

益智叶绿体基因组密码子使用偏好分析

李英英, 李香振, 晏小霞*

中国热带农业科学院热带作物品种资源研究所/农业农村部中药材生物学与栽培重点实验室/农业农村部热带农业野生植物基因资源鉴定评价中心/海南省热带药用植物工程研究中心/海南省热带作物资源遗传改良与创新重点实验室, 海口 571101

摘要: 益智 (*Alpinia oxyphylla* Miq.) 作为我国著名的“四大南药”之一, 具有抗氧化、抗炎、神经保护、抗肿瘤等现代药理活性, 但随着野生资源日益减少, 人工栽培面临病虫害、气候变化等压力, 亟需通过生物技术手段提升种质品质与抗逆性。本研究通过解析益智叶绿体基因组的密码子使用偏好性, 为其后续遗传工程和合成生物学应用提供理论依据。基于多重过滤条件筛选获得 51 条高质量的益智叶绿体基因组蛋白编码序列, 利用 CodonW 1.4.2 软件结合在线分析平台 CUSP 解析其特征。结果显示: (1) 益智叶绿体基因组的密码子存在显著的 A/U 末端偏好性: 密码子 GC 含量呈梯度递减 ($GC_1=45.69\%>GC_2=38.79\%>GC_3=27.50\%$), 90.6% 的高偏好密码子 ($RSCU>1$) 以 A/U 结尾; (2) 密码子偏好性整体偏弱: ENC 均值为 46.56 (67.78% 基因的 $ENC>45$), CAI 均值为 0.17; (3) 自然选择主导密码子进化: 中性绘图显示 GC_{12} 与 GC_3 无显著相关性, ENC-plot 中实际值显著偏离理论曲线, PR2-plot 揭示 GC_3 位点存在 A>T/C>G 的偏倚特性; (4) 确定了 10 个最优密码子。本研究首次揭示了益智叶绿体基因组保守的 A/U 偏好性规律及由自然选择主导的多因素协同调控机制, 所确定的最优密码子可作为益智叶绿体遗传工程的关键元件, 通过定向优化外源基因 (如药用成分合成酶编码基因) 的密码子, 提升其表达效率, 为益智品质改良与抗逆性育种提供技术支撑。

关键词: 益智; 叶绿体基因组; 密码子偏好性; 最优密码子

中图分类号: S567.239 文献标志码: A

Analysis on Codon Usage Bias of *Alpinia oxyphylla* Miq. Chloroplast Genome

LI Yingying, LI Xiangzhen, YAN Xiaoxia*

Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences / Key Laboratory of Biology and Cultivation of Herb Medicine (Haikou), Ministry of Agriculture and Rural Affairs / Tropical Wild Plant Gene Resource, Ministry of Agriculture and Rural Affairs / Hainan Provincial Engineering Research Center for Tropical Medicinal Plants / Key Laboratory of Tropical Crops Germplasm Resources Genetic Improvement and Innovation of Hainan Province, Haikou, Hainan 571101, China

Abstract: *Alpinia oxyphylla* Miq. renowned as one of the "Four Great Southern Medicinal Herbs" of China, possesses modern pharmacological activities including antioxidant, anti-inflammatory, neuroprotective, and anti-tumor effects. However, with its wild resources increasingly depleted and cultivated plants facing pressures from pests, diseases, and climate change, there is an urgent need to enhance germplasm quality and stress resistance through biotechnological approaches. This study aimed to analyze the codon usage bias in the chloroplast genome of *A. oxyphylla* Miq. providing a theoretical basis for its subsequent applications in genetic engineering and synthetic biology. Fifty-one high-quality protein-coding sequences from the *A. oxyphylla* chloroplast genome were screened based on multiple filtering criteria. Codon usage characteristics were analyzed using CodonW 1.4.2 software combined with the online platform CUSP. There existed significant A/U-ending preference. The GC content showed a decreasing gradient ($GC_1=45.69\%>GC_2=$

收稿日期 2025-07-21; 接受日期 2025-08-13

基金项目 热带作物物种品种资源保护项目 (No. 21250123); 物种品种资源保护费项目 (No. A120201); 中央级公益性科研院所基本科研业务费专项 (No. 1630032025012)。

作者简介 李英英 (1990—), 女, 硕士, 助理研究员, 研究方向: 药用植物资源研究与利用。*通信作者 (Corresponding author): 晏小霞 (YAN Xiaoxia), E-mail: yxx6232@163.com。

38.79%>GC₃=27.50%), with 90.6% of the preferred codons (RSCU>1) ending in A/U. There existed overall weak codon usage bias. The mean effective number of codons (ENC) was 46.56 (67.78% of genes had ENC>45), and the mean codon adaptation index (CAI) was 0.17. Natural selection dominated codon evolution. Neutrality plot analysis showed no significant correlation between GC₁₂ and GC₃. The actual values significantly deviated from the expected curve in the ENC-plot. The PR2-plot revealed a bias at the GC₃ position favoring A>T and C>G. Ten optimal codons were identified. This study is the first to reveal the conserved A/U-ending preference and the multifactor synergistic regulatory mechanism dominated by natural selection in the *A. oxyphylla* chloroplast genome. The identified optimal codons could serve as key elements for chloroplast genetic engineering in *A. oxyphylla*. Directional optimization of codon usage in exogenous genes (e.g., genes encoding enzymes for medicinal compound synthesis) using the codons is expected to enhance the expression efficiency, providing technical support for *A. oxyphylla* quality improvement and stress-resistant breeding.

Keywords: *Alpinia oxyphylla* Miq.; chloroplast genome; codon usage bias; optimal codon

DOI: 10.3969/j.issn.1000-2561.2025.11.003

叶绿体 (chloroplast) 作为光合作用的核心场所及多种中间代谢反应的参与者, 是植物与真核藻类细胞内不可或缺的细胞器^[1]。叶绿体具有半自主特性, 不仅主导光合作用并产生有机物, 还拥有独立的遗传物质复制、转录和翻译的能力^[2]。自 1986 年叶绿体基因组序列首次被解析后, 根据其结构保守性与母系遗传特征, 叶绿体基因组已经成为分子生物学领域探究物种演化和植物亲缘关系的重要工具^[3]。密码子 (codon) 作为遗传信息传递的组成单位, 是转录和翻译过程中的重要连接枢纽^[4], 在生物遗传信息传递中起着重要作用。除甲硫氨酸 (Met/AUG) 和色氨酸 (Trp/UGG) 在标准遗传密码表中只对应单一密码子外, 其他氨基酸均由 2 个或 2 个以上的同义密码子共同编码^[5]。蛋白质合成过程中存在一定的偏好性, 同义密码子并非等概率使用, 细胞会显著偏向于高频利用某些特定的密码子 (最优密码子 optimal codon), 这种现象被称为密码子使用偏倚 (codon usage bias, CUB)^[6]。CUB 的形成受多种因素调控, 例如基因表达丰度^[7]、基因编码区长度^[8]以及 DNA 复制起始位点^[9]等, 但突变压力与自然选择通常被认为是影响 CUB 形成的主要机制^[10-13]。因此, 深入解析密码子偏好性对于阐明物种进化历程及其系统分类关系具有重要的理论价值和应用意义^[14-15]。

益智 (*Alpinia oxyphylla* Miq.) 作为姜科山姜属多年生草本植物, 是我国应用广泛的传统中药材, 与槟榔、巴戟天和砂仁并称“四大南药”。其药用部位为成熟干燥的果仁, 习称“益智仁”。益智的主产区位于我国海南、广东及广西等省区, 其中海南为道地药材产区^[16]。解析益智密码子使用偏好性及最优密码子的确定, 可加深对其遗传

调控机制的认识, 为后续基因组优化工程及目标基因的分子设计、增强目标基因在益智体内的表达效能与稳定性奠定理论基础和技术支撑。

1 材料与方法

1.1 益智叶绿体基因组序列获取

本研究涉及的益智叶绿体基因组信息, 数据来源于 NCBI (<https://www.ncbi.nlm.nih.gov/>) 数据库, GenBank 登录号: KY985237.1, 益智样本为海南野生资源^[17]。该基因组全长 161 351 bp, 经自动化注释共收录 92 条蛋白编码序列 (CDS)。为避免冗余或低质量序列对后续系统发育及选择压力分析造成偏倚, 本研究借鉴田甜等^[18]提出的 CDS 过滤策略, 对初始数据实施多步质控: (1) 剔除重复序列; (2) 排除长度不足 300 bp 的序列; (3) 去除含有异常终止密码子的序列; (4) 剔除内部携带提前终止密码子的序列; (5) 剔除起始密码子为非标准 ATG 的序列。经上述筛选, 最终保留 51 条高质量、完整性良好的 CDS 作为本研究后续比较基因组学与进化分析的核心数据集。

1.2 密码子偏好性指标分析

基于 EMBOSS 在线平台的 CUSP 模块 (<https://www.bioinformatics.nl/emboss-explorer/>), 本研究对前期筛选的 51 条益智叶绿体高质量 CDS 序列进行密码子组成解析, 用以系统评估其密码子使用模式。分析内容包括计算密码子第一位 (GC₁)、第二位 (GC₂)、第三位 (GC₃) 位点及基因组整体 GC 含量 (GC_{all})^[19]。利用 CodonW 1.4.2 软件定量评估密码子偏好性, 生成 5 项核心指标: (1) 相对同义密码子使用度 (relative synonymous codon usage, RSCU), 该指标通过衡量特定密码子的实际使用频率与其理论期望频

率的比值,量化密码子使用的偏好程度。当 $RSCU = 1$ 时,表明该密码子无使用偏好; $RSCU > 1$ 表示该密码子属于高频使用; $RSCU < 1$ 则表示该密码子属于低频使用类型^[20]。(2) 密码子适应指数 (codon adaptation index, CAI), 该指数通过分析目标基因与高表达基因参照组在密码子使用模式上的相似程度,推断基因表达效率。其值域为 0~1, 数值增大指示密码子偏好性增强,通常也意味着更高的基因表达水平^[21]。(3) 有效密码子数 (effective number of codon, ENC), 是评估密码子偏好性强弱的核心参数,其数值与密码子偏好性呈负相关关系,较低的 ENC 值反映更强的密码子偏好性,常关联高表达基因;而较高的 ENC 值表明密码子使用趋于随机化,通常对应较低的基因表达水平^[22]。(4) 密码子偏好性指数 (codon bias index, CBI), 量化基因中最优密码子的富集水平^[23]。(5) 最优密码子使用频率 (frequency of optimal codons, Fop), 通过统计高表达基因集中最优密码子的使用比例,用以评估基因密码子优化程度^[23]。

采用 Pearson 相关检验 (双侧) 进行相关性分析;使用 SPSS 18.0 软件进行数据分析,使用 Origin 2021 软件制图。

1.3 中性绘图分析

本研究以密码子第三位碱基 GC 含量 (GC_3) 作为横坐标,以第一、二位碱基 GC 平均含量 (GC_{12}) 为纵坐标构建散点图,并进行线性回归以探究影响因素。若基因点集在对角线呈聚集分布,则表示突变压力主导密码子偏好性;若基因点呈现偏离对角线的离散特征,则反映自然选择是密码子使用模式的主要因素^[24-25]。在回归分析中:若回归系数趋近于 1,代表 GC_3 与 GC_{12} 存在显著相关性,表明密码子偏好性的主要决定因素是突变因素;反之,当回归系数逼近 0 时,则提示自然选择效应占主导地位。

1.4 ENC-plot 绘图分析

ENC 是评估基因密码子使用偏好性的关键指标之一,用于量化基因中实际使用的密码子的多样性。其理论值取值范围为 20~61,当 ENC 接近 20 时,表明基因偏好使用单一密码子,反映出极显著的密码子偏好性;当 ENC 接近 61 时,表示基因对各同义密码子的使用近乎均等,反应密码子使用模式接近随机分布,无明显偏好性^[26]。本研究采用 ENC-plot 分析方法,通过图形化展示密

码子偏好性特征并评估突变因素的影响。利用 Origin 2021 软件构建散点图:以 GC_3 值为横坐标,ENC 期望值为纵坐标,每个数据点对应益智叶绿体基因组中的一个基因。其中,标准曲线依据以下公式^[22]计算生成:

$$ENC = \frac{2 + GC_3 + 29}{GC_3^2 + (1 - GC_3)^2}$$

通过比较基因点与标准曲线的相对位置判断影响因素:若多数基因点分布于标准曲线上方或紧邻曲线分布,则表明密码子偏好性主要受突变因素驱动;反之,当基因点显著偏离标准曲线并主要位于其下方时,则提示自然选择等进化压力可能参与了密码子使用模式的形成^[27]。

1.5 PR2-plot 偏倚分析

PR2-plot 偏倚分析是通过计算各基因密码子第三位碱基 A、T、G、C 的相对含量,绘制以 $G_3/(G_3+C_3)$ 为 x 轴、 $A_3/(A_3+T_3)$ 为 y 轴的散点图 (每个散点对应单个基因) 进行分布解析。图中坐标系原点 (0.5, 0.5) 对应 A/T 平衡且 C/G 平衡的理论中性状态,位于此区域的基因未呈现明显密码子使用偏好性,其空间分布特征提示第三位碱基组成主要受基因组相关的突变压力调控;而偏离中心点的散点分布则从距离和方向 2 个维度量化了密码子第三位碱基的使用偏好性强度与模式。当散点显著偏离中心区域时,其空间位移特征可直接反映密码子偏好性形成过程中突变压力与自然选择等多种进化机制的共同作用^[28-29]。

1.6 最优密码子的确定

本研究以 ENC 值作为基因表达水平的度量依据,筛选益智叶绿体基因组的最优密码子。具体流程为:对全部叶绿体基因的 ENC 值进行排序,根据极端表达特征分组,分别从高表达组 (ENC 最低为 10%) 和低表达组 (ENC 最高为 10%) 提取目标基因 (高表达组基因包括 *rps18*、*rps8*、*ndhE*、*atpF*、*rpl14*, 低表达组基因包括 *petA*、*rpl20*、*ycf2*、*clpP*、*ndhC*) 用于后续分析。基于 CodonW 1.4.2 软件,分别计算 2 组基因的同义密码子相对使用频率 (RSCU)。通过构建差异分析模型,以高表达组与低表达组的平均 RSCU 差值 ($\Delta RSCU$) 作为筛选参数,建立最优密码子判定标准体系:当 $RSCU > 1$,同时 $\Delta RSCU \geq 0.08$ 时,确定其为候选的最优密码子;若同一氨基酸存在多个候选密码子,则进一步比较其在高表达组中的 RSCU 值,选取最大值对应的密码子作为该氨

氨酸的最优密码子^[30-31]。

2 结果与分析

2.1 益智叶绿体基因组中密码子使用方式分析

2.1.1 密码子碱基组成特征分析 为探究益智叶绿体基因组密码子碱基组成特征，通过使用 CodonW 1.4.2 软件结合 CUSP 在线分析工具，对 51 条益智叶绿体基因的 GC₁、GC₂、GC₃ 及 GC_{all} 含量进行定量分析，并同步计算 ENC 值以评估密码子偏好性强度。结果（表 1）显示：基因组整体 GC 含量（GC_{all}）范围为 35.56%~56.15%，平均值为 46.34%，其中，34 个基因（占总数的 66.67%）大于 45%。密码子第三位碱基组成分析中，GC₁、GC₂ 和 GC₃ 分布范围分别为 30.5%~56.5%、26.2%~57.6%、16.6%~50.4%，其均值呈现梯度差异[GC₁（45.69%）>GC₂（38.79%）>GC₃（27.50%）]。GC₁ 和 GC₂ 均值显著高于 GC₃，且

三者均低于 50%，这表明益智叶绿体基因显著偏好以 A/U 作为密码子终止密码子，且整体表现出强烈的 A/U 碱基组成偏倚特征。

ENC 值作为评价指标，可以反映密码子偏好性程度^[32]。本研究中益智叶绿体基因组的 51 个基因的 ENC 值分布区间为 36.56~56.15，均值为 46.56，其中 67.78%（34/51）的基因 ENC 值高于 45，表明益智叶绿体基因密码子偏好性整体较弱。进一步分析显示：CAI 指标（0.11~0.30，均值为 0.17）数值偏低，表明基因密码子使用模式与高表达参考基因集相似度较低；CBI 指标（-0.26~0.16，均值为-0.12）显示最优密码子实际使用频率低于碱基组成预测；Fop 值（0.25~0.51，均值为 0.34）处于较低区间，反映高表达基因中最优密码子占比较少。综合分析结果，益智叶绿体基因组密码子使用偏好性整体呈现弱化特征。

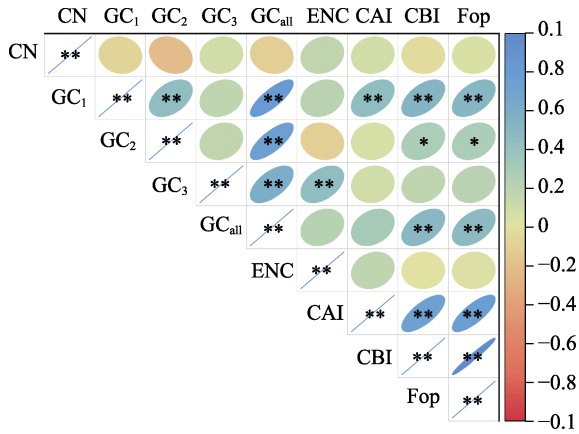
表 1 益智叶绿体基因组密码子碱基组成特征
Tab. 1 Codon nucleotide composition characteristics of the chloroplast genome of *A. oxyphylla* Miq.

基因 Gene	第 1 位碱 基含量 GC ₁ content/%	第 2 位碱 基含量 GC ₂ content/%	第 3 位碱 基含量 GC ₃ content/%	总碱基含量 GC _{all} content/%	有效密码子数 ENC content	密码子 适应指数 CAI	密码子 偏好指数 CBI	最优密码子 使用频率 Fop
<i>accD</i>	41.3	35.6	22.1	33.1	42.28	0.203	-0.196	0.323
<i>atpA</i>	55.7	39.6	26.6	40.7	46.25	0.195	-0.044	0.390
<i>atpB</i>	56.0	42.1	28.0	42.1	47.45	0.191	-0.039	0.387
<i>atpE</i>	52.2	40.4	29.4	40.7	47.25	0.160	-0.045	0.382
<i>atpF</i>	48.7	31.9	29.2	36.6	40.18	0.136	-0.170	0.313
<i>atpI</i>	48.4	36.3	25.0	36.7	44.89	0.174	-0.040	0.370
<i>ccsA</i>	30.5	38.6	27.2	32.1	47.55	0.145	-0.216	0.294
<i>cemA</i>	41.4	26.2	28.7	32.2	47.73	0.188	-0.058	0.374
<i>clpP</i>	53.4	36.9	32.6	41.1	54.61	0.145	-0.173	0.306
<i>matK</i>	37.2	29.3	23.4	30.0	48.24	0.148	-0.241	0.270
<i>ndhA</i>	44.1	39.7	22.1	35.4	42.03	0.124	-0.127	0.311
<i>ndhB</i>	41.5	39.4	31.5	37.5	46.35	0.161	-0.104	0.344
<i>ndhC</i>	48.8	34.7	31.4	38.6	56.15	0.219	-0.026	0.378
<i>ndhE</i>	39.2	32.3	26.4	33.0	39.21	0.130	-0.264	0.247
<i>ndhF</i>	35.5	37.0	19.8	30.8	40.64	0.134	-0.202	0.289
<i>ndhG</i>	40.7	35.0	21.4	32.6	44.02	0.140	-0.249	0.249
<i>ndhH</i>	49.3	35.0	27.4	37.2	51.42	0.160	-0.104	0.347
<i>ndhI</i>	39.8	35.9	23.8	33.3	49.64	0.181	-0.188	0.312
<i>ndhJ</i>	48.1	35.1	29.4	37.7	50.52	0.170	-0.180	0.307
<i>ndhK</i>	42.2	40.6	25.9	36.4	48.87	0.157	-0.167	0.315
<i>orf109</i>	39.1	32.7	30.9	34.6	45.31	0.136	-0.224	0.280
<i>petA</i>	54.5	35.5	28.1	39.4	52.18	0.188	-0.093	0.352
<i>petD</i>	48.6	40.0	28.6	39.3	45.70	0.178	-0.079	0.335
<i>psaA</i>	52.1	42.7	31.8	42.3	48.37	0.201	-0.077	0.369
<i>psaB</i>	48.4	43.0	29.6	40.4	46.37	0.181	-0.108	0.354
<i>psbA</i>	49.2	43.5	32.4	41.8	44.58	0.296	0.158	0.511

续表 1 益智叶绿体基因组密码子碱基组成特征
Tab. 1 Codon nucleotide composition characteristics of the chloroplast genome of *A. oxyphylla* Miq. (continued)

基因 Gene	第 1 位碱 基含量 GC ₁ content/%	第 2 位碱 基含量 GC ₂ content/%	第 3 位碱 基含量 GC ₃ content/%	总碱基含量 GC _{all} content/%	有效密码子数 ENC content	密码子 适应指数 CAI	密码子 偏好指数 CBI	最优密码子 使用频率 Fop
<i>psbB</i>	54.0	46.4	29.5	43.3	47.04	0.182	-0.076	0.370
<i>psbC</i>	53.2	46.4	32.9	44.2	45.25	0.179	-0.063	0.373
<i>psbD</i>	52.6	43.2	28.0	41.4	41.78	0.251	0.041	0.441
<i>rbcL</i>	56.5	42.8	29.1	42.9	47.47	0.264	0.057	0.457
<i>rpl14</i>	53.6	36.6	25.2	38.8	40.56	0.180	0	0.403
<i>rpl20</i>	35.0	42.0	30.1	35.7	52.42	0.135	-0.139	0.338
<i>rpl22</i>	34.1	31.9	19.7	28.5	44.86	0.160	-0.129	0.341
<i>rpoA</i>	44.0	34	22.5	33.6	50.01	0.138	-0.192	0.298
<i>rpoB</i>	48.7	38.2	26.5	37.8	46.33	0.155	-0.124	0.342
<i>rpoC1</i>	48.6	38.5	23.9	37.1	45.93	0.149	-0.121	0.336
<i>rpoC2</i>	44.1	36.0	24.6	34.9	46.59	0.147	-0.182	0.309
<i>rps11</i>	54.7	57.6	16.6	43.0	41.01	0.140	-0.158	0.316
<i>rps12</i>	52.6	48.3	23.8	41.6	44.89	0.136	-0.086	0.345
<i>rps14</i>	39.6	45.5	28.7	38.3	41.07	0.125	-0.170	0.313
<i>rps18</i>	32.1	41.9	27.7	33.9	35.56	0.110	-0.164	0.312
<i>rps2</i>	42.2	42.6	31.7	38.8	47.02	0.180	-0.114	0.347
<i>rps3</i>	42.9	31.6	24.2	33.0	44.40	0.145	-0.189	0.310
<i>rps4</i>	46.6	38.6	24.2	36.7	48.31	0.153	-0.047	0.374
<i>rps7</i>	51.3	45.5	22.5	40.0	46.04	0.189	-0.092	0.369
<i>rps8</i>	38.4	38.4	22.5	33.1	36.64	0.124	-0.030	0.391
<i>ycf1</i>	36.0	27.8	25.9	29.9	47.34	0.173	-0.118	0.356
<i>ycf2</i>	41.7	34.4	36.9	37.7	52.52	0.157	-0.137	0.338
<i>ycf3</i>	47.3	39.7	29.6	39.1	51.76	0.158	-0.199	0.321
<i>ycf4</i>	44.8	41.6	32.9	40.0	49.61	0.162	-0.033	0.385
<i>ycf68</i>	49.6	49.6	50.4	50.3	51.29	0.106	-0.149	0.319

2.1.2 相关性分析 为进一步解析益智叶绿体基因组密码子参数间的内在关联,本研究对各参数指标开展系统性相关性分析(图 1)。结果表明:GC_{all}与 GC₁、GC₂、GC₃均呈极显著正相关($P<0.01$),表明 GC_{all}的数值分布受到密码子 3 个位点 GC 含量的协同作用;GC₁与 GC₂呈极显著正相关($P<0.01$),表明密码子第 1 位与第 2 位碱基组成相似性较高,GC₁、GC₂均与 GC₃无显著相关性,表明第 3 位碱基与二者差异显著;ENC 与 GC₃呈极显著正相关($P<0.01$),但与 GC₁、GC₂无相关性,说明密码子使用偏好性主要受第 3 位碱基组成的调控;CAI 与 GC_{all}呈显著相关($P<0.05$),CAI 与 GC₁呈极显著相关($P<0.01$),CBI 和 Fop 分别与 GC_{all}、GC₁呈极显著正相关($P<0.01$);GC₂与 CBI、Fop 亦呈显著正相关($P<0.05$);密码子数(CN)与其他参数均无显著关联,表明密码子碱基组成对密码子数量分布的影响较小。



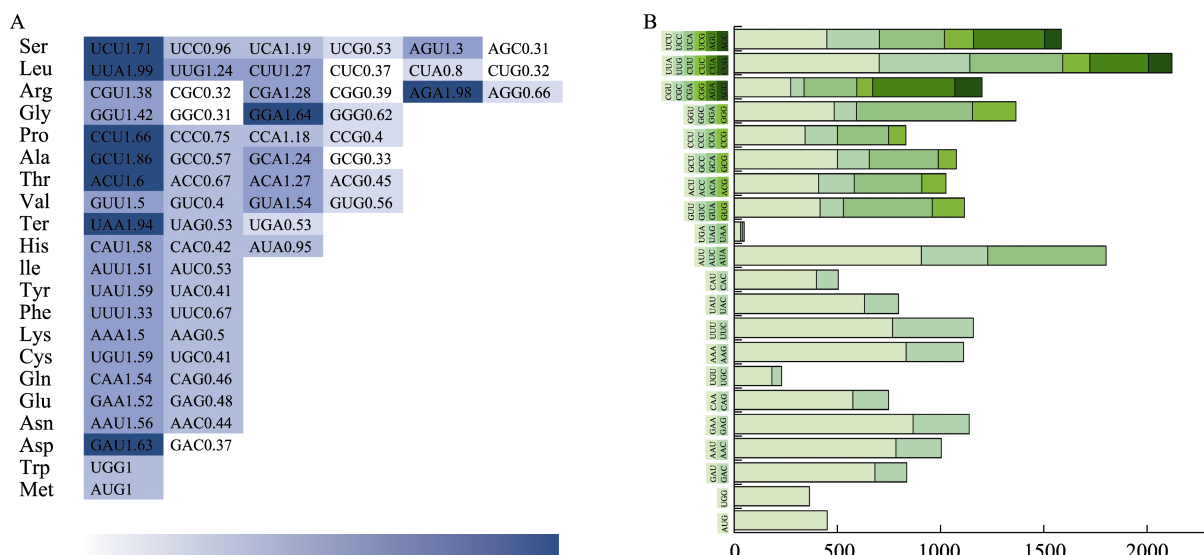
*表示显著相关($P<0.05$), **表示极显著相关($P<0.01$)。
* indicates significant correlation ($P<0.05$), ** indicates extremely significant correlation ($P<0.01$).

图 1 益智叶绿体基因组密码子中各参数的相关性分析
Fig. 1 Correlation analysis of parameters in *A. oxyphylla* Miq. chloroplast genome codon

2.1.3 相对同义密码子 RSCU 分析 RSCU 分析(图 2)显示,亮氨酸(Leu)的密码子 UUA 的 RSCU 值最高(1.99)。全基因组中共有 32 个密

码子的 RSCU 大于 1, 表明此类密码子存在使用偏好性, 可称为高频密码子。其中, 以 A 结尾的高频密码子占 40.6% (13/32), 以 U 结尾的密码子占 50.0% (16/32), 以 G 结尾的密码子占 9.4% (3/32), A 或 U 结尾的高频密码子合计占 90.6%。益智叶绿体基因组共包含 20 482 个密码子, 共编码 21 种氨基酸, 其中, 异亮氨酸 (Ile)

的密码子 AUU 使用频率最高, 出现次数为 906 次; 谷氨酸 (Glu) 的密码子 GAA 次之, 频次为 866 次。而终止密码子 (UAA、UAG、UGA) 的使用频率显著低于其他密码子, 出现次数分别为 33 次、9 次和 9 次。综合以上结果可知, 益智叶绿体基因组对以 A 或 U 结尾的密码子表现出显著偏好性。



A: 益智各氨基酸的同义密码子 RSCU 值; B: 益智各氨基酸的同义密码子数目。

A: RSCU values of synonymous codons for each amino acid in *Alpinia oxyphylla* Miq.; B: Number of synonymous codons for each amino acid in *Alpinia oxyphylla* Miq.

图2 益智各氨基酸的同义密码子RSCU值及数目

Fig. 2 Synonymous codon RSCU values and counts for each amino acid in *A. oxyphylla* Miq.

2.1.1.4 叶绿体基因结构分析 对 51 个叶绿体基因进行功能分类, 将其划分为六大类: 能量代谢相关基因、光合作用基因、遗传信息相关基因、代谢与修饰相关酶基因、膜运输与电子转运基因以及其他功能基因(表 2)。能量代谢相关基因中的 NADH 脱氢酶亚基基因数量最多, 该基因编码类囊体膜上的 NDH 复合体, 通过驱动环式电子传递合成 ATP 并减轻光损伤, 是能量平衡与胁迫响应的关键^[32]。其次为遗传信息相关基因中的核糖体小亚基蛋白基因, 其产物参与组装叶绿体 70S 核糖体, 在基质中执行叶绿体自身编码蛋白的翻译, 维持细胞器的半自主性^[33]。

2.2 益智叶绿体基因组密码子的中性绘图分析

基于中性绘图分析, 通过构建同义密码子 GC₁₂与 GC₃的散点图, 结合回归分析其相关性, 当回归系数接近 0 时, 表明密码子使用偏好性主

要受自然选择驱动；若回归系数接近 1，则表明主要由突变压力驱动^[18]。结果显示（图 3），GC₃ 含量分布区间为 16.60%~50.40%，GC₁₂ 含量分布区间为 38.45%~59.60%，二者回归系数为 0.2269。GC₁₂ 与 GC₃ 的相关系数为 0.0484，未达到统计学显著水平，表明益智叶绿体密码子偏好性形成主要受自然选择压力驱动，而非突变随机性影响。

2.3 益智叶绿体基因组密码子的 ENC-plot 绘图分析

益智叶绿体密码子基因组中多数基因的 ENC 值超过 35, 这初步表明益智整体密码子使用偏好性较弱。ENC-plot 绘图分析 ENC 值的分布特征及其与期望值之间的关系 (图 4), 结果表明益智基因分布分散, 少数基因的实际 ENC 与期望曲线接近, 多数基因的实际 ENC 值显著偏离期望曲线, 且主要分布在期望曲线下方, 表明多数益智叶绿体基因使用偏好性受到自然选择的影响。

表 2 益智叶绿体基因结构分析
Tab. 2 Structural analysis of chloroplast genome of *A. oxyphylla* Miq.

基因分类 Gene classification	基因分组 Gene grouping	基因名称 Gene name
能量代谢相关基因 Energy metabolism related genes	ATP 合成酶亚基 Subunits of ATP synthase	<i>atpA</i> 、 <i>atpB</i> 、 <i>atpE</i> 、 <i>atpF</i> 、 <i>atpI</i>
	NADH 脱氢酶亚基 Subunits of NADH dehydrogenase	<i>ndhA</i> 、 <i>ndhB</i> 、 <i>ndhC</i> 、 <i>ndhE</i> 、 <i>ndhF</i> 、 <i>ndhG</i> 、 <i>ndhH</i> 、 <i>ndhI</i> 、 <i>ndhJ</i> 、 <i>ndhK</i>
光合作用基因 Photosynthesis genes	细胞色素 b/f 复合体亚基 Subunits of cytochrome b/f complex	<i>petA</i> 、 <i>petD</i>
	光合系统 I 亚基 Subunits of photosystem I	<i>psaA</i> 、 <i>psaB</i>
	光合系统 II 亚基 Subunits of photosystem II	<i>psbA</i> 、 <i>psbB</i> 、 <i>psbC</i> 、 <i>psbD</i>
	RuBisCO 酶大亚基 Large subunit of RuBisCo	<i>rbcL</i>
遗传信息相关基因 Genetic information related genes	核糖体大亚基蛋白 Proteins of large ribosomal subunit	<i>rpl14</i> 、 <i>rpl20</i> 、 <i>rpl22</i>
	核糖体小亚基蛋白 Proteins of small ribosomal subunit	<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>
代谢与修饰相关酶基因 Metabolism and modification related enzyme genes	RNA 聚合酶亚基 Subunits of RNA polymerase	<i>rpoA</i> 、 <i>rpoB</i> 、 <i>rpoC1</i> 、 <i>rpoC2</i> 、 <i>accD</i>
	乙酰辅酶 A 羧化酶 Acetyl CoA carboxylase	
	蛋白酶 Protease	<i>clpP</i>
	成熟酶 Maturase	<i>matK</i>
膜运输与电子转运基因 Membrane transport and electron transfer genes	细胞色素 c 基因 c-type cytochrome synthesis gene	<i>ccsA</i>
	包膜膜蛋白 Envelope membrane protein	<i>cemA</i>
其他功能基因 Other functional genes	开放阅读框 Open Reading Frame	<i>orf109</i>
	保守假设性叶绿体开放阅读框 Conserved hypothetical chloroplast ORF	<i>ycf1</i> 、 <i>ycf2</i> 、 <i>ycf3</i> 、 <i>ycf4</i> 、 <i>ycf68</i>

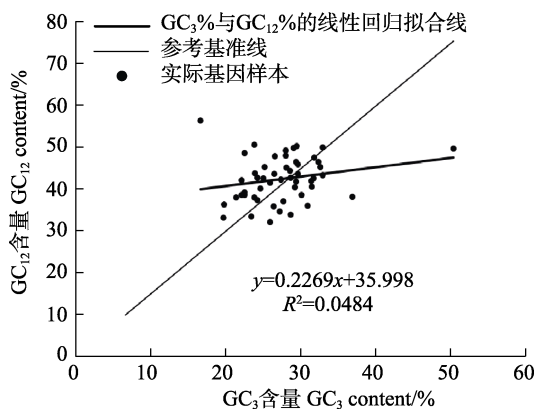


图 3 益智叶绿体基因组密码子使用的中性绘图分析
Fig. 3 Neutral plot analysis of codon usage in chloroplast genome of *A. oxyphylla* Miq.

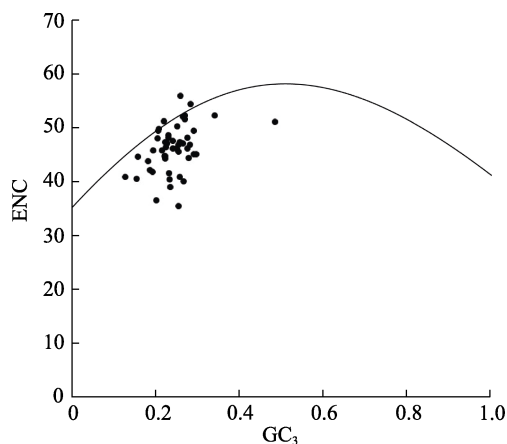


图 4 益智叶绿体基因组 ENC-plot 分析
Fig. 4 ENC-plot analysis of chloroplast genomes of *A. oxyphylla* Miq.

2.4 益智叶绿体基因组密码子的 PR2-plot 偏倚分析

PR2-plot 偏倚 (图 5) 分析显示, 候选基因在 PR2-plot 的 4 个象限中分布不均匀。当仅有少量基因位于图中轴线上时, 表明其密码子使用偏性主要受突变压力影响^[34]。益智叶绿体基因组中大多数基因分布于 y 轴中心线的上方, 表明密码子第三位腺嘌呤 (A) 的使用频率显著高于胸腺嘧啶 (T)。同时, 在 x 轴水平方向上, 位于左侧的基因数量明显多于右侧, 反映出密码子第三位胞嘧啶 (C) 的使用频率高于鸟嘌呤 (G)。候选基因在 4 个象限均有分布, 表明突变压力与选择压力共同作用于益智叶绿体密码子的使用偏好性。

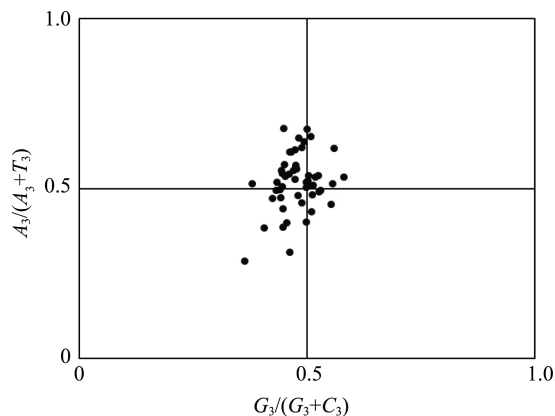


图 5 益智叶绿体基因组 PR2-plot 分析
Fig. 5 PR2-plot analysis in chloroplast genome of *A. oxyphylla* Miq.

2.5 益智叶绿体基因组最优密码子的确定

为鉴定益智叶绿体基因组的最优密码子，本研究通过构建益智叶绿体基因组高表达库和低表达库，分别计算 2 组基因的相对同义密码子使用度（RSCU）及差异值（ Δ RSCU），筛选高频密码子与高表达密码子。结果（表 3）表明：益智叶绿体基因组中共鉴定出 28 个高频密码子

（RSCU>1）和 25 个具有高表达优越性的密码子（ Δ RSCU> 0.08）。取交集后，最终确定 10 个密码子为最优密码子，分别为：CUU、GUU、GUA、UCA、CCU、ACA、GCA、UAU、UGU、AGA。经分析发现这 10 个最优密码子均以 A 或 U 碱基结尾，表明益智叶绿体基因组的最优密码子是显著倾向于使用以 A/U 结尾的同义密码子。

表 3 益智叶绿体基因组最优密码子分析
Tab. 3 Analysis of optimal codons in *A. oxyphylla* Miq. chloroplast genome

氨基酸 Amino acid	密码子 Codon	高表达基因 High-expressed gene		低表达基因 Low-expressed gene		同义密码子 相对使用度 Δ RSCU
		频率 Frequency	同义密码子相对使用度 RSCU	频率 Frequency	同义密码子相对使用度 RSCU	
苯丙氨酸 Phe	UUU	45	1.22	19	1.46	-0.24
	UUC*	29	0.78	7	0.54	0.24
亮氨酸 Leu	UUA	19	1.12	25	2.27	-1.15
	UUG	28	1.65	20	1.82	-0.17
	CUU***	31	1.82	11	1.00	0.82
	CUC*	10	0.59	5	0.45	0.14
	CUA	6	0.35	4	0.36	-0.01
	CUG**	8	0.47	1	0.09	0.38
异亮氨酸 Ile	AUU	51	1.47	42	1.68	-0.21
	AUC*	24	0.69	11	0.44	0.25
	AUA	29	0.84	22	0.88	-0.04
蛋氨酸 Met	AUG	31	1.00	16	1.00	0
缬氨酸 Val	GUU*	36	1.55	13	1.44	0.11
	GUC	9	0.39	4	0.44	-0.05
	GUA*	36	1.55	13	1.44	0.11
	GUG	12	0.52	6	0.67	-0.15
	UCU	25	1.50	11	1.47	0.03
丝氨酸 Ser	UCC	15	0.9	8	1.07	-0.17
	UCA***	24	1.44	3	0.40	1.04
	AGU	15	0.90	13	1.73	-0.83
	AGC*	9	0.54	2	0.27	0.27
	UCG	12	0.72	8	1.07	-0.35
	CCU**	20	1.54	4	1.23	0.31
脯氨酸 Pro	CCC	15	1.15	4	1.23	-0.08
	CCA	11	0.85	5	1.54	-0.69
	CCG**	6	0.46	0	0	0.46
	ACU	13	1.37	19	2.53	-1.16
苏氨酸 Thr	ACC*	6	0.63	3	0.40	0.23
	ACA***	11	1.16	4	0.53	0.63
	ACG**	8	0.84	4	0.53	0.31
	GCU	23	1.74	17	1.89	-0.15
丙氨酸 Ala	GCC	5	0.38	8	0.89	-0.51
	GCA***	22	1.66	7	0.78	0.88

续表 3 益智叶绿体基因组最优密码子分析
Tab. 3 Analysis of optimal codons in *A. oxyphylla* Miq. chloroplast genome (continued)

氨基酸 Amino acid	密码子 Codon	高表达基因 High-expressed gene		低表达基因 Low-expressed gene		同义密码子 相对使用度 Δ RSCU
		频率 Frequency	同义密码子相对使用度 RSCU	频率 Frequency	同义密码子相对使用度 RSCU	
酪氨酸 Tyr	GCG	3	0.23	4	0.44	-0.21
	UAU *	33	1.69	10	1.43	0.26
	UAC	6	0.31	4	0.57	-0.26
组氨酸 His	<u>CAU</u>	14	1.33	8	1.78	-0.45
	CAC**	7	0.67	1	0.22	0.45
谷氨酸 Gln	<u>CAA</u>	36	1.50	22	1.69	-0.19
	CAG*	12	0.50	4	0.31	0.19
天冬酰胺 Asn	<u>AAU</u>	45	1.55	37	1.57	-0.02
	AAC	13	0.45	10	0.43	0.02
赖氨酸 Lys	<u>AAA</u>	44	1.21	32	1.52	-0.31
	AAG**	29	0.79	10	0.48	0.31
天冬氨酸 Asp	<u>GAU</u>	41	1.58	18	1.64	-0.06
	GAC	11	0.42	4	0.36	0.06
谷氨酰胺 Glu	<u>GAA</u>	48	1.41	29	1.57	-0.16
	GAG*	20	0.59	8	0.43	0.16
半胱氨酸 Cys	UGU *	8	1.78	3	1.50	0.28
	UGC	1	0.22	1	0.50	-0.28
色氨酸 Trp	UGG	25	1.00	2	1.00	0
精氨酸 Arg	<u>CGU</u>	14	1.04	13	1.15	-0.11
	CGC	5	0.37	3	0.26	0.11
	<u>CGA</u>	15	1.11	27	2.38	-1.27
	CGG**	10	0.74	4	0.35	0.39
	AGA *	24	1.78	17	1.50	0.28
	AGG***	13	0.96	4	0.35	0.61
	<u>GGU</u>	23	1.31	13	1.44	-0.13
	GGC	4	0.23	2	0.22	0.01
甘氨酸 Gly	<u>GGA</u>	30	1.71	18	2.00	-0.29
	GGG**	13	0.74	3	0.33	0.41

注: *表示 $0.3 > \Delta \text{RSCU} \geq 0.08$, **表示 $0.5 > \Delta \text{RSCU} \geq 0.3$, ***表示 $\Delta \text{RSCU} \geq 0.5$, 下划线表示该叶绿体中基因组密码子的 RSCU 值 > 1 , 加粗字体标注的密码子为最优密码子。
Note: * Indicates $0.3 > \Delta \text{RSCU} \geq 0.08$, ** indicates $0.5 > \Delta \text{RSCU} \geq 0.3$, *** indicates $\Delta \text{RSCU} \geq 0.5$, the underlined codon indicates the RSCU value of the codon in the chloroplast genome is greater than 1, and the optimal codons are shown in bold.

3 讨论

近年来,随着植物基因组学的迅猛发展,大量的基因组数据被更新完善,这为密码子偏好性的深入探究奠定坚实基础。众所周知,DNA 由 4 种碱基构成,其特定的组成模式会对密码子的偏好性产生关键影响^[29]。作为遗传密码构成单位的密码子,在生物体内信息传递的进程中扮演着关键性角色。特别是在功能性基因的转录活动、翻译过程及其表达调控方面,在密码子使用偏好现象上也发挥着重要作用。基因转录准确度与效率、

mRNA 分子稳定性及翻译均受其显著影响,蛋白质合成速率与最终产物的数量并由此产生显著变化^[33]。密码子的偏好性受到多种因素的综合作用,自然选择是其中的重要因素之一,可以筛选出适应性强的相关基因;此外,基因突变的随机性和累积效应,也会不断扰动密码子偏好性的平衡,使其在漫长的进化历程中不断调整与重塑^[35-36]。
叶绿体基因组因其高度保守的结构和相对稳定的组成,已成为研究植物进化与适应机制的重要系统^[37]。叶绿体基因组密码子使用偏好性分析具

有多重意义,不仅有助于在分子层面阐明物种的起源、演化及其环境适应机制,同时为实现外源基因在叶绿体中的高效表达奠定理论基础^[38-39]。此外,叶绿体基因组密码子偏好性模式代表关键的进化信号,是解析物种间系统发育关系、探究基因进化规律及功能的重要依据^[40]。在植物叶绿体基因组密码子研究中,GC 含量发生的变化,被视为基因序列演化的重要驱动因子之一。鉴于密码子第 3 位作为简并位点,其碱基替换通常不改变所编码的氨基酸,因此,深入解析该位点的碱基组成特征及其动态变化规律,对于全面阐明植物叶绿体密码子使用偏好性的形成机制与进化动力具有不可替代的重要意义^[10, 33]。

本研究对益智叶绿体基因组密码子偏好性进行系统分析,首先,共筛选出 51 条 CDS 序列,明确其密码子的使用特征。对益智叶绿体基因组编码序列开展整体性密码子组成特征分析,得出如下结果:密码子的 GC 含量均低于 50%,呈现 $GC_1 > GC_2 > GC_3$ 。这一现象揭示益智密码子不同碱基位点的 GC 含量存在显著差异。此外,以 A 或 U 结尾的密码子类型呈现偏好性使用趋势在益智叶绿体基因组分析过程中被发现。密码子偏好以 A/U 结尾且 GC_3 含量最低的研究结果,与小黑麦 (*Triticale*)^[18]、甜龙竹 (*Dendrocalamus brandisii*)^[41]、豇豆 (*Vigna unguiculata*)^[42]、伊犁郁金香 (*Tulipa iliensis* Regel)^[37]、银白杨 (*Populus alba*)^[43] 等植物中的研究结果一致。这种密码子使用的偏好模式的形成,可能源于植物叶绿体基因组在长期进化过程中密码子使用偏好性的高度保守性,同时也与高等植物叶绿体基因普遍富含 A/T 碱基密切相关。为进一步验证此偏好性,本研究进一步对益智叶绿体基因组进行 RSCU 分析,结果显示:在益智 32 个 $RSCU > 1$ 的密码子中,以 A/U 结尾的有 29 个,占密码子总数的 90.6%,该结果进一步证实叶绿体基因组倾向于 A/U 结尾的密码子。

在密码子使用偏好性的量化分析中,常采用 ENC 与 CAI 这 2 个指标值^[44]。本研究发现,益智叶绿体基因组中 ENC 均值为 46.56,其中 ENC 大于 45 的占比为 67.78%,该结果表明益智叶绿体基因组的密码子使用偏好性总体较弱。各基因的 CAI 值介于 0.11~0.30 之间,平均值为 0.17,进一步印证了其密码子使用偏好性总体较弱的结论。

中性绘图分析结果发现, GC_{12} 和 GC_3 之间相关性不显著,表明自然选择在塑造益智叶绿体基因组密码子使用偏好性中发挥着主导作用。ENC-plot 分析发现,益智叶绿体基因组的多数 ENC 理论值与实际值之间存在显著差异,这表明自然选择是主导益智叶绿体基因组密码子使用模式的关键因素,这与巨龙竹 (*Dendrocalamus sinicus*)^[19]、灯盏花 (*Erigeron breviscapus*)^[36] 等的研究结果一致。在 PR2-plot 分析中,多数基因点在 PR2-plot 中偏离中心区域且分布较为分散, GC_3 的碱基 A 的使用频率显著高于 T, C 的使用频率显著高于 G,这表示益智叶绿体基因组密码子使用偏好性既受到自然选择的显著影响,也受基因突变的影响。

4 结论

本研究揭示了益智叶绿体基因组密码子使用偏好性的规律,结果表明益智叶绿体基因组高频密码子普遍呈现以 A/U 结尾的偏好性。通过 RSCU 分析,共筛选出 10 个高频且高表达的最优密码子,且这 10 个密码子均以 A/U 结尾。此结果从功能层面验证了 GC_3 含量偏低及整体 A/U 偏好性的发现,揭示 A/U 结尾密码子在益智叶绿体基因表达中的重要地位。本研究针对偏好性形成机制展开探讨,多维分析方法得以整合运用,包括中性绘图技术、ENC-plot 分析手段以及 PR2-plot 检测工具等,系统探讨了突变压力和自然选择等因素的共同作用。结果表明,自然选择是主导密码子使用模式的关键力量,而突变压力亦发挥协同作用。这种多因素协同调控模型,为全面理解植物叶绿体基因组密码子偏好性的进化与调控机制提供新视角。综合以上研究,益智叶绿体基因组中筛选的 10 个以 A/U 结尾的最优密码子,可直接应用于益智叶绿体遗传转化体系的设计。通过定向改造外源基因的密码子,适配宿主偏好性,可显著提升目标蛋白(如药用活性成分合成酶)的表达效率,为益智的品质改良和抗逆性增强提供技术路径。同时,该研究成果不仅为益智这一重要药用和经济作物的叶绿体基因组编辑奠定数据基础,其揭示的 A/U 偏好性保守规律及多因素调控机制,亦可为其他药用植物(尤其姜科、单子叶类群)的叶绿体遗传工程提供理论参照。

参考文献

- [1] JODI M, GEIR M S. Plastid division: evolution, mechanism and complexity[J]. *Annals of Botany*, 2007, 99(4): 565-579.
- [2] XIONG A S, PENG R H, ZHUANG J, GAO F, ZHU B, FU X Y, XUE Y, JIN X F, TIAN Y S, ZHAO W, YAO Q H. Gene duplication, transfer, and evolution in the chloroplast genome[J]. *Biotechnology Advances*, 2009, 27(4): 340-347.
- [3] HUA Z Y, TIAN D M, JIANG C, SONG S H, CHEN Z Y, ZHAO Y Y, JIN Y, HUANG L Q, ZHANG Z, YUAN Y. Towards comprehensive integration and curation of chloroplast genomes[J]. *Plant Biotechnology Journal*, 2022, 20(12): 2239-2241.
- [4] 赵月梅, 徐其碧, 杨贵清, 尹鑫, 徐林. 艾纳香叶绿体基因组密码子使用偏性分析[J]. *西部林业科学*, 2023, 52(3): 55-62, 77.
ZHAO Y M, XU Q B, YANG G Q, YIN X, XU L. Codon usage bias of chloroplast genome in *Blumea balsamifera*[J]. *Journal of West China Forestry Science*, 2023, 52(3): 55-62, 77. (in Chinese)
- [5] 洪森荣, 张牧彤, 徐子林, 张钦荣, 罗雨欣, 田文慧, 王心雨. ‘怀玉山’高山马铃薯叶绿体基因组特征及密码子使用偏好性分析[J]. *浙江农林大学学报*, 2024, 41(1): 92-103.
HONG S R, ZHANG M T, XU Z L, ZHANG Q R, LUO Y X, TIAN W H, WANG X Y. Chloroplast genome characteristics and codon usage preference of *Solanum tuberosum* var. *cormosus* ‘Huaiyushan’[J]. *Journal of Zhejiang A&F University*, 2024, 41(1): 92-103. (in Chinese)
- [6] 李红, 周梦丽, 王霞, 郭蕾, 王瑞红. 油点草叶绿体基因组密码子使用偏好性分析[J]. *浙江理工大学学报(自然科学)*, 2025, 53(3): 377-384.
LI H, ZHOU M L, WANG X, GUO L, WANG R H. Analysis of codon usage bias in the chloroplast genome of *Tricyrtis macropoda*[J]. *Journal of Zhejiang Sci-Tech University (Natural Sciences Edition)*, 2025, 53(3): 377-384. (in Chinese)
- [7] 罗永坚, 王茹, 赵仁菲, 卢新雄, 尹广鹏, 邓志军. 珙桐叶绿体基因组同义密码子使用偏好性分析[J]. *北京林业大学学报*, 2024, 46(3): 8-16.
LUO Y J, WANG R, ZHAO R F, LU X X, YIN G K, DENG Z J. Analysis of synonymous codon usage bias in the chloroplast genome of *Davidia involucreata*[J]. *Journal of Beijing Forestry University*, 2024, 46(3): 8-16. (in Chinese)
- [8] 石元豹, 汪贵斌, 杨晓明, 曹福亮. 银杏 WRKY 家族基因密码子使用偏向性分析[J]. *分子植物育种*, 2019, 17(5): 1503-1511.
SHI Y B, WANG G B, YANG X M, CAO F L. Analysis of codon usage bias of gene factors in *Ginkgo biloba* WRKY Family[J]. *Molecular Plant Breeding*, 2019, 17(5): 1503-1511. (in Chinese)
- [9] 孙月琪, 李密密, 周义峰. 珊瑚菜叶绿体基因组密码子使用偏性分析[J]. *植物资源与环境学报*, 2023, 32(6): 1-10.
SUN Y Q, LI M M, ZHOU Y F. Analysis on codon usage bias of chloroplast genome of *Glehnia littoralis*[J]. *Journal of Plant Resources and Environment*, 2023, 32(6): 1-10. (in Chinese)
- [10] 吴妙丽, 陈世品, 陈辉. 竹亚科叶绿体基因组的密码子使用偏性分析[J]. *森林与环境学报*, 2019, 39(1): 9-14.
WU M L, CHEN S P, CHEN H. Condon preference of chloroplast genome of *Bambusoideae*[J]. *Journal of Forest and Environment*, 2019, 39(1): 9-14. (in Chinese)
- [11] 杨涌, 付丽枝, 王萌, 王旋, 刘静, 于立洋, 王东升, 路艺, 张海娥. 板栗叶绿体基因组密码子使用偏好性分析[J]. *江苏农业科学*, 2024, 52(21): 40-46.
YANG Y, FU L Z, WANG M, WANG X, LIU J, YU L Y, WANG D S, LU Y, ZHANG H E. Analysis of codon usage bias in chloroplast genome of *Castanea mollissima*[J]. *Jiangsu Agricultural Sciences*, 2024, 52(21): 40-46. (in Chinese)
- [12] 杨嘉鹏, 朱紫乐, 范雅娟, 朱菲, 陈粤珺, 牛志韬, 丁小余. 三种石豆兰属药用植物的叶绿体基因组比较分析及其在物种鉴定中的意义[J]. *药学报*, 2020, 55(11): 2736-2745.
YANG J P, ZHU Z L, FAN Y J, ZHU F, CHEN Y J, NIU Z T, DING X Y. Comparative plastomic analysis of three *Bulbophyllum* medicinal plants and its significance in species identification[J]. *Acta Pharmaceutica Sinica*, 2020, 55(11): 2736-2745. (in Chinese)
- [13] 尹明华, 范馨月, 付悦欣, 胡慧玲. 广丰千金薯叶绿体基因组序列特征及密码子使用偏好性分析[J]. *中草药*, 2024, 55(21): 7444-7461.
YIN M H, FAN X Y, FU Y X, HU H L. Analysis of chloroplast genome sequence characteristics and codon usage bias in *Dioscorea opposita* Thunb. cv. *Guangfeng Qianjinshu*[J]. *Chinese Traditional and Herbal Drugs*, 2024, 55(21): 7444-7461. (in Chinese)
- [14] 钟才荣, 何斯敏, 方赞山, 张颖. 使君子科 3 种红树植物叶绿体基因组密码子使用偏好性分析[J]. *植物资源与环境学报*, 2024, 33(2): 1-12.
ZHONG C R, HE S M, FANG Z S, ZHANG Y. Analysis on codon usage bias in chloroplast genomes of three mangrove in *Combretaceae*[J]. *Journal of Plant Resources and Environment*, 2024, 33(02): 1-12. (in Chinese)
- [15] 余志伟, 陈垣, 郭凤霞, 董鹏斌, 王红燕. 菰蓝叶绿体基因组密码子使用偏好性分析[J]. *草地学报*, 2024, 32(8): 2374-2385.
YU Z W, CHEN Y, GUO F X, DONG P B, WANG H Y. Codon usage bias analysis of the chloroplast genome of *Isatis Indigotica*[J]. *Acta Agrestia Sinica*, 2024, 32(8): 2374-2385. (in Chinese)

- 2374-2385. (in Chinese)
- [16] 高炳森, 王朝, 何萍, 田建平, 张俊清, 魏娜. 野外保存对益智叶片基因组 DNA 的影响[J]. 贵州农业科学, 2015, 43(7): 19-22, 25.
- GAO B M, WANG C, HE P, TIAN J P, ZHANG J Q, WEI N. Effects of wild preserved on genomic DNA from *Alpinia oxyphylla* leaves[J]. Guizhou Agricultural Sciences, 2015, 43(7): 19-22, 25. (in Chinese)
- [17] GAO B, YUAN L, TANG T, HOU J, PAN K, WEI N. The complete chloroplast genome sequence of *Alpinia oxyphylla* Miq. and comparison analysis within the Zingiberaceae family[J]. PLoS One, 2019, 14(6): e0218817.
- [18] 田甜, 张银霞, 杜文华. 小黑麦叶绿体基因组密码子偏好性分析与基因功能相关性分析[J]. 草地学报, 2025, 33(3): 707-718.
- TIAN T, ZHANG Y X, DU W H. Correlation between codon usage patterns and gene functions in the chloroplast Genome of *Triticale*[J]. Acta Agrestia Sinica, 2025, 33(3): 707-718. (in Chinese)
- [19] 曹建杰, 辉朝茂, 向自乾, 蔡霖, 胡梦茹, 张寻梦, 刘蔚漪. 巨龙竹叶绿体基因组密码子偏好性分析[J]. 西南农业学报, 2025, 38(4): 1-12.
- CAO J J, HUI C M, XIANG Z Q, CAI L, HU M R, ZHANG X M, LIU W Y. Analysis of the condon usage bias in the chloroplast genome of *Dendrocalamus sinicus*[J]. Southwest China Journal of Agricultural Sciences, 2025, 38(4): 1-12. (in Chinese)
- [20] 王芳, 周盛茂, 赵雪利. 葫芦茶叶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), 2025, 45(1): 34-43.
- WANG F, ZHOU S M, ZHAO X L. Analysis of codon usage Bias in the chloroplast genome of *Tadehagi triquetrum*[J]. Journal of Southwest Forestry University (Natural Science), 2025, 45(1): 34-43. (in Chinese)
- [21] 吴宪明, 吴松峰, 任大明, 朱云平, 贺福初. 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, ZHU Y P, HE F C. The analysis method and progress in the study of codon bias[J]. Hereditas, 2007, 29(4): 420-426. (in Chinese)
- [22] WRIGHT F. The 'effective number of codons' used in a gene[J]. Gene, 1990, 87(1): 23-29.
- [23] 龚明贵, 白娜, 李影, 姜霞. 硬头黄竹叶绿体基因组密码子偏好性分析[J]. 江苏农业科学, 2024, 52(3): 67-72.
- GONG M G, BAI N, LI Y, JIANG X. Analysis of codon preference in chloroplast genome of *Bambusa rigida*[J]. Jiangsu Agricultural Sciences, 2024, 52(3): 67-72. (in Chinese)
- [24] KAWABE A, MIYASHITA N T. Patterns of codon usage bias in three dicot and four monocot plant species[J]. Genes & Genetic Systems, 2003, 78(5): 343-352.
- [25] 秦政, 郑永杰, 桂丽静, 谢谷艾, 伍艳芳. 樟树叶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 Z, ZHENG Y J, GUI L J, XIE G A, WU Y F. Codon usage bias analysis of chloroplast genome of camphora tree (*Cinnamomum camphora*)[J]. Guihaia, 2018, 38(10): 1346-1355. (in Chinese)
- [26] 刘汉梅, 赵耀, 顾勇, 张怀渝, 黄玉碧. 几种植物 waxy 基因的密码子用法特性分析[J]. 核农学报, 2010, 24(3): 476-481.
- LIU H M, ZHAO Y, GU Y, ZHANG H Y, HUANG Y B. Characterization of codon usage of waxy genes in several plants[J]. Journal of Nuclear Agricultural Sciences, 2010, 24(3): 476-481. (in Chinese)
- [27] 辛雅萱, 黎若竹, 李鑫, 陈丽琼, 唐军荣, 屈亚亚, 杨琳懿, 辛培尧, 李云芳. 杧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, CHEN L Q, TANG J R, QU Y Y, YANG L Y, XIN P Y, LI Y F. 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)
- [28] 张加强, 刘慧春, 王杰, 许雯婷, 周江华, 朱开元. 金银花大毛花叶绿体基因组密码子的偏好性分析[J]. 浙江农业学报, 2023, 35(4): 821-830.
- ZHANG J Q, LIU H C, WANG J, XU W T, ZHOU J H, ZHU K Y. Analysis on codon usage patterns of chloroplast genome of *Lonicera japonica* cv. Damaohua[J]. Acta Agriculturae Zhejiangensis, 2023, 35(4): 821-830. (in Chinese)
- [29] 李江平, 秦政, 国春策, 杨光耀, 张文根. 抽筒竹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 J P, QIN Z, GUO C C, YANG G Y, ZHANG W G. Codon bias in the chloroplast genome of *Gelidocalamus tessellatus*[J]. Journal of Bamboo Research, 2019, 38(2): 79-87. (in Chinese)
- [30] 王雪, 马军, 杨学文, 张跃华. 栽培型向日葵叶绿体基因组密码子使用偏好性分析[J]. 分子植物育种, 2024, 22(4): 1096-1102.
- WANG X, MA J, YANG X W, ZHANG Y H. Analysis of codon usage bias in chloroplast genome of cultivated *Helianthus annuus*[J]. Molecular Plant Breeding, 2024, 22(4): 1096-1102. (in Chinese)
- [31] 王媛媛, 杨美青. 蒙古韭叶绿体基因组密码子使用偏好性分析[J]. 分子植物育种, 2021, 19(04): 1084-1092.
- WANG Y Y, YANG M Q. Analysis of the codon usage bias

- in the chloroplast genome of *Allium mongolicum* regel[J]. Molecular Plant Breeding, 2021, 19(4): 1084-1092. (in Chinese)
- [32] 包国媛, 李文辛, 杨鑫光, 王雅琼. 海甜菜线粒体和叶绿体基因组密码子使用偏好性分析[J]. 江苏农业学报, 2023, 39(9): 1804-1817.
- BAO G Y, LI W X, YANG X G, WANG Y Q. Analysis of codon usage bias in mitochondrial and chloroplast genomes of *Beta vulgaris* subsp.[J]. Jiangsu Journal of Agricultural Sciences, 2023, 39(9): 1804-1817. (in Chinese)
- [33] 梁湘兰, 覃逸明, 孙晓波, 郭松. 藿香叶绿体基因组密码子偏好性分析[J]. 分子植物育种, 2023, 21(3): 809-818.
- LIANG X L, QIN Y M, SUN X B, GUO S. Analysis of codon usage bias in the chloroplast genome of *Agastache rugosa*[J]. Molecular Plant Breeding, 2023, 21(3): 809-818. (in Chinese)
- [34] 万美暄, 黄显俊, 李雪, 徐立. 须弥葛叶绿体基因组密码子使用偏好性分析[J]. 安徽农业科学, 2022, 50(21): 102-107.
- WAN M X, HUANG X J, LI X, XU L. Analysis of codon bias in the chloroplast genome of *Haymondia wallihii*[J]. Journal of Anhui Agricultural Sciences, 2022, 50(21): 102-107. (in Chinese)
- [35] 李江飞, 王瑜, 颜廷雨, 汪啟波, 陈诗, 蔡年辉, 许玉兰, 唐红燕. 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 J F, WANG Y, YAN T Y, WANG Q B, CHEN S, CAI N H, XU Y L, TANG H Y. Analysis on codon usage bias of *Keteleeria evelyniana* chloroplast genome[J]. Journal of Central South University of Forestry & Technology, 2022, 42(4): 30-39. (in Chinese)
- [36] 李显煌, 杨生超, 辛雅萱, 屈亚亚, 杨琳懿, 冯发玉, 范伟, 辛培尧. 灯盏花叶绿体基因组密码子偏好性分析[J]. 云南农业大学学报(自然科学), 2021, 36(3): 384-392.
- LI X H, YANG S C, XIN Y X, QU Y Y, YANG L Y, FENG F Y, FAN W, XIN P Y. Analysis of the codon usage bias of chloroplast genome in *Erigeron breviscapus* (Vant.) Hand-Mazz[J]. Journal of Yunnan Agricultural University (Natural Science), 2021, 36(3): 384-392. (in Chinese)
- [37] 刘伟强, 秦斗文, 史国民, 徐庭亮, 巨秀婷. 伊犁郁金香叶绿体基因组密码子偏好性分析[J]. 中国野生植物资源, 2024, 43(2): 39-45.
- LIU W Q, QIN D W, SHI G M, XU T L, JU X T. Analysis of the codon usage bias of chloroplast genome in *Tulipa iliensis* regel[J]. Chinese Wild Plant Resources, 2024, 43(2): 39-45. (in Chinese)
- [38] ZHANG L, GUO Y, LUO L, WANG Y P, DONG Z M, SUN S H, QIU L J. Analysis of nuclear gene codon bias on soybean genome and transcriptome[J]. Acta Agronomica Sinica, 2011, 37(6): 965-974.
- [39] ZHOU J H, DING Y Z, HE Y, CHU Y F, ZHAO P, MA L Y, WANG X J, LI X R, LIU Y S. The effect of multiple evolutionary selections on synonymous codon usage of genes in the *Mycoplasma bovis* genome[J]. PLoS One, 2017, 9(10): e108949.
- [40] 朱灵芝, 朱沛煌, 李荣, 陈好, 季孔庶. 马尾松 *PmDXR* 基因密码子偏好性分析[J]. 林业科学研究, 2021, 34(2): 102-113.
- ZHU L Z, ZHU P H, LI R, CHEN Y, JI K S. Analysis on codon bias of *PmDXR* gene in *Pinus massoniana* Lamb. [J]. Forest Research, 2021, 34(2): 102-113. (in Chinese)
- [41] 曹建杰, 辉朝茂, 向自乾, 蔡霖, 胡梦茹, 张寻梦, 刘蔚漪. 甜龙竹叶绿体基因组密码子偏好性分析[J]. 中国农业科技导报(中英文), (2025-06-17)[2025-07-11]. <https://doi.org/10.13304/j.nykjdb.2024.0692>.
- CAO J J, HUI C M, XIANG Z Q, CAI L, HU M R, ZHANG X M, LIU W Y. Analysis of codon usage bias in chloroplast genome of *Dendrocalamus brandisii*[J]. Journal of Agricultural Science and Technology, (2025-06-17)[2025-07-11]. <https://doi.org/10.13304/j.nykjdb.2024.0692>. (in Chinese)
- [42] 仇学文, 李丹, 甘玉迪, 杨有新, 程柳洋, 徐梦怡, 吴才君. 豇豆叶绿体基因组密码子使用偏好性分析[J]. 核农学报, 2023, 37(6): 1118-1131.
- QIU X W, LI D, GAN Y D, YANG Y X, CHENG L Y, XU M Y, WU C J. Analysis of codon usage bias in the cowpea chloroplast genome[J]. Journal of Nuclear Agricultural Sciences, 2023, 37(6): 1118-1131. (in Chinese)
- [43] ZHOU M, LONG W, LI X. Analysis of synonymous codon usage in chloroplast genome of *Populus alba*[J]. Journal of Forestry Research, 2008, 19(4): 293-297.
- [44] SHARP P M, LI W H. An evolutionary perspective on synonymous codon usage in unicellular organisms[J]. Journal of Molecular Evolution, 1986, 24(1/2): 28-38.