



草地学报
Acta Agrestia Sinica
ISSN 1007-0435, CN 11-3362/S

《草地学报》网络首发论文

题目：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
作者：田甜，张银霞，杜文华
收稿日期：2024-08-26
网络首发日期：2024-12-02
引用格式：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/OL]. 草地学报.
<https://link.cnki.net/urlid/11.3362.S.20241129.1741.004>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

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

田甜, 张银霞, 杜文华*

(甘肃农业大学草业学院/草业生态系统教育部重点实验室/甘肃省草业工程实验室/中-美草地畜牧业可持续发展研究中心, 甘肃 兰州 730070)

摘要: 为揭示小黑麦 (*Triticale*) 的密码子使用偏好性及其影响因素, 用 CodonW 1.4.2 和在线软件 CUSP (Relative synonymous codon usage, CUSP) 对小黑麦叶绿体基因组中 52 个编码序列 (Coding sequence, CDS) 特征进行分析。结果表明: 小黑麦叶绿体基因组密码子第三位主要偏向以 A/U 为主; 有效密码子数 (Effective number of codon, ENC) 值在 37.64~61.00 之间, 平均值为 47.31, 大于 45 的有 42 个; 密码子适应指数 (Codon adaptation index, CAI) 在 0.12~0.33 之间, 平均值为 0.17, 说明小黑麦叶绿体基因组密码子偏好性较弱; 16 个均以 A/U 结尾的密码子被筛选为最优密码子。中性绘图、ENC-GC₃S 和偏倚性绘图分析 (Parity Rule 2, PR2-plot) 结果表明多数功能基因的密码子使用偏好性受自然选择压力影响为主。小黑麦叶绿体基因组中与光合作用密切相关的基因可能受到了突变和选择的强烈作用, 使这些基因的密码子使用频率具有共同特征, 表现为密码子使用模式与基因功能的有较大的相关性。

关键词: 小黑麦; 叶绿体基因组; 密码子偏好性; 最优密码子

中图分类号: S519

Correlation between Codon Usage Patterns and Gene Functions in the Chloroplast Genome of *Triticale*

TIAN Tian, ZHANG Yin-xia, DU Wen-hua*

(College of Grassland Science, Gansu Agricultural University/Key Laboratory of Grassland Ecosystem, Gansu Agricultural

University, Ministry of Education/Sino-U.S. Centers for Grazingland Ecosystem Sustainability, Lanzhou, Gansu Province 730070, China)

Abstract To understand the codon usage bias (CUB) and the factor in *Triticale*. Based on

收稿日期: 2024-08-26; 修回日期: 2024-10-13

基金项目: 农业生物育种重大专项 (2023ZD0402605-02); 国家自然科学基金 (32260339); 甘肃省高等学校产业支撑计划 (2022CYZC-49); 甘肃省重大专项 (21ZD4NA012); 西藏重大专项 (XZ202101ZD003N) 资助

作者简介: 田甜 (1997-), 女, 苗族, 贵州铜仁人, 博士研究生, 从事草地生物多样性研究, E-mail:1138115730@qq.com;

*通信作者 Author for corresponding, E-mail: duwh@gasu.edu.cn.

CodonW 1.4.2 and CUSP(Relative synonymous codon usage) software to analyze the characteristic in chloroplast genome of triticales. Results showed that the $GC_1(46.83\%) > GC_2(39.62\%) > GC_3(30.28\%)$. This indicates a low GC content for the third base of codons, suggesting a preference end for A or U bases. ENC(Effective Number of Codon) ranged from 37.64 to 61.00, with an average of 47.31, and there were 42 ENC values > 45 , and CAI(Codon Adaptation Index) ranged from 0.12 to 0.33, with a mean of 0.17, It indicated that triticales codon preference was weak. Optimal codon analysis identified a total of sixteen optimal codons with a preference for ending with A or U bases. Neutral mapping, ENC- GC_3 s and PR2-plot(Parity Rule 2 plot) all showed that the preference of codon usage for the majority of functional genes was influenced by a combination of mutation and natural selection pressure, and the influence of natural selection was predominant. Photosynthesis-related genes in the Triticales chloroplast genome may undergo significant mutational and selective effects. As a result, codon usage in these genes commonly exhibits a strong correlation with their respective functions. In order to enhance our comprehension of the molecular mechanisms responsible for chloroplast gene expression and the role of codon usage patterns in biological evolution, further studies can be conducted. These studies could reveal more patterns of codon usage and biological significance. Furthermore, this knowledge can also be utilized to support the engineering applications of chloroplast genomes.

Key word: *Triticales*; Chloroplast genome, Codon usage bias, Optimal codon

叶绿体是一种半自主细胞器,通过光合作用将太阳能转化为生物能,参与了许多其他化合物的合成^[1-4],叶绿体基因组是其实现活跃的代谢功能的必要条件^[3-4]。单个叶绿体基因组的详细表征与比较叶绿体基因组学可以作为功能基因组研究的基础^[5],也可以指导叶绿体基因工程的试验^[6-7]。叶绿体基因转化具有外源基因表达效率高、稳定遗传、定点整合以及没有位置效应现象等优点^[8-9],可以用来开发分子标记、系统发育分析、种群水平研究、基因分型和定位等研究^[10]。密码子是生物体遗传信息和蛋白质翻译之间转化的桥梁,在实现遗传信息的传递和表达上起到了至关重要的作用。蛋白质翻译过程中,翻译为同种氨基酸的不同密码子称为同义密码子,同义密码子在组合频率上也有所差异,这种现象称为密码子偏好性

(Codon Usage bias, CUB)^[11]。同义密码子偏好性可能是由进化过程中基因上的不同因素造成的^[12]，是基因表达的一种微调方式^[13]。分析目标基因或受体植物的基因密码子使用模式，可以为构建植物成熟稳定的叶绿体转基因系统奠定基础^[14-15]。研究表明，密码子偏好性由自然选择和突变共同决定，前人提出了“选择—突变—漂移”假说和“中性理论”假说^[16]，前者认为密码子的使用偏倚性是由最优密码子的选择压力和非同义密码子的突变—漂移所决定，后者认为同义密码子的突变是中性选择，而与选择压力无关^[17-18]。除此之外，GC含量的高低、上下游依赖性核苷酸的组成、种群大小以及基因表达水平和长度等因素也会对密码子的偏倚性产生影响^[11]。同时，密码子的使用偏好会影响 mRNA 折叠、翻译延伸和蛋白折叠，从而对基因的表达水平产生不同程度的影响^[13,19]。

小黑麦 (*Triticale*) 是由小麦属 (*Triticum*) 和黑麦属 (*Secale*) 经过属间杂交和染色，体数加倍而人工育成的新品种^[20]。小黑麦杂种优势强，具有产量高、品质好以及抗逆性强等优势，是继苜蓿、饲用玉米后的在西北地区的优质饲草作物^[21]。同时，小黑麦越冬性能好，在高寒农牧交错区能够错季种植，保证了应季作物的正常播种，同时减少了冬春季节农田裸露带来的环境问题^[22-25]。小黑麦六倍体叶绿体基因组已完成测序 (GenBank 登录号: ON422331.1)，但其密码子偏好性研究尚未有见报道。明确小黑麦叶绿体基因的密码子偏好性是开展小黑麦叶绿体基因工程的基础之一，此工程在改良植物性状和提高抵抗病虫害等方面具有重要作用^[26]。培育高产、抗病性强、抗逆性强的小黑麦一直是遗传改良的目标，但受到小黑麦遗传背景复杂、倍性多样和传统育种周期长等问题的限制。分子育种一定程度上可以缓解这些问题，但外源基因的表达水平不稳定，会受到密码子偏好性的影响^[27]。另外，规律间隔成簇短回文重复序列 (Clustered regularly interspaced short palindromic repeats, CRISPR) 基因编辑技术中，Cas9蛋白的表达量在基因编辑中起着关键作用。Cas9主要通过sgRNA的引导定位到目标DNA序列进行特异性切割DNA双链从而在DNA断裂时引入突变实现基因的敲除、插入或替换，因此，Cas9的定位效率直接影响到编辑效率，而Cas9的定位效率与其表达量有关^[28-30]。Ma等人^[31]研究发现几个影响Cas9定位效率的因素：首先，不同启动子驱动的sgRNA靶向的位点突变率相似；其次，高GC含量（高于50%）具有更高的编辑效率，最后，所有碱基都具有相似的编辑效率。这一发现扩大了有效基因组编辑目标的选择范围，不限于U3启动子的A和U6启动子的G的常规靶标^[28]。研究表明，密码子的使用模式会影响基因的表

达^[13], Shan等^[32]对密码子进行优化了Cas9, 靶向水稻Os PDS基因的突变率达14.5%~20%, Zhang等^[33]使用了两种密码子类型的Cas9(h Sp Cas9, p Sp Cas9)研究靶向水稻除草剂抗性基因BEL基因, 结果表明, p Sp Cas9比h Sp Cas9的编辑效率高7.42倍, 差异可能与p Sp Cas9转基因植物中较高水平的Cas9蛋白积累相关。因此, Cas9的表达量影响其定位效率, 从而影响基因编辑效率。而当前尚未有任何报道关于Cas9在小黑麦叶绿体基因组密码子偏好性中的关系。因此, 通过研究小黑麦叶绿体基因组密码子的使用特征, 确定最优密码子, 并探讨植物叶绿体基因的密码子使用模式与基因功能的相关性, 该研究结果能够为提高基因在小黑麦中表达、预测及外源受体的选择提供一定参考, 同时, 在小黑麦基因编辑的研究中, 能够为提高小黑麦Cas9表达从而提高编辑效率提供一定的研究价值。

1 材料与方法

1.1 序列获取

从 NCBI 数据库下载小黑麦的叶绿体基因组序列(GenBank 登录号: ON422331.1)87 条, 剔除长度小于 300 bp 的序列、去除重复序列、去除非 ATG 作为起始密码子的序列、删掉异常终止密码子和 CDS 内部存在终止密码子的序列后, 得到 52 条序列用于后续分析。

1.2 密码子组成分析

用 CodonW 1.4.2 对小黑麦叶绿体基因组的 52 条 CDS 序列进行同义密码子相对使用度(Relative usage of synonymous codons, RSCU)、密码子适应指数(Codon adaptation index, CAI)、密码子偏好指数(Codon bias index, CBI)和最优密码子使用频率值(Frequency of optimal codons, FOP)分析; 使用在线软件 CUSP (<http://emboss.toulouse.inra.fr/cgi-bin/emboss/cusp>) 统计得到的 52 条序列的密码子数量(Codon Number, CN)、密码子不同位置碱基的 GC 含量, 包括第 1 位 GC 含量(GC₁)、第 2 位 GC 含量(GC₂)、第 3 位 GC 含量(GC₃)以及 GC 总含量(GC_{all}), 同时计算出 GC₁ 和 GC₂ 的平均值(GC₁₂)利用 R 语言进行相关性分析。

1.3 中性绘图

以 GC₃ 为横坐标, GC₁ 和 GC₂ 的平均值 GC₁₂ 为纵坐标绘制中性图。若密码子偏好性仅受到突变的影响, 则基因分布于对角线上; 若受到选择压力的影响更大, 则基因分布于对角线周围^[34]。同时, 可通过横纵坐标 GC₃ 和 GC₁₂ 的相关性来判断密码子偏好性

受到的影响因素，若两者显著相关，回归系数（对角线斜率）较大且接近于 1，则密码子受到的影响因素主要为突变，反之（回归系数接近于 0）则为选择压力^[35]。

1.4 ENC-plot 分析

有效密码子数 (Effective Number of Codon, ENC) 用于度量同义密码子的偏倚性，当 ENC 值=20 时，表示每个氨基酸只使用一个密码子，为极端偏好；ENC 值=61，表示密码子偏向随机使用，不存在使用偏好。以公式 (1) 为标准曲线^[36]，即期望值，每点代表一个基因。

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

每个基因与标准曲线(即期望值)之间的距离为判断影响密码子偏倚性因素的标准。当 ENC 实际值与期望值相差较大，即距离标准曲线较远时，偏倚性受选择压力的影响，反之（即 ENC 实际值距离标准曲线较近时）则主要由碱基组成决定，即受突变的影响较大。为了更好地反映 ENC 实际值与期望值之间的差异，可以通过计算 ENC 比值频率分布，以-0.05~0.05 的范围为界对差异进行量化分析。在此范围内则接近期望值，即密码子偏好性受到突变的影响较大；在此范围外则影响密码子偏好性的主要因素为选择压力。

1.5 PR2-plot 绘图分析

以 $A_3/(A_3+T_3)$ 为纵坐标，以 $G_3/(G_3+C_3)$ 为横坐标，绘制 PR2 偏倚平面图，每个基因的碱基组成在一个平面上表示，中心点表示密码子无偏好性使用时的状态 ($A=T, C=G$)，而其他各点与中心点的矢量距离表示偏倚的方向和程度^[37]，以此来判断突变压力和自然选择是否影响了 DNA 双链上核苷酸的组成。如果 4 种碱基的使用频率相似，则基因坐落于中线或靠近中心点，表示密码子偏好性只受到突变的影响；反之坐落于中心点四周或远离中心点，表示同时受到了突变和自然选择的影响。

1.6 最优密码子的确定

为了得到小黑麦叶绿体基因的最优密码子，以 ENC 值为参考指标，筛选前后各 10% 的基因构建高低基因表达库，通过 Codon W1.4.2 计算其 RSCU 值，并计算高低表达库之间的差值 ($\Delta RSCU$)，将 $\Delta RSCU \geq 0.08$ 且 $RSCU > 1$ 的密码子筛选出，确定为最优密码子^[38]。

2 结果与分析

2.1 密码子组成分析

小黑麦叶绿体基因组总长为 135 899 bp，从中得到 87 个 CDS（不含假基因），经筛选而得到 52 个 CDS 序列用于后续分析。结果如表 1 所示，52 个 CDS 的总密码子数为 17 152。总 GC 含量分布范围为 32.55%~49.20%，平均值为 38.98%。GC₁、GC₂ 和 GC₃ 含量分布范围分别为 32.75%~56.28%，29.87%~55.56%，22.22%~48.28%之间，平均值为 GC₁(47.43%)>GC₂(39.43%)>GC₃(30.09%)，表明密码子分布的不均匀性，GC 含量较低，主要以 A/U 结尾。其中，GC₂ 和 GC_{all} 的差异不大（38.98%和 39.43%），而 GC₃ 最小（30.09%），所有 GC 含量均在 50%以下，说明小黑麦叶绿体基因组密码子向 A /U 偏倚。

根据偏倚性分析结果（表 1），ENC 在 37.64~61.00 之间（平均值为 47.31），其中有 42 个基因的 ENC 值>45，表明小黑麦叶绿体基因组密码子偏好性相对较弱。CAI 范围为 0.12~0.33，平均值为 0.17，趋近于 0，表明小黑麦叶绿体基因组密码子偏好性弱。CBI 取值为-0.21~0.24，均值为-0.09，则最优密码子使用次数小于密码子平均使用次数；FOP 取值为 0.27~0.56，均值为 0.35，则最优密码子占同义密码子数量的比值较小。综合以上，表明小黑麦叶绿体基因组密码子使用偏好性较弱。

表 1 小黑麦叶绿体基因组密码子的碱基组成和表达特征
Table 1 Base composition and expression characteristics of codons in chloroplast genome of triticales

基因功能 Function	基因 Genes	密码子 数量 CN	第 1	第 2 位	第 3	第 1、2	总碱	有效	密码	密码	最优密 码子使 用频率 Fop
			位碱 基含 量 GC ₁	碱基含 量 GC ₂	位碱 基含 量 GC ₃	位碱 基平 均值 GC ₁₂	基含 量 GC _{all}	密码 子数 ENC	子适 应指 数 CAI	子偏 好指 数 CBI	
ATP 合成酶亚基	<i>atpA</i>	505.00	55.84	39.80	29.90	47.82	41.85	45.48	0.19	-0.05	0.39
Subunits of ATP	<i>atpB</i>	499.00	54.51	41.88	30.06	48.20	42.15	45.85	0.19	-0.05	0.38

synthase	<i>atpE</i>	138.00	51.45	38.41	37.68	44.93	42.51	51.48	0.18	-0.03	0.38
	<i>atpF</i>	184.00	47.83	34.24	32.07	41.04	38.04	50.12	0.16	-0.06	0.38
	<i>atpI</i>	248.00	47.98	36.29	31.45	42.14	38.58	49.63	0.17	-0.07	0.35
细胞色素 c 基因											
c-type cytochrome	<i>ccsA</i>	323.00	33.44	41.80	25.39	37.62	33.54	41.42	0.14	-0.21	0.29
synthesis gene											
包膜膜蛋白											
Envelope membrane											
	<i>cemA</i>	231.00	40.69	29.87	29.44	35.28	33.33	42.05	0.17	-0.13	0.33
protein											
蛋白酶											
Protease	<i>clpP</i>	217.00	52.07	38.25	36.87	45.16	42.40	50.36	0.18	-0.14	0.33
翻译起始因子											
Translation initiatio											
n	<i>infA</i>	114.00	44.74	35.09	38.60	39.92	39.47	61.00	0.16	-0.07	0.38
factor											
成熟酶											
Maturase	<i>matK</i>	513.00	39.96	31.58	28.85	35.77	33.46	46.83	0.17	-0.12	0.35
	<i>ndhA</i>	364.00	42.03	36.54	23.63	39.29	34.07	42.29	0.14	-0.10	0.33
	<i>ndhB</i>	511.00	41.88	39.33	32.49	40.61	37.90	45.66	0.16	-0.10	0.34
	<i>ndhC</i>	121.00	50.41	34.71	33.06	42.56	39.39	56.55	0.17	-0.08	0.34
NADH 脱氢酶亚基	<i>ndhD</i>	501.00	40.32	36.53	29.14	38.43	35.33	45.19	0.14	-0.15	0.30
Subunits of NADH	<i>ndhE</i>	102.00	41.18	32.35	26.47	36.77	33.33	50.13	0.14	-0.15	0.31
	<i>ndhF</i>	740.00	37.43	39.32	23.51	38.38	33.42	43.00	0.14	-0.15	0.32
dehydrogenase	<i>ndhG</i>	177.00	45.76	33.33	25.99	39.55	35.03	47.48	0.13	-0.20	0.27
	<i>ndhH</i>	394.00	51.02	34.52	27.92	42.77	37.82	46.12	0.15	-0.14	0.32
	<i>ndhI</i>	181.00	37.02	39.78	25.97	38.40	34.25	42.95	0.18	-0.12	0.36

	<i>ndhJ</i>	160.00	50.00	37.50	30.63	43.75	39.38	48.97	0.16	-0.17	0.32
	<i>ndhK</i>	246.00	41.46	43.09	29.27	42.28	37.94	49.55	0.15	-0.15	0.32
细胞色素 b/f 复合	<i>petA</i>	321.00	53.58	36.14	30.84	44.86	40.19	47.84	0.16	-0.11	0.34
体亚基	<i>petB</i>	235.00	47.66	41.70	33.19	44.68	40.85	47.29	0.19	-0.13	0.32
Subunits of cytochr											
ome b/f complex	<i>petD</i>	161.00	51.55	38.51	28.57	45.03	39.54	44.09	0.16	-0.15	0.28
光和系统 I 亚基	<i>psaA</i>	751.00	52.46	43.54	33.02	48.00	43.01	48.82	0.19	-0.10	0.36
Subunits of											
photosystem I	<i>psaB</i>	735.00	48.71	42.99	33.20	45.85	41.63	46.78	0.17	-0.12	0.35
光和系统 II 亚基	<i>psbA</i>	354.00	48.87	42.94	33.33	45.91	41.71	37.64	0.33	0.24	0.56
Subunits of	<i>psbB</i>	509.00	54.22	46.17	29.27	50.20	43.22	45.42	0.20	-0.07	0.38
photosystem II	<i>psbC</i>	474.00	53.59	44.73	32.91	49.16	43.74	45.20	0.19	-0.04	0.39
	<i>psbD</i>	354.00	52.82	43.79	33.05	48.31	43.22	46.27	0.23	0.04	0.44
RuBisCO 酶大亚基											
Large subunit of	<i>rbcL</i>	478.00	56.28	44.56	29.92	50.42	43.58	45.76	0.28	0.06	0.46
rubisco											
核糖体大亚基蛋白	<i>rpl14</i>	124.00	54.84	37.90	23.39	46.37	38.71	46.52	0.18	-0.01	0.40
Proteins of large	<i>rpl16</i>	137.00	51.09	54.74	27.01	52.92	44.28	39.12	0.12	-0.06	0.38
ribosomal subunit	<i>rpl20</i>	120.00	39.17	40.00	31.67	39.59	36.94	53.11	0.13	-0.16	0.32
	<i>rpl22</i>	149.00	40.27	36.91	36.24	38.59	37.81	47.95	0.17	-0.04	0.38
RNA 聚合酶亚基	<i>rpoA</i>	340.00	47.06	35.88	27.65	41.47	36.86	48.60	0.16	-0.13	0.33
Subunits of RNA	<i>rpoB</i>	1,077.00	49.49	38.07	30.18	43.78	39.24	48.89	0.16	-0.09	0.36
polymerase	<i>rpoC1</i>	684.00	49.56	38.16	30.41	43.86	39.38	49.82	0.15	-0.11	0.34
	<i>rpoC2</i>	1,480.00	47.91	36.28	29.93	42.10	38.04	49.68	0.16	-0.13	0.34
核糖体小亚基蛋白	<i>rps11</i>	144.00	51.39	55.56	22.22	53.48	43.06	42.37	0.17	-0.10	0.36
Proteins of small	<i>rps12</i>	125.00	52.80	47.20	27.20	50.00	42.40	45.14	0.15	-0.06	0.37
ribosomal subunit	<i>rps14</i>	104.00	42.31	43.27	30.77	42.79	38.78	45.83	0.15	-0.03	0.38

	<i>rps18</i>	171.00	32.75	38.60	26.32	35.68	32.55	45.80	0.14	-0.15	0.34
	<i>rps2</i>	237.00	40.08	40.51	29.96	40.30	36.85	49.94	0.16	-0.15	0.32
	<i>rps3</i>	240.00	42.50	31.67	26.67	37.09	33.61	47.17	0.20	-0.03	0.41
	<i>rps4</i>	202.00	47.52	37.13	27.23	42.33	37.29	52.64	0.16	-0.02	0.39
	<i>rps7</i>	157.00	49.68	45.22	23.57	47.45	39.49	47.11	0.16	-0.07	0.37
	<i>rps8</i>	137.00	41.61	41.61	25.55	41.61	36.25	44.76	0.12	-0.08	0.36
保守假设性叶绿体	<i>ycf2</i>	190.00	43.68	42.11	31.05	42.90	38.95	47.07	0.14	-0.18	0.31
开放阅读框	<i>ycf3</i>	150.00	43.33	35.33	40.00	39.33	39.56	52.67	0.14	-0.17	0.32
Conserved hypothet	<i>ycf4</i>	186.00	47.31	39.25	33.33	43.28	39.96	45.74	0.17	-0.03	0.38
ical chloroplast OR											
F	<i>ycf68</i>	145.00	53.79	45.52	48.28	49.66	49.20	53.04	0.12	-0.13	0.32
	平均值										
	Average	329.62	46.83	39.62	30.28	43.22	38.91	47.31	0.17	-0.09	0.35

各参数相关性分析表明（图 1），GC_{all} 与 GC₁、GC₂ 以及 GC₃ 呈极显著相关（0.01 水平），GC₁ 和 GC₂ 呈显著相关（0.05 水平），但 GC₃ 和 GC₁、GC₂ 相关性差，表明小黑麦叶绿体基因组密码子第 1、2 位碱基组成上较相似，而与第 3 位存在较大差异。ENC 与 GC₃ 呈极显著相关，而与 GC₂ 呈极显著负相关，与 GC₁ 没有相关性，表明密码子第 2 位和第 3 位对 ENC 具有偏性影响。密码子数 CN 与 ENC、GC₁、GC₂、GC₃ 以及 GC_{all} 均无相关性，说明小黑麦密码子碱基受偏性影响较小。

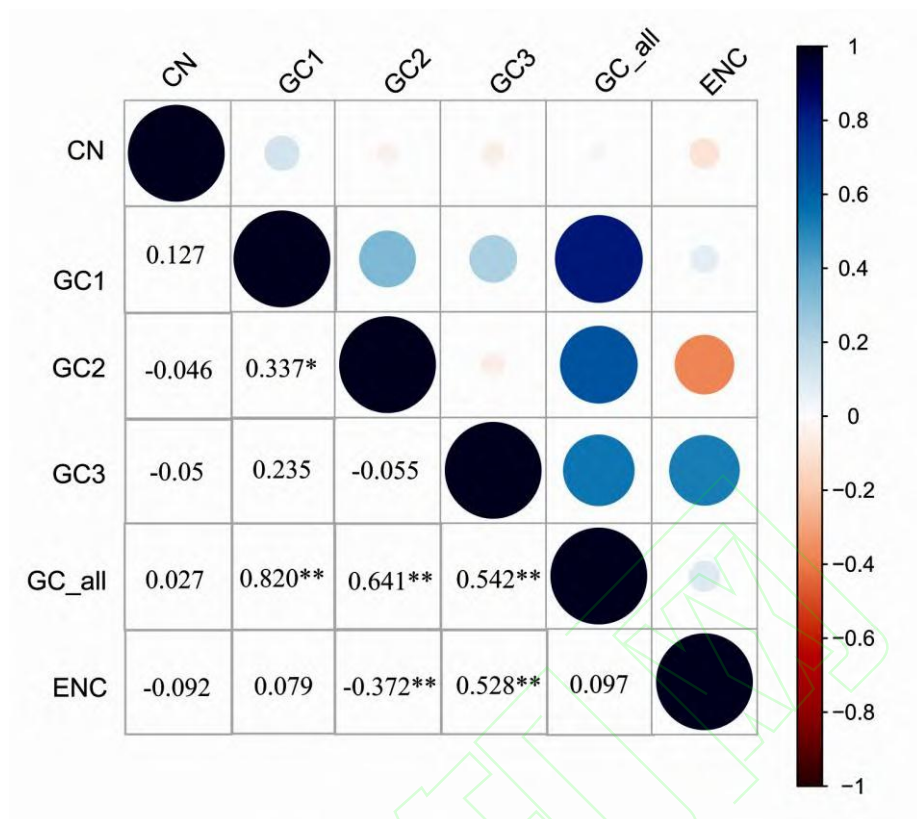


图 1 小黑麦叶绿体基因组密码子中各参数的相关性分析

Fig. 1 Correlation analysis of the parameters in triticales chloroplast genome codon

注: **表示 P 值 < 0.01 , 两者极显著相关。CN 表示密码子数量; GC_1 表示第 1 位 GC 含量; GC_2 表示第 2 位 GC 含量; GC_3 表示第 3 位 GC 含量; GC_{all} 表示 GC 总含量; ENC 表示有效密码子数

Note: ** indicates P value < 0.01 , both are highly significantly correlated. CN stands for condons number; GC1 stands for the GC content in the first position; GC2 stands for the GC content in the second position; GC3 stands for the GC content in the third position; GCall stands for GC total content; ENC stands for effective number of codon

密码子总体 ESCU 分析表明 (图 2), Leu、Ser 和 Arg 分别由多达 6 个密码子编码, Val、Pro、Thr、Ala 和 Gly 分别由四个密码子编码, 其余除了 Met 和 Trp 两种氨基酸分别由一个密码子编码以及 Lie 由 3 个密码子编码之外, 剩余所有的氨基酸均由两个密码子编码。

由图 3 可知, 编码 Ile 的密码子 AUU 数量最多, 有 711 个; 编码 Ter 的密码子 UGA 最少, 仅有 12 个。小黑麦叶绿体基因组中密码子 $RSCU > 1$ 的有 33 个, 其中 29 个以 A/U 结尾, $RSCU < 1$ 的有 31 个, 具有明显的 A/U 偏向性。

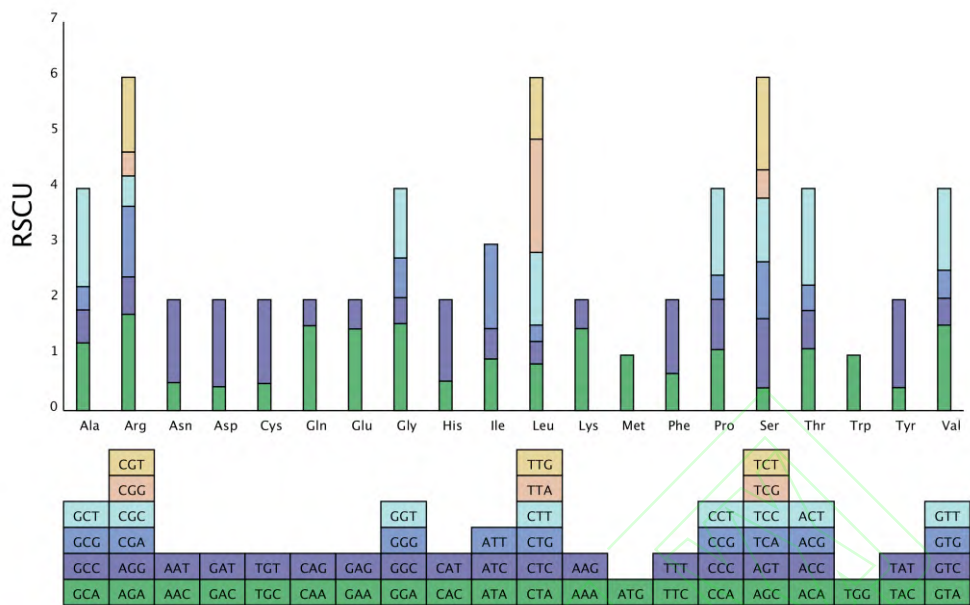


图 2 小黑麦叶绿体基因组中各氨基酸的 RSCU 分析

Fig. 2 RSCU analysis of the amino acids in triticale chloroplast genome

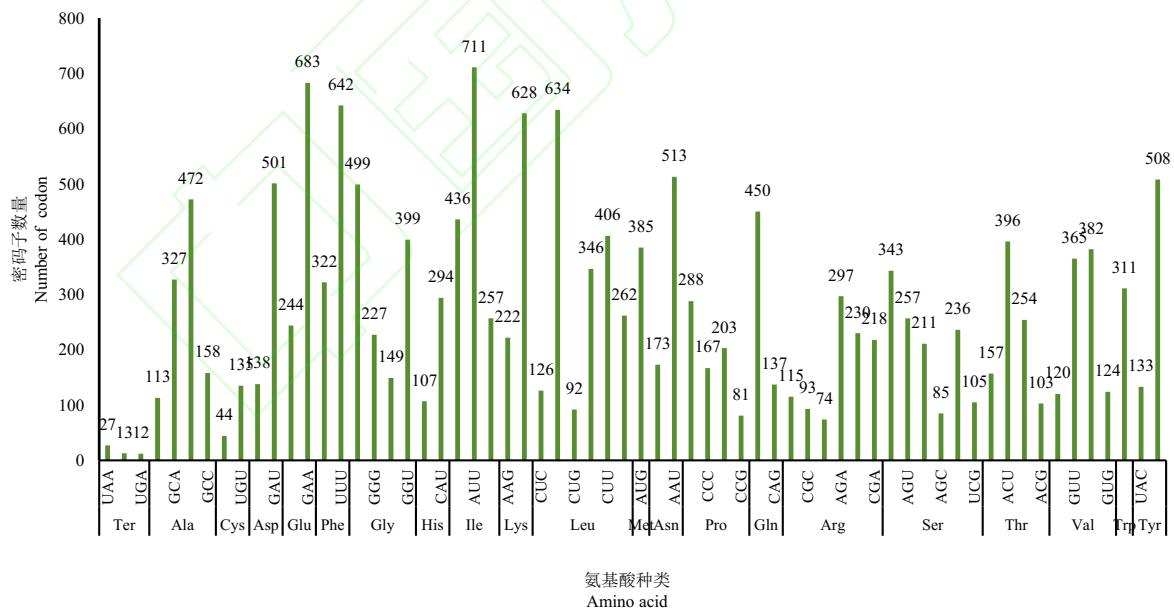


图 3 小黑麦叶绿体基因组中各氨基酸的密码子组成分析

Fig. 3 Codons composition analysis of the amino acids in chloroplast genome of triticale

2.2 中性绘图分析

通过中性绘图来判断小黑麦叶绿体基因组密码子偏好性的主要影响因素，结果如图 4。 GC_{12} 的取值范围在 0.3577~0.5348 之间， GC_3 取值范围在 0.2222~0.4828 之间，仅有 1 个基因在对角线位置下方，其余所有基因均位于对角线位置上方，说明主要的影响来源是自然选择； GC_{12} 与 GC_3 回归系数为 0.119，相关系数 R 为 0.0671，相关性弱，另外，GC 含量仅为 47.43，进一步表明自然选择是主要的影响因素。同时，涉及保守假设性叶绿体 ORF、核糖体大亚基蛋白和翻译起始因子 3 种功能中的 4 个基因位于对角线附近甚至下方，说明在一定程度上，个别基因受到突变的影响较大。

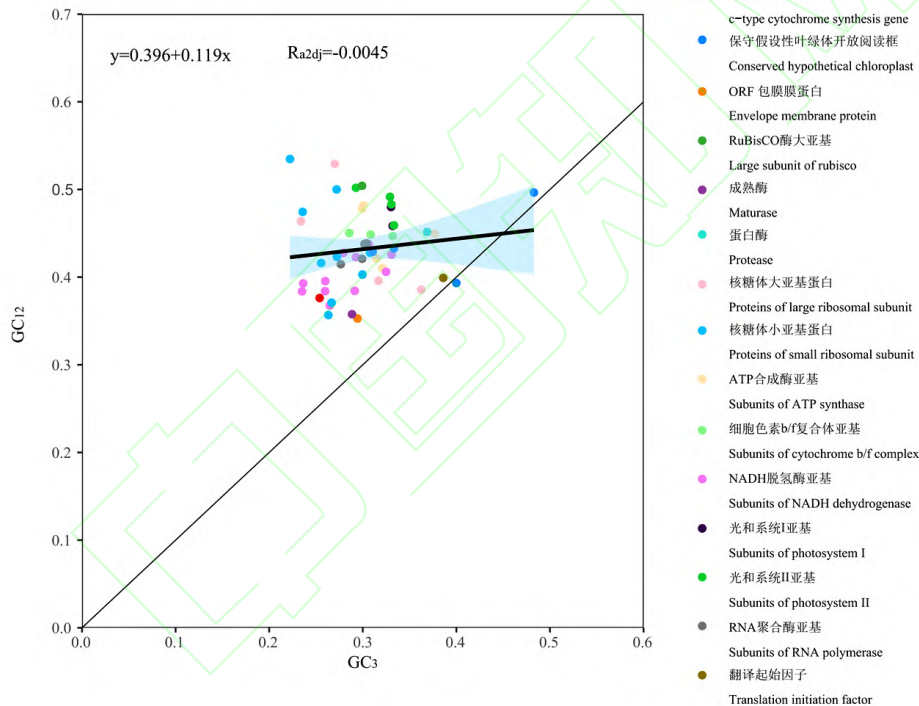


图 4 小黑麦叶绿体基因组中性绘图分析

Fig.4 Neutrality plot analysis in chloroplast genome of triticale

2.3 ENC-plot 分析

ENC-plot 绘图（图 5）分析发现，实际 ENC 值与期望值存在较大差异，除了涉及 NADH 脱氢酶亚基、核糖体大亚基蛋白和核糖体小亚基蛋白 3 种功能中的 4 个基因 ENC 值接近

期望值（标准曲线）以外，其他基因的 ENC 值均高于或低于期望值，即大多数与标准曲线的距离较远，说明自然选择是密码子偏好性的主要影响因素。后续通过统计小黑麦叶绿体基因 ENC 比值频率以更为准确地反映差异，结果表明（图 6），仅有 11 个基因分布于-0.05~0.05 之间，占比 21.15%，其余 41 个基因（占比 78.85%）全部在此范围外，距离 ENC 标准期望值较远，进一步说明小黑麦叶绿体基因组密码子偏好性的主要影响因素是选择压力。

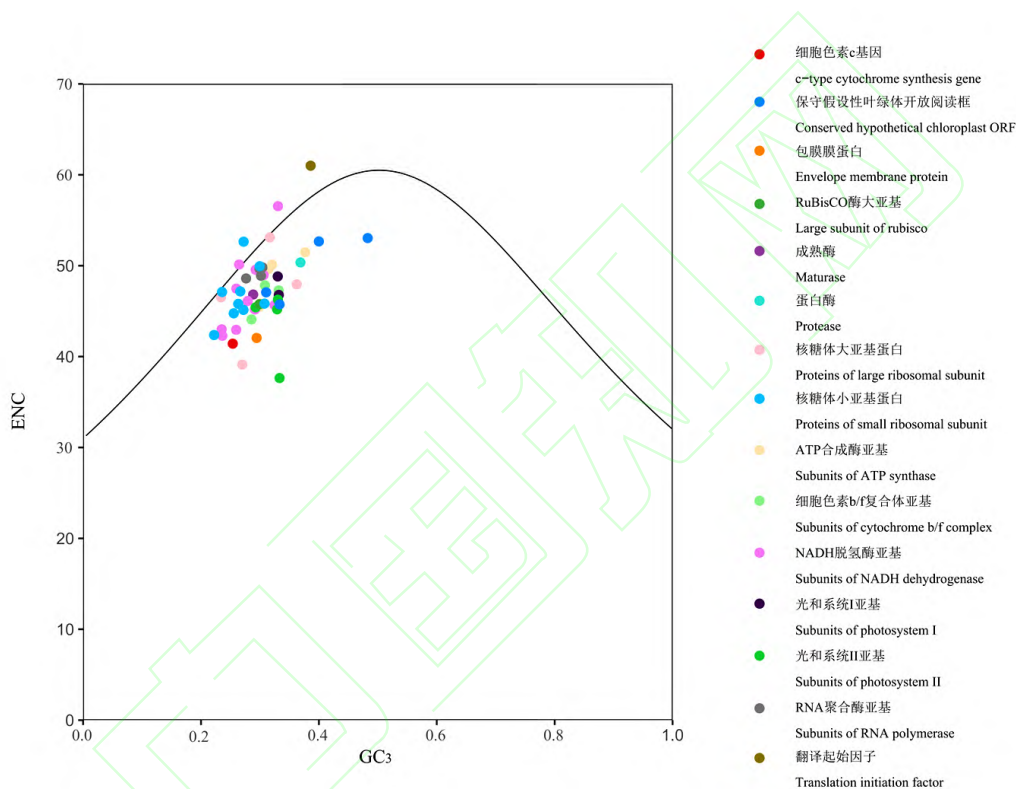


图 5 小黑麦叶绿体基因组 ENC-plot 绘图分析

Fig. 5 ENC-plot analysis in chloroplast genome of Triticale

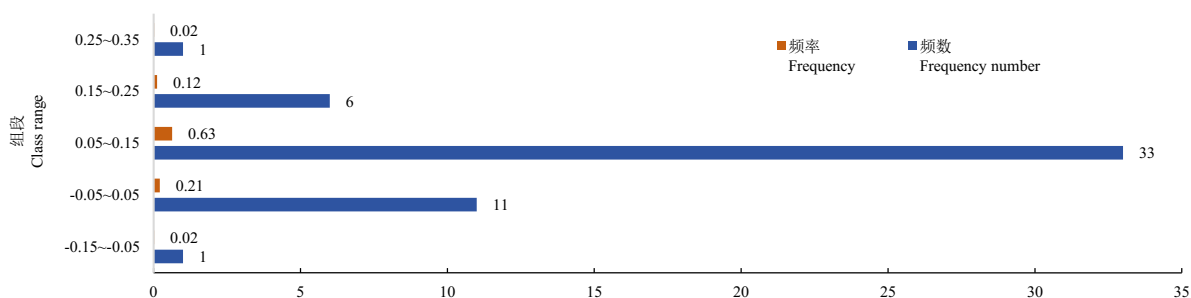


图 6 小黑麦叶绿体基因 ENC 比值频数分布

Fig. 6 Frequency distribution of ENC ratio in chloroplast genes of Triticale

2.4 PR2-plot 绘图分析

通过 PR2-plot 分析密码子偏倚性，当突变为唯一影响因素时，四种碱基在图中的表现为均衡，即集中于中心点，而当有选择压力同时存在时，根据影响大小而不同程度偏离中心点，向四周不均匀排布。小黑麦叶绿体基因组 PR2-plot 分析结果表明（图 7），涉及保守假设性叶绿体 ORF 和 RNA 聚合酶亚基两种功能中的 2 个基因接近于中心点，偏向程度小，表明其密码子偏好性受到突变的影响较多。其他大多数基因均坐落于平面图的下半部，尤其以右下方基因数量最多，说明在碱基出现的频率上 $G_3 > C_3$ ， $T_3 > A_3$ ；RuBisCO 酶大亚基全部基因落于左下方，即密码子偏倚性为 $C_3 > G_3$ ， $T_3 > A_3$ 。图中可看出各基因分布不均匀，以平面图左下和右下居多，说明 4 种碱基的使用频率存在一定的差异，且同时受到突变和选择的多重因素的影响。

整体而言，多数基因在 A_3/U_3 偏倚表现出了功能相关性：光合系统 I 亚基、光合系统 II 亚基、细胞色素 c 基因、NADH 脱氢酶亚基、成熟酶等基因的密码子都更偏好使用 U_3 ；RNA 聚合酶亚基的众多基因中，仅有一个未表现出明显的 A_3/U_3 偏倚，其余基因都更偏好使用 U_3 ；核糖体小亚基蛋白的众多基因中，多数都更偏好使用 A_3 。然而，在 C_3/G_3 偏倚方面，整体上未表现出明显的基因功能相关性。小黑麦叶绿体基因组密码子偏好性受到了突变和选择的共同影响，而这种影响可能与相应基因的功能是相关的。

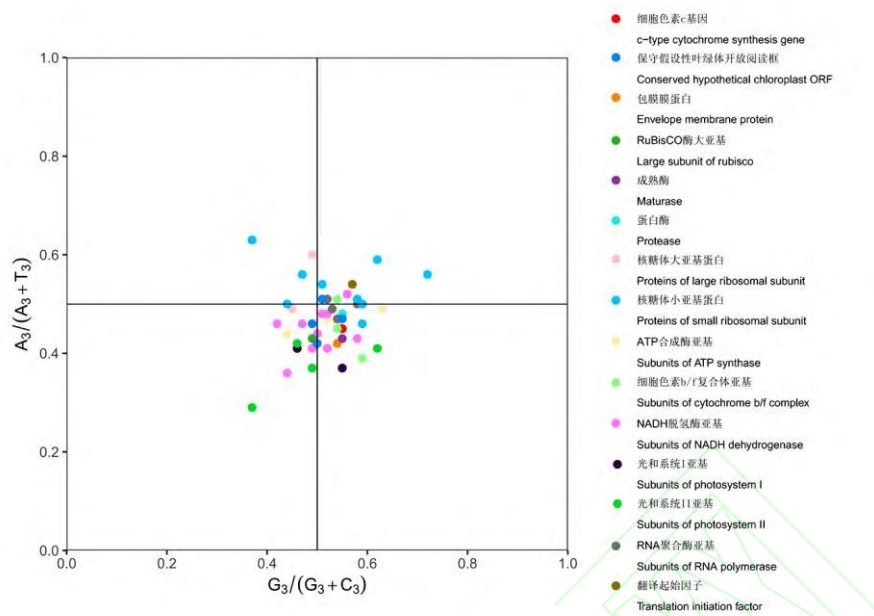


图 7 小黑麦叶绿体基因组 PR2-plot 分析

Fig. 7 PR2-plot analysis in chloroplast genome of Triticale

2.5 最优密码子的确定

通过构建高、低表达基因库，将 $\Delta\text{RSCU} \geq 0.08$ 且 $\text{RSCU} \geq 1$ 的密码子选出作为最优密码子，共得到 GCU、AGA 和 CGU 等 16 个最优密码子（GCU，AGA，CGU，CAA，GGU，CAU，AUU，UUA，AAA，UUU，CCU，AGU，UCU，ACA，ACU，GUU），全部以 A/U 结尾（表 2），这与 GC3、RSCU 分析结果一致。

表 2 小黑麦叶绿体基因组最优密码子分析

Table 2 Optimal codons analysis in chloroplast genome of Triticale

		高表达基因		低表达基因		
氨基酸	密码子	High-expressed genes		Low-expressed genes		同义密码子
		同义密码子		同义密码子		相对使用度
		频率	频率			
		相对使用度	相对使用度			
		Frequency	Frequency			
Amino acid	Codon		RSCU		RSCU	Δ RSCU
丙氨酸	GCA	28	1.29	12	1.37	-0.08

Ala	GCC	9	0.41	7	0.8	-0.39
	GCG	5	0.23	5	0.57	-0.34
	GCU***	45	2.07	11	1.26	0.81
	AGA**	19	1.84	15	1.48	0.36
	AGG	5	0.48	11	1.08	-0.6
精氨酸	CGA	10	0.97	13	1.28	-0.31
Arg	CGC*	4	0.39	2	0.2	0.19
	CGG	0	0	11	1.08	-1.08
	CGU***	24	2.32	9	0.89	1.43
天冬酰胺	AAC	17	0.58	9	0.58	0
Asn	AAU	42	1.42	22	1.42	0
天冬氨酸	GAC	4	0.28	4	0.32	-0.04
Asp	GAU	25	1.72	21	1.68	0.04
半胱氨酸	UGC**	3	0.43	0	0	0.43
Cys	UGU	11	1.57	2	2	-0.43
谷氨酰胺	CAA*	23	1.48	7	1.4	0.08
Gln	CAG	8	0.52	3	0.6	-0.08
谷氨酸	GAA	50	1.43	27	1.46	-0.03
Glu	GAG	20	0.57	10	0.54	0.03
	GGA	35	1.43	16	1.68	-0.25
甘氨酸	GGC*	11	0.45	3	0.32	0.13
Gly	GGG	7	0.29	10	1.05	-0.76
	GGU***	45	1.84	9	0.95	0.89
组氨酸	CAC	9	0.62	6	0.8	-0.18
His	CAU*	20	1.38	9	1.2	0.18
异亮氨酸	AUA	48	1.04	18	1.04	0
Ile	AUC	25	0.54	14	0.81	-0.27

	AUU*	66	1.42	20	1.15	0.27
	CUA	27	0.91	12	1.03	-0.12
	CUC	9	0.3	10	0.86	-0.56
亮氨酸	CUG	11	0.37	6	0.51	-0.14
Leu	CUU	37	1.25	16	1.37	-0.12
	UUA***	69	2.33	9	0.77	1.56
	UUG	25	0.84	17	1.46	-0.62
赖氨酸	AAA***	42	1.65	21	1.05	0.6
Lys	AAG	9	0.35	19	0.95	-0.6
蛋氨酸	AUG	33	1	16	1	0
Met						
苯丙氨酸	UUC	34	0.72	15	0.86	-0.14
Phe	UUU*	61	1.28	20	1.14	0.14
	CCA	12	0.91	8	1.19	-0.28
脯氨酸	CCC	11	0.83	8	1.19	-0.36
Pro	CCG	2	0.15	5	0.74	-0.59
	CCU***	28	2.11	6	0.89	1.22
	AGC	8	0.38	5	0.58	-0.2
	AGU*	27	1.28	9	1.04	0.24
丝氨酸	UCA	17	0.8	13	1.5	-0.7
Ser	UCC	19	0.9	13	1.5	-0.6
	UCG	10	0.47	4	0.46	0.01
	UCU***	46	2.17	8	0.92	1.25
	ACA*	23	1.16	7	1.04	0.12
苏氨酸	ACC	15	0.76	8	1.19	-0.43
Thr	ACG	8	0.41	6	0.89	-0.48
	ACU***	33	1.67	6	0.89	0.78

色氨酸	UGG	43	1	7	1	0
Trp						
酪氨酸	UAC	11	0.46	6	0.41	0.05
Tyr	UAU	37	1.54	23	1.59	-0.05
	GUA	30	1.52	14	1.56	-0.04
缬氨酸	GUC*	11	0.56	3	0.33	0.23
Val	GUG	4	0.2	7	0.78	-0.58
	GUU**	34	1.72	12	1.33	0.39

注: *表示 $\Delta\text{RSCU} \geq 0.08$; **表示 $\Delta\text{RSCU} \geq 0.30$; ***表示 $\Delta\text{RSCU} \geq 0.50$

Note: * indicates $\Delta\text{RSCU} \geq 0.08$; ** indicates $\Delta\text{RSCU} \geq 0.30$; *** indicates $\Delta\text{RSCU} \geq 0.50$

3 讨论

研究在获得了小黑麦的完整叶绿体基因组序列之后,探究了 52 个 CDS 的密码子使用特征及其与基因功能的关系。首先,将小黑麦叶绿体基因组中的蛋白编码序列作为一个整体进行分析而不考虑其功能基因时,密码子组成特征分析表明,GC 含量较低,且 GC1(47.43%)>GC2(39.43%)>GC3(30.09%),说明密码子分布不均,偏向于 A/U 结尾,这与前人在高粱 (*Sorghum bicolor* L.)、苇状羊茅 (*Festuca arundinacea* Schreb) 大麦 (*Hordeum vulgare* ssp. *vulgare*) 以及多年生黑麦草 (*Lolium perenne* L.) 等 23 种禾本科 (Gramineae) 植物中的研究结果一致^[39-40]。根据 RSCU 分析结果,小黑麦叶绿体基因组的编码氨基酸的密码子中,偏好使用和非偏好使用的密码子种类几乎持平,但总体偏好使用的密码子多以 A/U 结尾,结果与此前研究所报道的双子叶植物叶绿体密码子使用偏好特征一致^[14,41-42],符合“高等植物密码子倾向于使用 A/U 结尾”的假设^[43],说明植物叶绿体密码子的总体使用模式在进化过程中比较保守。密码子使用特征的数据主要应用于优化或重新设计基因,以便在异源系统中获得高表达或最佳表达,本研究确定了小黑麦叶绿体基因组的 16 个最优密码子,在进行基因工程研究时,针对最优密码子的优化将有利于调控相关基因的表达^[38]。

将密码子与基因功能相联合分析时,不同功能的基因使用的密码子具有较大的区别。研究表明,CAI 值越高,基因的密码子使用偏好性越强,在生物体内的表达水平越高^[44]。小黑

麦叶绿体基因组的 52 个 CDS 中, 各基因的 CAI 值在 0~1.0 之间, 平均值为 0.17, 密码子适应指数 CAI 总体偏小, 说明该基因组的密码子使用偏好性总体偏弱。相较而言, 小黑麦的 *rpl16*, *rps8*, *ycf68*, *ndhG* 和 *rpl20* 等基因的密码子使用偏好性较弱, 而 *psbB*, *rps3*, *psbD*, *rbcL* 和 *psbA* 等基因的密码子使用偏好性较强, 说明 *psbB*, *rps3*, *psbD* 等基因可能在小黑麦的叶绿体中表达水平较高。事实上, 光合作用功能是叶绿体的最重要的功能^[3], 这意味着与光合作用相关的 *psbB*, *rps3*, *psbD* 等基因必然有较高的表达水平。

对于各基因, ENC 的值介于 20 (当每个氨基酸只使用一个密码子时存在极端偏差) 和 61 (当所有密码子都被随机使用时) 之间, 小黑麦叶绿体 CDS 的 ENC 保持在较高水平 (均值为 47.31>45.00), 说明密码子使用的偏好性总体较弱^[45], 与 CAI 的分析结果一致。值得注意的是, ENC 值高达 61 的 *inFA* 属于翻译起始因子, 通过与核糖体、mRNA 和 tRNA 的相互作用, 帮助识别起始密码子并启动蛋白质合成, 但在小黑麦的叶绿体基因组上所有密码子都被随机使用, 说明该基因表达量低, 起到特定识别的能力较弱。中性绘图、ENC-GC_{3s} 和 PR2-plot 均显示绝大多数功能基因的密码子使用的偏好性受突变和自然选择压力的共同影响, 且以自然选择的影响为主。生物现有同义密码子使用模式的形成是突变、自然选择和基因翻译效率的遗传漂变之间平衡的结果^[46-47], 小黑麦叶绿体基因组中与光合作用密切相关的重要功能基因必然会受到突变和选择的强烈作用, 也许正是这些原因导致了小黑麦叶绿体基因组上与光合作用相关基因的密码子使用常表现出更鲜明的共同特征。

本研究结果可以为小黑麦叶绿体基因组在基因工程中的应用提供依据。密码子使用模式及其影响因素的分析结果表明, 小黑麦叶绿体基因, 尤其是与光合作用密切相关的基因, 可能是因受到选择突变和选择的强烈影响, 其密码子使用模式与基因功能的有较大的相关性。将来的研究可以通过结合生物信息学和系统生物学等技术手段, 进行大规模的叶绿体基因组密码子偏好性研究, 以探究叶绿体基因组密码子偏好性与基因表达调控之间的关系, 从而揭示更多的密码子使用模式规律和生物学意义, 以便更好地了解叶绿体基因表达的分子机制和密码子使用模式在生物演化中的作用和意义。

4 结论

通过对小黑麦叶绿体基因组密码子偏好性特征及其影响因素进行多角度分析, 筛选出 1

6 个最优密码子,明确了密码子偏好性较弱,但对 A/U 具有明显的偏倚性,造成偏倚性主要因素为选择压力。研究表明小黑麦叶绿体基因组密码子使用模式与功能基因之间具有较大的相关性,尤其是光合作用相关基因的密码子使用模式上,更偏向于使用 A₃/U₃。研究结果可为小黑麦基因工程提供理论基础,特别是在提高基因表达量的设计上或能提供研究基础。

参考文献

- [1] VIDI P A,KANWISCHER M,BAGINSKY S,et al. Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles[J]. *Journal of Biological Chemistry*, 2006, 281(16):11225-11234
- [2] LIPPOLD F,VOM DORP K,ABRAHAM M,et al. Fatty acid phytyl ester synthesis in chloroplasts of *Arabidopsis*[J]. *The Plant Cell*,2012,24(5):2001-2014
- [3] DANIELL H,JIN S X,ZHU X G,et al. Green giant—a tiny chloroplast genome with mighty power to produce high-value proteins:history and phylogeny[J]. *Plant Biotechnology Journal*,2021,19(3):430-447
- [4] DOBROGOJSKI J,ADAMIEC M,LUCIŃSKI R. The chloroplast genome:a review[J]. *Acta Physiologiae Plantarum*,2020,42(6):98
- [5] SUGIURA M. History of chloroplast genomics[J]. *Photosynthesis Research*,2003,76(1/2/3):371-377
- [6] GREVICH J J,DANIELL H. Chloroplast genetic engineering:recent advances and future perspectives[J]. *Critical Reviews in Plant Sciences*,2005,24(2):83-107
- [7] DANIELL H,LIN C S,YU M,et al. Chloroplast genomes:diversity,evolution,and applications in genetic engineering[J]. *Genome Biology*,2016,17(1):134
- [8] KWAK S Y,LEW T T S,SWEENEY C J,et al. Chloroplast-selective gene delivery and expression in planta using chitosan-complexed single-walled carbon nanotube carriers[J]. *Nature Nanotechnology*,2019,14(5):447-455
- [9] RUF S,HERMANN M,BERGER I J,et al. Stable genetic transformation of tomato plastids and expression of a

foreign protein in fruit[J]. Nature Biotechnology,2001,19(9):870-875

- [10] RAUBESON L A, PEERY R, CHUMLEY T W, et al. Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*[J]. BMC Genomics, 2007, 8: 174
- [11] 高守與, 李钰莹, 杨志青, 等. 白羊草叶绿体基因组密码子使用偏好性分析[J]. 草业学报, 2023, 32(7): 85-95
- [12] IKEMURA T. *Codon* usage and tRNA content in unicellular and multicellular organisms[J]. Molecular Biology and Evolution, 1985, 2(1): 13-34
- [13] QUAX T E F, CLAASSENS N J, SÖLL D, et al. *Codon* bias as a means to fine-tune gene expression[J]. Molecular Cell, 2015, 59(2): 149-161
- [14] WANG Z J, XU B B, LI B, et al. Comparative analysis of *Codon* usage patterns in chloroplast genomes of six *Euphorbiaceae* species[J]. PeerJ, 2020, 8: e8251
- [15] SHENG J J, SHE X, LIU X Y, et al. Comparative analysis of *Codon* usage patterns in chloroplast genomes of five *Miscanthus* species and related species[J]. PeerJ, 2021, 9: e12173
- [16] MAZUMDAR P, BINTI OTHMAN R, MEBUS K, et al. *Codon* usage and *Codon* pair patterns in non-grass monocot genomes[J]. Annals of Botany, 2017, 120(6): 893-909
- [17] CHEN S L, LEE W, HOTTES A K, et al. *Codon* usage between genomes is constrained by genome-wide mutational processes[J]. Proceedings of the National Academy of Sciences of the United States of America, 2004, 101(10): 3480-3485
- [18] 刘慧, 王梦醒, 岳文杰, 等. 糜子叶绿体基因组密码子使用偏性的分析[J]. 植物科学学报, 2017, 35(3): 362-371
- [19] 原晓龙, 李云琴, 张劲峰, 等. 降香黄檀叶绿体基因组密码子偏好性分析[J]. 广西植物, 2021, 41(4): 622-

- [20] 刘晶,赵明月,郭宏玉,等. 高寒牧区小黑麦研究进展[J]. 青海科技,2021,28(4):23-26
- [21] 王伟强,刘晶,田新会,等. '甘农 4 号'小黑麦品种在青海省不同区域的适应性评价[J]. 草地学报,2020,28(6):1626-1634
- [22] 裴亚斌,杜文华,刘汉成,等. 甘南高寒牧区 3 种饲草不同种植模式下的生产性能及经济效益[J]. 草业科学,2020,37(4):791-799
- [23] 杨晓,宋谦,余小亮,等. 陇东旱塬区秋播小黑麦与青贮玉米的复种效果[J]. 草业科学,2019,36(8):2127-2134
- [24] 杨永强,尹玉林,杜文华. 青藏高原川西北牧区一年两季饲草种植模式对土壤理化性质及酶活性的影响[J]. 草原与草坪,2024,44(02):57-66
- [25] 宋谦,马淑梅,张兰,等. 陇东地区秋播小黑麦茬后复种青贮玉米品种的筛选[J]. 中国饲料,2024(7):140-145
- [26] 倪典墨,李冉郡,郭忠静,等. 百合科植物 *psbA* 基因密码子偏好性分析[J]. 生物学杂志,2020,37(6):68-72
- [27] BIN PEK H, KLEMENT M, ANG K S, et al. Exploring *Codon* context bias for synthetic gene design of a thermostable invertase in *Escherichia coli*[J]. *Enzyme and Microbial Technology*, 2015, 75: 57-63
- [28] 时欢,林玉玲,赖钟雄,等. CRISPR/Cas9 介导的植物基因编辑技术研究进展[J]. 应用与环境生物学报,2018,24(03):640-650
- [29] DOUDNA J A, CHARPENTIER E. The new frontier of genome engineering with CRISPR-Cas9[J]. *Science*, 2014, 346(6213): 1258096
- [30] JIANG F, DOUDNA J A. CRISPR-Cas9 structures and mechanisms[J]. *Annual Review of Biophysics*, 2017, 46(1): 505-529

- [31] MA X,ZHANG Q,ZHU Q,et al. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants[J]. *Molecular Plant*,2015,8(8):1274-1284
- [32] SHAN Q,WANG Y,LI J,et al. Targeted genome modification of crop plants using a CRISPR-Cas system[J]. *Nature Biotechnology*,2013,31(8): 686-688
- [33] ZHANG H,ZHANG J S,WEI P L,et al. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation[J]. *Plant Biotechnology Journal*,2014,12(6):797-807
- [34] SUEOKA N. Directional mutation pressure and neutral molecular evolution[J]. *Proceedings of the National Academy of Sciences of the United States of America*,1988,85(8):2653-2657
- [35] 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
- [36] WRIGHT F. The 'effective number of codons' used in a gene[J]. *Gene*,1990,87(1):23-29
- [37] SUEOKA N. Translation-coupled violation of Parity Rule 2 in human genes is not the cause of heterogeneity of the DNA G+C content of third *Codon* position[J]. *Gene*,1999,238(1):53-58
- [38] WANG L J,ROOSSINCK M J. Comparative analysis of expressed sequences reveals a conserved pattern of optimal *Codon* usage in plants[J]. *Plant Molecular Biology*,2006,61(4/5):699-710
- [39] 洪森荣,林顺来,李盈萍,等. 甜高粱叶绿体基因组特征及密码子偏好性分析[J]. *草地学报*,2023,31(12):3636-3650
- [40] ZHANG Y R,NIE X J,JIA X O,et al. Analysis of *Codon* usage patterns of the chloroplast genomes in the Poaceae family[J]. *Australian Journal of Botany*,2012,60(5):461
- [41] 季凯凯,宋希强,陈春国,等. 木兰科叶绿体基因组的密码子使用特征分析[J]. *中国农业科技导报*,2020,22(11):52-62

- [42] YANG S W,LI G,LI H. Molecular characterizations of genes in chloroplast genomes of the genus *Arachis* L.(Fabaceae) based on the *Codon* usage divergence[J]. PLoS One,2023,18(3):e0281843
- [43] CAMPBELL W H,GOWRI G. *Codon* usage in higher plants,green algae,and cyanobacteria[J]. Plant Physiology,1990,92(1):1-11
- [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
- [45] SALIM H M W,CAVALCANTI A R O. Factors influencing *Codon* usage bias in genomes[J]. Journal of the Brazilian Chemical Society,2008,19(2):257-262
- [46] LIU Q P. Mutational bias and translational selection shaping the *Codon* usage pattern of tissue-specific genes in rice[J]. PLoS One,2012,7(10):e48295
- [47] PLOTKIN J B,KUDLA G. Synonymous but not the same:the causes and consequences of *Codon* bias[J]. Nature Reviews Genetics,2011,12(1):32-42