

小粒咖啡铁皮卡叶绿体基因组密码子偏好性分析

李亚麒¹, 黄家雄¹, 娄予强¹, 付兴飞¹, 王健敏², 李亚男¹,
吕玉兰¹, 李贵平¹, 程金焕^{1*}

(1. 云南省农业科学院 热带亚热带经济作物研究所, 云南 保山 678000; 2. 云南省农村科技服务中心, 云南 昆明 650021)

摘要:为了解铁皮卡叶绿体基因组密码子的使用特征及其成因,以筛选的 52 条蛋白编码序列为研究对象,利用 CodonW 和 CUSP 在线软件对其密码子使用特征进行系统分析。结果表明,基因组各位置的 GC 含量 GC₁、GC₂、GC₃ 分别为 47.47%、39.52%、27.92%,30 个高频密码子中,以 U 结尾占比 53.33%,以 A 结尾占比 43.33%,以 G 结尾占比 3.33%,说明铁皮卡叶绿体基因偏好使用 NNA 和 NNU 型密码子,尤其偏好使用 NNU 型密码子。有效密码子数(ENC)、密码子适应指数(CAI)、最优密码子频率(Fop)分别为 46.85、0.167、0.352,均揭示该基因组的密码子偏性较弱。中性绘图、ENC-plot、PR2-plot 以及对应分析揭示密码子的偏好性受到选择及其他因素共同作用,最终确定了 AUU、GUU 和 UCU 等 20 个最优密码子。表明铁皮卡叶绿体基因组密码子的偏性较弱,偏好 AU 结尾的密码子。密码子的使用偏好性受到突变及选择等多种作用共同影响。

关键词:小粒咖啡;铁皮卡;叶绿体基因组;密码子偏好性;最优密码子

中图分类号:S571.2

文献标志码:A

文章编号:1001-7461(2023)02-0092-08

Codon Usage Bias of the Chloroplast Genome in *Coffea arabica* 'Typica'

LI Ya-qi¹, HUANG Jia-xiong¹, LOU Yu-qiang¹, FU Xing-fei¹, WANG Jian-min², LI Ya-nan¹,
LÜ Yu-lan¹, LI Gui-ping¹, CHENG Jin-huan^{1*}

(1. Institute of Tropical and Subtropical Cash Crops, Yunnan Academy of Agricultural Sciences, Baoshan 678000, Yunnan, China;

2. Yunnan Rural Science and Technology Service Center, Kunming 650021, Yunnan, China)

Abstract: To dissect the codon usage bias of the chloroplast genome codons and contributing factors, 52 protein coding gene sequences from the chloroplast genome of *Coffea arabica* 'Typica' were studied. CodonW and CUSP online software were used to analyze the characteristics of its codon usage. It was found out that the GC₁, GC₂ and GC₃ contents were 47.47%, 39.52% and 27.92%, respectively, among the 30 high-frequency codons, 53.33% ended with U, 43.33% ended with A, and 3.33% ended with G, indicating that the chloroplast genes of *C. arabica* 'Typica' preferred to NNA and NNU codons, especially NNU codons. The effective codon number (ECN), codon adaptation index (CAI) and optimal codon frequency (Fop) were 46.85, 0.167 and 0.352, respectively, suggesting that the codon bias of the genome was weak. The Neutral plot analysis, ENC-plot analysis, PR2-plot analysis and correspondence analysis showed that codon usage bias was influenced by selection and other factors. Finally, 20 codons, including AUU, GUU and UCU, were identified as the optimal codons. In summary, the codon usage bias of *C. arabica* 'Typica' is weak, and prefers to the AU ending codons. Multiple factors contribute to the codon usage bias, such as selection and mutation.

Key words: *Coffea arabica*; "Typica"; chloroplast genome; codon bias; optimal codon

收稿日期:2022-03-11 修回日期:2022-09-20

基金项目:国家重点研发计划课题(2020YFD1001202-2);云南省科技厅创新引导与科技型企业培养计划(202104BI090015);国家咖啡标准化区域服务与推广平台(NBFW-17-2021);云南省重点研发计划课题(2018ZG015)。

第一作者:李亚麒。研究方向:咖啡栽培与利用。E-mail:1640949082@qq.com

* 通信作者:程金焕,副研究员。研究方向:咖啡加工。E-mail:jinhuancheng-2006@163.com

编码同一种氨基酸的所有密码子被定义为同义密码子,而这种同义密码子使用不均衡的现象称为密码子偏好性^[1]。生物界中普遍存在密码子使用偏好性,在原核生物、动物和植物中都得到了验证。起初学者们对密码子偏好性的研究主要集中在低等真核生物与原核生物中,后来逐步转移到高等植物及动物上^[2]。密码子偏好性的形成是多因素综合作用的结果,早期针对密码子偏好性的解释,主要有“选择-突变-漂移”和“中性理论”假说^[3-4]。随着人们对密码子使用偏好性研究的不断深入,发现密码子使用偏好性还与基因表达水平、tRNA 丰度和氨基酸序列组成等众多因素有关^[5-6]。但在这些因素中,普遍认为选择与突变起着关键作用,而两者中何者又占据主导是当前比较关注的点。研究表明,在不同物种中两者的主导作用不一致,如在籽粒苋(*Amaranthus hypochondriacus*)^[7]、乳油木(*Vitellaria paradoxa*)^[8]等一些物种中发现突变是主导因素。而在芒果(*Mangifera indica*)^[9]、滇楸(*Catalpa fargesii* Bur. f. *duclouxii*)^[10]等一些物种中自然选择起着主导作用。密码子使用偏好性能够反映物种或基因的起源,对基因表达、预测基因功能、基因家族分化等问题的研究意义重大^[11]。

咖啡,为茜草科咖啡属一种园艺性多年生的经济作物,与茶、可可并称为世界三大饮料^[12]。咖啡有大粒咖啡(*Coffea liberica*)、中粒咖啡(*Coffea robusta*)、小粒咖啡(*Coffea arabica*)之分,其中,小粒咖啡是世界性的主要栽培种,我国除海南以外以小粒种为主。铁皮卡(Typica)是最经典的优质小粒咖啡种,同时也是小粒种咖啡最重要的品种之一,很多改良种都源自此种^[13-14]。目前,有关铁皮卡的研究主要集中于病虫害^[15]、化学物质^[16]等层面。叶绿体基因组虽已测序完成^[17],但未见其密码子用法方面的报道。由于基因组庞大和结构复杂,植物基因组学研究面临着挑战。但随着新一代高通量测序技术的产生和发展为非模式植物基因组研究带来了新的机遇和契机。作为研究生命现象本质的科学,植物基因组学研究一直是植物生物学研究的热点和重点。与核基因组相比叶绿体基因组密码子使用模式的研究相对落后些^[18],近年来,关于叶绿体基因组密码子偏好性已在茶树(*Camellia sinensis*)^[5]、乳油木^[8]、滇楸^[10]等多种植物中开展。然而,铁皮卡叶绿体基因组学相关知识尤其是密码子使用偏好性仍不清楚。鉴于以上情况,本研究系统地研究了铁皮卡叶绿体基因组密码子的使用模式,以期铁皮卡基因进化研究及遗传育种提供理论依据。

1 材料与方法

1.1 材料

从 GenBank 下载完整的小粒咖啡铁皮卡叶绿体基因组(登录号:MK862266)。在 86 条 CDS (coding DNA sequence)中,选择长度>300 bp,以 ATG 开始,TAG、TAA、TGA 结尾,内部不存在终止密码子及重复序列的 52 条 CDS 进行分析。

1.2 方法

1.2.1 密码子组成分析 采用 Codon W 程序计算小粒咖啡铁皮卡叶绿体基因偏好性相关参数,包括有效密码子数(ENC)、同义密码子使用度(RSCU)、最优密码子频率(Fop)、密码子适应指数(CAI)、第 3 位密码子各碱基含量(T_{3s} 、 A_{3s} 、 G_{3s} 、 C_{3s})、总 GC 含量以及密码子第 1、第 2 和第 3 位上的含量(GC_1 、 GC_2 、 GC_3)及同义密码子第 3 位碱基的 GC 含量(GC_{3s})、 GC_{12} (GC_1 、 GC_2 的平均值)、芳香族氨基酸比例(Aromo)、基因编码序列长度(L)。并对这些指标进行相关分析,剖析彼此之间的相互反馈程度。CAI 介于 0~1,值越接近 0,表示偏好性越弱,和基因表达水平呈正相关^[19]。ENC 介于 20~61,值越大偏好性越弱。若基因受突变作用,则基因点分布在标准曲线上^[20]。Fop 值在 0~1 内,值越接近 1,偏好性越强^[21]。RSCU = 1,密码子使用无偏好性,RSCU > 1 表示密码子使用频率较高,RSCU > 1.6 表示密码子偏好性较强^[22-23]。

1.2.2 密码子偏好性形成因素分析 通过中性绘图、ENC-Plot、偏移分析(PR2-plot)确定密码子偏好性形成因素。其中,对于中性绘图分析而言,若 GC_{12} 与 GC_3 间相关关系达到显著水平,则突变影响密码子偏好性,反之选择影响其偏好性^[23]。ENC-Plot:计算理论 ENC 值,绘制标准曲线,若基因受突变作用,则基因点分布在标准曲线上^[10]。PR2-plot:图中中心点表示 $A = T$ 且 $C = G$,其余点代表该基因的偏倚程度^[8]。

理论 ENC 值(E_{NC})计算公式:

$$E_{NC} = 2 + G_{C3} + \frac{29}{G_{C3}^2 + (1 - G_{C3})^2} \quad (1)$$

1.2.3 最优密码子确定 根据 ENC 值进行大小排序,选取两端 10% 的基因建立高低偏好性库,同时满足高频密码子($RSCU \geq 1$)与高表达密码子($\Delta RSCU \geq 0.08$)确定为最优密码子^[9]。

2 结果与分析

2.1 铁皮卡密码子偏好性分析

由图 1 可知,铁皮卡叶绿体基因组中,密码子第

3 位 A、T、C、G 核苷酸碱基含量分别为 43.82%、47.25%、15.77%、16.23%。所有 CDS 密码子的 GC 含量为 38.31%，GC₁、GC₂、GC₃ 含量分别为 47.47%、39.52%、27.92%，可见密码子第 3 位的

GC 含量明显比第 1、第 2 位要低。因此，铁皮卡叶绿体基因倾向于 AU 结尾的密码子。ENC、CAI、FOP 平均值分别为 46.85、0.167、0.352，均揭示铁皮卡叶绿体基因密码子偏好性不强。

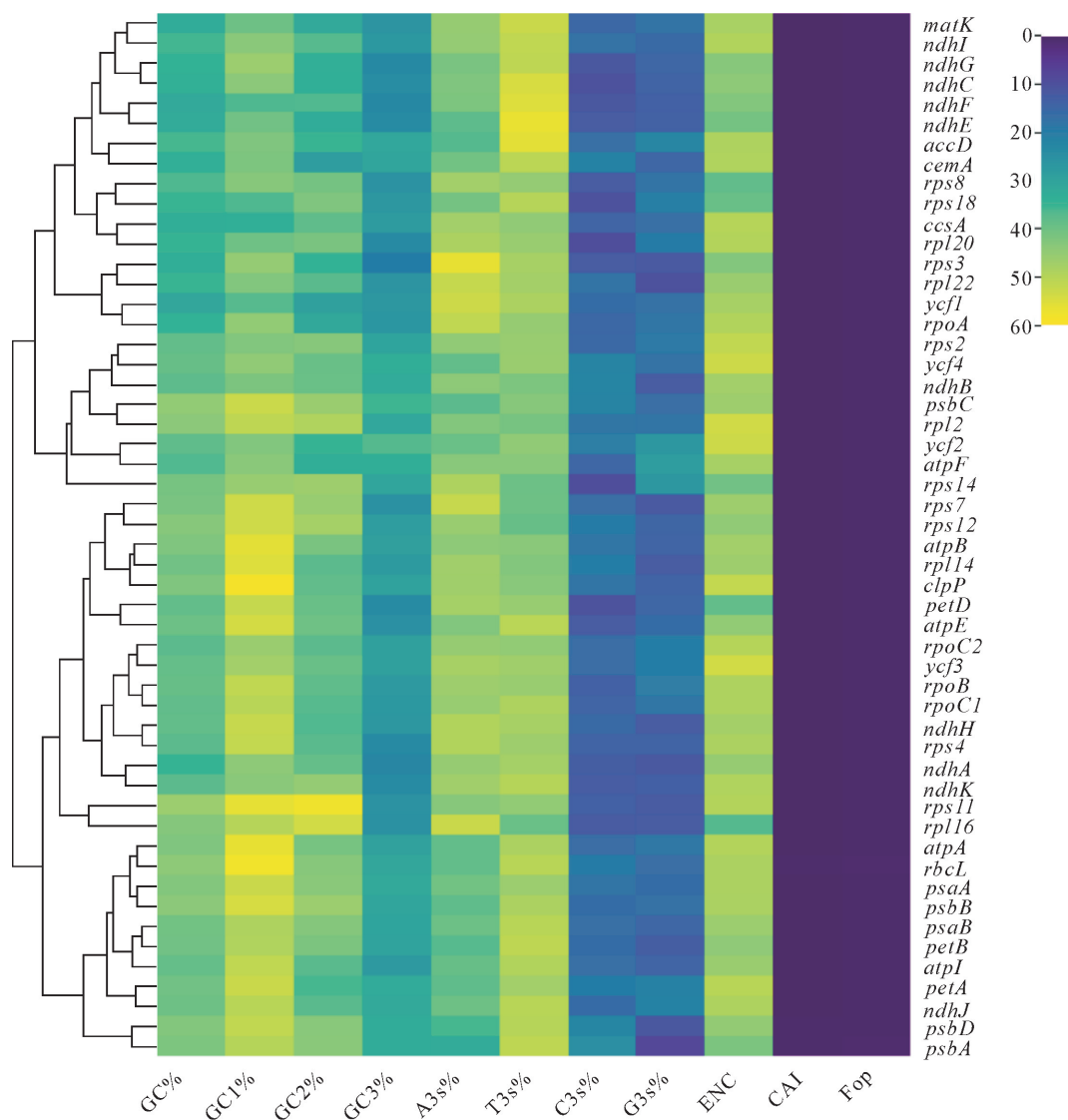


图 1 小粒咖啡铁皮卡密码子偏好性相关参数

Fig. 1 Codon usage bias parameters of *Coffea arabica* 'Typica'

相关性分析显示(表 1),GC₁、GC₂、GC₃ 与 GC 具有极显著的相关关系,GC₁、GC₂ 极显著相关,GC₃ 与 GC₁、GC₂ 无明显的相关关系,说明密码子第 1、第 2 位碱基组成较为相似,与第 3 位碱基具有差异,存在明显的位置效应。ENC 与 GC₃、C_{3s}、G_{3s}、基因长度呈显著或极显著的相关关系,与 GC₁、GC₂ 无显著的相关关系,说明第 3 位碱基对密码子偏好性的影响更显著,且密码子偏好性受到基因长度影响。CAI 与 GC、GC₁、GC₂、C_{3s}、T_{3s}、Fop、Aromo 呈显著或极显著的正相关关系,与 A_{3s}、G_{3s} 具有极显著的负相关关系。Fop 与 C_{3s}、GC₁、GC、GC₃、GC₂、CAI 呈显著或极显著的正相关关系,与 A_{3s}、G_{3s} 具有极显著的负相关关系。说明基因的表达水平和密码子使用

偏好性程度的大小,受到核苷酸组成的影响。

对铁皮卡基因组同义密码子相对使用频率进行分析。由图 2 可知,UUC、CUC、CUG 等 32 个密码子的使用偏好性较弱,出现频率较低(RSCU 值 < 1)。UUA、UCU、UAU、GAU、AGA、GCU 6 个密码子的 RSCU 值 > 1.6,表明这 6 个密码子有着较强偏好性。AUG、UGG 2 个密码子的 RSCU = 1,说明这 2 个密码子无偏好性。UUU、UUG、CUU 等 30 个高频密码子中,AU 结尾占比 96.67%,CG 结尾占比 3.33%。其中,以 U 结尾占比 53.33%,以 A 结尾占比 43.33%,以 G 结尾占比 3.33%,再次揭示铁皮卡叶绿体基因组偏爱 AU 结尾的密码子,尤其偏爱以 U 结尾。

表 1 小粒咖啡铁皮卡密码子偏好性参数相关分析

Table 1 Correlation coefficients between codon usage bias parameters of *C. arabica* ‘Typica’

因子	GC	GC ₁	GC ₂	GC ₃	A _{3s}	T _{3s}	C _{3s}	G _{3s}	ENC	CAI	Fop	Length	Aromo
GC	1												
GC ₁	0.830 **	1											
GC ₂	0.787 **	0.452 **	1										
GC ₃	0.483 **	0.191	0.121	1									
A _{3s}	-0.307 *	-0.158	-0.067	-0.582 **	1								
T _{3s}	-0.470 **	-0.301 *	-0.481 **	-0.210	-0.451 **	1							
C _{3s}	0.402 **	0.295 *	0.022	0.726 **	-0.427 **	-0.138	1						
G _{3s}	-0.067	-0.209	-0.181	0.434 **	-0.095	-0.132	-0.163	1					
ENC	0.121	0.098	-0.127	0.418 **	-0.076	-0.097	0.413 **	0.305 *	1				
CAI	0.368 **	0.436 **	0.021	0.379 **	-0.534 **	0.325 *	0.603 **	-0.367 **	0.078	1			
Fop	0.525 **	0.459 **	0.302 *	0.388 **	-0.408 **	-0.017	0.619 **	-0.376 **	-0.102	0.763 **	1		
Length	-0.117	-0.160	-0.252	0.307 *	-0.045	0.016	0.213	0.307 *	0.301 *	0.059	-0.017	1	
Aromo	-0.284 *	-0.325 *	-0.332 *	0.189	-0.350 *	0.498 **	0.166	-0.110	0.084	0.285 *	-0.029	0.176	1

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

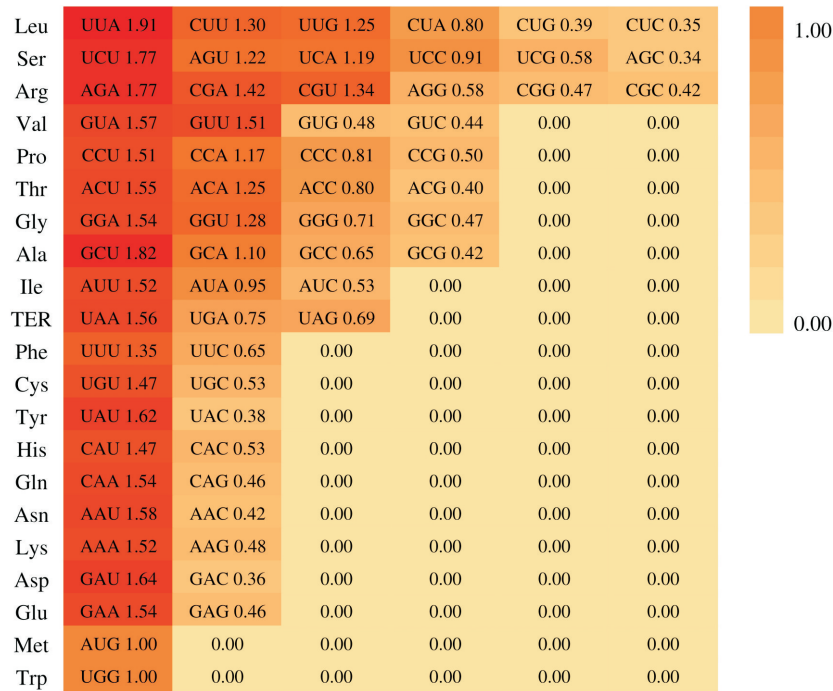


图 2 小粒咖啡铁皮卡叶绿体基因同义密码子相对使用度

Fig. 2 Usage of synonymous codons of the chloroplast genes in *C. arabica* ‘Typica’

2.2 中性绘图分析

中性分析主要揭示 GC₃ 和 GC₁₂ 之间的相互关系,能够反映铁皮卡叶绿体基因组密码子使用模式中自然选择和突变压力之间的均衡^[24]。由图 3 可知,该基因组中 GC₃ 含量的分布范围为 19.63%~36.68%,GC₁₂ 含量的分布范围为 33.01%~56.84%。GC₁₂ 和 GC₃ 间无明显的相关关系($r=0.185,P>0.05$)。表明密码子 3 个位置上的碱基组成不同,GC 含量水平高度保守,突变对密码子偏好性的影响较弱,而其他因素,如自然选择可能对密码子的偏好性影响较大。

2.3 ENC 绘图分析

从图 4 可以看出,大部分基因未落在期望曲线上,表明实际 ENC 值与预期 ENC 值差异较大。对 ENC 比值频率进行分析,由图 5 可知,在-0.05~0.05 内少数的 20 个基因与预期 ENC 值较为接近,说明这部分基因受突变的影响作用较大。在该区间范围外大部分的 32 个基因与预期值相差较大,说明这部分基因受自然选择的影响作用较大。可见,铁皮卡绝大部分基因的密码子偏好性受到选择作用的影响,少部分基因受突变作用的影响。

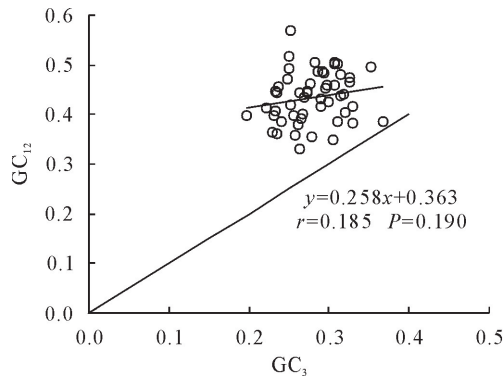


图3 小粒咖啡铁皮卡叶绿体基因组中性绘图分析
Fig. 3 Neutrality plot analysis in the chloroplast genes of *C. arabica* 'Typica'

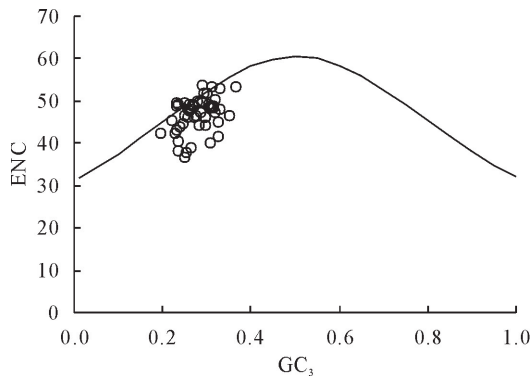


图4 小粒咖啡铁皮卡叶绿体基因组 ENC-plot 绘图分析
Fig. 4 ENC-plot analysis in the chloroplast genes of *C. arabica* 'Typica'

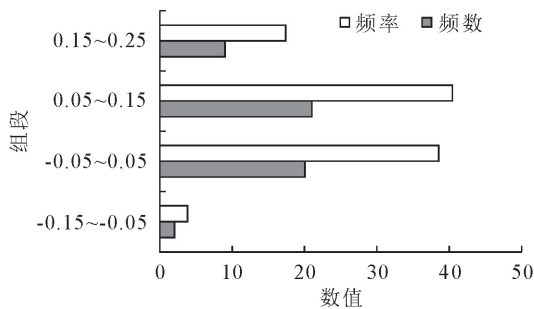


图5 小粒咖啡铁皮卡基因组有效密码子数比值频率
Fig. 5 Ratio frequency of ENC in the chloroplast genes of *C. arabica* 'Typica'

2.4 PR2 奇偶偏好分析

通过 PR2-plot 绘图探究 4 个密码子编码的氨基酸家族中第 3 位上嘌呤和嘧啶的分布情况 (图 6)。结果表明,第一、二、三、四象限分别有 12、10、13、17 个基因,分布不均匀,说明碱基 T 的使用频率高于 A,碱基 G 的使用频率高于 C。由此推测,铁皮卡叶绿体基因组密码子偏好性除了受到突变作用,可能还受到选择等其他因素作用。

2.5 最优密码子的确定

铁皮卡叶绿体基因组最优密码子如表 2 所示,

由表 2 可知, $\Delta\text{RSCU} \geq 0.08$ 的高表达密码子有 27 个,结合 30 个高频密码子,最终确定 UUU、UUA、UUG、UAA、UAU、CUU、AUU、UCA、CAA、GUU、UCU、ACU、ACA、UGU、CGU、GGU、GCU、AAA、GCA 和 AAU 20 个最优密码子,以 U、A、G 结尾分别有 12、7、1 个。

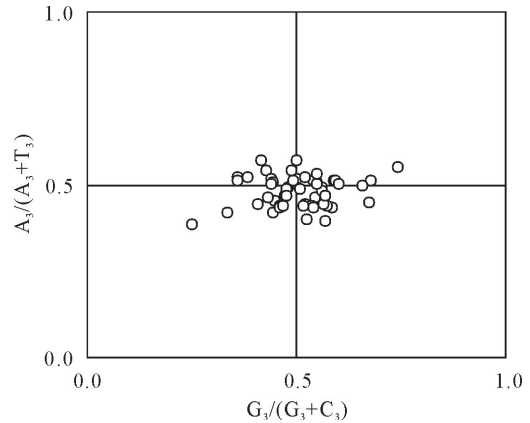


图6 小粒咖啡铁皮卡密码子奇偶偏好性分析
Fig. 6 Analysis of PR2-plot of codon usage of *C. arabica* 'Typica'

2.6 对应分析

对应分析表明 (图 7), 4 个向量轴分别解释了 18.12%、10.70%、6.33%、5.65% 的基因差异,累计解释率为 40.8%。4 个轴对密码子偏好性均有不同程度影响,其中,第 1 轴、第 2 轴对密码子偏好的影响较大,说明密码子偏好性受到多种因素作用。进一步进行相关性分析,发现 CAI 与 1 轴之间具有极显著负相关关系 ($r = -0.662, P < 0.00$), 2、4 轴与所分析因子无相关性,第 3 轴与 GC_3 ($r = -0.779, P < 0.00$)、ENC ($r = -0.478, P < 0.01$)、CAI ($r = -0.444, P < 0.00$) 间具有极显著的负相关关系。表明 CAI 是影响图 4 中轴 1 的关键因素, ENC、CAI、 GC_3 是影响轴 3 的主要因素,推断出基因表达水平、GC 含量对铁皮卡密码子偏好性存在影响。从图 4 中基因的分布情况可看出,核糖体基因、光合系统基因、遗传基因大致呈集中分布,表明这些基因的密码子使用模式较为接近。而其他蛋白基因、保守基因等分布较分散,表明此类基因的密码子偏好性是由多种因素引起的。综上,铁皮卡叶绿体基因组的密码子偏好性是多种因素共同作用的结果。

3 结论与讨论

3.1 结论

本研究基于铁皮卡叶绿体基因组开展了密码子组成、中性绘图、PR2-plot、ENC-plot 绘图等分析比

表 2 小粒咖啡铁皮卡基因组最优密码子
Table 2 Optimal codons in the genome of *C. arabica* ‘Typica’

氨基酸	密码子	高表达基因		低表达基因		Δ RSCU
		数目	RSCU	数目	RSCU	
Phe	<u>UUU</u> ***	21	8.60	97	7.00	1.60
Leu	<u>UUA</u> ***	20	8.07	55	6.90	1.17
	<u>UUG</u> ***	19	7.97	84	7.14	0.83
	<u>CUU</u> ***	13	9.68	80	6.96	2.72
Ile	<u>AUU</u> ***	27	8.42	108	6.04	2.38
Val	<u>GUU</u> ***	15	11.85	42	5.28	6.57
	<u>GUA</u>	15	5.69	44	8.76	-3.07
Ser	<u>UCU</u> ***	11	8.81	78	6.47	2.34
	<u>UCA</u> *	6	6.75	58	6.62	0.13
	<u>AGU</u>	7	4.72	51	7.44	-2.72
Pro	<u>CCU</u>	9	4.31	48	7.82	-3.51
	<u>CCA</u>	10	4.31	39	5.91	-1.60
Thr	<u>ACU</u> ***	13	6.53	40	5.94	0.59
	<u>ACA</u> ***	13	7.86	42	6.05	1.81
Cys	<u>UGU</u> ***	5	8.00	23	5.31	2.69
TER	<u>UAA</u> ***	4	12.00	2	6.00	6.00
Tyr	<u>UAU</u> ***	14	10.00	92	7.90	2.10
His	<u>CAU</u>	10	8.00	57	8.07	-0.07
Gln	<u>CAA</u> ***	12	7.47	83	6.62	0.85
Asn	<u>AAU</u> ***	21	8.30	139	7.52	0.78
Lys	<u>AAA</u> ***	36	8.69	117	6.94	1.75
Asp	<u>GAU</u>	9	4.60	136	8.45	-3.85
Glu	<u>GAA</u>	21	7.10	119	7.33	-0.23
Arg	<u>CGU</u> ***	14	8.19	34	6.61	1.58
	<u>CGA</u>	24	7.66	54	9.66	-2.00
	<u>AGA</u>	25	9.50	64	9.06	0.44
Gly	<u>GGU</u> ***	22	8.57	43	5.12	3.45
	<u>GGA</u>	16	8.08	75	8.95	-0.87
Ala	<u>GCU</u> ***	15	10.21	45	7.34	2.87
	<u>GCA</u> ***	15	6.34	27	4.91	1.43

注:*. Δ RSCU ≥ 0.08 , *. *. Δ RSCU ≥ 0.3 , *. *. *. Δ RSCU ≥ 0.5 ,下划线表示高频密码子。

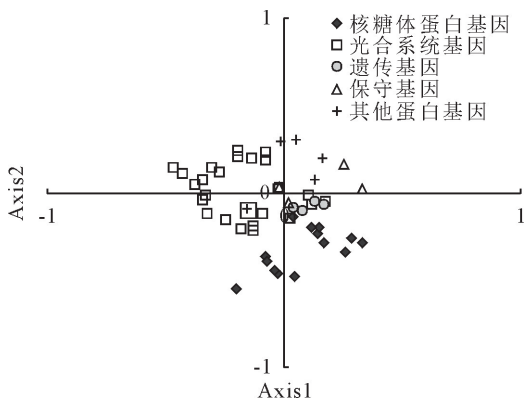


图 7 小粒咖啡铁皮卡叶绿体基因对应分析

Fig. 7 Chloroplast genes correspondence analysis of *C. arabica* ‘Typica’ genome

较,发现铁皮卡叶绿体基因偏好使用 NNA 和 NNU 型密码子,尤其偏好使用 NNU 型密码子;密码子的偏好性受到选择和其他因素共同影响;最终确定铁皮卡叶绿体基因组最优密码子 20 个,以 U、A、G 结

尾分别有 12、7、1 个。该研究结果为今后进一步开展铁皮卡基因的进化研究和遗传育种有重要指导作用。

3.2 讨论

基因密码子的使用偏好性体现了其在特定自然选择压力及突变压力下的进化方式,是物种对基因组环境和自然进化压力的适应性选择^[25]。本研究对铁皮卡叶绿体基因的密码子使用偏好性进行分析,发现该基因中密码子第 3 位的 A、T、C、G 含量分别为 43.82%、47.25%、15.77%、16.23%。GC₁、GC₂、GC₃ 含量分别为 47.47%、39.52%、27.92%,说明铁皮卡叶绿体基因倾向于 AU 结尾的密码子,即偏好使用 NNA 和 NNU 型密码子。屈亚亚等^[23]的研究结果表明香花枇杷(*Eriobotrya fragrans*)叶绿体基因组偏爱 AU 结尾的密码子,与本研究结果类似。原晓龙等^[26]的西南桦(*Betula alnoides*)叶

绿体基因组密码子偏好性结果也支持这一结论。先前研究表明,单子叶植物具有以 GC 结尾的偏好性,双子叶植物具有以 AU 结尾的偏好性^[27]。如杨国锋等^[18]研究发现蒺藜苜蓿(*Medicago truncatula*)叶绿体基因组密码子偏好以 AU 结尾。原晓龙等^[28]研究也发现云南蓝果树(*Nyssa yunnanensis*)叶绿体基因组偏好使用 AU 结尾的密码子。本研究显示铁皮卡叶绿体基因倾向于 AU 结尾的密码子,符合大部分双子叶植物密码子的使用特征,这类植物可能具有高度相似的密码子使用模式^[29]。本研究还发现,52 条 CDS 基因的 ENC 平均值为 46.85,均>35。CAI、FOP 值分别为 0.167、0.352,均<1。说明该基因的密码子偏好性较弱,表达水平不高。在芒果^[9]、柿(*Diospyros kaki*)中^[30]也存在类似情况。

密码子使用偏好性的形成并非只受单一因素作用,选择与突变是 2 个较为关键的因素。陆奇丰等^[31]研究揭示,梧桐(*Firmiana simplex*)基因组密码子使用偏好性受选择与突变共同作用。冯瑞云等^[7]研究表明,籽粒苋密码子偏好性主要受到突变的影响。牛元等^[20]研究表明,大花香水月季(*Rosa odorata* var. *gigantea*)密码子使用模式受选择作用的影响较突变作用要显著。在本文中,通过 ENC 绘图、中性绘图、PR2 奇偶偏移分析表明,铁皮卡密码子偏好性的形成同样受到选择与突变等的共同作用,但受选择的影响更多。同时基因长度也对密码子偏好性产生了影响,与原晓龙等^[28]研究结果类似,但与梁晓静等^[32]研究结果存在一定差异,说明不同植物的密码子偏好的影响因素有所不同。对应分析表明,铁皮卡叶绿体基因密码子使用偏好性受到选择、突变、基因表达水平等多种因子的共同作用。遗传基因、核糖体基因、光合系统基因整体上分布相对集中,有着相似的密码子使用偏好性,与李江飞等^[10]研究结果类似。

参考文献:

- [1] LIU Q P, DOU S J, JI Z J, *et al.* Synonymous codon usage and gene function are strongly related in *Oryza sativa* [J]. *Biosystems*, 2005, 80(2): 123-131.
- [2] 吴宪明, 吴松峰, 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.
WU X M, WU S F, REN D M, *et al.* The analysis method and progress in the study of codon bias [J]. *Hereditas*, 2007, 29(4): 420-426. (in Chinese)
- [3] ARIS-BROSOU S, BIELAWSKI J P. Large-scale analyses of synonymous substitution rates can be sensitive to assumptions about the process of mutation [J]. *Gene*, 2006, 378(1): 58-64.
- [4] BERNARDI G, BERNARDI G. Compositional constraints and genome evolution [J]. *Journal of Molecular Evolution*, 1986, 24(1-2): 1-11.
- [5] 周承哲, 朱晨, 李小桢, 等. 茶树密码子使用偏性分析方法及其研究进展[J]. *分子植物育种*, 2020, 18(5): 1480-1488.
ZHOU C Z, ZHU C, LI X Z, *et al.* Codon usage bias analysis methods of *Camellia sinensis* and its research progress [J]. *Molecular Plant Breeding*, 2020, 18(5): 1480-1488. (in Chinese)
- [6] 张晓峰, 薛庆中. 水稻和拟南芥 NBS-LRR 基因家族同义密码子使用偏好的比较[J]. *作物学报*, 2005(5): 596-602.
ZHANG X F, XUE Q Z. Synonymous codon bias of NBS-LRR gene family in rice and *Arabidopsis* [J]. *Acta Agronomica Sinica*, 2005(5): 596-602. (in Chinese)
- [7] 冯瑞云, 梅超, 王慧杰, 等. 籽粒苋叶绿体基因组密码子偏好性分析[J]. *中国草地学报*, 2019, 41(4): 8-15.
FENG R Y, MEI C, WANG H J, *et al.* Analysis of codon usage in the chloroplast genome of *Grain Amaranth* (*Amaranthus hypochondriacus* L.) [J]. *Chinese Journal of Grassland*, 2019, 41(4): 8-15. (in Chinese)
- [8] 原晓龙, 李云琴, 张劲峰, 等. 乳油木叶绿体基因组密码子偏好性分析[J]. *分子植物育种*, 2020, 18(17): 5658-5664.
YUAN X L, LI Y Q, ZHANG J F, *et al.* Codon usage bias analysis of chloroplast genome in *Vitellaria paradoxa* [J]. *Molecular Plant Breeding*, 2020, 18(17): 5658-5664. (in Chinese)
- [9] 辛雅萱, 黎若竹, 李鑫, 等. 芒果叶绿体基因组密码子使用偏好性分析[J]. *中南林业科技大学学报*, 2021, 41(9): 148-156, 165.
XIN Y X, LI R Z, LI X, *et al.* Analysis on codon usage bias of chloroplast genome in *Mangifera indica* [J]. *Journal of Central South University of Forestry & Technology*, 2021, 41(9): 148-156, 165. (in Chinese)
- [10] 李江飞, 李熙颜, 王瑜, 等. 滇楸叶绿体基因组密码子偏好性分析[J]. *基因组学与应用生物学*, 2021, 41(4): 843-853.
LI J F, LI X Y, WANG Y, *et al.* Analysis on codon usage bias of chloroplast genome in *Catalpa fargesii* Bur. f. *duclouxii* [J]. *Genomics and Applied Biology*, 2021, 41(4): 843-853. (in Chinese)
- [11] 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.) *Raesch* [J]. *Chinese Journal of Applied & Environmental Biology*, 2017, 23(5): 945-951. (in Chinese)
- [12] 吴建, 蒋快乐, 时玲, 等. 云南小粒种咖啡初加工工艺发展现状及趋势[J]. *中国农机化学报*, 2021, 42(11): 205-213.
WU J, JIANG K L, SHI L, *et al.* Development status and trend of Yunnan *Arabica* coffee primary processing [J]. *Journal of Chinese Agricultural Mechanization*, 2021, 42(11): 205-213. (in Chinese)
- [13] 黄家雄, 李贵平, 杨世贵. 咖啡种类及优良品种简介[J]. *农村实用技术*, 2009(1): 42-43.
HUANG J X, LI G P, YANG S G. Introduction of coffee types and superior varieties [J]. *Practical Rural Technology*, 2009(1): 42-43. (in Chinese)
- [14] 梁伟红, 李玉萍, 刘燕群, 等. 我国咖啡产业发展现状及对策研究[J]. *福建农业科技*, 2014(6): 71-76.

- LIANG W H, LI Y P, LIU Y Q, *et al.* Research on present development situation and countermeasures of coffee industry [J]. Fujian Agricultural Science and Technology, 2014(6): 71-76. (in Chinese)
- [15] 付兴飞, 李贵平, 黄家雄, 等. 云南省3个咖啡产区小粒咖啡病虫害危害调查分析[J]. 热带农业科学, 2020, 40(3): 67-75.
- FU X F, LI G P, HUANG J X, *et al.* Investigation and analysis on the diseases and insect pest of *Coffea arabica* in three coffee producing areas of Yunnan Province [J]. Chinese Journal of Tropical Agriculture, 2020, 40(3): 67-75. (in Chinese)
- [16] 刘兴勇, 陈兴连, 林涛, 等. 烘焙程度对小粒咖啡5-羟甲基糠醛生成的影响[J]. 中国食品学报, 2022, 22(1): 324-331.
- LIU X Y, CHEN X L, LIN T, *et al.* Effects of roasting degree on formation of 5-hydroxymethylfurfural contents in *Coffea arabica* [J]. Journal of Chinese Institute of Food Science and Technology, 2022, 22(1): 324-331. (in Chinese)
- [17] MIN J, KIM Y, XI H, *et al.* The complete chloroplast genome of coffee tree, *Coffea arabica* L. 'Typica' (Rubiaceae) [J]. Mitochondrial DNA Part B, 2019, 4(2): 2240-2241.
- [18] 杨国锋, 苏昆龙, 赵怡然, 等. 蕹菜苜蓿叶绿体基因组密码子偏好性分析[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. (in Chinese)
- [19] 孙继峰, 韩太利, 徐立功, 等. 萝卜ICE1密码子使用偏性分析[J]. 核农学报, 2018, 32(3): 477-485.
- SUN J F, HAN T L, XU L G, *et al.* Analysis of codon bias of the ICE1 in *Radish* [J]. Journal of Nuclear Agricultural Sciences, 2018, 32(3): 477-485. (in Chinese)
- [20] 牛元, 徐琼, 王尚德, 等. 大花香水月季叶绿体基因组密码子使用偏性分析[J]. 西北林学院学报, 2018, 33(3): 123-130.
- NIU Y, XU Q, WANG Y D, *et al.* An analysis on codon usage bias of chloroplast genome of *Rosa odorata* var. *gigantea* [J]. Journal of Northwest Forestry University, 2018, 33(3): 123-130. (in Chinese)
- [21] STENICO M, LLOYD A T, SHARP P M. Codon usage in *Caenorhabditis elegans*: delineation of translational selection and mutational biases [J]. Nucleic Acids Research, 1994, 22(13): 2437-2466.
- [22] 吴彦庆, 李志远, 赵大球, 等. 几种植物花分生组织决定基因APETALA1密码子使用模式比较[J]. 园艺学报, 2017, 44(3): 504-514.
- WU Y Q, LI Z Y, ZHAO D Q, *et al.* Comparative analysis of codon usage patterns in APETALA1 gene across several plant species [J]. Acta Horticulturae Sinica, 2017, 44(3): 504-514. (in Chinese)
- [23] 屈亚亚, 辛静, 冯发玉, 等. 香花枇杷质体基因组序列密码子偏好性分析[J]. 西北林学院学报, 2021, 36(4): 138-144, 158.
- QU Y Y, XIN J, FENG F Y, *et al.* Codon usage bias in chloroplast genome of *Eriobotrya fragrans* Champ. ex Benth [J]. Journal of Northwest Forestry University, 2021, 36(4): 138-144, 158. (in Chinese)
- [24] 李秀璋, 宋辉, 张宗豪, 等. 甘肃内生真菌基因组密码子使用的偏好性分析[J]. 草业学报, 2020, 29(5): 67-77.
- LI X Z, SONG H, ZHANG Z H, *et al.* Analysis of codon usage bias in the genome of *Epichloë gansuensis* [J]. Acta Prataculturae Sinica, 2020, 29(5): 67-77. (in Chinese)
- [25] 赖瑞联, 林玉玲, 钟春水, 等. 龙眼生长素受体基因TIR1密码子偏好性分析[J]. 园艺学报, 2016, 43(4): 771-780.
- LAI R L, LIN Y L, ZHONG C S, *et al.* Analysis of codon bias of auxin receptor gene TIR1 in *Dimocarpus longan* [J]. Acta Horticulturae Sinica, 2016, 43(4): 771-780. (in Chinese)
- [26] 原晓龙, 李云琴, 王毅, 等. 西南桦叶绿体基因组密码子偏好性分析[J]. 基因组学与应用生物学, 2020, 39(12): 5758-5764.
- YUAN X L, LI Y Q, WANG Y, *et al.* Analysis of codon usage in the chloroplast genome of *Betula alnoides* [J]. Genomics and Applied Biology, 2020, 39(12): 5758-5764. (in Chinese)
- [27] MURRAY E E, LOTZER J, EBERLE M. Codon usage in plant genes [J]. Nucleic Acids Res, 1989, 17(2): 477-498.
- [28] 原晓龙, 康洪梅, 王毅. 云南蓝果树叶绿体基因组密码子偏好性分析[J]. 西北林学院学报, 2020, 35(4): 26-31, 124.
- YUAN X L, KANG H M, WANG Y. Codon usage bias analysis of chloroplast genome in *Nyssa yunnanensis* [J]. Journal of Northwest Forestry University, 2020, 35(4): 26-31, 124. (in Chinese)
- [29] 喻凤, 韩明. 紫花苜蓿叶绿体基因组密码子偏好性分析[J]. 广西植物, 2021, 41(12): 2069-2076.
- YU F, HAN M. Analysis of codon usage bias in the chloroplast genome of alfalfa (*Medicago sativa*) [J]. Guihaia, 2021, 41(12): 2069-2076. (in Chinese)
- [30] 傅建敏, 索玉静, 刘慧敏, 等. 柃属植物叶绿体蛋白质编码基因密码子用法[J]. 经济林研究, 2017, 35(2): 38-44.
- FU J M, SUO Y J, LIU H M, *et al.* Analysis on codon usage in the chloroplast protein-coding genes of *Diospyros* spp. [J]. Nonwood Forest Research, 2017, 35(2): 38-44. (in Chinese)
- [31] 陆奇丰, 骆文华, 黄至欢. 两种梧桐叶绿体基因组密码子使用偏性分析[J]. 广西植物, 2020, 40(2): 173-183.
- LU Q F, LUO W H, HUANG Z H. Codon usage bias of chloroplast genome from two species of *Firmiana* Marsili [J]. Guihaia, 2020, 40(2): 173-183. (in Chinese)
- [32] 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 X J, ZHU C S, LI K X, *et al.* Genes codon bias of transcriptome in *Cinnamomum camphora* [J]. Guihaia, 2021, 41(12): 2077-2083. (in Chinese)