

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

钟才荣¹, 何斯敏², 方赞山¹, 张颖^{1,2,①}

[1. 海南省林业科学研究院(海南省红树林研究院), 海南 海口 571129; 2. 岭南师范学院 湛江市红树林生态系统保护与修复重点实验室
红树林珍稀濒危物种保护与利用工程技术研究中心, 广东 湛江 524048]

摘要: 为探究使君子科(Combretaceae)红树植物叶绿体基因组密码子使用偏好性及其影响因子,对拉关木(*Laguncularia racemosa* C. F. Gaertn.)、红榄李[*Lumnitzera littorea* (Jack) Voigt]、榄李(*Lumnitzera racemosa* Willd.)叶绿体基因组的 52 个共有蛋白质编码序列(CDS)进行密码子第 1、第 2、第 3 位碱基的 GC 含量(分别为 GC₁、GC₂、GC₃)、密码子总 GC 含量(GC_{all})、有效密码子数(ENC)和相对同义密码子使用度(RSCU)分析;在此基础上,对每种植物的 GC₁、GC₂、GC₃、GC_{all}和 ENC 值进行相关性分析,并对每种植物叶绿体基因组进行中性绘图、ENC-plot、PR2-plot 和最优密码子分析;此外,还基于 *rbcL* 基因序列和叶绿体基因组共有 CDS 的 RSCU 值对使君子科 11 种植物进行聚类分析。结果表明:3 种植物叶绿体基因组 GC_{all} 值的均值为 38.0%~38.2%,GC₁、GC₂ 和 GC₃ 值的均值均低于 50%,均表现为 GC₁ 值的均值最大、GC₃ 值的均值最小,且 GC₃ 值的均值明显低于 GC₁ 和 GC₂ 值的均值,并与二者呈不显著正相关;ENC 值的均值为 47.2~47.3,且 3 种植物叶绿体基因组中 ENC 值大于 45 的基因占比均高于 70%。3 种植物叶绿体基因组中的高频密码子均有 29 个,以 A 或 U 结尾的高频密码子占比达 96.6%。中性绘图、ENC-plot 和 PR2-plot 分析结果表明 3 种植物叶绿体基因组密码子使用偏好性主要受自然选择的影响。拉关木、红榄李和榄李叶绿体基因组的最优密码子分别有 8、6 和 10 个,且多以 A 或 U 结尾;共有最优密码子有 4 个,分别为 CGA、AGA、GGA、ACA。聚类分析结果表明红榄李和榄李均首先聚在一起,说明二者的亲缘关系较近。综上所述,3 种红树植物叶绿体基因组密码子偏好以 A 或 U 结尾,且密码子使用偏好性主要受自然选择的影响。

关键词: 红树植物; 叶绿体基因组; 密码子使用偏好性; 最优密码子; 自然选择

中图分类号: Q943.2; Q949.762.1 文献标志码: A 文章编号: 1674-7895(2024)02-0001-12

DOI: 10.3969/j.issn.1674-7895.2024.02.01

Analysis on codon usage bias of chloroplast genomes of three species of mangroves in Combretaceae ZHONG Cairong¹, HE Simin², FANG Zanshan¹, ZHANG Ying^{1,2,①} [1. Hainan Academy of Forestry (Hainan Academy of Mangrove), Haikou 571129, China; 2. Zhanjiang Key Laboratory of Conservation and Restoration of Mangrove Ecosystem, Engineering and Technological Research Center in Protection and Utilization of Mangrove Rare and Endangered Species, Lingnan Normal University, Zhanjiang 524048, China], *J. Plant Resour. & Environ.*, 2024, 33(2): 1-12

Abstract: To investigate the codon usage bias of chloroplast genomes of mangroves in Combretaceae and its affecting factors, GC contents of the first, second, and third bases of codon (GC₁, GC₂, and GC₃, respectively), total GC content of codon (GC_{all}), effective number of codon (ENC), and relative synonymous codon usage (RSCU) of 52 common protein coding sequences (CDS) of chloroplast genomes of *Laguncularia racemosa* C. F. Gaertn., *Lumnitzera littorea* (Jack) Voigt, and *Lumnitzera racemosa*

收稿日期: 2023-07-25

基金项目: 海南省省属科研院所技术创新专项(KYYS-2021-13); 国家自然科学基金项目(41776148)

作者简介: 钟才荣(1976—),男,海南海口人,本科,高级工程师,主要从事红树植物保育与生态修复工作。

①通信作者 E-mail: Zhangyingred@lingnan.edu.cn

引用格式: 钟才荣, 何斯敏, 方赞山, 等. 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.

Willd. were analyzed; on the basis, the correlations between GC_1 , GC_2 , GC_3 , GC_{all} , and ENC values of each species were analyzed, and neutral mapping, ENC-plot, PR2-plot, and optimal codon analyses were performed for chloroplast genomes of each species; in addition, cluster analysis was conducted for 11 species in Combretaceae based on *rbcL* gene sequences and RSCU values of common CDS of chloroplast genomes. The results show that the averages of GC_{all} values of chloroplast genomes of the three species are 38.0%–38.2%, the averages of GC_1 , GC_2 , and GC_3 values are all lower than 50%, it is appeared in all of the three species that the average of GC_1 value is the largest and that of GC_3 value is the smallest, and the average of GC_3 value is obviously lower than those of GC_1 and GC_2 values and has non-significant positive correlations with them; the averages of ENC values are 47.2–47.3, and the percentages of genes with ENC value greater than 45 in chloroplast genomes of the three species are all greater than 70%. There are 29 high frequency codons in chloroplast genomes of the three species, and the high frequency codons ended with A or U account for 96.6%. The results of neutral mapping, ENC-plot, and PR2-plot analyses show that the codon usage bias of chloroplast genomes of the three species is mainly affected by natural selection. There are 8, 6, and 10 optimal codons in chloroplast genomes of *Laguncularia racemosa*, *Lumnitzera littorea*, and *Lumnitzera racemosa*, respectively, and most of them are ended with A or U; there are 4 common optimal codons, namely CGA, AGA, GGA, and ACA, respectively. The cluster analysis results show that *Lumnitzera littorea* and *Lumnitzera racemosa* are clustered together first, indicating that their relationship is close. In conclusion, the codons of chloroplast genomes of the three species of mangroves prefer to end with A or U, and the codon usage bias is mainly affected by natural selection.

Key words: mangrove; chloroplast genome; codon usage bias; optimal codon; natural selection

红树林生态系统是一种陆地向海洋过渡的独特森林生态系统,在促淤护岸、防风消浪、净化生态环境、维护生物多样性等方面具有重要价值^[1]。受气候变化和人类活动的影响,红树林生态系统遭到不同程度的破坏,其脆弱性和不稳定性非常突出^[1-2]。保护濒危红树植物对于红树林生态系统的稳定和可持续发展具有重要意义。中国分布的使君子科(Combretaceae)红树植物只有3种,分别为对叶榄李属(*Laguncularia* C. F. Gaertn.)的拉关木(*L. racemosa* C. F. Gaertn.)以及榄李属(*Lumnitzera* Willd.)的红榄李〔*L. littorea* (Jack) Voigt〕和榄李(*L. racemosa* Willd.)^[3]。拉关木为1999年从墨西哥拉巴斯市引入中国的外来红树植物,被广泛用于红树林分布区的修复,现已广泛分布于海南岛及华南沿海许多不同生境中,具有生长快速、繁殖能力强的特点^[4]。红榄李的自然分布地非常狭窄,仅限于海南省的铁炉港,为国家I级重点保护野生植物,在国际自然保护联盟(International Union for Conservation of Nature, IUCN)的濒危等级标准中属于极危(critically endangered, CR)^[1],并且该种对温度极为敏感,是中国红树植物分布最南界的物种^[5]。榄李常与红榄李相伴而生,在海南、广东、广西、香港和台湾均有较大面积的自然分布区,福建有引种^[6]。研究发现,这3种红树植物对潮间带生境的适应性明显不同,榄李在潮间带的分

布最广泛,拉关木在适生区的速生能力极强,而红榄李的适应能力在3种红树植物中最弱^[4-5,7-12]。对这3种红树植物叶绿体基因组信息进行解析,有助于探究叶绿体基因组对潮间带生境中红树植物不同适应能力的作用,并可为红树林生态恢复提供理论支撑^[13]。

三联体密码子简称密码子(codon),是mRNA与蛋白质之间传递遗传信息的重要纽带,对生物体遗传信息的传递具有重要意义。受自然选择和碱基突变等因子的影响,不同基因、不同物种的密码子使用频率往往存在明显差异^[14]。同一物种或同一基因频繁倾向于使用1个或1个以上的特定同义密码子,这种现象被称为密码子使用偏好性(codon usage bias)。密码子使用偏好性普遍存在于植物的基因组及线粒体和叶绿体基因组中,能够反映密码子在基因编码中的不均匀使用,在基因调控和分子进化上发挥重要作用^[15]。密码子使用偏好性是物种对生存环境和自然进化适应性选择的结果,一般来说,同一物种或亲缘关系较近的物种的密码子使用偏好性相似^[16]。目前对叶绿体基因组密码子使用偏好性的研究主要集中在分子进化和翻译调控等方面^[16]。叶绿体基因组的密码子使用偏好性和核苷酸组成可为基于叶绿体基因组的基因工程研究提供理论基础^[17],研究者还可参照叶绿体基因组中已知的偏好密码子对异源表达

的基因进行修饰, 从而提高基因的转化效率和表达效率^[18]。

叶绿体是绿色植物完成光合作用的重要场所。植物叶绿体具有独特的基因组, 能够进行半自主复制和基因表达等生命过程^[19]。在中国分布的 29 种红树植物中, 虽然已有 20 多个种类完成了叶绿体基因组测序工作^[20-25], 但明确叶绿体基因组密码子使用偏好性的只有秋茄 (*Kandelia obovata* Sheue, H. Y. Liu et J. Yong)^[19]。虽然通过叶绿体基因工程技术可有效提高植物的抗逆性, 但直接转化叶绿体基因可能会出现转化效率低和基因沉默等问题, 因此, 有必要对不同植物的叶绿体基因组密码子使用偏好性进行分析, 进而提高外源基因的转化效率和表达效率^[19]。

分子生物学证据表明: 使君子科的榄李属和对叶榄李属在进化上来自同一祖先^[26]。笔者基于项目组前期研究获得的拉关木、红榄李和榄李叶绿体基因组信息^[20], 对 3 种红树植物叶绿体基因组密码子使用偏好性进行分析, 明确 3 种植物叶绿体基因组的最优密码子, 并探究影响 3 种植物叶绿体基因组密码子使用偏好性的因子, 从而为使君子科红树植物分子植物育种研究提供理论依据。

1 材料和方法

1.1 材料

在 GenBank 数据库中下载拉关木、红榄李、榄李的叶绿体基因组序列 (登录号分别为 MH551145、MG182696、MH551146)。为了降低误差, 对 3 种植物叶绿体基因组的蛋白质编码序列 (CDS) 进行筛选, 剔除长度小于 300 bp 的序列, 且仅保留 1 个重复序列^[19], 筛选出 52 个共有 CDS, 分别为 *accD*、*atpA*、*atpB*、*atpE*、*atpF*、*atpI*、*ccsA*、*cemA*、*clpP*、*matK*、*ndhA*、*ndhB*、*ndhC*、*ndhD*、*ndhE*、*ndhF*、*ndhG*、*ndhH*、*ndhI*、*ndhJ*、*ndhK*、*petA*、*petB*、*petD*、*psaA*、*psaB*、*psbA*、*psbB*、*psbC*、*psbD*、*rbcL*、*rpl2*、*rpl14*、*rpl20*、*rpl22*、*rpoA*、*rpoB*、*rpoC1*、*rpoC2*、*rps2*、*rps3*、*rps4*、*rps7*、*rps8*、*rps11*、*rps12*、*rps14*、*rps18*、*ycf1*、*ycf2*、*ycf3*、*ycf4*。

1.2 方法

1.2.1 密码子相关参数计算 使用 EMBOSS 软件包中的 CUSP 程序 (<http://www.bioinformatics.nl/emboss-explorer/>) 计算每种植物叶绿体基因组的密码子总 GC 含量 (GC_{all}); 统计每种植物叶绿体基因组密

码子第 1、第 2、第 3 位碱基的 GC 含量 (分别为 GC_1 、 GC_2 、 GC_3); 利用 CodonW 1.4.2 软件计算每种植物叶绿体基因组的有效密码子数 (effective number of codon, ENC)、相对同义密码子使用度 (relative synonymous codon usage, RSCU)。ENC 值可用于判断密码子使用偏好性强弱, 变化范围为 20~61, ENC 值大于 45, 表示密码子使用偏好性较弱^[15]。RSCU 值为 1, 表示密码子无使用偏好性; RSCU 值大于 1, 表示密码子使用偏好性强; RSCU 值小于 1, 表示密码子使用偏好性弱^[18]。使用 Rstudio (R core team, 2022) 软件对每种植物叶绿体基因组的 GC_{all} 、 GC_1 、 GC_2 、 GC_3 、ENC 值进行相关性分析。

1.2.2 中性绘图分析 利用 EXCEL 2021 软件获得 3 种植物叶绿体基因组密码子第 1 和第 2 位碱基 GC 含量的平均值 (GC_{12}); 以每个基因的 GC_3 值为横坐标、 GC_{12} 值为纵坐标, 使用 Rstudio (R core team, 2022) 软件绘制二维散点图 (每个圆点代表 1 个基因), 分析 GC_3 与 GC_{12} 值间的相关性^[9]。

1.2.3 ENC-plot 分析 以每个基因的 GC_3 值为横坐标、ENC 值为纵坐标, 使用 Rstudio (R core team, 2022) 软件绘制二维散点图, 根据孙月琪等^[9]的方法计算标准曲线, 根据基因的分布情况分析造成密码子使用偏好性的原因。在没有自然选择作用的情况下, 基因会分布在标准曲线上或附近。基因越靠近标准曲线, 表示其密码子使用偏好性受碱基突变的影响越大; 基因越远离标准曲线, 表示其密码子使用偏好性受自然选择的影响越大。

1.2.4 PR2-plot 分析 计算基因密码子第 3 位碱基的 A、T (U, 下同)、C、G 含量 (分别为 A_3 、 T_3 、 C_3 、 G_3); 以基因的 $G_3/(G_3+C_3)$ 为横坐标、 $A_3/(A_3+T_3)$ 为纵坐标, 使用 Rstudio (R core team, 2022) 软件绘制二维散点图。基因离中心点越近, 表示 4 种碱基的使用频率相等或接近, 密码子使用偏好性仅受碱基突变的影响; 反之, 则表示密码子使用偏好性受自然选择和碱基突变的共同影响^[9]。

1.2.5 最优密码子分析 RSCU 值大于 1 的密码子为高频密码子。根据 ENC 值对基因进行升序排序, 利用排列在前 10% 和后 10% 的基因分别建立高表达组和低表达组。将高表达组和低表达组的 RSCU 差值 ($\Delta RSCU$) 大于或等于 0.08 的密码子定义为高表达密码子, 将既是高频密码子又是高表达密码子的密码子定义为最优密码子^[9]。

1.2.6 系统发育分析

1.2.6.1 基于 *rbcL* 基因序列的系统发育分析 从 GenBank 数据库中下载使君子科植物 *Quisqualis indicum* Linn.、长柱使君子 [*Q. littorea* (Engl.) Exell]、*Combretum kraussii* Hochst.、拉关木、红榄李、榄李、*Terminalia phillyreifolia* (Van Heurck et Müll. Arg.) Gere et Boatwr、榄仁 (*T. catappa* Linn.)、诃子 (*T. chebula* Retz.)、千果榄仁 (*T. myriocarpa* Van Heurck et Müll. Arg.)、小叶榄仁 (*T. neotaliala* Capuron H. Perrier) 叶绿体基因组的 *rbcL* 基因序列, 登录号分别为 MT700490、MT700495、MH551145、MG182696、MH551146、MT700496、MT700489、MN736954、OM202511、MN251603、MT242598。基于上述基因序列, 使用 Mega X 软件中的最大似然法 (maximum likelihood) 构建系统发育树, 采用 1 000 次 bootstrap 检验各分支的置信度。

1.2.6.2 基于叶绿体基因组共有 CDS 的 RSCU 值的系统发育分析 从 GenBank 数据库中下载上述使君子科 11 种植物叶绿体基因组的 CDS, 剔除长度小于 300 bp 的序列, 且仅保留 1 个重复序列, 筛选出 44 个共有 CDS; 利用 CodonW 1.4.2 软件计算每种植物共

有 CDS 的 RSCU 值; 根据计算结果, 使用 Rstudio (R core team, 2022) 软件绘制系统发育树。

2 结果和分析

2.1 叶绿体基因组密码子组成分析

2.1.1 密码子 GC 含量和有效密码子数 (ENC) 分析 对 3 种红树植物叶绿体基因组的 52 个共有 CDS 进行分析, 结果 (表 1) 显示: 拉关木、红榄李、榄李叶绿体基因组密码子总 GC 含量 (GC_{all}) 的均值分别为 38.0%、38.2%、38.1%。红榄李叶绿体基因组密码子第 1、第 2、第 3 位碱基的 GC 含量 (分别为 GC_1 、 GC_2 、 GC_3) 的均值分别为 47.0%、39.2% 和 28.5%, 拉关木 GC_1 、 GC_2 和 GC_3 值的均值分别为 46.8%、39.1%、28.2%, 榄李 GC_1 、 GC_2 和 GC_3 值的均值分别为 46.8%、39.2% 和 28.4%。值得注意的是, 3 种植物 GC_1 、 GC_2 和 GC_3 值的均值均低于 50%, 均表现为 GC_1 值的均值最大、 GC_3 值的均值最小, 且 GC_3 值的均值明显低于 GC_1 和 GC_2 值的均值, 说明 3 种植物叶绿体基因组密码子偏好使用 A 或 U 结尾, 且密码子的第 3 位碱基较第 1 和第 2 位碱基更偏好使用 A 或 U。

表 1 使君子科 3 种红树植物叶绿体基因组的密码子 GC 含量和有效密码子数分析¹⁾

Table 1 Analyses on GC content of codon and effective number of codon in chloroplast genomes of three species of mangroves in Combretaceae¹⁾

基因 Gene	拉关木 <i>Laguncularia racemosa</i>					红榄李 <i>Lumnitzera littorea</i>					榄李 <i>Lumnitzera racemosa</i>				
	$GC_1/\%$	$GC_2/\%$	$GC_3/\%$	$GC_{all}/\%$	ENC	$GC_1/\%$	$GC_2/\%$	$GC_3/\%$	$GC_{all}/\%$	ENC	$GC_1/\%$	$GC_2/\%$	$GC_3/\%$	$GC_{all}/\%$	ENC
<i>accD</i>	39.4	37.8	29.3	35.5	47.2	38.9	37.5	29.7	35.4	46.8	39.0	37.6	29.8	35.4	47.0
<i>atpA</i>	55.1	39.8	24.6	39.8	45.2	55.7	39.6	26.8	40.7	48.1	55.1	40.6	26.5	40.8	48.1
<i>atpB</i>	56.1	41.9	28.7	42.2	47.0	56.0	41.8	28.2	42.0	46.5	56.7	42.2	28.6	42.5	46.9
<i>atpE</i>	53.7	40.3	30.6	41.5	59.8	53.0	40.3	29.9	41.0	59.8	53.0	40.3	29.9	41.0	59.8
<i>atpF</i>	45.4	33.0	35.1	37.8	46.6	46.2	34.3	35.5	38.7	44.0	45.4	32.4	34.6	37.5	45.7
<i>atpI</i>	49.6	37.9	27.4	38.3	45.4	50.4	37.9	27.4	38.6	47.0	50.0	38.3	27.0	38.4	46.7
<i>ccsA</i>	33.4	36.9	22.5	30.9	39.0	34.4	36.9	24.1	31.8	40.2	34.4	36.9	23.4	31.6	39.9
<i>cemA</i>	39.6	27.8	27.0	31.5	44.1	39.6	28.3	29.1	32.3	47.6	39.6	28.3	29.6	32.5	47.7
<i>clpP</i>	57.1	36.2	29.1	40.8	49.5	57.7	36.2	29.1	41.0	50.7	57.7	36.2	29.1	41.0	50.7
<i>matK</i>	40.0	29.8	26.8	32.2	50.3	40.0	30.5	28.2	32.9	51.8	40.0	29.9	28.3	32.8	52.0
<i>ndhA</i>	41.8	37.6	22.3	33.9	40.5	41.8	37.6	23.9	34.4	41.6	41.5	37.6	24.2	34.4	41.5
<i>ndhB</i>	42.1	38.8	31.3	37.4	47.5	41.9	38.8	31.3	37.3	47.2	41.9	38.8	31.3	37.3	47.2
<i>ndhC</i>	47.1	33.1	27.3	35.8	47.4	48.8	33.1	28.1	36.6	47.7	48.8	33.1	28.1	36.6	47.7
<i>ndhD</i>	40.4	36.7	26.8	34.6	45.6	40.2	37.3	28.6	35.4	48.1	40.4	37.3	28.6	35.4	47.8
<i>ndhE</i>	38.2	33.3	26.5	32.7	43.9	38.2	33.3	27.5	33.0	45.0	39.2	33.3	27.5	33.3	46.0
<i>ndhF</i>	37.1	35.7	21.4	31.4	43.0	38.0	35.6	22.0	31.9	42.9	37.8	35.5	22.0	31.8	42.8
<i>ndhG</i>	44.1	33.9	21.5	33.2	43.8	44.1	34.5	21.5	33.3	45.2	44.1	34.5	21.5	33.3	45.0
<i>ndhH</i>	51.0	36.8	28.2	38.7	50.0	51.3	36.6	27.9	38.6	49.0	51.3	36.6	28.2	38.7	49.3
<i>ndhI</i>	43.6	37.2	23.8	34.9	46.5	43.4	38.0	22.3	34.5	44.8	43.4	38.0	22.3	34.5	45.4
<i>ndhJ</i>	50.9	38.4	28.3	39.2	48.1	52.8	38.4	28.9	40.0	50.5	52.8	38.4	28.9	40.0	50.5
<i>ndhK</i>	43.4	45.6	25.2	38.1	49.9	43.0	44.3	26.3	37.9	49.3	43.4	44.3	26.3	38.0	50.0

续表1 Table 1 (Continued)

基因 Gene	拉关木 <i>Laguncularia racemosa</i>					红榄李 <i>Lumnitzera littorea</i>					榄李 <i>Lumnitzera racemosa</i>				
	GC ₁ /%	GC ₂ /%	GC ₃ /%	GC _{all} /%	ENC	GC ₁ /%	GC ₂ /%	GC ₃ /%	GC _{all} /%	ENC	GC ₁ /%	GC ₂ /%	GC ₃ /%	GC _{all} /%	ENC
<i>petA</i>	52.0	36.8	31.8	40.2	50.9	52.0	35.8	32.4	40.1	53.7	51.4	35.8	35.8	39.6	51.7
<i>petB</i>	49.1	41.2	31.9	40.7	42.4	48.6	40.9	31.4	40.3	44.0	49.1	41.2	31.5	40.6	42.8
<i>petD</i>	51.6	38.5	24.2	38.1	40.4	50.9	38.4	23.9	37.7	41.4	51.6	38.5	24.2	38.1	41.0
<i>psaA</i>	52.3	43.4	32.1	42.6	49.3	52.5	43.4	32.8	42.9	50.0	52.3	43.4	32.5	42.7	49.6
<i>psaB</i>	48.3	43.0	31.7	41.0	46.9	48.3	43.0	31.0	40.8	46.8	48.2	43.0	30.9	40.7	46.6
<i>psbA</i>	50.0	43.5	31.4	41.6	48.0	50.0	43.5	31.9	41.8	39.7	50.0	43.5	31.4	41.6	39.6
<i>psbB</i>	54.2	45.8	26.7	42.2	45.5	54.2	45.8	27.3	42.4	45.8	54.2	45.8	27.5	42.5	45.9
<i>psbC</i>	54.0	46.2	32.7	44.3	48.7	54.4	46.2	33.3	44.7	49.5	54.4	46.2	33.3	44.7	49.5
<i>psbD</i>	52.3	43.2	31.4	42.3	43.8	52.3	43.2	31.1	42.2	43.0	52.3	43.2	30.5	42.0	42.7
<i>rbcL</i>	57.4	43.1	29.0	43.1	46.2	57.4	43.3	29.8	43.5	47.5	57.1	43.3	29.6	43.4	47.5
<i>rpl2</i>	50.6	48.4	33.5	44.1	54.7	50.6	48.4	33.8	44.2	54.5	50.2	48.4	33.1	43.9	54.2
<i>rpl14</i>	53.7	38.2	25.2	39.0	44.4	53.7	38.2	23.6	38.5	43.7	53.7	38.2	23.6	38.5	43.7
<i>rpl20</i>	36.4	40.5	24.8	33.9	48.3	36.4	39.8	24.6	33.6	46.2	36.4	39.8	24.6	33.6	46.2
<i>rpl22</i>	42.4	33.9	26.7	34.3	48.6	45.0	35.1	27.2	35.8	43.5	45.1	37.0	26.5	36.2	41.6
<i>rpoA</i>	43.5	32.7	26.5	34.2	48.5	44.1	33.2	26.5	34.6	48.8	44.4	33.2	26.5	34.7	48.8
<i>rpoB</i>	50.1	38.1	27.4	38.6	48.9	50.4	38.3	27.2	38.7	48.3	50.5	38.3	27.0	38.6	48.2
<i>rpoC1</i>	51.0	38.1	30.0	39.7	51.4	50.4	38.0	29.0	39.1	50.0	50.5	38.1	29.0	39.2	50.2
<i>rpoC2</i>	45.4	37.5	28.2	37.0	49.2	45.3	37.3	28.3	36.9	49.6	45.2	37.3	28.2	36.9	49.5
<i>rps2</i>	44.3	42.6	26.6	37.8	45.5	44.3	41.8	26.6	37.6	45.2	44.3	42.2	26.6	37.7	45.6
<i>rps3</i>	48.0	33.8	24.2	35.3	46.8	48.4	33.8	24.2	35.5	45.9	48.4	33.8	24.7	35.6	46.9
<i>rps4</i>	51.5	38.1	26.2	38.6	48.8	49.5	38.1	26.2	38.0	49.3	49.5	38.1	26.2	38.0	49.3
<i>rps7</i>	52.6	45.5	26.3	41.5	44.7	52.6	45.5	26.3	41.5	44.7	52.6	45.5	26.3	41.5	44.7
<i>rps8</i>	39.3	40.0	26.7	35.3	41.3	38.5	39.3	26.7	34.8	47.1	39.3	40.0	25.9	35.1	42.9
<i>rps11</i>	53.2	56.8	22.3	44.1	45.5	54.0	56.8	20.1	43.7	45.2	54.0	56.8	21.6	44.1	45.4
<i>rps12</i>	52.9	50.4	29.4	44.3	51.2	52.1	50.4	30.3	44.3	51.7	52.1	50.4	30.3	44.3	51.7
<i>rps14</i>	43.6	46.5	32.7	40.9	39.7	42.6	46.5	32.7	40.6	37.6	42.6	46.5	32.7	40.6	37.6
<i>rps18</i>	35.3	40.2	28.4	34.6	37.4	35.3	40.2	29.4	35.0	38.4	35.3	40.2	29.4	35.0	38.4
<i>ycf1</i>	36.5	28.3	25.4	30.1	48.0	40.6	31.8	34.9	35.8	47.6	36.4	28.9	25.5	30.2	47.6
<i>ycf2</i>	42.1	34.3	37.3	37.9	54.0	42.0	34.4	37.6	38.0	54.0	38.0	34.3	37.6	38.0	54.0
<i>ycf3</i>	47.9	38.5	33.7	40.0	57.7	47.3	38.5	32.5	39.5	58.7	47.3	38.5	32.5	39.5	57.4
<i>ycf4</i>	44.3	40.0	35.7	40.0	55.2	43.8	40.5	34.6	39.6	56.6	43.8	40.5	35.1	39.8	57.5
\bar{X}	46.8	39.1	28.2	38.0	47.2	47.0	39.2	28.5	38.2	47.3	46.8	39.2	28.4	38.1	47.2

¹⁾ GC₁: 密码子第1位碱基的GC含量 GC content of the first base of codon; GC₂: 密码子第2位碱基的GC含量 GC content of the second base of codon; GC₃: 密码子第3位碱基的GC含量 GC content of the third base of codon; GC_{all}: 密码子总GC含量 Total GC content of codon; ENC: 有效密码子数 Effective number of codon. \bar{X} : 均值 Average.

由表1还可见:拉关木、红榄李和榄李叶绿体基因组ENC值的变化范围和均值均较为接近,变化范围分别为37.4~59.8、37.6~59.8和37.6~59.8,均值分别为47.2、47.3和47.2。3种植物叶绿体基因组中ENC值大于45的基因分别有38、37和38个,占比分别为73.1%、71.2%和73.1%,说明供试3种植物叶绿体基因组密码子使用偏好性较弱。

2.1.2 相关性分析 对3种红树植物叶绿体基因组的GC₁、GC₂、GC₃、GC_{all}和ENC值进行相关性分析,结果(表2)显示:GC₁与GC₂值呈极显著($P<0.01$)正相

关,而GC₁和GC₂与GC₃值均呈不显著正相关,表明3种植物叶绿体基因组密码子第1和第2位碱基组成差异较小,第3位碱基与第1和第2位碱基组成差异较大。ENC与GC₃值呈极显著正相关,表明3种植物叶绿体基因组密码子第3位碱基的组成对密码子使用偏好性的影响非常大;ENC与GC_{all}值呈显著($P<0.05$)正相关,表明3种植物叶绿体基因组密码子的总GC含量对密码子使用偏好性的影响较大。

2.1.3 相对同义密码子使用度(RSCU)分析 去除3个终止密码子(UAG、UAA、UGA)和2个起始密码

子(AUG、UGG),对剩余59个密码子的RSCU值进行分析,结果(表3)显示:拉关木、红榄李和榄李叶绿体基因组密码子的RSCU值分别为0.34~2.06、0.33~2.01和0.34~2.03。3种植物叶绿体基因组亮氨酸(Leu)密码子UUA的RSCU值均最大,丝氨酸(Ser)密码子AGC的RSCU值均最小。值得注意的是,除Leu密码子UUG的RSCU值大于1外,3种植物叶绿体基因组中以G或C结尾的密码子的RSCU值均小于1,表明3种植物叶绿体基因组中以G或C结尾的密码子使用偏好性弱。3种植物叶绿体基因组中RSCU值大于1的密码子均有29个,表明这些密码子的使用偏好性较强,为高频密码子。其中,以A结尾的高频密码子有12个,以G结尾的高频密码子有1个,以U结尾的高频密码子有16个,以A或U结尾的高频密码子占比为96.6%,表明3种植物叶绿体基因组密码子偏好以A或U结尾。

2.2 中性绘图分析

中性绘图分析结果(图1)显示:拉关木、红榄李和榄李叶绿体基因组密码子第1和第2位碱基GC含量的均值(GC_{12})分别为32.4%~55.0%、32.8%~56.3%和32.7%~55.4%,密码子第3位碱基的GC含量(GC_3)分别为21.4%~37.3%、20.1%~37.6%和21.5%~37.6%。从图1可见,3种植物叶绿体基因组的大部分基因远离标准曲线,说明3种植物叶绿体基因组密码子使用偏好性受自然选择影响。拉关木、红榄李和榄李叶绿体基因组中性绘图的回归系数(即

表2 使君子科3种红树植物叶绿体基因组密码子GC含量和有效密码子数的相关性分析¹⁾

Table 2 Correlation analysis on GC content of codon and effective number of codon in chloroplast genomes of three species of mangroves in Combretaceae¹⁾

指标 Index	拉关木指标间的相关系数 Correlation coefficient between indexes of <i>Laguncularia racemosa</i>				
	GC ₁	GC ₂	GC ₃	GC _{all}	ENC
GC ₁	1.000				
GC ₂	0.468 **	1.000			
GC ₃	0.194	0.142	1.000		
GC _{all}	0.840 **	0.779 **	0.504 **	1.000	
ENC	0.302	0.016	0.499 **	0.339 *	1.000
指标 Index	红榄李指标间的相关系数 Correlation coefficient between indexes of <i>Lumnitzera littorea</i>				
	GC ₁	GC ₂	GC ₃	GC _{all}	ENC
GC ₁	1.000				
GC ₂	0.430 **	1.000			
GC ₃	0.050	0.007	1.000		
GC _{all}	0.823 **	0.748 **	0.397 **	1.000	
ENC	0.261	-0.049	0.386 **	0.272 *	1.000
指标 Index	榄李指标间的相关系数 Correlation coefficient between indexes of <i>Lumnitzera racemosa</i>				
	GC ₁	GC ₂	GC ₃	GC _{all}	ENC
GC ₁	1.000				
GC ₂	0.461 **	1.000			
GC ₃	0.091	0.067	1.000		
GC _{all}	0.823 **	0.775 **	0.430 **	1.000	
ENC	0.235	-0.063	0.419 **	0.253 *	1.000

¹⁾ GC₁: 密码子第1位碱基的GC含量 GC content of the first base of codon; GC₂: 密码子第2位碱基的GC含量 GC content of the second base of codon; GC₃: 密码子第3位碱基的GC含量 GC content of the third base of codon; GC_{all}: 密码子总GC含量 Total GC content of codon; ENC: 有效密码子数 Effective number of codon. *: P<0.05; **: P<0.01.

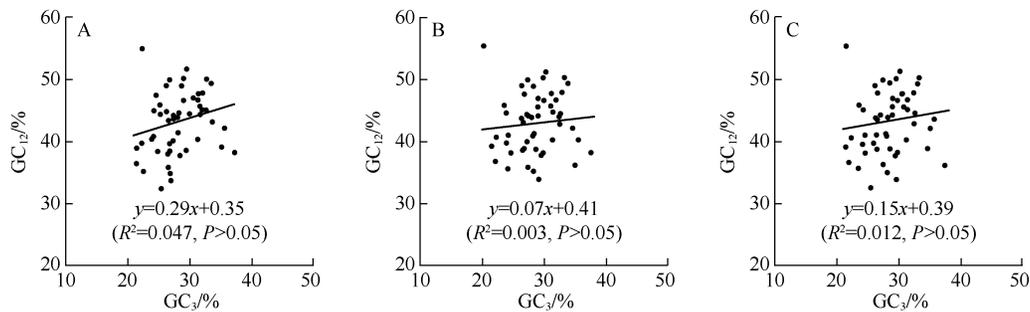
表3 使君子科3种红树植物叶绿体基因组的相对同义密码子使用度(RSCU)分析

Table 3 Analysis on relative synonymous codon usage (RSCU) of chloroplast genomes of three species of mangroves in Combretaceae

氨基酸 Amino acid	密码子 Codon	不同植物的RSCU值 RSCU value of different species			氨基酸 Amino acid	密码子 Codon	不同植物的RSCU值 RSCU value of different species			
		拉关木 <i>Laguncularia racemosa</i>	红榄李 <i>Lumnitzera littorea</i>	榄李 <i>Lumnitzera racemosa</i>			拉关木 <i>Laguncularia racemosa</i>	红榄李 <i>Lumnitzera littorea</i>	榄李 <i>Lumnitzera racemosa</i>	
Ala	GCU	1.89	1.89	1.88	Asp	GAU	1.63	1.62	1.63	
	GCC	0.59	0.59	0.59		GAC	0.37	0.38	0.37	
	GCA	1.09	1.09	1.10		Cys	UGU	1.53	1.52	1.50
	GCG	0.43	0.43	0.43			UGC	0.47	0.48	0.50
Arg	CGU	1.40	1.41	1.39	Gln	CAA	1.55	1.53	1.55	
	CGC	0.36	0.39	0.37		CAG	0.45	0.47	0.45	
	CGA	1.45	1.40	1.40	Glu	GAA	1.50	1.49	1.50	
	CGG	0.43	0.45	0.43		GAG	0.50	0.51	0.50	
Arg	AGA	1.74	1.70	1.77	Gly	GGU	1.30	1.31	1.29	
	AGG	0.62	0.66	0.64		GGC	0.43	0.43	0.43	
Asn	AAU	1.56	1.54	1.55		GGA	1.64	1.61	1.63	
	AAC	0.44	0.46	0.45	GGG	0.63	0.65	0.65		

续表3 Table 3 (Continued)

氨基酸 Amino acid	密码子 Codon	不同植物的 RSCU 值 RSCU value of different species			氨基酸 Amino acid	密码子 Codon	不同植物的 RSCU 值 RSCU value of different species		
		拉关木 <i>Laguncularia racemosa</i>	红榄李 <i>Lumnitzera littorea</i>	榄李 <i>Lumnitzera racemosa</i>			拉关木 <i>Laguncularia racemosa</i>	红榄李 <i>Lumnitzera littorea</i>	榄李 <i>Lumnitzera racemosa</i>
		His	CAU	1.52			1.50	1.50	Ser
	CAC	0.48	0.50	0.50		UCU	1.71	1.72	1.70
Ile	AUU	1.50	1.50	1.49		UCC	0.93	0.96	0.97
	AUC	0.61	0.62	0.60		UCA	1.18	1.11	1.16
	AUA	0.89	0.88	0.90		UCG	0.58	0.57	0.56
Leu	UUA	2.06	2.01	2.03		AGU	1.26	1.31	1.28
	UUG	1.12	1.14	1.13		AGC	0.34	0.33	0.34
	CUU	1.25	1.24	1.26	Thr	ACU	1.58	1.58	1.57
	CUC	0.38	0.39	0.39		ACC	0.71	0.76	0.74
	CUA	0.83	0.85	0.84		ACA	1.23	1.19	1.22
	CUG	0.36	0.36	0.35		ACG	0.48	0.47	0.47
Lys	AAA	1.52	1.48	1.51	Tyr	UAU	1.64	1.63	1.64
	AAG	0.48	0.52	0.49		UAC	0.36	0.37	0.36
Phe	UUU	1.33	1.30	1.33	Val	GUU	1.46	1.46	1.47
	UUC	0.67	0.70	0.67		GUC	0.50	0.50	0.51
Pro	CCU	1.65	1.64	1.63		GUA	1.48	1.48	1.47
	CCC	0.70	0.70	0.70		GUG	0.56	0.56	0.55
	CCA	1.13	1.09	1.11					



GC₁₂: 密码子第 1 和第 2 位碱基 GC 含量的均值 Average of GC content of the first and second bases of codon; GC₃: 密码子第 3 位碱基的 GC 含量 GC content of the third base of codon.

A: 拉关木 *Laguncularia racemosa* C. F. Gaertn.; B: 红榄李 *Lumnitzera littorea* (Jack) Voigt; C: 榄李 *Lumnitzera racemosa* Willd.

图 1 使君子科 3 种红树植物叶绿体基因组的中性绘图分析

Fig. 1 Neutral mapping analysis on chloroplast genomes of three species of mangroves in Combretaceae

标准曲线斜率)分别为 0.29、0.07 和 0.15,表明 3 种植物叶绿体基因组的 GC₁₂与 GC₃值均呈正相关,但相关性均不显著。以上结果表明:3 种植物叶绿体基因组密码子第 1 和第 2 位碱基与第 3 位碱基的组成差异明显。

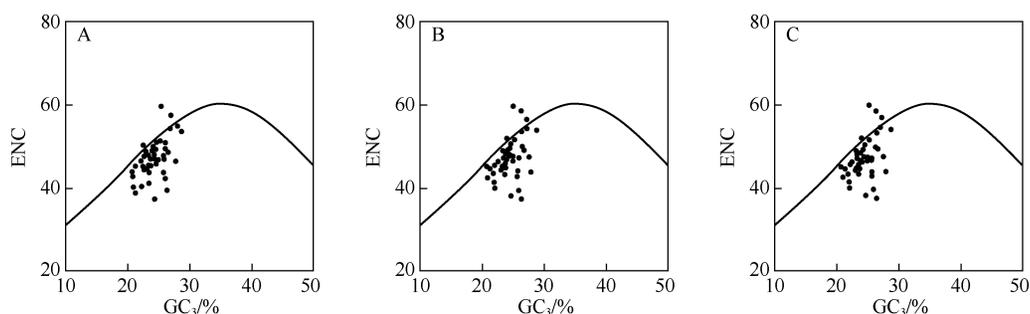
2.3 ENC-plot 分析

ENC-plot 分析结果(图 2)显示:仅少数基因分布在标准曲线上或附近,大多数基因分布在标准曲线下方,且距离标准曲线较远,说明 3 种植物叶绿体基

因组密码子使用偏好性受碱基突变影响较小,受自然选择影响较大。

2.4 PR2-plot 分析

PR2-plot 分析结果(图 3)显示:3 种植物叶绿体基因在 4 个区域内的分布不均匀,且多数基因距离中心点较远,说明 3 种植物叶绿体基因组密码子使用偏好性受自然选择和碱基突变的共同影响。拉关木、红榄李和榄李叶绿体基因组中 G₃/(G₃+C₃)小于 0.5 的基因分别有 22、23 和 24 个,A₃/(A₃+T₃)小于 0.5 的

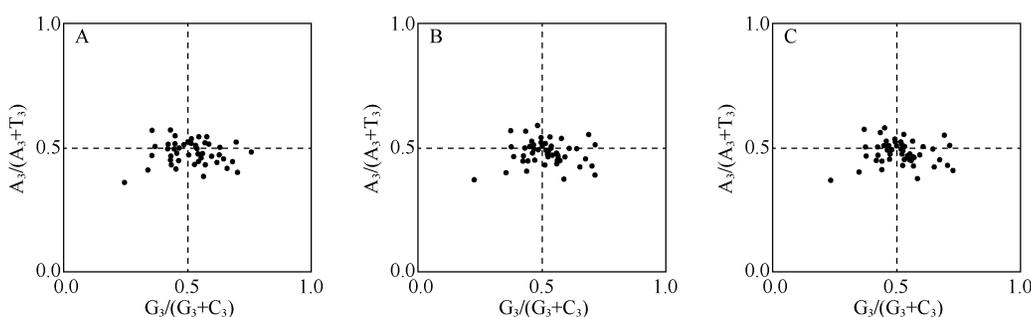


ENC: 有效密码子数 Effective number of codon; GC₃: 密码子第3位碱基的GC含量 GC content of the third base of codon.

A: 拉关木 *Laguncularia racemosa* C. F. Gaertn.; B: 红榄李 *Lumnitzera littorea* (Jack) Voigt; C: 榄李 *Lumnitzera racemosa* Willd.

图2 使君子科3种红树植物叶绿体基因组的ENC-plot分析

Fig. 2 ENC-plot analysis on chloroplast genomes of three species of mangroves in Combretaceae



A₃, T₃, C₃, G₃: 分别为密码子第3位碱基的A、T、C和G含量 Contents of A, T, C and G of the third base of codon, respectively.

A: 拉关木 *Laguncularia racemosa* C. F. Gaertn.; B: 红榄李 *Lumnitzera littorea* (Jack) Voigt; C: 榄李 *Lumnitzera racemosa* Willd.

图3 使君子科3种红树植物叶绿体基因组的PR2-plot分析

Fig. 3 PR2-plot analysis on chloroplast genomes of three species of mangroves in Combretaceae

基因分别有31、34和33个,表明3种植物叶绿体基因组密码子第3位碱基使用G的频率高于C,使用T的频率高于A。

2.5 最优密码子分析

最优密码子分析结果(表4)显示:拉关木有8个最优密码子,分别为GCA、CGA、AGA、GGA、UUG、

表4 使君子科3种红树植物叶绿体基因组的最优密码子分析¹⁾

Table 4 Optimal codon analysis on chloroplast genomes of three species of mangroves in Combretaceae¹⁾

氨基酸 Amino acid	拉关木 <i>Laguncularia racemosa</i>						红榄李 <i>Lumnitzera littorea</i>						榄李 <i>Lumnitzera racemosa</i>					
	C	n1	RSCU1	n2	RSCU2	ΔRSCU	C	n1	RSCU1	n2	RSCU2	ΔRSCU	C	n1	RSCU1	n2	RSCU2	ΔRSCU
Ala	GCU	23	1.80	40	2.25	-0.45	GCU	28	1.70	42	2.40	-0.70	GCU	48	1.79	42	2.40	-0.61
	GCC	7	0.55	4	0.23	0.32	GCC	15	0.91	4	0.23	0.68	GCC	21	0.79	4	0.23	0.56
	GCA *	15	1.18	17	0.96	0.22	GCA	14	0.85	16	0.91	-0.06	GCA *	27	1.01	16	0.91	0.10
	GCG	6	0.47	10	0.56	-0.09	GCG	9	0.55	8	0.46	0.09	GCG	11	0.41	8	0.46	-0.05
Arg	CGU	15	0.60	23	2.51	-1.91	CGU	16	1.14	22	2.40	-1.26	CGU	30	0.83	22	2.40	-1.57
	CGC	11	0.44	2	0.22	0.22	CGC	5	0.36	2	0.22	0.14	CGC	16	0.44	2	0.22	0.22
	CGA *	35	1.39	11	1.20	0.19	CGA *	21	1.50	11	1.20	0.30	CGA *	50	1.38	11	1.20	0.18
	CGG	18	0.72	2	0.22	0.50	CGG	5	0.36	3	0.33	0.03	CGG	22	0.61	3	0.33	0.28
AGA *	AGA *	45	1.79	11	1.20	0.59	AGA *	27	1.93	11	1.20	0.73	AGA *	63	1.74	11	1.20	0.54
	AGG	27	1.07	6	0.65	0.42	AGG	10	0.71	6	0.65	0.06	AGG *	36	1.00	6	0.65	0.35
Asn	AAU	106	1.51	39	1.44	0.07	AAU *	51	1.59	38	1.38	0.21	AAU *	143	1.55	39	1.42	0.13

续表4 Table 4 (Continued)

氨基酸 Amino acid	拉关木 <i>Laguncularia racemosa</i>						红榄李 <i>Lumnitzera littorea</i>						榄李 <i>Lumnitzera racemosa</i>					
	C	n1	RSCU1	n2	RSCU2	ΔRSCU	C	n1	RSCU1	n2	RSCU2	ΔRSCU	C	n1	RSCU1	n2	RSCU2	ΔRSCU
Asp	AAC	34	0.49	15	0.56	-0.07	AAC	13	0.41	17	0.62	-0.21	AAC	42	0.45	16	0.58	-0.13
	GAU	109	1.65	18	1.64	0.01	GAU	24	1.55	17	1.62	-0.07	GAU	132	1.62	17	1.62	0.00
Cys	GAC	23	0.35	4	0.36	-0.01	GAC	7	0.45	4	0.38	0.07	GAC	31	0.38	4	0.38	0.00
	UGU	17	1.17	6	1.50	-0.33	UGU	8	1.23	6	1.50	-0.27	UGU	23	1.15	6	1.50	-0.35
Gln	UGC	12	0.83	2	0.50	0.33	UGC	5	0.77	2	0.50	0.27	UGC	17	0.85	2	0.50	0.35
	CAA	60	1.38	21	1.68	-0.30	CAA	19	1.19	22	1.63	-0.44	CAA	75	1.34	22	1.63	-0.29
Glu	CAG	27	0.62	4	0.32	0.30	CAG	13	0.81	5	0.37	0.44	CAG *	37	0.66	5	0.37	0.29
	GAA	80	1.22	36	1.60	-0.38	GAA	61	1.51	37	1.64	-0.13	GAA	113	1.28	37	1.64	-0.36
Gly	GAG	51	0.78	9	0.40	0.38	GAG	20	0.49	8	0.36	0.13	GAG	63	0.72	8	0.36	0.36
	GGU	23	0.97	31	1.72	-0.75	GGU	31	1.43	31	1.75	-0.32	GGU	46	1.14	31	1.72	-0.58
His	GGC	9	0.38	13	0.72	-0.34	GGC	7	0.32	12	0.68	-0.36	GGC	14	0.35	12	0.67	-0.32
	GGA *	39	1.64	24	1.33	0.31	GGA *	33	1.52	23	1.30	0.22	GGA *	65	1.6	24	1.33	0.27
	GGG	24	1.01	4	0.22	0.79	GGG	16	0.74	5	0.28	0.46	GGG	37	0.91	5	0.28	0.63
	CAU	48	1.52	19	1.46	0.06	CAU	15	1.36	20	1.48	-0.12	CAU	59	1.55	20	1.48	0.07
Ile	CAC	15	0.48	7	0.54	-0.06	CAC	7	0.64	7	0.52	0.12	CAC	17	0.45	7	0.52	-0.07
	AUU	74	1.30	49	1.65	-0.35	AUU	44	1.33	50	1.69	-0.36	AUU	105	1.27	51	1.74	-0.47
Leu	AUC	44	0.77	18	0.61	0.16	AUC	23	0.70	17	0.57	0.13	AUC	65	0.79	16	0.55	0.24
	AUA	53	0.93	22	0.74	0.19	AUA	32	0.97	22	0.74	0.23	AUA	78	0.94	21	0.72	0.22
	UUA	39	0.91	55	2.73	-1.82	UUA	34	2.02	54	2.63	-0.61	UUA	59	1.11	55	2.66	-1.55
	UUG *	68	1.59	22	1.09	0.50	UUG	19	1.13	24	1.17	-0.04	UUG *	81	1.53	24	1.16	0.37
Lys	CUU *	60	1.40	22	1.09	0.31	CUU	16	0.95	22	1.07	-0.12	CUU *	71	1.34	22	1.06	0.28
	CUC	28	0.65	3	0.15	0.50	CUC	11	0.65	4	0.20	0.45	CUC	33	0.62	4	0.19	0.43
	CUA	39	0.91	16	0.79	0.12	CUA	15	0.89	16	0.78	0.11	CUA	50	0.94	16	0.77	0.17
	CUG	23	0.54	3	0.15	0.39	CUG	6	0.36	3	0.15	0.21	CUG	24	0.45	3	0.15	0.30
Phe	AAA	93	1.26	30	1.50	-0.24	AAA	43	1.39	28	1.44	-0.05	AAA	121	1.29	28	1.44	-0.15
	AAG	55	0.74	10	0.50	0.24	AAG	19	0.61	11	0.56	0.05	AAG	67	0.71	11	0.56	0.15
Pro	UUU	83	1.06	44	1.28	-0.22	UUU	32	1.19	43	1.28	-0.09	UUU	95	1.07	43	1.30	-0.23
	UUC	74	0.94	25	0.72	0.22	UUC	22	0.81	24	0.72	0.09	UUC	83	0.93	23	0.70	0.23
Ser	CCU	28	1.27	17	1.48	-0.21	CCU	16	1.52	17	1.48	0.04	CCU	37	1.24	17	1.48	-0.24
	CCC	17	0.77	7	0.61	0.16	CCC	10	0.95	6	0.52	0.43	CCC	25	0.84	6	0.52	0.32
	CCA	27	1.23	17	1.48	-0.25	CCA	6	0.57	17	1.48	-0.91	CCA	32	1.08	17	1.48	-0.40
	CCG	16	0.73	5	0.43	0.30	CCG	10	0.95	6	0.52	0.43	CCG	25	0.84	6	0.52	0.32
Thr	UCU	67	1.70	24	1.67	0.03	UCU	19	1.54	25	1.70	-0.16	UCU	78	1.65	25	1.70	-0.05
	UCC	48	1.22	8	0.56	0.66	UCC	10	0.81	9	0.61	0.20	UCC	52	1.10	9	0.61	0.49
	UCA	46	1.17	18	1.26	-0.09	UCA	11	0.89	16	1.09	-0.20	UCA	53	1.12	17	1.16	-0.04
	UCG	29	0.74	7	0.49	0.25	UCG	11	0.89	9	0.61	0.28	UCG	36	0.76	8	0.55	0.21
Tyr	AGU	35	0.89	24	1.67	-0.78	AGU *	19	1.54	21	1.43	0.11	AGU	50	1.06	21	1.43	-0.37
	AGC	11	0.28	5	0.35	-0.07	AGC	4	0.32	8	0.55	-0.23	AGC	14	0.30	8	0.55	-0.25
	ACU	27	1.16	27	1.89	-0.73	ACU	12	0.87	28	2.00	-1.13	ACU	38	1.16	28	2.00	-0.84
	ACC	19	0.82	17	1.19	-0.37	ACC	16	1.16	16	1.14	0.02	ACC	30	0.92	16	1.14	-0.22
Val	ACA *	31	1.33	13	0.91	0.42	ACA *	20	1.45	12	0.86	0.59	ACA *	41	1.25	12	0.86	0.39
	ACG	16	0.69	0	0.00	0.69	ACG	7	0.51	0	0.00	0.51	ACG	22	0.67	0	0.00	0.67
	UAU	62	1.57	27	1.54	0.03	UAU	33	1.57	26	1.53	0.04	UAU	85	1.53	26	1.53	0.00
	UAC	17	0.43	8	0.46	-0.03	UAC	9	0.43	8	0.47	-0.04	UAC	26	0.47	8	0.47	0.00
Val	GUU *	30	1.41	18	1.24	0.17	GUU	15	1.02	19	1.33	-0.31	GUU	39	1.25	19	1.31	-0.06
	GUC	16	0.75	6	0.41	0.34	GUC	11	0.75	6	0.42	0.33	GUC	22	0.70	6	0.41	0.29
	GUA	21	0.99	30	2.07	-1.08	GUA	23	1.56	28	1.96	-0.40	GUA	38	1.22	29	2.00	-0.78
	GUG	18	0.85	4	0.28	0.57	GUG	10	0.68	4	0.28	0.40	GUG	26	0.83	4	0.28	0.55

¹⁾ C: 密码子 Codon; n1: 高表达组的密码子数量 Codon number of high expression group; RSCU1: 高表达组的相对同义密码子使用度 Relative synonymous codon usage of high expression group; n2: 低表达组的密码子数量 Codon number of low expression group; RSCU2: 低表达组的相对同义密码子使用度 Relative synonymous codon usage of low expression group; ΔRSCU: 高表达组与低表达组相对同义密码子使用度的差值 Difference of relative synonymous codon usage between high expression group and low expression group. *: 最优密码子 Optimal codon.

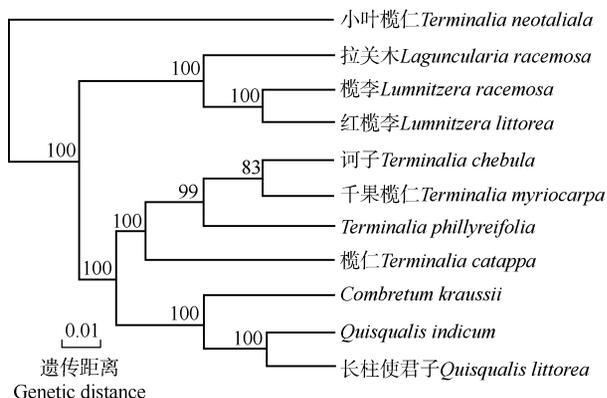
CUU、ACA、GUU,其中以 A 或 U 结尾的最优密码子有 7 个,以 G 或 C 结尾的最优密码子仅 1 个;红榄李有 6 个最优密码子,分别为 CGA、AGA、AAU、GGA、AGU、ACA,并且,这些最优密码子均以 A 或 U 结尾;榄李有 10 个最优密码子,分别为 GCA、CGA、AGA、AGG、AAU、CAG、GGA、UUG、CUU、ACA,其中以 A 或 U 结尾的最优密码子有 7 个,以 G 或 C 结尾的最优密码子有 3 个。可见,3 种植物的最优密码子多以 A 或 U 结尾。

由表 4 还可见:拉关木、红榄李和榄李叶绿体基因组的共有最优密码子有 4 个,分别为 CGA、AGA、GGA 和 ACA,其中,CGA 和 AGA 编码精氨酸(Arg),GGA 编码甘氨酸(Gly),ACA 编码苏氨酸(Thr)。值得注意的是,这些共有最优密码子的第 3 位碱基均为 A,说明 3 种植物叶绿体基因组的共有最优密码子模式为“NNA”。

2.6 系统发育分析

基于 *rbcL* 基因序列构建使君子科 11 种植物的系统进化树,结果(图 4)显示:红榄李和榄李首先聚在一起,再与拉关木聚为一个小支,说明红榄李与同属植物榄李的亲缘关系最近,并且这 2 种植物与拉关木的亲缘关系较近。

基于叶绿体基因组 44 个共有蛋白质编码序列的 RSCU 值对使君子科 11 种植物进行聚类分析,结果(图 5)显示:红榄李和榄李也首先聚在一起,但拉关木与这 2 种植物的亲缘关系与图 4 明显不同。



分支上的数据为自展支持率 The datums on the branches are bootstrap values.

图 4 基于 *rbcL* 基因序列的使君子科 3 种红树植物与同科其他植物的系统发育树

Fig. 4 Phylogenetic tree of three species of mangroves in Combretaceae and other species of the same family based on *rbcL* gene sequence

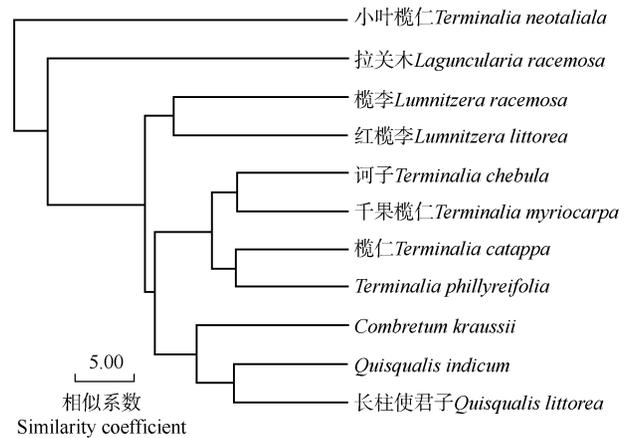


图 5 基于叶绿体基因组共有蛋白质编码序列相对同义密码子使用度的使君子科 3 种红树植物与同科其他植物的系统发育树

Fig. 5 Phylogenetic tree of three species of mangroves in Combretaceae and other species of the same family based on relative synonymous codon usage of common protein coding sequence of chloroplast genome

3 讨论和结论

研究发现,植物基因组中蛋白质编码序列的 GC 含量在同一物种的个体间较为保守,一般高于 40%;叶绿体基因组蛋白质编码序列的 GC 含量一般低于 40%,可能独立于核基因组起源和进化^[27]。供试使君子科 3 种红树植物叶绿体基因组密码子总 GC 含量(GC_{all})的均值为 38.0%~38.2%,均低于 40%,说明拉关木、红榄李和榄李叶绿体基因组均偏好使用 A 或 U 结尾的密码子,研究者在忍冬属(*Lonicera* Linn.) 24 种植物^[27]、梔子(*Gardenia jasminoides* Ellis)^[28]和睡莲属(*Nymphaea* Linn.) 7 种植物^[29]的叶绿体基因组密码子研究中也得到相同的结果,说明高等植物叶绿体基因组密码子偏好以 A 或 U 结尾是较为保守的特性。植物叶绿体基因组密码子 3 个碱基位置的 GC 含量一般为第 1 位碱基的 GC 含量(GC_1)最高、第 2 位碱基的 GC 含量(GC_2)居中、第 3 位碱基的 GC 含量(GC_3)最低^[19,27-29]。供试使君子科 3 种红树植物叶绿体基因组的 GC_1 、 GC_2 和 GC_3 值的均值也表现出相同的规律,说明 3 种植物叶绿体基因组的第 3 位碱基较第 1 和第 2 位碱基更偏好以 A 或 U 结尾。由于植物叶绿体基因组密码子第 3 位碱基受自然选择影响高于第 1 和第 2 位碱基,故可将 GC_3 值作为衡量密码子使用偏好性的重要指标^[29]。研究发现,有效密码子数(ENC)越大,表示该密码子的使用偏好性越弱^[9,15]。拉关木、红榄李和榄李叶绿体基因组 ENC

值的均值均大于 45,且 ENC 值大于 45 的基因占比均在 70%以上,表明这 3 种植物叶绿体基因组密码子使用偏好性较弱,赵森等^[19]在红树植物秋茄中也得到相同的结果,研究者在非红树植物白羊草 [*Bothriochloa ischaemum* (Linn.) Keng]^[14]、滇楸 (*Catalpa fargesii* Bureau)^[15]和杉木属 (*Cunninghamia* R. Br. ex A. Rich.)^[18]等的叶绿体基因组密码子使用偏好性研究中也得到类似的结果,可见,潮间带生境不是影响红树植物叶绿体基因组密码子偏好性较弱的惟一原因^[21]。

相关研究表明:植物叶绿体基因组密码子使用偏好性由碱基突变和自然选择相互作用的强度和方向决定;当 GC₃ 值在一定范围内时,若 GC₁ 和 GC₃ 值之间、GC₂ 和 GC₃ 值之间的关系没有达到显著水平,则密码子使用偏好性主要受自然选择的影响^[27]。相关性分析和中性绘图分析结果均表明:拉关木、红榄李和榄李叶绿体基因组密码子的 GC₁ 和 GC₂ 与 GC₃ 值的相关性均较弱,表明碱基突变不是影响拉关木、红榄李和榄李叶绿体基因组密码子使用偏好性的主要因子,自然选择才是影响 3 种植物叶绿体基因组密码子使用偏好性的主要因子。在无自然选择的情况下,基因多分布在 ENC-plot 标准曲线上或附近^[27],拉关木、红榄李和榄李叶绿体基因组的多数基因分布于标准曲线的下方,且距离标准曲线较远,表明这 3 种植物叶绿体基因组密码子使用偏好性受碱基突变的影响较小,受自然选择的影响较大,而红树植物秋茄的叶绿体基因组密码子使用偏好性主要受碱基突变的影响^[19],可见,造成不同种类植物叶绿体基因组密码子使用偏好性的因子不尽相同。

在碱基突变和自然选择的双重作用下,植物会形成最优密码子^[29]。拉关木、红榄李和榄李叶绿体基因组的最优密码子分别有 8、6 和 10 个,均明显高于红树植物秋茄叶绿体基因组的最优密码子数(仅 3 个)^[19]。供试 3 种植物中,拉关木为外来种,在中国红树林适生区表现出极强的速生能力^[3];红榄李为濒危红树植物,在红树林适生区表现出最弱的适生性^[1];榄李则为中国分布最广泛的红树植物^[6]。至于同科红树植物叶绿体基因组的最优密码子数是否能够较好地反映植物对环境的适应能力尚待进一步的实验研究。另外,供试 3 种植物叶绿体基因组的共有最优密码子均以 A 结尾,符合高等植物叶绿体基因组密码子大部分最优密码子以 A 或 U 结尾的规

律^[30-33],并说明高等植物叶绿体基因组的进化相对保守^[19,30,34]。在后续研究中,可利用最优密码子提高目的基因的表达效率^[31,35-36],有助于濒危红树植物的分子育种工作,对濒危红树植物种质资源保育及生态修复亦有重要意义。

Ruang-Areerate 等^[23]认为,物种间的亲缘关系越近,其密码子使用模式越相近。本研究结果显示:使君子科 11 种植物基于 *rbcl* 基因序列和叶绿体基因组共有蛋白质编码序列(CDS)的相对同义密码子使用度(RSCU)构建的系统发育树不尽相同,红榄李和榄李率先聚在一起,但与拉关木的亲缘关系不一致。基于 *rbcl* 基因序列的聚类结果与项目组前期使用拉关木、红榄李和榄李叶绿体基因组序列构建的系统发育树聚类结果一致^[20],但在其他植物类似研究中存在基于 *rbcl* 基因序列的聚类结果与基于共有 CDS 的 RSCU 值的聚类结果不一致的问题^[19]。因此,在较大的分类单元(如科水平)中,基于共有 CDS 的 RSCU 值构建的系统发育树只能作为传统分类的补充^[32],不宜单独使用。

综上所述,供试 3 种红树植物叶绿体基因组密码子使用偏好性较为接近,均偏好以 A 或 U 结尾,共有最优密码子均以 A 结尾;并且,3 种植物叶绿体基因组密码子使用偏好性主要受自然选择的影响。

参考文献:

- [1] 张颖,陈光程,钟才荣.中国濒危红树植物研究与恢复现状[J].应用海洋学报,2021,40(1):142-153.
- [2] 范航清,王文卿.中国红树林保育的若干重要问题[J].厦门大学学报(自然科学版),2017,56(3):323-330.
- [3] ZHANG S J, LIU Y Q, WANG B Y, et al. Unraveling molecular mechanisms underlying low-temperature adaptation in *Laguncularia racemosa*[J]. Plant Physiology and Biochemistry, 2023, 199: 107747.
- [4] 刘强,张颖,钟才荣,等.外来红树植物拉关木入侵性研究[J].湖北农业科学,2019,58(21):60-64,67.
- [5] 郝露露,柯明思,朱奕秀,等.低温胁迫下红榄李(*Lumnitzera littorea*) DEAD-box RNA 解旋酶基因的表达分析[J].热带海洋学报,2022,41(6):44-55.
- [6] 王文卿,王瑁.中国红树林[M].北京:科学出版社,2007:26-27.
- [7] 陈健辉,缪绅裕,黄惠芳,等.五种红树植物叶片结构的比较研究[J].广西植物,2018,38(5):655-664.
- [8] 向敏,刘强,李妮亚,等.引进红树拉关木和两种乡土红树离子平衡及光合作用的比较研究[J].广西植物,2016,36(4):387-396.
- [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好性分析[J].植物资源与环境学报,2023,32(6):1-10.

- [10] XIA X H, HAO L L, SUN Y F, et al. Unravelling chilling-stress resistance mechanisms in endangered mangrove plant *Lumnitzera littorea* (Jack) Voigt[J]. *Marine Environmental Research*, 2023, 192: 106210.
- [11] 李燕华, 杨勇, 张颖. 铁炉港红树植物红榄李和榄李各器官矿质元素含量分析[J]. *湿地科学*, 2016, 14(3): 433-438.
- [12] 邓传远, 林鹏, 郭素枝. 榄李属(*Lumnitzera*)红树植物的木材解剖学研究[J]. *厦门大学学报(自然科学版)*, 2004, 43(3): 406-411.
- [13] HE Z W, FENG X, CHEN Q, et al. Evolution of coastal forests based on a full set of mangrove genomes[J]. *Nature Ecology and Evolution*, 2022, 6: 738-749.
- [14] 高守奥, 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.
- [15] 李江飞, 李熙颜, 王瑜, 等. 滇楸叶绿体基因组密码子偏好性分析[J]. *基因组学与应用生物学*, 2022, 41(4): 843-853.
- [16] YU C H, DANG Y K, ZHOU Z P, et al. Codon usage influences the local rate of translation elongation to regulate co-translational protein folding[J]. *Molecular Cell*, 2015, 59: 744-754.
- [17] 赵明, 张宏斌, 李伟, 等. 祁连圆柏叶绿体基因组序列特征分析[J]. *植物资源与环境学报*, 2023, 32(3): 1-11.
- [18] 郑薇玮, 陈金慧, 郝兆东, 等. 杉木等5种针叶树叶叶绿体密码子偏好性分析[J]. *分子植物育种*, 2016, 14(5): 1091-1097.
- [19] 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.
- [20] ZHANG Y, LI H L, ZHONG J D, et al. Chloroplast genome sequences and comparative analysis of Combretaceae mangroves with related species[J]. *BioMed Research International*, 2020, 2020: 5867673.
- [21] ZHANG Y, ZHANG J W, YONG Y, et al. Structural and comparative analysis of the complete chloroplast genome of a mangrove plant; *Scyphiphora hydrophyllacea* Gaertn. f. and related Rubiaceae species[J]. *Forests*, 2019, 10: 1000.
- [22] XU X M, SHEN Y J, ZHANG Y C, et al. A comparison of 25 complete chloroplast genomes between sister mangrove species *Kandelia obovata* and *Kandelia candel* geographically separated by the South China Sea[J]. *Frontiers in Plant Science*, 2023, 13: 1075353.
- [23] RUANG-AREERATE P, KONGKACHANA W, NAKTANG C, et al. Complete chloroplast genome sequences of five *Bruguiera* species (Rhizophoraceae): comparative analysis and phylogenetic relationships[J]. *PeerJ*, 2021, 9: e12268.
- [24] ASAF S, KHAN A L, Numan M, et al. Mangrove tree (*Avicennia marina*): insight into chloroplast genome evolutionary divergence and its comparison with related species from family Acanthaceae[J]. *Scientific Reports*, 2021, 11: 3586.
- [25] RUANG-AREERATE P, YOOCCHA T, KONGKACHANA W, et al. Comparative analysis and phylogenetic relationships of *Ceriops* species (Rhizophoraceae) and *Avicennia lanata* (Acanthaceae): insight into the chloroplast genome evolution between middle and seaward zones of mangrove forests[J]. *Biology*, 2022, 11: 383.
- [26] TAN F X, SHI S H, ZHONG Y, et al. Phylogenetic relationships of Combretaceae (Combretaceae) inferred from plastid, nuclear gene and spacer sequences[J]. *Journal of Plant Research*, 2002, 115: 475-481.
- [27] 王杰敏, 马东来, 韩晓伟, 等. 24份忍冬属材料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*, 2023, 32(3): 12-23.
- [28] 覃兴化, 张麒功, 汪其双, 等. 椴子叶绿体基因组密码子偏好性分析[J]. *福建农林大学学报(自然科学版)*, 2023, 52(5): 632-639.
- [29] 毛立彦, 黄秋伟, 龙凌云, 等. 7种睡莲属植物叶绿体基因组密码子偏好性分析[J]. *西北林学院学报*, 2022, 37(2): 98-107.
- [30] 舒军霞, 杨林, 周涛, 等. 刺柏属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.
- [31] 耿晓姝, 贾魏, 陈佳宁, 等. 金花茶叶绿体基因组密码子偏好性分析[J]. *分子植物育种*, 2022, 20(7): 2196-2203.
- [32] 赵春丽, 彭丽云, 王晓, 等. 莴菜 *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.
- [33] 王占军, 吴子琦, 王朝霞, 等. 3个茶树品种 *WOX* 基因家族的精华及密码子偏好性比较[J]. *南京林业大学学报(自然科学版)*, 2022, 46(2): 71-80.
- [34] 赵振宁, 孙浩田, 宋雨茹, 等. 山楂属植物叶绿体基因组特征与密码子偏好性分析[J]. *江苏农业学报*, 2023, 39(2): 504-517.
- [35] 刘潮, 韩利红, 盛巧, 等. 8种植物类甜蛋白家族基因进化及密码子特征分析[J]. *江苏农业科学*, 2022, 50(10): 44-51.
- [36] 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.

(责任编辑: 佟金凤)