

梁山慈竹叶绿体基因组密码子偏好性分析

魏亚楠¹, 龚明贵¹, 白娜¹, 苏佳杰¹, 姜霞²

(1. 河南科技大学 食品与生物工程学院, 河南 洛阳 471023; 2. 贵州省林业科学研究院, 贵州 贵阳 550005)

摘要: 【目的】探究梁山慈竹 *Dendrocalamus farinosus* 叶绿体基因组密码子偏好性使用形式, 分析影响梁山慈竹密码子使用偏好性的主要原因, 并确定最优密码子, 为竹亚科 Bambusoideae 植物叶绿体基因组学研究提供参考依据。【方法】根据 GenBank 登录号 MZ681865.156 从美国国家生物技术信息中心 (NCBI) 数据库中下载 85 条梁山慈竹叶绿体基因序列, 利用 CodonW、CUSP 及 R 语言软件分析有效密码子数 (ENC)、适应指数 (CAI) 和同义密码子相对使用度 (RSCU) 等指标, 对 RSCU 进行对应性分析, 并根据 ENC 和 RSCU 对密码子进行排序。【结果】梁山慈竹叶绿体基因组密码子的鸟嘌呤 (G) 和胞嘧啶 (C) 所占的比率 (GC 比率) 平均为 39.48%, 且 GC1(47.69%)>GC2(39.70%)>GC3(31.05%), 密码子末位碱基偏好以 A/U 结尾; ENC 多数在 35 以上, CAI 为 16.6%, 其密码子偏好性较弱; 中性绘图分析、ENC-plot 分析和 PR2-plot 分析表明: 影响梁山慈竹叶绿体基因组密码子偏好性的主要原因是自然选择。共有 GCA、GCU、UUC 及 GGU 等 18 个密码子被鉴定为梁山慈竹叶绿体基因组的最优密码子。【结论】自然选择是梁山慈竹叶绿体基因组密码子偏好性的主要因素, 并筛选出梁山慈竹叶绿体基因组有 GCU、GAU 以及 GGU 等 18 个最优密码子。图 5 表 5 参 30

关键词: 梁山慈竹; 密码子偏好性; 叶绿体基因组; 最优密码子

中图分类号: S795.5; S722 文献标志码: A 文章编号: 2095-0756(2024)00-0001-10

Analysis of codon preference in chloroplast genome of *Dendrocalamus farinosus*

WEI Ya'nan¹, GONG Minggui¹, BAI Na¹, SU Jiajie¹, JIANG Xia²

(1. College of Food and Bioengineering, Henan University of Science and Technology, Luoyang 471023, Henan, China; 2. Guizhou Academy of Forestry, Guiyang 550005, Guizhou, China)

Abstract: [Objective] The objective is to explore the preferred usage patterns of chloroplast genome codon in *Dendrocalamus farinosus*, analyze the main reasons affecting the codon usage preference of *D. farinosus*, and determine the optimal codon, so as to provide reference for chloroplast genomics research in Bambusoideae plants. [Method] According to the GenBank login number MZ681865.156, 85 chloroplast gene sequences of *D. farinosus* were downloaded from National Center for Biotechnology Information (NCBI) database in the United States. CodonW, CUSP and R language software were used to analyze the effective number of codons (ENC), adaptation index (CAI) and relative synonymous codon usage (RSCU). Correspondence analysis of RSCU was performed and the codons were sorted based on ENC and RSCU values. [Result] The average ratio of guanine (G) and cytosine (C) in the chloroplast genome codon (GC ratio) was 39.48%, with GC1 (47.69%)>GC2 (39.70%)>GC3 (31.05%), and the last codon base preferred to end in A/U. The majority of ENC value was above 35, and CAI value was 0.167, so the codon preference was weak. Neutral plot analysis,

收稿日期: 2023-10-08; 修回日期: 2024-03-06

基金项目: 贵州省林业局资助项目 (黔林科合 [2019] 06 号); 国家自然科学基金资助项目 (31860178)

作者简介: 魏亚楠 (ORCID: 0009-0001-4831-8988), 从事资源微生物研究。E-mail: weiyankfc@163.com。通信作者: 龚明贵 (ORCID: 0000-0001-6863-5115), 副教授, 博士, 从事资源微生物研究。E-mail: gongminggui@163.com。姜霞 (ORCID: 0009-0006-0720-4622), 副研究员, 从事森林生态学研究。E-mail: 43644135@qq.com

ENC-plot and PR2-plot analysis showed that natural selection was the main factor affecting the codon preference of the chloroplast genome of *D. farinosus*. A total of 18 codons, including GCA, GCU, UUC, and GGU, were identified as the optimal codons for the chloroplast genome of *D. farinosus*. [Conclusion] Natural selection is the main factor contributing to the codon preference of the chloroplast genome in *D. farinosus*, and 18 optimal codons such as GCU, GAU, and GGU are screened for the chloroplast genome of *D. farinosus*. [Ch, 5 fig. 5 tab. 30 ref.]

Key words: *Dendrocalamus farinosus*; codon preference; chloroplast genome; optimal codons

密码子是识别和传递生物体遗传信息、联系蛋白质与 DNA 之间的重要桥梁,在生物体遗传和变异中起着至关重要的作用^[1]。编码同一氨基酸的不同密码子被称为同义密码子。由于基因突变和自然选择的影响,某些同义密码子在蛋白质翻译过程中往往被高频使用,被称为密码子的使用偏好性^[2-3]。物种的生物学功能与密码子偏好性密切相关,密码子偏好性不仅可以影响生物编码基因蛋白质的合成速率和翻译速率^[4],还会影响蛋白质结构、折叠程度和 mRNA 的合成^[5]。研究表明:同一物种或亲缘关系相近的物种,具有相似的密码子偏好使用模式^[6],通过分析物种的密码子偏好性可以衡量物种之间的基因表达量,进而探究物种之间亲属关系^[7]。通过密码子偏好性的研究,能够更好地阐明物种进化过程中基因的表达规律^[8],为利用基因工程技术改良物种目标基因提供参考依据^[9]。

梁山慈竹 *Dendrocalamus farinosus* 属竹亚科 Bambusoideae 牡竹属 *Dendrocalamus*, 又名大叶竹和瓦灰竹,是中国西南地区重要的经济作物^[10],生长速度快,适应性强,竹笋效益高,属于优良的笋竹两用竹种,与硬头黄竹 *Bambusa rigida* 都属于竹编和制浆造纸的优质原料^[11]。针对梁山慈竹叶绿体基因组密码子使用偏好性的研究鲜见报道。为了更好地挖掘和利用梁山慈竹的潜在经济价值,本研究以梁山慈竹叶绿体基因组序列为研究对象,分析其密码子偏好性使用模式,探究并总结其相关表达基因的密码子偏好性,以期分析影响梁山慈竹叶绿体基因组密码子偏好性的主要因素,并筛选出最优密码子,为后续梁山慈竹叶绿体基因工程改造等研究提供理论基础。

1 材料与方法

1.1 叶绿体基因组序列的获取

根据 GenBank 登录号 MZ681865.156 在美国国家生物技术信息中心 (NCBI) 数据库中搜索并下载梁山慈竹叶绿体基因组序列,共有 85 条 CDS 序列。序列重复或小于 300 bp 会对密码子偏好性指标的测定产生影响^[12]。对基因序列进行筛选,剔除序列长度小于 300 bp 且重复的序列,获取起始密码子为 ATG,终止密码子为 TAG、TGA 和 TAA 的序列,最终获得 51 条 CDS 序列作为后续分析的样本序列。

1.2 密码子组成分析

运用 CodonW1.4.2 (<http://sourceforge.net/projects/codonw>) 和 EMBOSS (<http://imed.med.ucm.es/EMBOSS/>) 计算有效密码子数 (ENC)、适应指数 (CAI)、密码子偏性指数 (CBI)、最优密码子频率 (FOP) 以及密码子第 3 位核苷酸 A、T、C、G 的含量 (分别记为 A3、T3、C3、G3)。利用 ENC 判断密码子偏好性程度,ENC > 35 说明密码子偏好性比较弱;反之,说明偏好性强^[13]。通过 CUSP 软件分析并获得密码子鸟嘌呤 (G) 和胞嘧啶 (C) 所占的比率 (GC 比率) 及 GC 平均比率 (GC_{all}),使用 SPSS 25.0 软件对梁山慈竹密码子各位置的 GC 比率与 ENC 进行相关分析。

1.3 相对同义密码子使用度分析

运用 CodonW1.4.2 对同义密码子相对使用度 (RSCU) 进行分析,即该密码子的实际使用频率与其理论使用频率的比值^[14]。当 RSCU 大于 1 时,同义密码子中偏好使用该密码子,被称为高频密码子;当 RSCU 等于 1 时,密码子无偏好性;当 RSCU 小于 1 时,密码子使用偏好性较弱^[15]。

1.4 中性绘图分析

中性绘图分析是对影响密码子使用偏好性的关键因素进行分析,X轴为 GC3,Y轴为 GC1 和 GC2 的平均值,绘制二维散点图对 GC3 和 GC12 (各基因 GC1 和 GC2 的平均值) 的相关性进行分析 (GC1、GC2、GC3 分别代表第 1、2、3 位密码子的 GC 比例)。若回归系数接近 1,代表 GC3 和 GC12 显

著相关，碱基组成没有差异，说明突变是决定密码子偏好性的主要因素；若回归系数接近 0，则代表自然选择是主要因素。

1.5 ENC-plot 绘图分析

ENC-plot 绘图分析表现密码子的使用偏好性受到突变和自然选择的影响程度。使用 Python 3.7 进行 ENC-plot 绘图分析，构建散点图，横纵坐标分别为 GC3、ENC，并绘制 ENC 的标准曲线。基因位点靠近或在标准曲线上，表明突变是决定密码子偏好性的主要因素，若基因位点和标准曲线距离很大，则说明偏好性主要由自然选择决定。

1.6 PR2-plot 偏倚分析

PR2-plot 分析表明基因中密码子的第 3 位碱基的构成情况。计算密码子碱基中第 3 位上 4 种碱基 A、T、C、G 比例， $G3/(G3+C3)$ 为 X 轴， $A3/(A3+T3)$ 为 Y 轴，绘制 PR2-plot 散点图，中心点为碱基比例 A=T、C=G 时的值，代表处于此区域的密码子并无使用偏好性^[16]。

1.7 最优密码子分析

将 51 条基因升序排列后的 ENC 前后两端 10% 的基因建立高、低表达基因库。通过 CodonW 软件计算 2 个表达库中密码子的 RSCU 和 Δ RSCU，同时满足高频密码子 ($RSCU > 1$) 和高表达密码子 (Δ RSCU ≥ 0.08) 的为最优密码子^[17]。

1.8 梁山慈竹和其他几种生物密码子偏好性比较分析

在 Codon Usage Database(<http://www.kazusa.or.jp/codon/>) 下载异源表达宿主和植物代表类群，包括巨龙竹 *D. farinosus*、粉麻竹 *D. sinicus*、小叶龙竹 *D. pulverulentus*、硬头黄竹、大肠埃希菌 *Escherichia coli*、烟草 *Nicotiana tabacum*、拟南芥 *Arabidopsis thaliana* 和酿酒酵母 *Saccharomyces cerevisiae* 等物种基因组密码子的使用频率，与梁山慈竹基因组密码子使用频率比值进行比较分析，当梁山慈竹密码子使用频率比其他生物的比值 ≥ 2.0 或 ≤ 0.5 时，说明该物种与梁山慈竹的同义密码子的使用偏好性差异较大，当比值不在上述范围内时，表明这 2 个物种对该密码子的偏好性较接近。

1.9 对应分析

将叶绿体基因如表 1 所示进行功能分类，使用 CodinW 软件，选择对应分析计算样本中各个基因的 RSCU，将分析结果分布在 59 维向量空间中，分析指标间的对应性。

表 1 梁山慈竹叶绿体基因结构分析

Table 1 Structural analysis of the chloroplast genome of *D. farinosus*

基因分类	基因分组	基因名称
光合系统基因	光系统 I 基因	<i>psaA</i> 、 <i>psaB</i> 、 <i>psbA</i> 、 <i>psbC</i> 、 <i>psbD</i> 、 <i>psbB</i>
	光系统 II 基因	<i>petA</i> 、 <i>petB</i> 、 <i>petD</i>
	细胞色素 b ₆ /f 复合体基因	<i>atpA</i> 、 <i>atpB</i> 、 <i>atpE</i> 、 <i>atpF</i> 、 <i>atpI</i>
	三磷酸腺苷合成酶基因	<i>ndhA</i> 、 <i>ndhB</i> 、 <i>ndhC</i> 、 <i>ndhD</i> 、 <i>ndhE</i> 、 <i>ndhF</i> 、 <i>ndhG</i> 、 <i>ndhH</i> 、 <i>ndhI</i> 、 <i>ndhJ</i> 、 <i>ndhK</i>
遗传系统基因	烟酰胺腺嘌呤二核苷酸氧化还原酶基因	<i>rbcl</i>
	二磷酸核酮糖羧化酶大亚基基因	<i>rpoA</i> 、 <i>rpoB</i> 、 <i>rpoC1</i> 、 <i>rpoC2</i>
	RNA 聚合酶亚基基因	<i>rps2</i> 、 <i>rps3</i> 、 <i>rps4</i> 、 <i>rps7</i> 、 <i>rps8</i> 、 <i>rps11</i> 、 <i>rps12</i> 、 <i>rps14</i> 、 <i>rps18</i>
	核糖体蛋白小亚基基因	<i>rpl2</i> 、 <i>rpl14</i> 、 <i>rpl16</i> 、 <i>rpl20</i> 、 <i>rpl22</i>
其他基因	成熟酶 K 基因	<i>matK</i>
	膜蛋白基因	<i>cemA</i>
	细胞色素合成基因	<i>ccsA</i>
	酪蛋白分解蛋白酶基因	<i>clpP</i>
未知功能基因	假定叶绿体阅读框	<i>ycf2</i> 、 <i>ycf3</i> 、 <i>infA</i>

2 结果与分析

2.1 密码子的碱基组成分析

分析梁山慈竹叶绿体基因组 CDS 序列的碱基组成：梁山慈竹的 4 种碱基所对应的同义密码子的第 3 位碱基比例 (T3s、A3s、C3s、G3s) 分别为 45.28%、42.07%、18.13%、17.96%，T3s 和 A3s 远高于 G3s 和 C3s，表明梁山慈竹叶绿体基因组密码子的第 3 位碱基以 A/U 结尾为主。梁山慈竹的 ENC 为

50.40, CAI为16.6%,第3位同义密码子的GC比率(GC3_s)为28.1%,表明其叶绿体基因组密码子偏好性较弱。

梁山慈竹叶绿体基因组密码子的GC平均比率为39.48%,且GC1(47.69%)>GC2(39.70%)>GC3(31.05%)。ENC为39.04~61.00,均值为49.51,GC比率在基因密码子上并没有均匀分布(表2)。ENC和密码子3个位置GC比率的相关分析(表3)结果发现:ENC与GC3比率显著相关,与GC1、GC2不显著相关,说明密码子使用偏好性形成过程中GC3的影响作用大于GC1、GC2。

表2 梁山慈竹叶绿体基因组各基因密码子相关参数统计

Table 2 Statistics of codon related parameters of various genes in the chloroplast genome of *D. farinosus*

基因	GC比率/%				ENC	CAI	FOP	基因	GC比率/%				ENC	CAI	FOP
	GC	GC1	GC2	GC3					GC	GC1	GC2	GC3			
<i>rps12</i>	41.87	52.00	47.20	26.40	44.85	0.140	0.341	<i>rps18</i>	33.53	34.50	39.77	26.32	39.04	0.147	0.333
<i>psbA</i>	42.56	49.72	42.94	35.03	41.33	0.313	0.532	<i>rpl20</i>	36.11	38.33	40.83	29.17	50.97	0.112	0.298
<i>matK</i>	34.44	40.82	32.42	30.08	49.49	0.166	0.329	<i>clpP</i>	43.01	52.53	38.25	38.25	52.37	0.175	0.337
<i>psbD</i>	44.44	53.39	43.50	36.44	48.99	0.242	0.456	<i>psbB</i>	44.01	54.42	45.97	31.63	50.73	0.190	0.380
<i>psbC</i>	44.66	53.59	44.73	35.63	48.91	0.183	0.386	<i>petB</i>	41.06	48.93	41.20	33.05	47.31	0.191	0.333
<i>rpoB</i>	39.19	49.81	38.01	29.74	49.69	0.153	0.353	<i>petD</i>	40.37	50.93	39.13	31.06	49.46	0.161	0.305
<i>rpoC1</i>	39.87	49.93	38.07	31.63	52.77	0.156	0.347	<i>rpoA</i>	37.06	46.18	35.59	29.41	49.94	0.151	0.311
<i>rpoC2</i>	38.95	49.01	36.64	31.18	52.29	0.154	0.333	<i>rps11</i>	43.52	50.69	56.25	23.61	44.33	0.174	0.396
<i>rps2</i>	38.40	40.51	40.93	33.76	52.55	0.168	0.338	<i>infA</i>	40.35	43.86	35.96	41.23	61.00	0.181	0.409
<i>atp1</i>	38.84	47.58	36.29	32.66	50.55	0.163	0.353	<i>rps8</i>	36.50	41.61	41.61	26.28	46.62	0.122	0.374
<i>atpF</i>	38.27	47.62	35.45	31.75	53.17	0.147	0.353	<i>rpl14</i>	38.71	54.84	37.10	24.19	51.90	0.181	0.392
<i>atpA</i>	42.06	56.01	39.96	30.12	49.96	0.182	0.385	<i>rpl16</i>	44.76	52.14	53.57	28.57	39.41	0.115	0.354
<i>rps14</i>	39.42	39.42	46.15	32.69	41.73	0.135	0.384	<i>rps3</i>	33.47	43.75	31.67	25.00	48.03	0.193	0.402
<i>psaB</i>	41.81	48.71	43.13	33.61	49.34	0.172	0.350	<i>rpl22</i>	37.56	41.33	36.67	34.67	47.48	0.188	0.415
<i>psaA</i>	43.68	51.80	43.28	35.95	52.07	0.198	0.373	<i>rpl2</i>	44.56	51.77	48.58	33.33	53.33	0.143	0.361
<i>ycf3</i>	39.69	47.40	38.15	33.53	55.45	0.156	0.343	<i>ndhB</i>	38.16	42.07	39.33	33.07	46.71	0.156	0.348
<i>rps4</i>	37.13	47.52	37.13	26.73	49.59	0.169	0.386	<i>rps7</i>	39.49	49.68	45.22	23.57	48.31	0.164	0.373
<i>ndhJ</i>	39.38	49.38	36.88	31.88	51.48	0.176	0.356	<i>ndhF</i>	34.19	37.84	38.92	25.81	46.19	0.144	0.321
<i>ndhK</i>	38.60	41.70	43.72	30.36	51.91	0.159	0.329	<i>ccsA</i>	33.64	33.74	41.10	26.07	45.60	0.152	0.307
<i>ndhC</i>	39.67	50.41	36.36	32.33	48.75	0.177	0.345	<i>ndhD</i>	36.19	40.72	36.93	30.94	48.98	0.133	0.314
<i>atpE</i>	42.51	52.17	39.13	36.23	59.51	0.167	0.405	<i>ndhE</i>	33.33	41.18	32.35	26.47	59.06	0.144	0.316
<i>atpB</i>	42.62	53.91	41.68	32.26	47.43	0.192	0.381	<i>ndhG</i>	34.46	44.07	32.77	26.55	45.77	0.125	0.250
<i>rbcL</i>	44.14	57.11	43.93	31.38	50.19	0.271	0.454	<i>ndhI</i>	34.99	37.57	38.67	28.73	52.09	0.171	0.345
<i>ycf4</i>	41.22	48.39	39.78	35.48	47.14	0.162	0.385	<i>ndhA</i>	33.98	42.42	36.36	23.14	44.35	0.140	0.321
<i>cemA</i>	33.62	41.99	27.71	31.17	55.91	0.176	0.342	<i>ndhH</i>	37.82	50.76	34.77	27.92	49.95	0.155	0.322
<i>petA</i>	40.29	53.58	35.2	32.09	51.12	0.155	0.331								

表3 梁山慈竹叶绿体基因组中各基因参数的相关性分析

Table 3 Correlation analysis of various gene parameters in the chloroplast genome of *D. farinosus*

参数	GC1	GC2	GC3	ENC	CAI	CBI	FOP	GC3s	GC
GC1	1								
GC2	0.300*	1							
GC3	0.265	-0.009	1						
ENC	0.142	-0.425**	0.389**	1					
CAI	0.409**	0.076	0.370**	0.012	1				
CBI	0.438**	0.272	0.322*	-0.092	0.774**	1			
FOP	0.402**	0.312*	0.341*	-0.064	0.797**	0.965**	1		
GC3s	0.271	-0.029	0.946**	0.445**	0.330*	0.330*	0.370**	1	
GC	0.814**	0.673**	0.525**	0.010	0.407**	0.512**	0.518**	0.499**	1

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

2.2 相对同义密码子使用度分析

梁山慈竹叶绿体基因组中共包含 18110 个密码子 (表 4)，总计编码 20 个氨基酸，密码子数为 12~705 个，其中密码子 UGA 共有 12 个，密码子含量最多的是编码谷氨酸的 GAA，共有 705 个。梁山慈竹叶绿体基因组蛋白编码序列 RSCU 分析表明：氨基酸含量较高的有亮氨酸 (Leu) 和精氨酸 (Arg)，均为 6 个密码子编码，编码精氨酸的是 UUA、UUG、CUU、CUC、CUA 和 CUG；编码亮氨酸的有 AGA、AGG、CGU、CGC、CGA 和 CGG；除此之外，蛋氨酸 (Met) 和色氨酸 (Trp) 均只有 1 个密码子编码，分别是 AUG 和 UGG，其余氨基酸密码子编码个数分别为 2~4 个。

表 4 梁山慈竹叶绿体基因组蛋白编码序列 RSCU 分析

Table 4 RSCU of protein coding region in the chloroplast of *D. farinosus*

氨基酸	密码子	数量	RSCU	氨基酸	密码子	数量	RSCU	氨基酸	密码子	数量	RSCU	氨基酸	密码子	数量	RSCU
Phe	UUU*	644	1.29	Tyr	UAU*	532	1.59	Ser	UCU*	343	1.58	Cys	UGU*	151	1.53
Phe	UUC	351	0.71	Tyr	UAC	137	0.41	Ser	UCC*	260	1.19	Cys	UGC	47	0.47
Leu	UUA*	634	1.94	TER	UAA*	28	1.56	Ser	UCA*	222	1.02	Arg	AGA*	322	1.75
Leu	UUG*	362	1.11	TER	UAG	14	0.78	Ser	UCG	119	0.55	Arg	AGG	119	0.64
Leu	CUU*	420	1.29	TER	UGA	12	0.67	Ser	AGU*	273	1.25	Arg	CGU*	261	1.41
Leu	CUC	138	0.42	Trp	UGG*	328	1.00	Ser	AGC	89	0.41	Arg	CGC	95	0.51
Leu	CUA	295	0.90	Gln	CAA*	477	1.53	Thr	ACU*	403	1.68	Arg	CGA*	234	1.27
Leu	CUG	107	0.33	Gln	CAG	148	0.47	Thr	ACC	181	0.75	Arg	CGG	76	0.41
Ile	AUU*	740	1.48	Glu	GAA*	705	1.46	Thr	ACA*	259	1.08	Gly	GGU*	421	1.24
Ile	AUC	295	0.59	Glu	GAG	263	0.54	Thr	ACG	116	0.48	Gly	GGC	145	0.43
Ile	AUA	461	0.92	Lys	AAA*	647	1.44	Ala	GCU*	493	1.73	Gly	GGA*	538	1.58
Met	AUG*	416	1.00	Lys	AAG	253	0.56	Ala	GCC	172	0.60	Gly	GGG	259	0.76
Val	GUU*	382	1.47	Asp	GAU*	522	1.54	Ala	GCA*	343	1.20	Pro	CCU*	286	1.48
Val	GUC	126	0.49	Asp	GAC	155	0.46	Ala	GCG	135	0.47	Pro	CCC*	196	1.01
Val	GUA*	390	1.50	His	CAU*	311	1.47	Asn	AAU*	528	1.48	Pro	CCA*	209	1.08
Val	GUG	139	0.54	His	CAC	112	0.53	Asn	AAC	187	0.52	Pro	CCG	84	0.43

说明：*. RSCU大于1为高频密码子。

RSCU 反应密码子使用度的高低，大于 1 说明该密码子的使用度较高。梁山慈竹叶绿体基因组 RSCU 大于 1 的密码子数目为 34 个 (分别为 UUU、UUA、UUG、CUU、AUU、AUG、GUU、GUA、UCU、UCC、UCA、AGU、ACU、ACA、GCU、GCA、AAU、UAU、UAA、UGG、CAA、GAA、AAA、GAU、CAU、UGU、AGA、CGU、CGA、GGU、GGA、CCU、CCC 和 CCA)，即筛选出了 34 个高频密码子，其中以 A、U、C、G 结尾的密码子分别有 13、16、2 和 1 个，说明密码子偏好以 A 和 U 结尾，RSCU 较高的 3 个密码子为 UUU(1.94)、CUA(1.73) 和 UCU(1.75)。

2.3 中性绘图分析

中性绘图分析量化自然选择和突变压力之间的关系，阐明 3 个密码子位置之间的联系。结果表明：横坐标 GC3 的数值为 23.14%~41.23%，纵坐标 GC12 的数值为 39.04%~61.00%(图 1)。梁山慈竹的 Pearson 相关系数为 0.17，呈正相关关系，数据拟合后的回归系数为 0.1868，决定系数 (R^2) 较小，为 0.0282，GC12 和 GC3 的相关性不显著，说明其叶绿体基因组密码子偏好性受自然选择影响较大。

2.4 ENC-plot 绘图分析

结果显示 (图 2)：ENC 分布并不紧密，少量分布在标准曲线附近，还有个别分布在标准曲线上侧，位点的 ENC 均大于 35，与预期 ENC 值有差距。说明梁山慈竹密码子偏好性较弱且自然选择和突变都对其偏好性有影响。由于落在标准曲线下方的基因点数量比较多，所以梁山慈竹基因组密码子使用偏好性主要受自然选择的影响。

2.5 PR2-plot 绘图分析

结果显示 (图 3)：基因位点在平面图 4 个区域内分布并不均匀，在 $A3/(A3+T3)<0.5$ 和 $G3/(G3+C3)>$

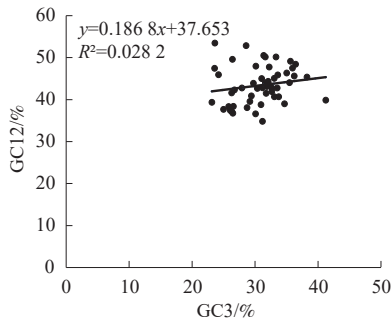


图 1 中性绘图分析

Figure 1 Analysis of neutrality plot

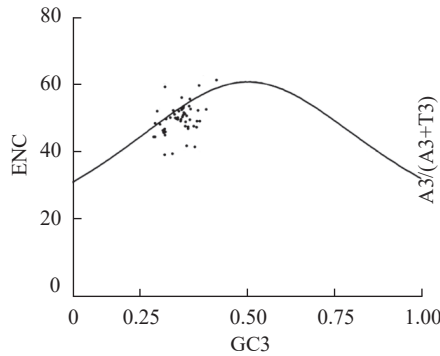


图 2 ENC-plot 分析

Figure 2 Analysis of ENC-plot

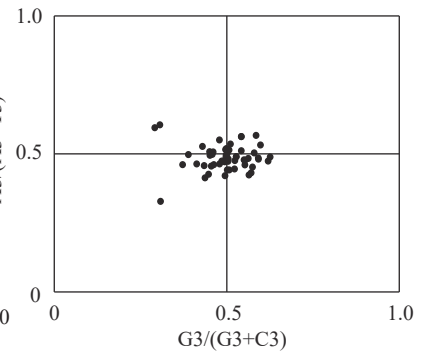


图 3 PR2-plot 分析

Figure 3 Analysis of PR2-plot

0.5 区域范围内分布最多。表明第 3 位碱基使用频率为：T>A、G>C，梁山慈竹叶叶绿体基因组密码子的第 3 位碱基在选择上具有偏好性，同时说明其密码子使用偏好性主要受自然选择的影响。

2.6 最优密码子的确定

对梁山慈竹的 ENC 进行升序排列，前 10% 为高表达基因，即 *rps18*、*rpl16*、*psbA*、*rps14*、*rps11*，后 10% 为低表达基因，即 *ycf3*、*cemA*、*ndhE*、*atpE*、*infA*。梁山慈竹的 RSCU 和 ΔRSCU 表明 (表 5)：梁山慈竹叶叶绿体基因组有 32 个高频密码子，筛选出 GCA、GCU 等 25 个高表达密码子，最终确定 18 个密码子作为梁山慈竹叶叶绿体基因组的最优密码子，分别为 UAA、GCA、GCU、UUC、GCU、AAA、CUU、UUA、CCA、CCU、CAA、AGA、CGU、AGU、UCC、ACU、GUA、GUU。其中 16 个以 A/U 结尾，2 个以 C 结尾。

2.7 梁山慈竹和其他几种生物密码子偏好性比较分析

将梁山慈竹基因组密码子使用频率与巨龙竹、粉麻竹、小叶龙竹、硬头黄竹、大肠埃希菌、烟草、拟南芥和酿酒酵母等物种的基因组密码子使用频率进行比较 (图 4)。结果显示：梁山慈竹与巨龙竹、粉麻竹、小叶龙竹和硬头黄竹的密码子使用频率为 0.5~2.0，说明它们的密码子使用偏好性相似，推测具有亲缘关系的禾本科 Gramineae 牡竹属植物叶绿体基因组密码子偏好性相似；在大肠埃希菌、烟草、拟南

氨基酸 密码子	密码子偏好性比较																密码子氨基酸	
	Df/Ds	Df/Dp	Df/Db	Df/Br	Df/Ec	Df/Nt	Df/At	Df/Sc	Df/Ds	Df/Dp	Df/Db	Df/Br	Df/Ec	Df/Nt	Df/At	Df/Sc		
Phe UUU	1.03	1.00	1.00	0.98	0.85	0.55	0.54	0.87	0.99	1.02	1.02	0.99	0.71	0.43	0.48	0.50	UAU Tyr	54.00
UUC	0.98	1.01	1.01	1.02	0.59	0.46	0.49	0.42	1.00	1.00	0.99	0.97	0.84	1.34	1.20	0.43	UAC	
Leu UUA	1.00	1.00	1.00	1.01	0.53	0.24	0.25	0.46	0.92	0.98	0.97	0.97	2.40	2.40	2.40	4.36	UAA TER	
UUG	0.99	0.99	0.99	0.98	2.09	1.23	1.39	1.74	1.03	0.99	0.99	0.98	53.68	17.89	23.01	0.00	UAG	
CUU	0.96	1.03	1.03	1.10	0.23	0.12	0.11	0.30	0.99	0.99	0.99	1.03	2.01	1.54	1.68	1.89	CAU His	
CUC	1.04	1.01	1.01	1.00	0.80	1.07	0.91	1.13	0.97	1.02	1.02	0.99	0.84	1.43	1.38	0.59	CAC	
CUA	1.03	0.98	0.98	1.10	2.06	0.59	0.58	0.49	1.01	1.01	1.02	1.01	1.25	0.71	0.77	1.07	CAA Gln	
CUG	1.01	1.01	1.00	1.03	0.55	3.81	3.61	2.12	1.00	0.98	0.98	0.94	0.22	0.71	0.70	0.37	CAG	
Ile AUU	1.00	1.01	1.01	1.03	1.32	1.00	1.05	2.89	1.01	1.00	1.00	1.03	0.68	0.36	0.41	0.66	AAU Asn	
AUC	0.99	1.00	1.00	1.02	0.57	0.79	0.80	1.17	1.02	0.97	0.96	1.00	0.24	0.40	0.34	0.21	AAC	
AUA	0.98	1.00	1.00	0.97	3.60	0.81	0.80	1.05	1.04	0.97	0.97	0.90	0.12	0.11	0.10	0.21	AAA Lys	
Met AUG	1.01	1.00	1.01	0.99	1.31	1.46	1.63	1.24	1.03	1.02	1.02	1.02	1.35	1.02	0.75	0.62	AAG	
Val GUU	1.00	1.00	1.00	1.01	1.61	1.45	1.30	1.27	0.97	1.00	0.97	1.04	0.16	0.16	0.17	0.13	GAU Asp	
GUC	0.97	1.01	1.01	1.00	0.52	1.07	0.90	0.46	0.99	0.99	1.00	1.00	0.79	1.79	1.48	0.55	GAC	
GUA	1.01	1.02	1.01	0.99	1.29	0.66	0.83	0.56	1.01	1.01	1.01	1.00	0.31	0.31	0.31	0.41	GAA Glu	
GUG	1.02	0.99	0.99	0.99	0.88	2.86	1.95	1.11	0.95	0.98	0.98	0.95	0.73	0.94	0.75	1.01	GAG	
Ser UCU	1.03	1.02	1.02	0.99	0.69	0.27	0.26	0.31	0.94	1.00	1.01	1.03	1.24	0.81	0.79	0.75	UGU Cys	
UCC	1.03	1.00	1.00	1.01	1.92	1.33	1.32	7.75	0.98	1.01	1.02	1.00	3.03	6.47	5.24	4.41	UGC	
UCA	1.00	1.00	1.00	1.01	3.09	1.67	1.55	2.32	1.00	0.98	0.99	1.00	13.37	14.70	12.25	0.00	UGA TER	
UCG	1.04	1.00	0.99	0.98	1.81	1.99	2.31	1.70	1.00	0.98	0.98	1.02	0.64	0.57	0.62	0.51	UGG Trp	
Pro CCU	1.00	0.99	1.00	1.01	5.67	2.39	2.39	2.83	0.97	1.00	1.00	0.95	0.32	0.53	0.53	0.62	CGU Arg	
CCC	1.02	1.00	1.00	0.99	6.49	4.98	5.51	3.95	1.00	1.01	1.00	1.02	1.08	5.63	5.63	2.45	CGC	
CCA	1.01	1.01	1.01	1.00	1.77	1.23	1.27	0.89	0.98	0.98	0.98	0.98	5.70	1.51	1.97	3.23	CGA	
CCG	1.01	1.00	1.00	1.01	0.73	2.92	3.15	3.89	0.96	1.01	1.00	1.05	1.19	1.47	1.60	1.31	CGG	
Thr ACU	1.02	1.03	1.03	0.90	0.84	0.38	0.37	0.27	0.99	0.97	0.98	0.95	0.84	0.53	0.54	0.58	AGU Ser	
ACC	1.01	1.03	1.02	1.04	0.26	0.60	0.63	0.69	1.02	1.02	1.01	1.00	1.35	3.99	3.53	3.37	AGC	
ACA	0.96	1.00	1.00	1.00	2.83	1.52	1.32	1.31	1.00	1.00	1.00	0.98	6.00	0.99	1.03	1.12	AGA Arg	
ACG	1.00	0.99	0.99	0.99	2.32	6.44	7.24	3.28	1.01	0.97	0.97	0.95	4.14	1.10	1.11	0.64	AGG	
Ala GCU	1.00	1.01	1.00	0.99	1.30	0.77	0.76	0.58	0.99	1.00	1.01	1.05	1.19	1.23	1.29	1.52	GGU Gly	
GCC	0.99	1.00	1.00	1.00	0.94	2.41	2.54	2.02	1.02	1.01	1.04	0.99	0.08	0.28	0.32	0.15	GGC	
GCA	1.00	1.03	1.03	1.01	0.50	0.66	0.64	0.36	1.02	1.04	1.04	1.02	0.12	0.04	0.04	0.09	GGA	
GCG	1.01	0.99	0.99	0.98	0.91	5.04	3.75	1.55	0.88	1.05	0.99	0.92	0.08	0.08	0.09	0.12	GGG	

Df. 梁山慈竹; Ds. 巨龙竹; Dp. 粉麻竹; Db. 小叶龙竹; Br. 硬头黄竹; Ec. 大肠埃希菌; Nt. 烟草; At. 拟南芥; Sc. 酿酒酵母。

图 4 梁山慈竹与其他物种密码子偏好性比较

Figure 4 Comparison of codon preference between *D. farinosus* and other species

表 5 梁山慈竹叶绿体基因组各氨基酸的 RSCU 分析及最优密码子分析

Table 5 RSCU analysis and optimal codon analysis of amino acids in chloroplast genome of *D. farinosus*

氨基酸	密码子	基因组 RSCU	高表达 RSCU	低表达 RSCU	Δ RSCU	氨基酸	密码子	基因组 RSCU	高表达 RSCU	低表达 RSCU	Δ RSCU
Ter	<u>UAA***</u>	1.560 0	1.800 0	1.200 0	0.600 0	Met	AUG	1.000 0	1.000 0	1.000 0	0
	UAG	0.780 0	0.600 0	1.200 0	-0.600 0	Asn	AAC*	0.520 0	0.893 6	0.625 0	0.268 6
	UGA	0.670 0	0.600 0	0.600 0	0		<u>AAU</u>	1.480 0	1.106 4	1.375 0	-0.268 6
Ala	<u>GCA**</u>	1.200 0	1.200 0	0.734 7	0.465 3	Pro	<u>CCA**</u>	1.080 0	0.800 0	0.500 0	0.300 0
	GCC	0.600 0	0.457 1	0.653 1	-0.196 0		<u>CCC</u>	1.010 0	0.800 0	1.166 7	-0.366 7
	GCG	0.470 0	0.228 6	0.734 7	-0.506 1		CCG	0.430 0	0.444 4	1.000 0	-0.555 6
	<u>GCU*</u>	1.730 0	2.114 3	1.877 6	0.236 7		<u>CCU***</u>	1.480 0	1.955 6	1.333 3	0.622 3
Cys	UGC**	0.470 0	0.400 0	0	0.400 0	Gln	<u>CAA*</u>	1.530 0	1.500 0	1.368 4	0.131 6
	<u>UGU</u>	1.530 0	1.600 0	2.000 0	-0.400 0		CAG	0.470 0	0.500 0	0.631 6	-0.131 6
Asp	GAC*	0.460 0	0.500 0	0.411 8	0.088 2	Arg	<u>AGA*</u>	1.750 0	1.723 4	1.534 9	0.188 5
	<u>GAU</u>	1.540 0	1.500 0	1.588 2	-0.088 2		AGG	0.640 0	0.319 1	0.837 2	-0.518 1
Glu	<u>GAA</u>	1.460 0	1.189 2	1.578 9	-0.389 7		<u>CGA</u>	1.270 0	1.276 6	1.395 3	-0.118 7
	GAG**	0.540 0	0.810 8	0.421 1	0.389 7		CGC	0.510 0	0.319 1	0.837 2	-0.518 1
Phe	<u>UUC**</u>	1.290 0	1.041 7	0.650 0	0.391 7		CGG	0.410 0	0.319 1	0.279 1	0.040 0
	UUU	0.710 0	0.958 3	1.350 0	-0.391 7		<u>CGU***</u>	1.410 0	2.042 6	1.116 3	0.926 3
Gly	<u>GGA</u>	1.580 0	1.253 7	1.818 2	-0.564 5	Ser	AGC	0.410 0	0.384 6	0.470 6	-0.086 0
	GGC	0.430 0	0.417 9	0.484 8	-0.066 9		<u>AGU**</u>	1.250 0	1.846 2	1.411 8	0.434 4
	GGG	0.760 0	0.119 4	0.363 6	-0.244 2		<u>UCA</u>	1.020 0	0.615 4	1.058 8	-0.443 4
	<u>GGU***</u>	1.240 0	2.209 0	1.333 3	0.875 7		<u>UCC***</u>	1.190 0	1.769 2	0.941 2	0.828 0
His	CAC**	0.530 0	0.941 2	0.571 4	0.369 8		UCG	0.550 0	0.153 8	0.705 9	-0.552 1
	<u>CAU</u>	1.470 0	1.058 8	1.428 6	-0.369 8		<u>UCU</u>	1.580 0	1.230 8	1.411 8	-0.181 0
Ile	AUA	0.920 0	0.850 7	0.949 4	-0.098 7	Thr	<u>ACA</u>	1.080 0	1.181 8	1.176 5	0.005 3
	AUC*	0.590 0	0.626 9	0.531 6	0.095 3		ACC	0.500 0	0.818 2	1.058 8	-0.240 6
	<u>AUU</u>	1.480 0	1.522 4	1.519 0	0.003 4		ACG	0.480 0	0.363 6	0.588 2	-0.224 6
Lys	<u>AAA**</u>	1.440 0	1.471 7	1.155 6	0.316 1		<u>ACU**</u>	1.680 0	1.636 4	1.176 5	0.459 9
	AAG	0.560 0	0.528 3	0.844 4	-0.316 1	Val	<u>GUA***</u>	1.500 0	1.767 4	1.257 1	0.510 3
Leu	CUA	0.900 0	0.833 3	1.295 5	-0.462 2		GUC	0.490 0	0	1.028 6	-1.028 6
	CUC	0.420 0	0	0.545 5	-0.545 5		GUG	0.540 0	0.372 1	0.342 9	0.029 2
	CUG	0.330 0	0.250 0	0.477 3	-0.227 3		<u>GUU**</u>	1.470 0	1.860 5	1.371 4	0.489 1
	<u>CUU*</u>	1.290 0	1.333 3	1.227 3	0.106 0	Trp	UGG	1.000 0	1.000 0	1.000 0	0
	<u>UUA***</u>	1.940 0	2.166 7	1.022 7	1.144 0	Tyr	UAC**	0.410 0	0.521 7	0.166 7	0.355 0
	<u>UUG</u>	1.110 0	1.416 7	1.431 8	-0.015 1		<u>UAU</u>	1.590 0	1.478 3	1.833 3	-0.355 0

说明: 高频密码子(RSCU>1.00)带下划线; *, Δ RSCU \geq 0.08; **, Δ RSCU \geq 0.3; ***, Δ RSCU \geq 0.5; 加粗的密码子表示最优密码子。

芥和酿酒酵母的密码子使用比值中筛选 \geq 2.0 或 \leq 0.5 的密码子, 分别有 28 和 15、15、14 个, 表明梁山慈竹与这些物种在同义密码子的偏好性上有一定的差异。

2.8 对应分析

将梁山慈竹的 51 个叶绿体基因的基因功能分为光合系统基因、遗传系统基因、其他基因和未知功能基因四大类, 在计算 RSCU 的基础上将各个基因分布到 59 维的向量空间。对应分析结果(图 5)显示: 前 4 个向量轴分别存在 18.3%、16.8%、15.6% 和 15.4% 的差异, 前 4 向量轴累计差异为 66.1%, 4 个轴对密码子均有不同程度的影响; 第 1 轴的值大于其他轴, 说明第 1 轴对梁山慈竹叶绿体基因组密码子偏好性的影响较大。对第 1 轴与 CAI、CBI、FOP、ENC 和 GC3s 等指数进行进一步的相关分析发现: 梁山慈竹基因在第 1 轴上的坐标值与 CAI($r=-0.001 7$, $P<0.01$)、CBI($r=0.099 0$, $P<0.01$)、

FOP($r=0.0830$, $P<0.01$)、ENC($r=0.1120$, $P<0.01$)、GC3s($r=-0.1450$, $P<0.01$)间具有极显著的相关关系,其中CAI和GC3s第1轴具有负相关关系,表明基因组密码子的偏好性不止受单一因素的影响,自然选择、基因突变均有可能影响梁山慈竹基因组密码子使用偏好性^[18]。

3 讨论

本研究对梁山慈竹叶绿体基因组密码子进行使用偏好性分析,筛选出51条CDS序列,分析表明:GC1>GC2>GC3,密码子在3个位置上的分布并不均匀,密码子偏好使用以A或U结尾的碱基,且梁山慈竹叶绿体基因组的ENC均值为49.51,表明其叶绿体基因组密码子使用偏好性较弱。这与乳油木 *Vitellaria paradoxa*^[19] 和二乔玉兰 *Magnolia soulangeana*^[20] 等植物叶绿体基因组密码子偏好性相似。

对梁山慈竹叶绿体基因组密码子进行中性绘图、ENC-plot分析、PR2-plot分析和对应分析。在中性绘图分析中,回归系数为0.4128,说明密码子偏好性更多受到自然选择的影响;在ENC-plot分析中,多数基因离标准曲线距离较远,实际ENC和预期ENC有差距,表明该部分基因的密码子偏好性主要受自然选择的影响;在PR2-plot绘图分析中,大部分基因位于平面图的右下方,即T>A、G>C,表明其密码子的使用更多受自然选择的影响。综上所述,影响梁山慈竹叶绿体基因组密码子偏好性的主要原因是自然选择。该研究结果与巨桉 *Eucalyptus grandis*^[21]、灰毛浆果楸 *Cipadessa cinerascens*、酸枣 *Ziziphus jujuba* var. *spinosa*^[22] 和云南油杉 *Keteleeria evelyniana*^[23] 等叶绿体基因组密码子偏好性研究结果基本一致;但在对4种蔷薇科Rosaceae果树^[24]和银白杨 *Populus alba*^[25] 的研究中发现:突变是影响密码子偏好性的主要因素。这说明密码子的使用偏好性受自然选择或基因突变因素影响。基于RSCU的对应分析表明:梁山慈竹的密码子使用变异原因除了突变和自然选择之外,还有其他的因素,这其中光合系统基因和遗传系统基因分布相对集中,各类基因密码子使用偏好性较为接近。该结论与木薯 *Manihot esculenta*^[26] 和高山松 *Pinus densata*^[27] 的研究结果一致。密码子使用频率比较结果显示:梁山慈竹与禾本科牡竹属的植物密码子偏好性相似,在基因选择外源系统表达时,可以选择密码子偏好性差异相对较小的酿酒酵母,在选择大肠埃希菌、烟草和拟南芥作为外源表达宿主时,需要根据密码子使用偏好性进行碱基优化,从而使基因在宿主体内更好地表达。

最优密码子分析表明:梁山慈竹叶绿体基因组有GCU、GAU以及GGU等18个最优密码子,最优密码子大部分以A或U结尾。该结果与抽筒竹 *Gelidocalamus tessellatus*^[28] 和毛竹 *Phyllostachys edulis*^[29] 叶绿体基因组最优密码子分析结果一致,这可能与亲缘关系相近,但不同物种之间叶绿体基因组进化过程中的相对保守性有关系^[21]。通过筛选获取梁山慈竹偏好使用密码子,可进一步对目标基因进行密码子优化,提高梁山慈竹的竹笋产量和造纸纤维含量,以及利用新一代精准基因编辑工具CRISPR/Cas9优化梁山慈竹密码子,从而改造梁山慈竹基因组编辑的Cas9基因,提高该基因在梁山慈竹中的表达水平^[30]。

4 结论

本研究通过分析梁山慈竹叶绿体基因组的CDS序列,对梁山慈竹的叶绿体基因组进行生物信息学分析,筛选出梁山慈竹叶绿体基因组有GCU、GAU以及GGU等18个最优密码子。研究表明:影响梁山慈竹密码子偏好性的主要因素是自然选择。研究结果为后续在分子层面上利用基因工程开发梁山慈竹优良资源提供参考。

5 参考文献

[1] 周涛,杨林,舒军霞,等.3种含笑属植物叶绿体基因组密码子偏好性分析[J].西部林业科学,2022,51(3):91-100.

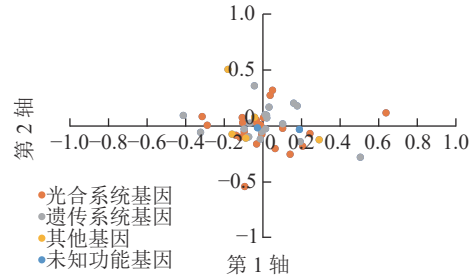


图5 梁山慈竹基因组密码子RSCU的对应性分析

Figure 5 Correspondence Analysis on RSCU of *D. farinosus*

- ZHOU Tao, YANG Lin, SHU Junxia, *et al.* Analysis of codon bias in the chloroplast genome of three of *Michelia* spp. species [J]. *Journal of West China Forestry Science*, 2022, **51**(3): 91 – 100.
- [2] ANDARGIE M, ZHU Congyi. Genome-wide analysis of codon usage in sesame (*Sesamum indicum* L.) [J/OL]. *Heliyon*, 2022, **8**(1): e8687[2023-09-08]. doi: 10.1016/j.heliyon.2021.e08687.
- [3] 孟祎, 李菁, 杜少兵, 等. 17种鼠李科植物的叶绿体基因组特征和密码子偏好性分析[J/OL]. 分子植物育种, 2023-08-22[2023-09-08]. <https://link.cnki.net/urlid/46.1068.S.20230822.1011.002>.
- MENG Yi, LI Jing, DU Shaobing, *et al.* Analysis of chloroplast genome characteristics and codon preference of 17 species of Rhamnaceae [J/OL]. *Molecular Plant Breeding*, 2023-08-22[2023-09-08]. <https://link.cnki.net/urlid/46.1068.S.20230822.1011.002>.
- [4] 辛雅萱, 黎若竹, 李鑫, 等. 杠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 Yaxuan, LI Ruozhu, LI Xin, *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.
- [5] PARVATHY S T, UDAYASURIYAN V, BHADANA V. Codon usage bias [J]. *Molecular Biology Reports*, 2022, **49**(1): 539 – 565.
- [6] CHAKRABORTY S, YENGGHOM S, UDDIN A. Analysis of codon usage bias of chloroplast genes in *Oryza* species [J/OL]. *Planta*, 2020, **252**(4): 67[2023-09-08]. doi: 10.1007/s00425-020-03470-7.
- [7] LI Changle, ZHOU Ling, NIE Jiangbo, *et al.* Codon usage bias and genetic diversity in chloroplast genomes of *Elaeagnus* species (Myrtiflorae: Elaeagnaceae) [J]. *Physiology and Molecular Biology of Plants*, 2023, **29**(2): 239 – 251.
- [8] WU Peng, XIAO Wenqi, LUO Yingyong, *et al.* Comprehensive analysis of codon bias in 13 *Ganoderma* mitochondrial genomes [J/OL]. *Frontiers in Microbiology*, 2023, **14**: 1170790[2023-09-08]. doi: 10.3389/fmicb.2023.1170790.
- [9] GENG Xiaoshan, HUANG Ning, ZHU Yuling, *et al.* Codon usage bias analysis of the chloroplast genome of cassava [J]. *South African Journal of Botany*, 2022, **151**: 970 – 975.
- [10] 王身昌, 胡尚连, 曹颖, 等. 梁山慈竹高通量转录组测序及差异表达基因分析[J]. 华北农学报, 2016, **31**(3): 65 – 71.
- WANG Shenchang, HU Shanglian, CAO Ying, *et al.* High-throughput RNA-seq and analysis on differential expressed gene from *Dendrocalamus farinosus* [J]. *Acta Agriculturae Boreali-Sinica*, 2016, **31**(3): 65 – 71.
- [11] 姜霞, 周华, 晏玉莹, 等. 梁山慈竹和硬头黄竹秆形结构和地上生物量分配特性[J]. 贵州林业科技, 2023, **51**(1): 39 – 43.
- JIANG Xia, ZHOU Hua, YAN Yuying, *et al.* Culm structure and above-ground biomass allocation of *Dendrocalamus farinosus* and *Bambusa rigida* [J]. *Guizhou Forestry Science and Technology*, 2023, **51**(1): 39 – 43.
- [12] JIANG Yue, DENG Fei, WANG Hualin, *et al.* An extensive analysis on the global codon usage pattern of baculoviruses [J]. *Archives of Virology*, 2008, **153**(12): 2273 – 2282.
- [13] 原晓龙, 王毅, 张劲峰. 灰毛浆果楝叶绿体基因组密码子使用特征分析[J]. 森林与环境学报, 2020, **40**(2): 195 – 202.
- YUAN Xiaolong, WANG Yi, ZHANG Jinfeng. Characterization of codon usage in *Cipadessa cinerascens* chloroplast genome [J]. *Journal of Forest and Environment*, 2020, **40**(2): 195 – 202.
- [14] MAZUMDER G A, UDDIN A, CHAKRABORTY S. Analysis of codon usage bias in mitochondrial CO gene among platyhelminthes [J/OL]. *Molecular & Biochemical Parasitology*, 2021, **245**: 111410[2023-09-08]. doi: 10.1016/j.molbiopara.2021.111410.
- [15] 秦政, 郑永杰, 桂丽静, 等. 樟树叶绿体基因组密码子偏好性分析[J]. 广西植物, 2018, **38**(10): 1346 – 1355.
- QIN Zheng, ZHENG Yongjie, GUI Lijing, *et al.* Codon usage bias analysis of chloroplast genome of camphora tree (*Cinnamomum camphora*) [J]. *Guihaia*, 2018, **38**(10): 1346 – 1355.
- [16] BEGUM N S, CHAKRABORTY S. Influencing elements of codon usage bias in Birnaviridae and its evolutionary analysis [J/OL]. *Virus Research*, 2022, **310**: 198672[2023-09-08]. doi: 10.1016/j.virusres.2021.198672.
- [17] 唐晓芬, 陈莉, 王玉韬. 密码子使用偏性量化方法研究综述[J]. 基因组学与应用生物学, 2013, **32**(5): 660 – 666.
- TANG Xiaofen, CHEN Li, MA Yutao. Review and prospect of the principle and methods quantifying codon usage bias [J]. *Genomics and Applied Biology*, 2013, **32**(5): 660 – 666.
- [18] 陆奇丰, 骆文华. 广西秋海棠叶绿体基因组密码子偏好性分析[J/OL]. 分子植物育种, 2023, 2023-09-05[2023-09-08]. <https://link.cnki.net/urlid/46.1068.S.20230905.0920.002>.

- LU Qifeng, LUO Wenhua. Analysis of codon usage bias in chloroplast genome of *Begonia guangxiensis* [J/OL]. *Molecular Plant Breeding*, 2023, 2023-09-05 [2023-09-08]. <https://link.cnki.net/urlid/46.1068.S.20230905.0920.002>.
- [19] 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 Xiaolong, LI Yunqin, ZHANG Jinfeng, *et al.* Codon usage bias analysis of chloroplast genome in *Vitellaria paradoxa* [J]. *Molecular Plant Breeding*, 2020, **18**(17): 5658 – 5664.
- [20] 沈莲文, 田金红, 王玉昌, 等. 2种玉兰属植物叶绿体基因组密码子偏好性分析[J]. 西南林业大学学报(自然科学), 2023, **43**(2): 44 – 53.
SHEN Lianwen, TIAN Jinhong, WANG Yuchang, *et al.* Analysis of codon usage bias (CUB) in the chloroplast genomes of 2 *Yulania* species [J]. *Journal of Southwest Forestry University (Natural Sciences)*, 2023, **43**(2): 44 – 53.
- [21] 王鹏良, 吴双成, 杨利平, 等. 巨桉叶绿体基因组密码子偏好性分析[J]. 广西植物, 2019, **39**(12): 1583 – 1592.
WANG Pengliang, WU Shuangcheng, YANG Liping, *et al.* Analysis of codon bias of chloroplast genome in *Eucalyptus grandis* [J]. *Guihaia*, 2019, **39**(12): 1583 – 1592.
- [22] 胡晓艳, 许艳秋, 韩有志, 等. 酸枣叶绿体基因组密码子使用偏性分析[J]. 森林与环境学报, 2019, **39**(6): 621 – 628.
HU Xiaoyan, XU Yanqiu, HAN Youzhi, *et al.* Codon usage bias analysis of the chloroplast genome of *Ziziphus jujuba* var. *spinosa* [J]. *Journal of Forest and Environment*, 2019, **39**(6): 621 – 628.
- [23] 李江飞, 王瑜, 颜廷雨, 等. 云南油杉叶绿体基因组密码子偏好性分析[J]. 中南林业科技大学学报, 2022, **42**(4): 30 – 39.
LI Jiangfei, WANG Yu, YAN Tingyu, *et al.* Analysis on codon usage bias of *Keteleeria evelyniana* chloroplast genome [J]. *Journal of Central South University of Forestry & Technology*, 2022, **42**(4): 30 – 39.
- [24] 刘兴跃, 何仲坚, 邱毅敏. 4种蔷薇科果树叶绿体基因组密码子偏好性分析[J]. 分子植物育种, 2022, **20**(16): 5299 – 5308.
LIU Xingyue, HE Zhongjian, QIU Yimin. Analysis of codon bias in the chloroplast genome of four Rosaceae fruit trees [J]. *Molecular Plant Breeding*, 2022, **20**(16): 5299 – 5308.
- [25] ZHOU Meng, LONG Wei, LI Xia. Analysis of synonymous codon usage in chloroplast genome of *Populus alba* [J]. *Journal of Forestry Research*, 2008, **19**(4): 293 – 297.
- [26] 王占军, 丁亮, 蔡倩文, 等. 3种木薯全基因组的密码子偏好性模式与变异来源比较[J]. 应用与环境生物学报, 2021, **27**(4): 1013 – 1021.
WANG Zhanjun, DING Liang, CAI Qianwen, *et al.* Comparison of codon preference patterns and variation sources in *Manihot esculenta* Crantz genomes [J]. *Chinese Journal of Applied and Environmental*, 2021, **27**(4): 1013 – 1021.
- [27] 李江飞, 李亚麒, 唐军荣, 等. 高山松叶绿体基因组密码子偏好性模式[J]. 生物学杂志, 2023, **40**(1): 52 – 59.
LI Jiangfei, LI Yaqi, TANG Junrong, *et al.* Comparison of codon preference patterns in the chloroplast genome of *Pinus densata* [J]. *Journal of Biology*, 2023, **40**(1): 52 – 59.
- [28] 李江平, 秦政, 国春策, 等. 抽筒竹叶叶绿体基因组的密码子偏好性分析[J]. 竹子学报, 2019, **38**(2): 79 – 87.
LI Jiangping, QIN Zheng, GUO Chunce, *et al.* Codon bias in the chloroplast genome of *Gelidocalamus tessellatus* [J]. *Journal of Bamboo Research*, 2019, **38**(2): 79 – 87.
- [29] 黄笑宇, 许在恩, 郭小勤. 基于全基因组的毛竹同义密码子使用偏好性分析[J]. 浙江农林大学学报, 2017, **34**(1): 120 – 128.
HUANG Xiaoyu, XU Zai'en, GUO Xiaoqin. Synonymous codon bias of *Phyllostachys edulis* [J]. *Journal of Zhejiang A&F University*, 2017, **34**(1): 120 – 128.
- [30] MA Xingliang, ZHU Qinlong, CHEN Yuanling, *et al.* CRISPR/Cas9 platforms for genome editing in plants: developments and applications [J]. *Molecular Plant*, 2016, **9**(7): 961 – 974.