlington: Jones & Bartlett Learning.
- LAI RL, FENG X, CHEN J, et al., 2019. Codon usage bias of *Canarium album* (Lour.) R. transcriptome and its influence factors [J]. J Nucl Agric Sci, 33(1): 31–38 [赖瑞联, 冯新, 陈瑾, 等, 2019. 橄榄转录组密码子使用偏好性及其影响因素 [J]. 核农学报, 33(1): 31–38.]
- LIU QH, ZHANG CJ, CHENG YH, et al., 2016. Molecular cloning, characterization, and functional analysis of *CcBBM* gene from camphor tree (*Cinnamomum camphora* L.) [J]. Trees, 30: 1033–1043.
- LIU S, 2019. Population genetic structure of *Cinnamomum camphora* in China [D]. Nanchang: Jiangxi Agricultural University [刘爽, 2019. 中国樟树群体遗传结构研究 [D]. 南昌: 江西农业大学.]
- LU YS, PENG C, CHEN Z, et al., 2018. Analysis of codon usage bias in *Clausena lansium* through annotated transcriptome data [J]. Mol Plant Breed, 18(18): 5904–5913 [陆育生, 彭程, 陈喆, 等, 2018. 黄皮转录组密码子使用偏好性分析 [J]. 分子植物育种, 18(18): 5904–5913.]
- NELSON DL, COX MM, 2017. Lehninger principles of biochemistry [M]. New York: W. H. Freeman and Company.
- QIN Z, ZHENG YJ, GUI LJ, et al., 2018. Codon usage bias analysis of chloroplast genome of camphora tree [J]. Guihaia, 38(10): 1346–1355 [秦政, 郑永杰, 桂丽静, 等, 2018. 樟树叶叶绿体基因组密码子偏好性分析 [J]. 广西植物, 38(10): 1346–1355.]
- SHI XP, 2009. Enhancement of somatic embryogenesis and genetic transformation with *Barnase* and *PaFT* genes in *Cinnamomum camphora* L. [D]. Wuhan: Huazhong Agricultural University [施雪萍, 2009. 樟树体细胞胚再生体系的优化和转化 *Barnase*、*PaFT* 基因的研究 [D]. 武汉: 华中农业大学.]
- WANG CX, 2009. Study on *Cinnamomum camphora* Presl new germplasm with cold resistance through genetic engineering [D]. Tai'an: Shandong Agricultural University [王长宪, 2009. 利用基因工程创造香樟抗寒树种新种质技术研究 [D]. 泰安: 山东农业大学.]
- WANG GL, FANG HY, 2018. Plant genetic engineering [M]. Beijing: Science Press. [王美林, 方宏筠, 2014. 植物基因工程 [M]. 北京: 科学出版社.]
- WANG JF, HUANG TH, AN JC, et al., 2019. A comparative study on wood structure of timber and oil-used and timber-used *Cinnamomum camphora* [J]. J W China For Sci, 48(6): 15–20 [王军锋, 黄腾华, 安家成, 等, 2019. 材用和油材两用人工林樟树木材构造对比研究 [J]. 西部林业科学, 48(6): 15–20.]
- WANG K, AN JC, ZHU CS, et al., 2019. Relationship between leaf anatomical structure and drought resistance of different chemotypes oil-used camphor trees [J]. J S Agric, 50(11): 2525–2531. [王坤, 安家成, 朱昌叁, 等, 2019. 不同化学型油用樟树叶解剖结构特征及其抗旱特性 [J]. 南方农业学报, 50(11): 2525–2531.]
- WANG PL, WU SC, YANG LP, et al., 2019. Analysis of codon bias of chloroplast genome in *Eucalyptus grandis* [J]. Guihaia, 39(12): 1583–1592. [王鹏良, 吴双成, 杨利平, 等, 2019. 巨桉叶绿体基因组密码子偏好性分析 [J]. 广西植物, 39(12): 1583–1592.]
- WANG PL, YANG LP, WU HY, et al., 2018. Codon bias of chloroplast genome in *Camellia oleifera* [J]. Guihaia, 38(2): 135–144. [王鹏良, 杨利平, 吴红英, 等, 2018. 普通油茶叶绿体基因组密码子偏好性分析 [J]. 广西植物, 38(2): 135–144.]
- YOUNG R, PURTON S, 2016. Codon reassignment to facilitate genetic engineering and biocontainment in the chloroplast of *Chlamydomonas reinhardtii* [J]. Plant Biotechnol J, 14: 1251–1260.
- ZHENG H, YU MY, PU CJ, et al., 2020a. Cloning and expression analysis of 3-hydroxy-3-methylglutaryl coenzyme A reductase (*CcHMGR*) genes in *Cinnamomum camphora* (L.) Presl [J]. Acta Pharm Sin, 55(1): 152–159 [郑汉, 虞慕瑶, 濮春娟, 等, 2020a. 香樟 3-羟基-3-甲基戊二酰辅酶 A 还原酶 (*CcHMGRs*) 基因家族的克隆及表达分析 [J]. 药学学报, 55(1): 152–159.]
- ZHENG H, YU MY, PU CJ, et al., 2020b. Cloning and expression analysis of 5-phosphomevalonate kinase gene (*CcPMK*) in *Cinnamomum camphora* [J]. Chin J Chin Mat Med, 45(1): 78–84. [郑汉, 虞慕瑶, 濮春娟, 等, 2020b. 香樟甲羟戊酸-5-磷酸激酶基因 *CcPMK* 的克隆和表达分析 [J]. 中国中药杂志, 45(1): 78–84.]
- ZHENG YJ, WU YF, LI J, et al., 2018. Identification of NBS-LRR-like disease-resistant genes in *Cinnamomum camphora* and clone of two *CcRNL* genes [J]. Biotechnol Bull, 34(2): 142–149 [郑永杰, 伍艳芳, 李江, 等, 2018. 樟树 NBS-LRR 类抗病基因家族分析与 *CcRNL* 基因克隆 [J]. 生物技术通报, 34(2): 142–149.]
- ZHOU ZP, DANG YK, ZHOU M, et al., 2016. Codon usage is an important determinant of gene expression levels largely through its effects on transcription [J]. Proc Natl Acad Sci USA, 26: e6117–e6125.
- ZHU SG, XU CF, 2016. Biochemistry [M]. 4th ed. Beijing: Science Press. [朱圣庚, 徐长发, 2016. 生物化学 [M]. 4 版. 北京: 高等教育出版社.]