

朱 艳, 韩长志. 普通核桃与深纹核桃中叶绿体基因组密码子使用特征对比[J]. 江苏农业学报, 2024, 40(8) : 1507-1520.
doi: 10.3969/j.issn.1000-4440.2024.08.015

普通核桃与深纹核桃中叶绿体基因组密码子使用特征对比

朱 艳, 韩长志
(西南林业大学林学院/云南省森林灾害预警与控制重点实验室, 云南 昆明 650224)

摘要: 本研究旨在对比分析普通核桃与深纹核桃中叶绿体基因组密码子的使用偏好性,对普通核桃与深纹核桃中各 53 条蛋白质编码区序列开展预测分析,明确上述 2 种核桃叶绿体基因组中的最优密码子,为未来开展核桃物种的亲缘关系研究提供理论依据。利用 Codon W 1.4.2 和在线软件 CUSP 等分析普通核桃与深纹核桃中的叶绿体基因组密码子,获得相对同义密码子使用度(RSCU)、有效密码子数(ENC)、G+C 含量等参数,再进行中性绘图分析、有效密码子数分析及奇偶偏好性分析。普通核桃与深纹核桃中叶绿体基因组的密码子适应指数均为 0.167, ENC 均在 45.00 以上,表明密码子偏好性弱。普通核桃与深纹核桃中叶绿体基因组密码子的 G+C 含量整体表现为:基因中所有密码子第 1 位碱基 G+C 的含量(GC_1)>基因中所有密码子第 2 位碱基 G+C 的含量(GC_2)>基因中所有密码子第 3 位碱基 G+C 的含量(GC_3)。进一步分析发现,第 3 位碱基的使用频率表现为 T 大于 A, G 大于 C。从普通核桃中筛选出 14 个最优密码子,从深纹核桃中筛选出 17 个最优密码子,其中共有 23 个最优密码子在第 3 个位置偏好以 A/U 结尾。总之,普通核桃和深纹核桃中叶绿体基因组密码子偏好性主要受到自然选择的影响,本研究结果可以为进一步探索核桃抗性遗传基因的改良和表达提供理论基础和数据支撑。

关键词: 普通核桃; 深纹核桃; 叶绿体; 基因组; 同义密码子; 偏好性

中图分类号: S664.1 **文献标识码:** A **文章编号:** 1000-4440(2024)08-1507-14

Comparison of codon usage characteristics of chloroplast genomes in *Juglans regia* and *Juglans sigillata*

ZHU Yan, HAN Changzhi
(College of Forestry, Southwest Forestry University/Key Laboratory of Forest Disaster Warning and Control of Yunnan Province, Kunming 650224, China)

Abstract: The aim of the study was to compare and analyze the usage preference of chloroplast genome codons in *Juglans regia* and *Juglans sigillata*, and to predict and analyze 53 protein coding sequences (CDS) in *Juglans regia* and *Juglans sigillata*, so as to clarify the optimal codons in the chloroplast genomes of the above two walnuts and provide a theoretical basis for the future research of the genetic relationship of walnut species. Codon W 1.4.2 and online software CUSP were used to analyze the chloroplast genome codons in *Juglans regia* and *Juglans sigillata*, and the relative synonymous codon usage (RSCU), effective number of codons (ENC), G+C content and other parameters were obtained. Neutrality plot analysis, effective number of codons analysis (ENC-plot), and parity preference analysis (PR2-plot) were performed. The codon adaptation index of chloroplast genomes in *Juglans regia* and *Juglans sigillata* was 0.167, and the ENC was above 45.00, indicating that the codon preference was weak. The G+C content of chloroplast genome codons in *Juglans regia* and *Juglans sigillata* was as follows: the

收稿日期:2023-09-06
基金项目:云南省研究生导师团队建设项目(2022100);云南省“兴滇英才支持计划”青年人才专项(YNWR-QNBJ-2020-188)
作者简介:朱 艳(2000-),女,云南昭通人,硕士研究生,主要从事森林保护相关研究。(E-mail)508034093@qq.com
通讯作者:韩长志,(E-mail)swfuhcz@163.com

content of G+C at the first base of all codons in the gene (GC_1)>the content of G+C at the second base of all codons in the gene (GC_2)>the content of G+C at the third base of all codons in the gene (GC_3). Further analysis showed that the usage frequency of the third base was T>A, G>C. A total of 14 optimal codons were screened in *Juglans regia*, and 17 optimal codons were screened in *Juglans sigillata*. Among them, 23 optimal codons preferred the A/U end at the third position. In conclusion, the codon preference of chloroplast genome in *Juglans regia* and *Juglans sigillata* is mainly affected by natural selection, which provides theoretical basis and data support for further exploration of the improvement and expression of walnut resistance genes.

Key words: *Juglans regia*; *Juglans sigillata*; chloroplast; genome; synonymous codon; bias

生物体内的各种蛋白质由 20 种氨基酸组成,其中色氨酸、甲硫氨酸只有 1 个由三核苷酸组成的密码子,其他氨基酸有 2~6 个由三核苷酸组成的密码子^[1]。通过对密码子开展分析能够辅助识别自然界中的遗传信息^[2],目前共有 64 个密码子,18 种氨基酸能够利用其简并性实现不同密码子编码同一氨基酸^[3]。不同物种、同一物种不同组织或基因中的密码子存在不均衡使用的情况,即同义密码子偏好性(SCUB)^[4-6]。最优密码子是在物种高表达基因中频繁出现、使用频率最高的密码子,而部分很少使用甚至不会出现的密码子称为稀有密码子或者非优化密码子^[3]。SCUB 一般会受到基因突变、自然选择、基因漂移等多种因素影响^[7],对其开展研究,有助于了解物种之间的亲缘关系、遗传相似性、基因水平转移、基因分子进化及系统发育等,进而促进外源基因的表达和转基因作物研究的发展^[8-10]。叶绿体是绿色植物和藻类能进行光合作用的细胞器^[11],具有基因组小、拷贝数多、结构简单、进化保守等诸多优点,因此,叶绿体作为研究对象被广泛应用于 DNA 条形码分子标记、系统进化以及基因工程等研究领域^[8,12]。前人对不同物种叶绿体基因组中 SCUB 展开了分析工作,如肉苁蓉(*Cistanche deserticola* Ma)^[13]、茶树(*Camellia sinensis*)^[14]、二乔玉兰[*Yulania* × *soulangeana* (Soul.-Bod.) D. L. Fu]^[15]、蒜头果(*Malania oleifera* Chun & S. K. Lee)^[16]、榕树(*Ficus microcarpa* L. f.)^[17]和圆柏(*Juniperus chinensis* Roxb.)^[18]。

作为世界四大干果之一,核桃中含有大量的营养物质^[19],自古以来就深受消费者的喜爱^[20],同时,作为“木本油料之王”,其在木材建筑、经济效益方面发挥的作用显著^[21]。深纹核桃(*Juglans sigillata* D.)又被称为云南核桃、南方核桃^[22],主要

分布于中国西南地区和西藏自治区雅鲁藏布江中上游地区,生长于海拔 1 300~3 300 m 的山坡或山谷林中。普通核桃(*Juglans regia* L.)是胡桃属植物,主要盛产于中国北方。目前有关上述 2 种核桃的研究工作主要体现在分子生物学方面,涉及种质资源、表型研究和遗传多样性等^[22-24],然而,关于核桃叶绿体 DNA 的研究较少。本研究拟以深纹核桃、普通核桃中叶绿体基因组的高通量测序结果为基础,通过对比分析其叶绿体基因组蛋白质编码区序列(CDS)的碱基组成,利用中性绘图分析、有效密码子数分析(ENC-plot)及奇偶偏好性分析(PR2-plot)等深入剖析影响密码子偏好性的主要因素,并筛选获得叶绿体基因组的最优密码子,以期为进一步开展 2 种核桃的基因组研究、种质创新研究提供数据支撑。

1 材料与方法

1.1 试验材料

从美国国家生物技术信息中心(NCBI)数据库中获取普通核桃和深纹核桃的叶绿体基因组 CDS 序列,前者 GenBank 登录号为 KT963008.1,利用 EditSeq 计算 G+C 含量为 36.23%;后者 GenBank 登录号为 MF167465.1, G+C 含量为 36.02%。通过在线软件(<https://irscope.shinyapps.io/Chloroplast/>)绘制普通核桃和深纹核桃的叶绿体基因组图谱。

1.2 试验方法

1.2.1 密码子碱基组成分析 使用 Codon W 1.4.2 处理数据得到密码子适应指数(CAI)、有效密码子数(ENC)、相对同义密码子使用度(RSCU),通过 CUSP 分析获得基因中所有密码子第 1 位碱基、第 2 位碱基、第 3 位碱基 G+C 的含量(GC_1 、 GC_2 、 GC_3),最后使用 IBM SPSS Statistics、Excel、Origin 等分析软件分析并作图^[25]。

1.2.2 中性绘图分析 分别以 GC_{12} (GC_1 和 GC_2 的

平均值)为纵坐标,GC₃为横坐标,绘制散点图,分析密码子使用偏好性的影响因素^[26],进而对密码子不同位置的碱基使用性开展相关性分析^[27]。

1.2.3 ENC-plot 分析 分别以 GC₃作为 X 轴变量,ENC 作为 Y 轴变量绘制散点图,通过公式(1)计算出各基因的标准 ENC 并绘制标准曲线。ENC 的理论取值为 20~61,其值越接近 20 表明偏好性越强,反之,其值越接近 61 表明偏好性越弱,若散点与标准曲线接近,表示密码子使用偏好性的主要影响因素是基因突变,若散点远离标准曲线,说明密码子使用偏好性的主要影响因素是自然选择^[28]。为了更好地评估 ENC 实际观察值与 ENC 期望值之间的差异,根据公式(2)计算 ENC 比值并进行分析。

$$ENC = 2 + GC_3 + 29 / [GC_3^2 + (1 - GC_3)^2] \quad (1)$$

$$ENC \text{ 比值} = (ENC_{\text{exp}} - ENC_{\text{obs}}) / ENC_{\text{exp}} \quad (2)$$

式中,ENC 表示基因中使用有效密码子数;GC₃表示基因中所有密码子的第 3 位碱基 G+C 的含量;ENC_{exp}表示 ENC 期望值;ENC_{obs}表示 ENC 实际观察值。

1.2.4 PR2-plot 分析 分别以 G₃为分子,G₃与 C₃的和为分母,所得比值为 X 轴变量;以基因中所有密码子第 3 位碱基 A 的含量(A₃)为分子,A₃与基因中所有密码子第 3 位碱基 T 的含量(T₃)的和为分母,所得比值为 Y 轴变量,绘制散点图^[29]。中心点位置表示 A=T、G=C,即密码子不具有偏好性,其他点到中心点的距离表示该基因的偏倚程度。

1.2.5 最优密码子的确定 以密码子的 ENC 作为分析偏好性的标准,选取排序排列后 ENC 最低和最高两端各 10% 的基因,建立基因库。分别计算 2 组的 RSCU 值和 ΔRSCU 值(高表达基因的 RSCU 值-低表达基因的 RSCU 值)^[30],确定 2 种核桃叶绿体基因组的高频密码子(RSCU>1.00)和高表达密码子(ΔRSCU≥0.08),最优密码子要同时满足 RSCU>1.00、ΔRSCU≥0.08 这 2 个条件^[31]。

2 结果与分析

2.1 密码子碱基组成分析

普通核桃和深纹核桃的叶绿体基因组呈双链环状,为典型的四分体结构(图 1、图 2),普通核桃大单拷贝区(LSC)为 89 872 bp,小单拷贝区(SSC)为 18 423 bp,反向重复区包括 IRA 和 IRB,均为 26 036 bp;深纹核桃 LSC 为 89 871 bp,SSC 为

18 412 bp,IRA 为 26 034 bp,IRB 为 26 034 bp。普通核桃的叶绿体基因组一共注释到 136 个基因,包括 88 个蛋白质编码基因,40 个 tRNA 基因和 8 个 rRNA 基因,其中光合作用相关基因包括 6 类[5 个光系统 I 基因、13 个光系统 II 基因、12 个还原型辅酶 I (NADH)-脱氢酶基因、6 个细胞色素 b/f 复合物基因、6 个 ATP 合酶基因、1 个 RuBisCO 亚基基因],自我复制的相关基因包括 5 类(15 个小核糖体蛋白基因、11 个大核糖体蛋白基因、4 个 RNA 聚合酶基因、40 个 tRNA 基因、8 个 rRNA 基因),其他类型基因 15 个。深纹核桃的叶绿体基因组一共注释到 131 个基因,包括 86 个蛋白质编码基因,37 个 tRNA 基因和 8 个 rRNA 基因,其中光合作用相关基因包括 6 类[5 个光系统 I 基因、14 个光系统 II 基因、12 个 NADH-脱氢酶基因、6 个细胞色素 b/f 复合物基因、6 个腺嘌呤核苷三磷酸(ATP)合酶基因、1 个 RuBisCO 亚基基因],自我复制的相关基因包括 5 类(15 个小核糖体蛋白基因、10 个大核糖体蛋白基因、4 个 RNA 聚合酶基因、37 个 tRNA 基因、8 个 rRNA 基因),其他类型基因 13 个。

利用 GeneiousPrime 对 2 种核桃的叶绿体基因组序列进行筛选,删除长度过短或重复的 CDS,选择典型基因序列,最终普通核桃和深纹核桃各获得 53 条 CDS。进一步分析,普通核桃密码子第 3 位碱基上 T、C、A、G 的含量分别为 47.27%、15.55%、43.92%、16.21%;深纹核桃密码子第 3 位碱基上 T、C、A、G 的含量分别为 47.35%、15.50%、44.05%、16.10%,上述 2 种核桃中 T₃和 A₃高于 G₃和 C₃(表 1)。

普通核桃的 ENC 范围为 37.68~58.08 个,深纹核桃的 ENC 范围为 37.85~58.90 个,进一步分析发现,共有 74 条(普通核桃 37 条、深纹核桃 37 条)叶绿体基因组序列的有效密码子数大于 45.00 个,表明 2 种核桃叶绿体密码子使用偏好性较弱,从表 2 可以看出,GC 在不同位置、不同基因中未均匀分布。

对密码子各参数进行相关性分析,结果(表 3)表明,在普通核桃和深纹核桃中,ENC 与 GC₃极显著相关($P < 0.01$),与 GC₁、GC₂相关性不显著,说明 GC₃对密码子偏好性存在较强影响;GC₁、GC₂与 GC₃相关性不显著,说明 3 个位置的碱基组成差异较大。



Table 1 G+C content, effective number of codons (ENC) and codon adaptation index at different positions of chloroplast genes in *Juglans regia* and *Juglans sigillata*

T₃: 基因中所有密码子第 3 位碱基 T 的含量; C₃: 基因中所有密码子第 3 位碱基 C 的含量; A₃: 基因中所有密码子第 3 位碱基 A 的含量; G₃: 基因中所有密码子第 3 位碱基 G 的含量; GC₁: 基因中所有密码子第 1 位碱基 G+C 的含量; GC₂: 基因中所有密码子第 2 位碱基 G+C 的含量; GC₃: 基因中所有密码子第 3 位碱基 G+C 的含量; GC_{all}: 基因中所有密码子第 1 位、第 2 位、第 3 位碱基 G+C 含量的平均值; ENC: 有效密码子数; CAI: 密码子适应指数。

表 2 普通核桃和深纹核桃叶绿体基因组密码子在不同基因中的 G+C 含量及有效密码子数

Table 2 G+C content and effective number of codons in the chloroplast genome in different genes of *Juglans regia* and *Juglans sigillata*

基因	普通核桃					深纹核桃				
	GC ₁ (%)	GC ₂ (%)	GC ₃ (%)	GC _{all} (%)	ENC(个)	GC ₁ (%)	GC ₂ (%)	GC ₃ (%)	GC _{all} (%)	ENC(个)
<i>accD</i>	42.08	34.27	28.06	34.80	49.70	42.08	34.27	28.06	34.80	49.70
<i>atpA</i>	55.31	39.76	24.41	39.83	47.10	55.31	39.76	24.41	39.83	47.10
<i>atpB</i>	56.50	42.68	26.42	41.87	44.32	56.50	42.68	26.42	41.87	44.32
<i>atpE</i>	48.51	39.55	31.34	39.80	49.73	48.51	39.55	31.34	39.80	49.73
<i>atpF</i>	45.41	32.43	30.81	36.22	51.92	45.41	32.43	30.81	36.22	51.92
<i>atpI</i>	48.79	36.69	23.39	36.29	43.79	48.79	36.69	23.39	36.29	43.79
<i>ecsA</i>	33.86	36.99	25.71	32.19	46.62	33.86	36.99	25.71	32.19	46.62
<i>cemA</i>	38.26	28.26	33.04	33.19	39.76	38.26	28.26	33.04	33.19	39.76
<i>clpP</i>	58.38	37.56	28.93	41.62	51.74	57.87	37.06	28.93	41.29	51.57
<i>matK</i>	38.93	31.62	28.46	33.00	50.95	38.93	31.62	28.46	33.00	50.95
<i>ndhA</i>	43.13	37.91	21.15	34.06	43.86	43.13	37.91	21.15	34.06	43.86
<i>ndhB</i>	41.33	38.79	31.77	37.30	48.45	41.33	38.79	31.77	37.30	48.45
<i>ndhC</i>	46.28	33.88	23.97	34.71	47.60	46.28	33.88	23.97	34.71	47.60
<i>ndhD</i>	40.24	36.69	27.61	34.85	46.79	40.12	36.93	27.74	34.93	46.76
<i>ndhE</i>	41.18	33.33	22.55	32.35	50.14	41.18	33.33	22.55	32.35	50.14
<i>ndhF</i>	35.85	34.77	23.32	31.31	43.82	35.85	34.77	23.32	31.31	43.82
<i>ndhG</i>	43.50	34.46	29.38	35.78	47.18	43.50	34.46	29.38	35.78	47.18
<i>ndhH</i>	50.00	35.79	26.90	37.56	47.32	50.00	35.79	26.90	37.56	47.32
<i>ndhI</i>	41.32	36.53	20.96	32.94	40.60	41.32	36.53	20.96	32.94	40.60
<i>ndhJ</i>	51.57	37.74	28.93	39.41	46.78	51.57	37.74	28.93	39.41	46.78
<i>ndhK</i>	45.37	44.93	26.43	38.91	51.95	45.37	44.93	26.43	38.91	51.95
<i>petA</i>	52.48	36.02	30.75	39.75	51.11	52.48	36.02	30.75	39.75	51.11
<i>petB</i>	48.61	41.67	28.70	39.66	40.29	48.61	41.67	28.70	39.66	40.29
<i>petD</i>	50.93	39.13	22.98	37.68	37.68	50.91	39.39	22.42	37.57	37.85
<i>psaA</i>	51.93	43.54	31.16	42.21	49.67	51.93	43.54	31.16	42.21	49.67
<i>psaB</i>	48.71	42.86	29.52	40.36	47.55	48.71	42.86	29.52	40.36	47.55
<i>psbA</i>	49.72	43.50	33.05	42.09	41.19	49.72	43.50	33.05	42.09	41.19
<i>psbC</i>	53.69	45.49	32.17	43.78	47.03	53.80	45.99	30.80	43.53	45.83
<i>psbD</i>	52.54	43.22	31.07	42.28	46.60	52.54	43.22	31.07	42.28	46.60
<i>rbcL</i>	58.18	43.89	28.99	43.69	47.69	58.40	43.70	28.78	43.63	47.50
<i>rpl14</i>	53.66	37.40	23.58	38.21	47.19	53.66	37.40	23.58	38.21	47.19
<i>rpl16</i>	50.00	53.68	26.47	43.38	43.83	50.00	53.68	26.47	43.38	43.83
<i>rpl2</i>	51.22	47.04	31.01	43.09	52.54	51.27	48.00	31.64	43.64	52.68
<i>rpl20</i>	33.90	41.53	22.03	32.49	40.81	33.90	41.53	22.03	32.49	40.81
<i>rpl22</i>	35.33	37.72	23.95	32.33	44.95	35.33	37.72	23.95	32.33	44.95
<i>rpoA</i>	46.20	31.91	23.71	33.94	48.96	46.20	31.91	23.71	33.94	48.96
<i>rpoB</i>	50.14	38.00	26.52	38.22	47.69	50.14	38.00	26.52	38.22	47.69
<i>rpoC1</i>	50.73	37.32	28.86	38.97	50.02	50.81	37.34	28.84	39.00	49.98
<i>rpoC2</i>	46.03	36.03	27.09	36.38	48.80	46.03	36.03	27.09	36.38	48.80
<i>rps11</i>	51.80	56.83	25.90	44.84	44.09	51.80	56.83	25.90	44.84	44.09
<i>rps12</i>	51.20	48.80	31.20	43.73	48.54	51.20	48.80	31.20	43.73	48.54
<i>rps14</i>	43.56	47.52	30.69	40.59	38.92	43.56	47.52	30.69	40.59	38.92
<i>rps18</i>	37.25	41.18	25.49	34.64	38.22	37.25	41.18	25.49	34.64	38.22
<i>rps2</i>	42.62	40.51	27.43	36.85	47.30	42.62	40.51	27.43	36.85	47.30
<i>rps3</i>	48.64	34.55	22.27	35.15	44.59	48.64	34.55	22.27	35.15	44.59
<i>rps4</i>	50.50	37.62	25.74	37.95	49.10	50.50	37.62	25.74	37.95	49.10
<i>rps7</i>	52.56	44.87	25.00	40.81	45.12	52.56	44.87	25.00	40.81	45.12
<i>rps8</i>	40.74	39.26	29.63	36.54	49.73	40.74	39.26	29.63	36.54	49.73

续表2 Continued2

基因	普通核桃					深纹核桃				
	GC ₁ (%)	GC ₂ (%)	GC ₃ (%)	GC _{all} (%)	ENC(个)	GC ₁ (%)	GC ₂ (%)	GC ₃ (%)	GC _{all} (%)	ENC(个)
<i>ycf1</i>	40.00	29.87	31.17	33.68	52.66	34.67	26.43	23.57	28.22	46.51
<i>ycf2</i>	41.54	34.22	36.50	37.42	53.30	41.54	34.22	36.50	37.42	53.30
<i>ycf3</i>	46.15	38.46	31.95	38.85	58.08	46.15	38.46	31.36	38.66	58.90
<i>ycf4</i>	43.24	40.54	32.43	38.74	51.37	43.24	40.54	32.43	38.74	51.37
<i>psi</i>	54.81	46.37	28.29	43.16	46.51					
<i>psbB</i>						53.27	45.51	31.02	43.27	46.95

GC_{all}:基因中所有密码子第1位、第2位、第3位碱基G+C含量的平均值;GC₁:基因中所有密码子第1位碱基G+C的含量;GC₂:基因中所有密码子第2位碱基G+C的含量;GC₃:基因中所有密码子第3位碱基G+C的含量;ENC:有效密码子数。

表3 密码子不同位置上G+C含量、数量与有效密码子数的相关性分析

Table 3 Correlation analysis of G+C content, number and effective number of codons at different positions of codons

变量	普通核桃					深纹核桃				
	GC ₁	GC ₂	GC ₃	GC _{all}	ENC	GC ₁	GC ₂	GC ₃	GC _{all}	ENC
GC ₁	1.000					1.000				
GC ₂	0.426 **	1.000				0.449 **	1.000			
GC ₃	0.129	0.087	1.000			0.190	0.178	1.000		
GC _{all}	0.823 **	0.775 **	0.440 **	1.000		0.828 **	0.794 **	0.501 **	1.000	
ENC	0.130	-0.209	0.416 **	0.102	1.000	0.153	-0.161	0.394 **	0.126	1.000
N	-0.159	-0.123	0.304 *	-0.054	0.044	-0.031	0.015	0.321 **	0.089	0.089

**表示极显著相关($P<0.01$);*表示显著相关($P<0.05$)。N:密码子数;GC_{all}:基因中所有密码子第1位、第2位、第3位碱基G+C含量的平均值;GC₁:基因中所有密码子第1位碱基G+C的含量;GC₂:基因中所有密码子第2位碱基G+C的含量;GC₃:基因中所有密码子第3位碱基G+C的含量;ENC:有效密码子数。

表4 普通核桃和深纹核桃各氨基酸的相对同义密码子使用度(RSCU)

Table 4 Relative synonymous codon usage (RSCU) of each amino acid in *Juglans regia* and *Juglans sigillata*

氨基酸	普通核桃			深纹核桃		
	密码子	密码子数(个)	RSCU	密码子	密码子数(个)	RSCU
Phe	UUU	668	1.06	UUU	727	1.06
	UUC	589	0.94	UUC	643	0.94
Leu	UUA	483	1.00	UUA	532	1.05
	UUG	709	1.47	UUG	736	1.45
	CUU	494	1.02	CUU	514	1.01
	CUC	377	0.78	CUC	391	0.77
	CUA	467	0.97	CUA	490	0.96
Ile	CUG	373	0.77	CUG	384	0.76
	AUU	599	1.03	AUU	654	1.02
	AUC	508	0.87	AUC	542	0.84
	AUA	636	1.09	AUA	730	1.14
	AUG	608	1.00	AUG	642	1.00
Val	GUU	345	1.11	GUU	359	1.11
	GUC	221	0.71	GUC	234	0.72
	GUA	377	1.22	GUA	398	1.23
Ser	GUG	297	0.96	GUG	306	0.94
	UCU	236	1.26	UCU	245	1.24
	UCC	198	1.06	UCC	212	1.07
	UCA	310	1.66	UCA	332	1.67
Pro	UCG	196	1.05	UCG	202	1.02
	CCU	134	0.75	CCU	138	0.74
	CCC	140	0.78	CCC	145	0.78

续表4 Continued4

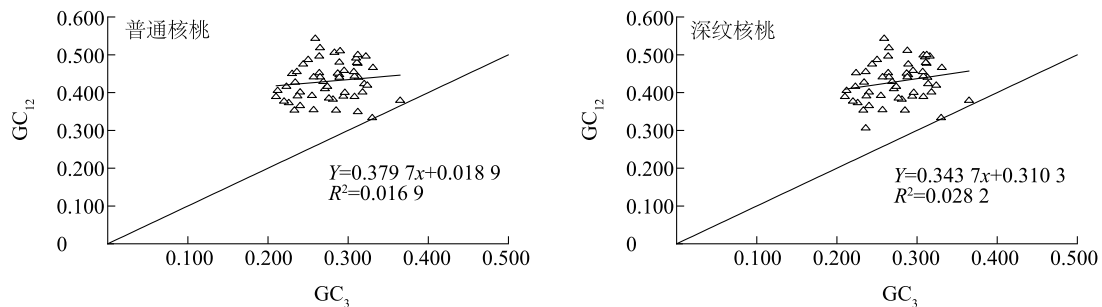
氨基酸	普通核桃			深纹核桃		
	密码子	密码子数(个)	RSCU	密码子	密码子数(个)	RSCU
Thr	CCA	226	1.26	CCA	239	1.28
	CCG	218	1.21	CCG	226	1.21
	ACU	119	0.77	ACU	131	0.77
	ACC	131	0.84	ACC	136	0.80
	ACA	241	1.55	ACA	271	1.60
Ala	ACG	131	0.84	ACG	139	0.82
	GCU	71	0.74	GCU	74	0.74
	GCC	89	0.93	GCC	91	0.91
	GCA	117	1.22	GCA	126	1.26
Tyr	GCG	106	1.11	GCG	110	1.10
	UAU	515	1.18	UAU	560	1.19
	UAC	360	0.82	UAC	378	0.81
TER	UAA	519	1.02	UAA	593	1.07
	UAG	610	1.20	UAG	635	1.15
His	CAU	239	1.18	CAU	256	1.18
	CAC	167	0.82	CAC	177	0.82
Gln	CAA	396	1.08	CAA	424	1.09
	CAG	336	0.92	CAG	354	0.91
Asn	AAU	469	1.16	AAU	526	1.17
	AAC	339	0.84	AAC	373	0.83
Lys	AAA	709	1.11	AAA	874	1.17
	AAG	574	0.89	AAG	622	0.83
Asp	GAU	270	1.25	GAU	288	1.27
	GAC	162	0.75	GAC	165	0.73
Glu	GAA	418	1.11	GAA	451	1.13
	GAG	335	0.89	GAG	346	0.87
Cys	UGU	203	1.12	UGU	213	1.11
	UGC	161	0.88	UGC	170	0.89
TER	UGA	398	0.78	UGA	428	0.78
Trp	UGG	391	1.00	UGG	398	1.00
Arg	CGU	56	0.36	CGU	59	0.35
	CGC	72	0.46	CGC	76	0.46
	CGA	159	1.02	CGA	165	0.99
	CGG	186	1.19	CGG	194	1.17
Ser	AGU	103	0.55	AGU	113	0.57
	AGC	80	0.43	AGC	86	0.43
Arg	AGA	258	1.65	AGA	289	1.74
	AGG	207	1.32	AGG	215	1.29
Gly	GGU	162	0.82	GGU	170	0.83
	GGC	130	0.66	GGC	132	0.65
	GGA	221	1.12	GGA	234	1.15
	GGG	275	1.40	GGG	281	1.38

Phe: 苯丙氨酸; Leu: 亮氨酸; Ile: 异亮氨酸; Met: 甲硫氨酸; Val: 缬氨酸; Ser: 丝氨酸; Pro: 脯氨酸; Thr: 苏氨酸; Ala: 丙氨酸; Tyr: 酪氨酸; His: 组氨酸; Gln: 谷氨酰胺; Asn: 天冬酰胺; Lys: 赖氨酸; Asp: 天冬氨酸; Glu: 谷氨酸; Cys: 半胱氨酸; Trp: 色氨酸; Arg: 精氨酸; Gly: 甘氨酸; TER: 终止密码子。RSCU: 相对同义密码子使用度。

2.2 中性绘图分析

图3显示,普通核桃叶绿体基因组密码子的 GC_{12} 、 GC_3 的取值范围分别为0.333~0.543、0.210~0.365,深纹核桃叶绿体基因组密码子的 GC_{12} 、 GC_3 的取值范围分别为0.306~0.543、0.210~0.365,由此可见,2种核桃的叶绿体基因组密码子分布基本重合,说明普通核桃和深纹核桃中叶绿体基因组密码子的G+C含量接近。普通核桃 GC_{12} 与 GC_3 的回

归系数为0.379 7,决定系数为0.016 9;深纹核桃 GC_{12} 与 GC_3 的回归系数为0.343 7,决定系数为0.028 2。2种核桃回归系数均不接近1.000 0,说明G+C在第1位、第2位、第3位这3个位置的碱基组成差异大, GC_{12} 与 GC_3 相关性不显著。大部分基因位于对角线上方,综合相关性分析和中性绘图分析的结果,2种核桃叶绿体基因组密码子的使用偏好性主要受自然选择的影响。



GC_1 : 基因中所有密码子第1位碱基G+C的含量; GC_2 : 基因中所有密码子第2位碱基G+C的含量; GC_{12} : GC_1 和 GC_2 的平均值; GC_3 : 基因中所有密码子第3位碱基G+C的含量; R^2 : 决定系数。

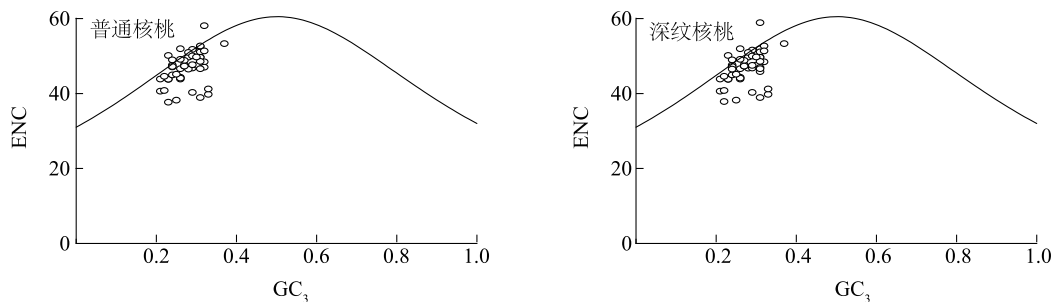
图3 中性绘图分析结果

Fig.3 Results of neutral plot analysis

2.3 ENC-plot 绘图分析

ENC-plot 分析结果(图4)表明,少数基因位于标准曲线上,该部分基因的ENC实际观察值与ENC期望值差异较小;大部分基因远离标准曲线,其ENC实际观察值与ENC期望值相差较大。进一步分析基因与标准曲线的距离,并对ENC比值进行计

算(表5),发现普通核桃和深纹核桃叶绿体密码子基因组在-0.049~0.049区间的基因频率为0.377,ENC实际观察值与ENC期望值差距较大。上述结果表明普通核桃和深纹核桃叶绿体密码子使用偏好性主要受到自然选择的影响。



ENC: 有效密码子数; GC_3 : 基因中所有密码子第3位碱基G+C的含量。

图4 有效密码子数(ENC)-plot 分析结果

Fig.4 Analysis results of effective number of codons (ENC)-plot

2.4 PR2-plot 分析

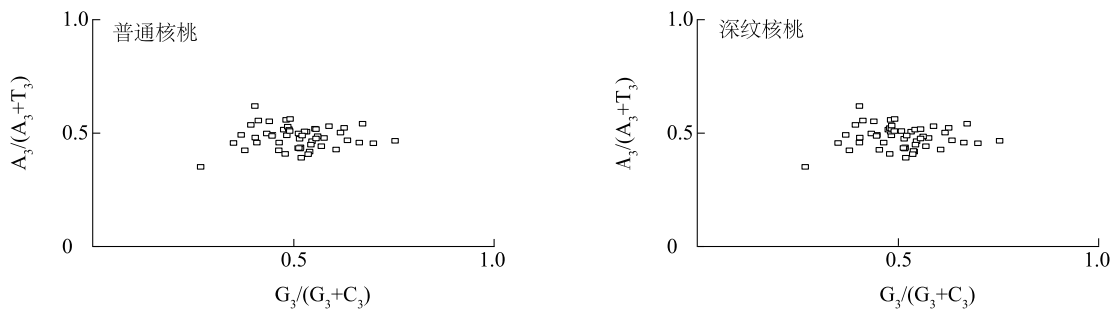
密码子偏好性仅受基因突变影响时,第3位碱基A/T、G/C具有相等的概率;受到自然选择影响时,A/T、G/C的概率不均等。对上述CDS进行碱基奇偶性分析,结果(图5)表明,2种核桃叶绿体基因在4个区

域的基因位点呈现不均匀分布,在 $A_3/(A_3+T_3) < 0.5$ 和 $G_3/(G_3+C_3) > 0.5$ 区域分布得最多,表示密码子在第3位碱基的选择上具有偏好性,使用频率表现为 $T > A$ 、 $G > C$ 。上述结果说明,普通核桃和深纹核桃叶绿体基因组密码子偏好性受到自然选择的影响。

表 5 有效密码子数 (ENC) 比值分布

Table 5 Distribution of effective number of codons (ENC) ratio

普通核桃				深纹核桃			
组段	组中值	频数	频率	组段	组中值	频数	频率
-0.150~-0.050	-0.1	2	0.038	-0.150~-0.050	-0.1	2	0.038
-0.049~0.049	0	20	0.377	-0.049~0.049	0	20	0.377
0.050~0.150	0.1	25	0.472	0.050~0.150	0.1	25	0.472
0.151~0.249	0.2	4	0.075	0.151~0.249	0.2	4	0.075
0.250~0.350	0.3	2	0.038	0.250~0.350	0.3	2	0.038
合计		53	1.000	合计		53	1.000



A₃: 基因中所有密码子第 3 位碱基 A 的含量; T₃: 基因中所有密码子第 3 位碱基 T 的含量; G₃: 基因中所有密码子第 3 位碱基 G 的含量; C₃: 基因中所有密码子第 3 位碱基 C 的含量。

图 5 PR2-plot 分析结果

Fig.5 Results of PR2-plot analysis

2.5 最优密码子的确定

对 5 个高表达基因 *petD*、*rps18*、*rps14*、*cemA*、*petB* 以及 5 个低表达基因 *atpF*、*ndhK*、*rpl2*、*ycf2*、*ycf3* 中的同义密码子进行计算,结果(表 6)表明,普通核桃中有 24 个密码子(其中 11 个以 U 结尾,6 个以 A 结尾,3 个以 G 结尾,4 个以 C 结尾)的 ΔRSCU (高表达基因的 RSCU-低表达基因的 RSCU) ≥ 0.08 ,深纹核桃中有 26 个密码子(其中 11 个以 U 结尾,8 个以 A 结尾,3 个以 G 结尾,4 个以 C 结尾) $\Delta\text{RSCU} \geq 0.08$ 。结合表 4,普通核桃和深纹核桃中 RSCU >

1.00 的密码子均为 32 个。同时满足 RSCU > 1.00 和 $\Delta\text{RSCU} \geq 0.08$ 这 2 个条件的密码子为最优密码子,从普通核桃中筛选出 UUU、AUU、UUG、GUU、CUU、UCC、UCG、CCG、GCA、UAU、CAA、AAA、GAA、UGU 共 14 个最优密码子;从深纹核桃中筛选出 UCC、UCA、UCG、CCG、UUU、GCA、UUA、UAU、UUG、CAA、CUU、AAA、AUU、GAA、AUA、UGU、GUU 共 17 个最优密码子,上述共有 23 个最优密码子在第 3 个位置偏好以碱基 A/U 结尾。

表 6 普通核桃和深纹核桃相对同义密码子使用度 (RSCU) 及最优密码子

Table 6 Relative synonymous codon usage (RSCU) of genes and the optimal codons in *Juglans regia* and *Juglans sigillata*

氨基酸	密码子	普通核桃					氨基酸	密码子	深纹核桃				
		高表达基因		低表达基因		Δ RSCU			高表达基因		低表达基因		Δ RSCU
		数量	RSCU _h	数量	RSCU _l				数量	RSCU _h	数量	RSCU _l	
Phe	UUU	33	8.11	97	6.28	1.83	Phe	UUU	33	8.11	97	6.28	1.83
	UUC	13	1.89	90	3.72	-1.83		UUC	13	1.89	90	3.72	-1.83
Leu	UUA	39	11.55	68	10.05	1.50	Leu	UUA	39	11.55	68	10.14	1.41
	UUG	25	8.10	78	6.62	1.48		UUG	25	8.10	77	6.41	1.69

续表6 Continued6

普通核桃							深纹核桃						
氨基酸	密码子	高表达基因		低表达基因		Δ RSCU	氨基酸	密码子	高表达基因		低表达基因		Δ RSCU
		数量	RSCU _h	数量	RSCU _l				数量	RSCU _h	数量	RSCU _l	
Ile	<u>CUU</u>	20	5.52	73	5.08	0.44	Ile	<u>CUU</u>	20	5.52	73	5.12	0.40
	CUC	5	1.30	29	1.98	-0.68		CUC	5	1.30	29	1.99	-0.69
	CUA	11	2.54	47	4.42	-1.88		CUA	11	2.54	47	4.48	-1.94
	CUG	4	1.00	25	1.85	-0.85		CUG	4	1.00	25	1.86	-0.86
	<u>AUU</u>	36	8.89	110	7.47	1.42		<u>AUU</u>	36	8.47	109	7.38	1.09
	AUC	13	1.73	54	3.16	-1.43		AUC	14	2.00	54	3.19	-1.19
	AUA	19	4.38	82	4.37	0.01		<u>AUA</u>	20	4.53	82	4.42	0.11
Met	AUG	23	5.00	59	5.00	0	Met	AUG	23	5.00	59	5.00	0
Val	<u>GUU</u>	17	8.00	39	4.67	3.33	Val	<u>GUU</u>	17	8.00	38	4.49	3.51
	GUC	5	1.15	20	2.60	-1.45		GUC	5	1.15	20	2.63	-1.48
	GUA	24	5.45	37	7.12	-1.67		GUA	25	5.70	37	7.22	-1.52
Ser	GUG	7	5.40	25	5.61	-0.21	Ser	GUG	6	5.15	25	5.65	-0.50
	UCU	13	5.19	76	7.85	-2.66		UCU	13	5.19	75	7.69	-2.50
	<u>UCC</u>	10	6.36	55	4.57	1.79		<u>UCC</u>	10	6.36	55	4.71	1.65
	UCA	9	6.96	65	6.99	-0.03		<u>UCA</u>	9	6.96	64	6.68	0.28
Pro	<u>UCG</u>	6	2.66	35	2.45	0.21	Pro	<u>UCG</u>	6	2.66	35	2.50	0.16
	CCU	15	5.53	42	5.61	-0.08		CCU	16	5.59	42	5.61	-0.02
	CCC	15	7.46	26	3.52	3.94		CCC	15	7.28	26	3.52	3.76
	CCA	7	2.44	38	8.67	-6.23		CCA	8	2.55	38	8.67	-6.12
	<u>CCG</u>	8	4.56	20	2.19	2.37		<u>CCG</u>	8	4.56	20	2.19	2.37
Thr	ACU	24	9.26	46	6.20	3.06	Thr	ACU	24	9.26	45	5.94	3.32
	ACC	6	3.40	30	5.33	-1.93		ACC	6	3.40	30	5.47	-2.07
	ACA	12	6.16	48	6.60	-0.44		ACA	12	6.16	48	6.71	-0.55
	ACG	4	1.17	20	1.87	-0.70		ACG	4	1.17	20	1.89	-0.72
Ala	GCU	16	9.36	44	7.61	1.75	Ala	GCU	16	9.36	44	7.61	1.75
	GCC	2	0.62	20	4.23	-3.61		GCC	2	0.62	20	4.23	-3.61
	<u>GCA</u>	17	8.74	32	6.41	2.33		<u>GCA</u>	17	8.74	32	6.41	2.33
	GCG	4	1.29	11	1.74	-0.45		GCG	4	1.29	11	1.74	-0.45
Tyr	<u>UAU</u>	22	9.10	96	7.15	1.95	Tyr	<u>UAU</u>	22	9.10	96	7.15	1.95
	UAC	4	0.90	26	2.85	-1.95		UAC	4	0.90	26	2.85	-1.95
TER	UAA	2	6.00	2	6.00	0	TER	UAA	2	6.00	2	6.00	0
	UAG	2	6.00	3	9.00	-3.00		UAG	2	6.00	3	9.00	-3.00
His	CAU	11	5.73	61	7.69	-1.96	His	CAU	11	5.73	61	7.69	-1.96
	CAC	6	2.27	17	2.31	-0.04		CAC	6	2.27	17	2.31	-0.04
Gln	<u>CAA</u>	19	9.17	85	6.48	2.69	Gln	<u>CAA</u>	19	9.17	84	6.35	2.82
	CAG	2	0.83	41	3.52	-2.69		CAG	2	0.83	41	3.65	-2.82
Asn	AAU	22	7.73	153	8.25	-0.52	Asn	AAU	22	7.73	153	8.25	-0.52
	AAC	7	2.27	45	1.75	0.52		AAC	7	2.27	45	1.75	0.52
Lys	<u>AAA</u>	35	7.38	128	7.17	0.21	Lys	<u>AAA</u>	35	7.38	127	7.16	0.22
	AAG	14	2.62	66	2.83	-0.21		AAG	14	2.62	66	2.84	-0.22
Asp	GAU	18	7.80	135	8.47	-0.67	Asp	GAU	18	7.80	134	8.38	-0.58
	GAC	6	2.20	30	1.53	0.67		GAC	6	2.20	30	1.62	0.58

续表6 Continued6

普通核桃							深纹核桃								
氨基酸	密码子	高表达基因		低表达基因		Δ RSCU	氨基酸	密码子	高表达基因		低表达基因		Δ RSCU		
		数量	RSCU _h	数量	RSCU _l				数量	RSCU _h	数量	RSCU _l			
Glu	GAA	31	7.64	124	7.48	0.16	Glu	GAA	31	7.64	122	7.46	0.18		
	GAG	7	2.36	62	2.52	-0.16		Cys	GAG	7	2.36	61	2.54	-0.18	
Cys	UGU	9	8.00	24	6.29	1.71	Cys		UGU	9	8.00	24	6.29	1.71	
	UGC	0	0	13	1.71	-1.71		TER	UGC	0	0	13	1.71	-1.71	
TER	UGA	1	3.00	0	0	3.00	TER		UGA	1	3.00	0	0	3.00	
Trp	UGG	18	4.00	57	5.00	-1.00	Trp	UGG	18	4.00	57	5.00	-1.00		
Arg	CGU	20	17.77	31	4.51	13.26	Arg	CGU	20	17.77	31	4.55	13.22		
	CGC	1	0.67	13	1.30	-0.63		Ser	CGC	1	0.67	13	1.30	-0.63	
	CGA	10	4.44	51	8.56	-4.12			Arg	CGA	10	4.44	51	8.71	-4.27
	CGG	1	0.38	23	3.50	-3.12				Ser	CGG	1	0.38	22	3.08
Ser	AGU	17	8.54	50	7.22	1.32	Ser		AGU		17	8.54	50	7.46	1.08
	AGC	1	0.30	15	0.90	-0.60		Arg	AGC	1	0.30	15	0.95	-0.65	
Arg	AGA	10	5.09	65	7.92	-2.83	Arg		AGA	10	5.09	65	8.04	-2.95	
	AGG	4	1.67	34	4.19	-2.52		Gly	AGG	4	1.67	34	4.31	-2.64	
Gly	GGU	22	7.84	39	4.19	3.65	Gly		GGU	22	7.84	40	4.55	3.29	
	GGC	5	1.50	13	1.91	-0.41			GGA	GGC	5	1.50	13	1.87	-0.37
	GGA	17	7.81	69	8.86	-1.05				GGG	GGA	17	7.81	69	8.59
	GGG	4	2.84	38	5.03	-2.19		GGG	4		2.84	38	4.99	-2.15	

加下划线的密码子表示最优密码子。Phe: 苯丙氨酸; Leu: 亮氨酸; Ile: 异亮氨酸; Met: 甲硫氨酸; Val: 缬氨酸; Ser: 丝氨酸; Pro: 脯氨酸; Thr: 苏氨酸; Ala: 丙氨酸; Tyr: 酪氨酸; His: 组氨酸; Gln: 谷氨酰胺; Asn: 天冬酰胺; Lys: 赖氨酸; Asp: 天冬氨酸; Glu: 谷氨酸; Cys: 半胱氨酸; Trp: 色氨酸; Arg: 精氨酸; Gly: 甘氨酸; TER: 终止密码子。RSCU: 相对同义密码子使用度; RSCU_h: 高表达基因的 RSCU; RSCU_l: 低表达基因的 RSCU; Δ RSCU: 高表达基因的 RSCU-低表达基因的 RSCU。

3 讨论

密码子作为生物体内重要的遗传信息,在核糖核酸翻译为蛋白质的过程中发挥着重要作用。前人研究发现,在自然环境中,生物体中同义密码子的使用频率逐渐适应宿主的生长环境,并出现不能均衡使用的情况^[32-33]。密码子使用偏好性受不同因素(基因序列碱基组成、自然选择等)影响^[34],但一般认为自然选择是影响密码子使用偏好性的主要因素^[35]。前人研究多将第3位碱基的G+C含量作为评价密码子偏好性的关键参数^[36]。ENC能反映同义密码子非均衡使用的偏好情况,是最具参考价值的参数,ENC理论取值范围为20~61^[37]。

普通核桃和深纹核桃的叶绿体基因组密码子中第1位、第2位和第3位碱基的组成都影响着叶绿体CDS的使用偏好。第3位碱基G+C含量总体小于第1位、第2位碱基,第3位碱基A、T含量均高于G、C含量,表明第3个位置上的碱基偏好使用A/T,这与南欧大戟(*Euphorbia peplus* L.)^[33]、蒲桃属

(*Syzygium* Gaertn.)^[38]、杜鹃(*Rhododendron simsii* Planch.)^[39]的密码子使用偏好性相同。碱基奇偶性分析结果表明,在第3个位置上碱基使用频率表现为T>A、G>C,说明2种核桃叶绿体基因组密码子使用偏好性受自然选择的影响,这与对木兰^[40]、柿树^[41]、蔷薇科果树^[42]、紫花苜蓿^[43]等植物的分析结果相同;赵森等^[44]分析秋茄,得出C大于G、T大于A的结果,说明秋茄叶绿体基因密码子使用模式受到自然选择的影响;竹节参叶绿体基因组密码子的偏好分析结果表明,基因突变是影响其密码子偏好形成的主要因素^[45],这与本研究结果不一致,说明不同植物密码子使用偏好性受到的主要影响因素不尽相同。

4 结论

基于获得的普通核桃、深纹核桃叶绿体基因组序列,进行碱基组成、中性绘图分析、PR2-plot、RSCU值及ENC-plot分析比较,明确2种核桃拥有UUU、UUG、CUU、AUU、GUU、UCC、UCG、CCG、GCA、UAU、

CAA、AAA、GAA、UGU 共 14 个共同的最优密码子,多数以 A 或者 U 结尾,反映了相同分类单位、亲缘关系近的生物密码子偏好使用类似的碱基结尾^[46-47]。基因突变和自然选择共同影响普通核桃、深纹核桃叶绿体基因组密码子的使用偏好性,但自然选择为主要影响因素。普通核桃中有 14 个最优密码子,深纹核桃中有 17 个最优密码子,且最优密码子更倾向于以 A/U 结尾,本研究得到 2 种核桃叶绿体基因组表达的重要信息,对之后进行核桃物种进化及种质资源的研究有着重要作用,同时也为预测核桃未知蛋白质的基因、基因工程表达载体的设计等提供了一定的数据支持。

参考文献:

- [1] NOVOA E M, PAVON-ETERNOD M, PAN T, et al. A role for tRNA modifications in genome structure and codon usage [J]. Cell, 2012, 149(1): 202-213.
- [2] 赵春丽,彭丽云,王 晓,等. 苋菜 *AtGAI* 基因密码子偏好性与进化分析[J]. 中国农业大学学报, 2019, 24(12): 10-22.
- [3] 辛雅萱,黎若竹,李 鑫,等. 杜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.
- [4] 胡桂兵,张上隆,徐昌杰,等. 不同种类柑橘的密码子用法分析[J]. 华南农业大学学报, 2006, 27(1): 13-16.
- [5] HERSHBERG R, PETROV D A. Selection on codon bias[J]. Annual Review of Genetics, 2008, 42: 287-299.
- [6] 刘慧敏,乌云塔娜,杜红岩. 杜仲转录组基因密码子使用偏好性分析[J]. 北方园艺, 2016(13): 85-89.
- [7] XU C, CAI X N, CHEN Q Z, et al. Factors affecting synonymous codon usage bias in chloroplast genome of *Oncidium Gower Ramsey* [J]. Evolutionary Bioinformatics, 2011, 7: 271-278.
- [8] 王 茹,罗永坚,刘相泉,等. 水杉叶绿体基因组同义密码子使用偏好性[J]. 湖北民族大学学报(自然科学版), 2022, 40(2): 121-128.
- [9] SINGH N D, DAVIS J C, PETROV D A. X-linked genes evolve higher codon bias in *Drosophila* and *Caenorhabditis* [J]. Genetics, 2005, 171(1): 145-155.
- [10] PARVATHY S T, UDAYASURIYAN V, BHADANA V. Codon usage bias[J]. Molecular Biology Reports, 2022, 49(1): 539-565.
- [11] SHAW J, LICKY E B, SCHILLING E E, et al. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III [J]. American Journal of Botany, 2007, 94(3): 275-288.
- [12] 李西文,胡志刚,林小涵,等. 基于 454FLX 高通量技术的厚朴叶绿体全基因组测序及应用研究[J]. 药学学报, 2012, 47(1): 124-130.
- [13] 冯 展,江 媛,郑 燕,等. 肉苁蓉属植物叶绿体基因组密码子偏好性分析[J]. 中草药, 2023, 54(5): 1540-1550.
- [14] 谭淳月,刘 勇,赖章凤,等. 茶树氧甲基转移酶基因密码子偏好性分析[J]. 江西农业大学学报, 2023, 45(3): 652-662.
- [15] 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.
- [16] 原晓龙,刘 音,康洪梅,等. 蒜头果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), 2021, 41(3): 15-22.
- [17] 夏 晞,彭劲谕,王大玮,等. 3 种榕属叶绿体基因组密码子偏好性分析[J]. 西北林学院学报, 2022, 37(5): 88-94.
- [18] 舒军霞,杨 林,周 涛,等. 刺楸属 4 种药用植物叶绿体基因组密码子偏好性分析[J]. 中草药, 2022, 53(23): 7507-7515.
- [19] 孟 佳,方晓璞,史宣明,等. 我国核桃产业发展现状、问题与建议[J]. 中国油脂, 2023, 48(1): 84-86, 103.
- [20] 郭从善. 核桃及其加工与利用[J]. 粮油食品科技, 1999(5): 23-25.
- [21] 祝友朋,韩长志,熊 智. 核桃细菌性黑斑病菌分泌蛋白质的理化性质及特征分析[J]. 江苏农业学报, 2019, 35(2): 295-301.
- [22] 谢 蕾,李贤忠,宁德鲁,等. 深纹核桃种质资源评价与挖掘研究进展[J]. 林业科技通讯, 2021(11): 26-31.
- [23] 覃 阳,全绍文,周 丽,等. 基于转录组测序的核桃 (*Juglans regia* L.) SSR 标记开发[J]. 分子植物育种, 2019, 17(20): 6736-6742.
- [24] 余务刚,肖良俊,吴 涛,等. 不同类型深纹核桃的表型多样性研究[J]. 西部林业科学, 2022, 51(6): 31-38, 78.
- [25] NOVEMBRE J A. Accounting for background nucleotide composition when measuring codon usage bias[J]. Molecular Biology and Evolution, 2002, 19(8): 1390-1394.
- [26] 陆奇丰,骆文华,黄至欢. 两种梧桐叶绿体基因组密码子使用偏性分析[J]. 广西植物, 2020, 40(2): 173-183.
- [27] 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.
- [28] YANG X, LUO X N, CAI X P. Analysis of codon usage pattern in *Taenia saginata* based on a transcriptome dataset[J]. Parasit Vectors, 2014, 7: 527.
- [29] SUEOKA N. Near homogeneity of PR2-bias fingerprints in the human genome and their implications in phylogenetic analyses[J]. Journal of Molecular Evolution, 2001, 53(4/5): 469-476.
- [30] 唐晓芬,陈 莉,马玉韬. 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.
- [31] SHARP P M, LI W H. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications[J]. Nucleic Acids Research, 1987, 15(3): 1281-1295.
- [32] 韩利红,代冬琴,赵明玉,等. 双孢蘑菇基因组密码子的偏好性分析[J]. 菌物学报, 2021, 40(3): 603-615.
- [33] 余 涛,蒲 芬,管 芹,等. 南欧大戟叶绿体基因组密码子偏好性分析[J]. 江苏农业科学, 2023, 51(15): 35-41.

- [34] ROMERO H, ZAVALA A, MUSTO H. Codon usage in *Chlamydia trachomatis* is the result of strand-specific mutational biases and a complex pattern of selective forces[J]. Nucleic Acids Research, 2000, 28(10):2084-2090.
- [35] 吴学俊,梁湘兰,易子群,等. 湖北山楂叶绿体基因组密码子偏好性分析[J]. 分子植物育种, 2023, 21(8):2508-2516.
- [36] 冯德江,蔡华雅,刘翔,等. 马铃薯 X 病毒外壳蛋白的表达水平与变偶密码子使用频率的相关性[J]. 微生物学报, 2003, 43(5):569-576.
- [37] 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):622-630.
- [38] 杨秀瑶,张梦洁,尹拓,等. 蒲桃属叶绿体基因组特征及密码子偏好性分析[J]. 西南农业学报, 2023, 36(9):1869-1880.
- [39] 万露露,范敦锦,王中焯,等. 杜鹃叶绿体基因组特征及密码子偏好性分析[J]. 江苏农业科学, 2023, 51(13):54-59.
- [40] 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.
- [41] 傅建敏,索玉静,刘慧敏,等. 柿属植物叶绿体蛋白质编码基因密码子用法[J]. 经济林研究, 2017, 35(2):38-44.
- [42] 刘兴跃,何仲坚,邱毅敏. 四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.
- [43] 喻凤,韩明. 紫花苜蓿叶绿体基因组密码子偏好性分析[J]. 广西植物, 2021, 41(12):2069-2076.
- [44] 赵森,邓力华,陈芬. 秋茄叶绿体基因组密码子使用偏好性分析[J]. 森林与环境学报, 2020, 40(5):534-541.
- [45] 梁娥,齐敏杰,丁延庆,等. 竹节参转录组使用密码子偏好性分析[J]. 江苏农业科学, 2019, 47(2):59-63.
- [46] 李显煌,杨生超,辛雅萱,等. 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.
- [47] 王婧,王天翼,王罗云,等. 沙枣叶绿体全基因组序列及其使用密码子偏性分析[J]. 西北植物学报, 2019, 39(9):1559-1572.

(责任编辑:王妮)