

陈佳莉,代洪苇,周丽娟,等. 茶树 *CsTLRs* 基因密码子偏好性与进化分析[J]. 江苏农业学报,2025,41(5):996-1008.
doi:10.3969/j.issn.1000-4440.2025.05.018

茶树 *CsTLRs* 基因密码子偏好性与进化分析

陈佳莉^{1,2}, 代洪苇^{1,2}, 周丽娟³, 童华荣^{1,2}, 袁连玉^{1,2}

(1.西南大学食品科学学院,重庆 400715; 2.川渝共建特色食品重庆市重点实验室,重庆 400715; 3.重庆北碚农业发展有限公司,重庆 400715)

摘要: *TLR* 基因具有调控植物毛状体生长发育的作用,而叶片毛状体(茸毛)是茶树[*Camellia sinensis* (L.)]的重要形态特征之一,在茶树的生理、生态和茶叶品质方面均发挥着重要作用。本研究利用 CodonW、EMBOSS 等软件,通过中性绘图、奇偶偏好性分析(PR2-plot)等多种分析方法,对茶树 *CsTLRs* 基因的密码子参数、亲缘关系以及异源表达受体等方面进行分析。结果表明,茶树 *CsTLRs* 基因密码子使用偏好性较弱,偏好性形成主要来源于自然选择,筛选出 20 个最优密码子,多以 A/U 结尾;编码序列(CDS)和相对同义密码子使用度(RSCU)聚类分析结果都显示,茶树 *CsTLRs* 基因与杨树、黄瓜亲缘关系更近;烟草、拟南芥和番茄可作为 *CsTLRs* 异源转化的植物受体,酿酒酵母可作为微生物异源表达受体。综上,本研究结果明确了 *CsTLRs* 基因在调控茶树叶片毛状体发育时的密码子使用特征,为后续 *CsTLRs* 基因的功能研究提供了重要参考。

关键词: 茶树; 叶片毛状体; *CsTLRs* 基因; 密码子偏好性; 进化分析

中图分类号: S571.1 文献标识码: A 文章编号: 1000-4440(2025)05-0996-13

Codon bias and evolutionary analysis of *CsTLRs* in tea plant

CHEN Jiali^{1,2}, DAI Hongwei^{1,2}, ZHOU Lijuan³, TONG Huarong^{1,2}, YUAN Lianyu^{1,2}

(1.College of Food Science, Southwest University, Chongqing 400715, China; 2.Chongqing Key Laboratory of Speciality Food Co-Built by Sichuan and Chongqing, Chongqing 400715, China; 3.Chongqing Beibei Agricultural Development Co., Ltd., Chongqing 400715, China)

Abstract: The *TLR* gene has the function of regulating the growth and development of plant trichomes. Leaf trichome is one of the important morphological characteristics of *Camellia sinensis* (L.), which plays an important role in the physiology and ecology of tea plant and tea quality. In this study, we conducted a comprehensive analysis of tea plant *CsTLRs* genes, including codon parameters, genetic relationship and heterologous expression hosts. This was achieved using multiple analytical approaches including neutral plot analysis, parity rule 2 plot (PR2-plot) analysis, implemented through softwares such as CodonW and EMBOSS. The results showed that the codon usage bias of *CsTLRs* genes was weak, and the bias was primarily dominated by natural selection. There were 20 optimal codons in *CsTLRs* genes, mostly ending with A/U. Both cluster analyses of the coding sequences (CDS) and relative synonymous codon usage (RSCU) revealed that *CsTLRs* genes in tea plant

were phylogenetically closer to those in poplar (*Populus*) and cucumber (*Cucumis sativus*). For heterologous expression of *CsTLRs* genes from tea plant, tobacco (*Nicotiana tabacum*), Arabidopsis (*Arabidopsis thaliana*), and tomato (*Solanum lycopersicum*) could be selected as genetic transformation recipients, while the expression in microorganisms was better in *Saccharomyces cerevisiae*. In summary, this study analyzed the codon usage bias of *CsTLRs* genes in tea plants, elucidating their characteristic codon

收稿日期:2025-02-10

基金项目:重庆市技术创新与应用发展重大专项(CSTB2022TIAD-CUX0021);重庆市农业农村委员会特色经济作物提质增效项目(F2023757);2024年宜宾市农业创新能力建设项目(2024NYCX012)

作者简介:陈佳莉(2004-),女,重庆北碚人,本科,研究方向为茶树分子生物学。(E-mail) charli135@163.com

通讯作者:袁连玉,(E-mail) yuanlianyu@swu.edu.cn

usage patterns during the regulation of leaf trichome development. These findings would provide references for future functional studies of *CsTLRs* genes.

Key words: tea plant; leaf trichomes; *CsTLRs* genes; codon bias; evolutionary analysis

密码子由 mRNA 上编码氨基酸的 3 个核苷酸序列组成。生物体中共有 64 种密码子,其中有 3 种终止密码子不编码氨基酸,其余的 61 种密码子分别编码 20 种氨基酸,因此存在多个密码子编码同一种氨基酸的情况,即存在同义密码子^[1]。然而,同义密码子在不同基因或基因组中的利用程度不同,这种现象被称为密码子使用偏好性^[2-3]。密码子使用偏好性在动植物、细菌中普遍存在,这种偏好性虽然不会改变氨基酸序列,但会影响 RNA 加工、蛋白质翻译和折叠等过程^[4-5]。研究密码子偏好性可以更加深入地了解生物体的遗传特性,根据宿主密码子偏好性优化和改造外源基因的密码子,提升目的基因的表达效率和稳定性,为基因功能的研究提供重要支撑。

茶树叶片茸毛是叶片上的毛状体组织,由表皮细胞向外突出形成,属于非分泌腺不分支的单细胞,形态呈细长的银色丝状,常位于茶树的幼嫩芽叶表面,并随着叶片成熟而逐渐脱落^[6]。茶树叶片茸毛的存在增强了茶树抵御外界环境胁迫的能力,毛状体作为植物的第一道物理屏障,在保护植物免受过度太阳辐射、增加蒸腾作用、增强抗冻性和防御病虫害等方面有着重要作用^[7-9]。此外,茶毫(成品茶中的茶树叶片茸毛)是影响茶叶品质的关键因素。研究表明,茶毫富含氨基酸、茶多酚等功能性成分,这些物质不仅能赋予茶汤独特的鲜爽滋味和香气特征,还能提升茶叶的营养价值。在加工过程中,毫毛的数量及其显隐程度已经成为茶叶感官品质评价的重要标准之一,如碧螺春、都匀毛尖等名优茶均以“白毫显露”为重要特征。消费者通常更倾向于选择外观优美、口感鲜爽的茶产品,茶毫的存在可以提高茶叶的商品价值^[10-13]。

TLRs (TrichomeLess regulators) 是一类调控植物毛状体发育的转录因子,目前已在黄花蒿 (*Artemisia annua*) 中鉴定了 3 个成员:TLR1、TLR2、TLR3,它们在调控毛状体密度方面扮演了重要角色^[14-15]。TLR1 和 TLR2 可通过降低赤霉素水平负向调节毛状体密度,对毛状体的形成和青蒿素的生物合成有抑制作用,当这 2 个转录因子的功能被削弱时,毛状体数量增加,青蒿素水平也随之上升^[14]。TLR3 负调控毛状

体的密度和分支,同时,TLR3 与 TAR1、CycTL 互作调节黄花蒿表皮角质层发育和毛囊形态,对毛状体起始阶段和角质层生物合成尤为关键^[15]。目前关于茶树叶片毛状体的分子调控研究还较少,尚不足以揭示茶树叶片茸毛发育的机制。Liu 等^[16] 研究发现,CsWD40 与 2 个 bHLH 转录因子 (CsGL3 和 CsTT8)、2 个 MYB 转录因子 (CsAN2 和 CsMYB5e) 互作,并且 CsWD40 基因在拟南芥 *ttg1* 突变体中过表达能够恢复毛状体和种皮的发育,与 Sun 等^[17] 的研究结果相互印证。Li 等^[18] 研究发现,CsMYB1 与 CsGL3、CsWD40 有相互作用,并形成 MYB-bHLH-WD40 三聚体复合物,进而激活毛状体调节基因 *CsGL2* 和 *CsCPC*。刘任坚等^[19] 研究发现 CsHLH024/CsHLH133 和 CsTTG1 蛋白在植物中互作,它们的启动子能够在叶片组织中驱动下游基因表达进而发挥功能。茶树毛状体的形成机制复杂,对于 *CsTLRs* 调控茶树叶片毛状体发育的研究还有待补充。

本研究基于茶树全基因组数据库,利用 CodonW、EMBOSS 等软件对茶树 *CsTLRs* 基因的密码子参数和密码子使用偏好性进行分析,以期为后续研究 *CsTLRs* 在茶树叶片毛状体发育中的具体功能以及确定合适的异源表达宿主提供重要参考。

1 材料与方法

1.1 基因组序列获取

从美国国家生物技术信息中心 (NCBI) 数据库 (<https://www.ncbi.nlm.nih.gov/>) 获取黄花蒿转录因子 TLR1、TLR2、TLR3 的编码序列 (Coding sequence, CDS)、蛋白质序列及注释文件,并在茶树基因组 TPIA 数据库 (<http://tpia.teaplant.org>) 中利用本地 BLAST 在“舒茶早”基因组数据库中获取茶树 *CsTLRs* 基因的 CDS 和蛋白质序列,剔除长度小于 300 bp 的序列与重复序列,最终获得 25 条符合条件的编码序列用于后续研究。

1.2 密码子相关参数分析

利用 CodonW、EMBOSS (<https://www.bioinformatics.nl/emboss-explorer/>) 和生信云网站 (<http://112.86.217.82:9919>) 等对 25 条 CDS 序列进行分

析,获取有效密码子数 (Effective number of codon, ENC)、各位点 G+C 含量 [GC_1 (密码子第 1 位碱基 G+C 含量); GC_2 (密码子第 2 位碱基 G+C 含量); GC_3 (密码子第 3 位碱基 G+C 含量); GC_{all} (总 G+C 含量)]、密码子第 3 位碱基 A、T、C、G 含量 (A_3 、 T_3 、 C_3 、 G_3)、相对同义密码子使用度 (Relative synonymous codon usage, $RSCU$)、密码子适应指数 (Codon adaptation index, CAI)、最优密码子使用频率 (Frequency of optimal codons, FOP) 等数据。使用 SPSS 软件对茶树 $CsTLRs$ 基因组密码子偏好性参数进行相关性分析。

1.3 中性绘图分析

计算各密码子 GC_1 和 GC_2 的平均值,定义为 GC_{12} ,然后以 GC_3 为横坐标, GC_{12} 为纵坐标,绘制散点图并分析它们之间的相关性。如果 GC_3 和 GC_{12} 之间具有显著的相关性,说明密码子在 3 个位置上的碱基组成差异较小,基因突变是主要的影响因素;若两者之间没有显著相关性,则说明碱基 G+C 组成高度保守,密码子使用的偏好性更多地受到自然选择的影响^[20]。

1.4 奇偶偏好性分析 (PR2-plot)

奇偶偏好性分析用于揭示基因中 4 种碱基含量之间的相互关系,从而深入分析基因组中密码子使用偏好的影响因素。以 $G_3/(G_3+C_3)$ 为横坐标, $A_3/(A_3+T_3)$ 为纵坐标绘制 PR2-plot 图。中心点表示碱基数量 $A=T$ 、 $C=G$,证明在此条件下密码子没有明显的使用偏好,仅受到突变的影响。偏离中心点的其他位置则反映了密码子使用的偏倚程度与方向^[21]。通过分析点的偏离情况,可以推断基因中第 3 位密码子嘌呤 (A 和 G) 与嘧啶 (C 和 T) 之间的相对含量及其可能存在的关系。

1.5 基于 GC_3 的有效密码子数分析 (ENC -plot)

以 GC_3 为横坐标, ENC 为纵坐标,绘制基因散点图,同时绘制标准曲线,标准曲线绘制公式参照文献 [22]。分析基因点在图中相对于标准曲线的位置,若基因点接近标准曲线,表明突变可能主导密码子使用模式;若基因点远离标准曲线,则表明自然选择对密码子使用偏好产生更大影响。同时为纠正 ENC -plot 分析造成的误差,需要结合 $ENC_{比值}$ 频数对误差进行量化分析。 $ENC_{比值} = (ENC_{理论} - ENC_{实际}) / ENC_{理论}$,其中 $ENC_{理论}$ 为密码子偏好性仅由碱基组成决定的预期有效密码子数, $ENC_{实际}$ 为密码子实际使用频率得出的有效密码子数,当 $ENC_{比值}$ 的绝对值大于 0.05,表明偏

好性受自然选择影响更大。

1.6 相对同义密码子使用度 ($RSCU$) 分析

$RSCU$ 的计算去除了氨基酸组成对密码子使用偏好性的影响,可用于衡量基因组中密码子使用的偏好。如果 $RSCU$ 等于 1.00,则密码子使用没有偏好性; $RSCU$ 大于 1.00 表明该密码子的使用频率较高,反之亦然^[23]。

1.7 最优密码子分析

按 ENC 升序排列基因,两端各取 3 个基因,分别构建高偏性库和低偏性库。计算偏性库之间的 $RSCU$ 差值 ($\Delta RSCU$),以 $\Delta RSCU$ 大于 0.08 的密码子为具有高表达潜力的优越密码子, $RSCU$ 大于 1.00 的密码子为高频密码子。当同时满足 $\Delta RSCU$ 大于 0.08 和 $RSCU$ 大于 1.00,该密码子被定义为最优密码子,具有高偏好性和高表达的潜力^[24]。

1.8 系统聚类分析

通过 NCBI 数据库获取与植物毛状体相关研究物种的 CDS,包括番茄 (*Solanum lycopersicum*)、烟草 (*Nicotiana tabacum*)、杨树 (*Populus*)、黄瓜 (*Cucumis sativus*)、拟南芥 (*Arabidopsis thaliana*)、棉花 (*Gossypium hirsutum*)、大豆 (*Glycine max*) 和水稻 (*Oryza sativa*)。在 ClustalW 网站 (<https://www.genome.jp/tools-bin/clustalw>) 采用邻接法 (Neighbor-joining) 构建系统进化树,并通过 EVOLVIEW (<https://evolgenius.info/evolview-v2/#login>) 对进化树进行可视化分析。同时以各物种 $RSCU$ 为变量,使用 SPSS 19.0 离差平方和法,设定欧式平方距离并进行偏好性聚类分析。

1.9 外源受体分析

通过 EMBOSS 网站获取茶树 $CsTLRs$ 基因的密码子使用频率,在 Codon Usage Database 网站 (<https://www.kazusa.or.jp/codon/>) 获取烟草 (*Nicotiana tabacum*)、番茄 (*Solanum lycopersicum*)、拟南芥 (*Arabidopsis thaliana*)、大肠杆菌 (*Escherichia coli*)、酿酒酵母 (*Saccharomyces cerevisiae*)、根癌农杆菌 (*Agrobacterium tumefaciens*) 的基因组密码子使用频率。计算茶树与各物种密码子使用频率的比值进而选择合适的外源受体。比值为 0.50~2.00 表示密码子偏性差异较小,适合作为异源表达受体^[25]。

2 结果与分析

2.1 茶树 $CsTLRs$ 基因密码子相关参数分析

2.1.1 密码子的碱基组成分析 对密码子的碱基组

成分析发现,茶树 *CsTLRs* 基因的总 G+C 含量 (GC_{all}) 为 39.20%~56.82%, 平均值为 44.94%; 密码子第 1 位碱基 G+C 含量 (GC_1) 为 38.83%~53.18%, 平均值为 46.82%; 密码子第 2 位碱基 G+C 含量 (GC_2) 为 32.73%~54.04%, 平均值为 42.20%; 密码子第 3 位碱基 G+C 含量 (GC_3) 为 36.68%~63.64%, 平均值为 45.81% (表 1)。以上分析结果表明,密码子组成更偏向碱基 A/U, 密码子不同位置碱基 G+C 含量不同, 其中 GC_1 含量平均值最高, GC_2 含量平均值最低, 各位点碱基组成不存在明显偏好性。

2.1.2 有效密码子数 (*ENC*)、密码子适应指数 (*CAI*) 及最优密码子使用频率 (*FOP*) 分析 *ENC* 反映了一个基因整体上的密码子使用偏好性, 其范围通常为 20.00~61.00。*ENC* 与密码子使用偏好性之间呈负相

关, *ENC* 越接近 20.00, 表示密码子偏好性越强; 越接近 61.00, 则表明密码子偏好性越弱^[26]。茶树 *CsTLRs* 基因 *ENC* 范围为 41.69~61.00, 平均值 53.06, 基因 *ENC* 都大于 40.00, 这表明茶树 *CsTLRs* 基因的密码子使用偏好性较弱 (表 1)。

CAI 是用来衡量某个基因所使用的密码子与该物种高表达基因密码子使用模式之间的相适度^[27], 其取值范围为 0~1.00, 随着 *CAI* 的增大, 基因的密码子使用偏好性趋向于物种特有的高表达基因模式。茶树 *CsTLRs* 基因组 *CAI* 范围为 0.13~0.22, 平均值 0.18 (表 1), 表明该基因密码子适应性较弱。*FOP* 指最优密码子在同义密码子中的占比, 其数值越大, 密码子使用偏好性越强^[26]。茶树 *CsTLRs* 基因 *FOP* 范围为 0.33~0.45, 平均值 0.39, 偏好性较弱。

表 1 茶树 *CsTLRs* 基因密码子 G+C 含量及相关偏好性参数

Table 1 The G+C content and related codon bias parameters of *CsTLRs* genes in tea plant

基因编号	G+C 含量 (%)				<i>ENC</i>	<i>CAI</i>	<i>FOP</i>
	GC_1	GC_2	GC_3	GC_{all}			
CSS0027686	42.71	38.19	36.68	39.20	49.72	0.16	0.33
CSS0033811	42.35	43.37	39.80	41.84	41.69	0.15	0.34
CSS0028457	41.58	40.10	44.06	41.91	53.73	0.17	0.39
CSS0026328	42.94	40.49	42.94	42.13	54.56	0.13	0.34
CSS0016632	46.78	43.39	40.34	43.50	54.55	0.17	0.37
CSS0023014	46.78	43.39	40.68	43.62	54.33	0.17	0.37
CSS0028845	49.20	44.37	49.20	47.59	53.83	0.19	0.40
CSS0034420	47.81	40.35	50.00	46.05	61.00	0.17	0.37
CSS0050347	52.78	54.04	63.64	56.82	51.81	0.16	0.41
CSS0050180	38.83	44.47	49.48	44.26	54.55	0.18	0.41
CSS0023113	46.84	42.28	48.95	46.02	56.08	0.18	0.40
CSS0027976	40.49	46.93	51.22	46.21	54.09	0.16	0.39
CSS0049096	45.46	39.14	53.31	45.97	58.86	0.20	0.42
CSS0046429	46.19	39.45	52.83	46.16	59.23	0.20	0.42
CSS0007159	44.60	32.73	45.32	40.89	54.35	0.20	0.38
CSS0032956	49.15	39.59	39.93	42.89	49.91	0.18	0.38
CSS0005060	48.81	39.25	41.30	43.12	50.95	0.18	0.39
CSS0018453	49.49	40.27	39.93	43.23	51.99	0.18	0.39
CSS0035631	49.75	45.27	41.29	45.44	47.12	0.17	0.39
CSS0040328	48.16	46.94	44.49	46.53	57.02	0.20	0.45
CSS0014780	49.83	45.05	57.68	50.85	52.16	0.22	0.45
CSS0026012	53.18	43.18	51.36	49.24	49.95	0.20	0.44
CSS0048639	47.42	39.21	42.86	43.16	49.52	0.17	0.33
CSS0041393	49.47	38.95	37.37	41.93	54.90	0.18	0.40
CSS0022757	49.78	44.59	40.69	45.02	50.66	0.15	0.37

GC_1 : 密码子第 1 位碱基 G+C 含量; GC_2 : 密码子第 2 位碱基 G+C 含量; GC_3 : 密码子第 3 位碱基 G+C 含量; GC_{all} : 总 G+C 含量; *ENC*: 有效密码子数; *CAI*: 密码子适应指数; *FOP*: 最优密码子使用频率。

2.1.3 密码子偏好性参数相关性分析 对茶树 *CsTLRs* 基因密码子偏好性参数进行相关性分析(表 2)。 GC_{all} 分别与 GC_1 、 GC_2 、 GC_3 呈极显著正相关($P < 0.01$)。 GC_2 和 GC_3 呈显著相关($P < 0.05$)， GC_1 与

GC_2 、 GC_3 无显著相关性,表明基因密码子第 2 位和第 3 位碱基组成有相似性,且与第 1 位碱基组成有较大差异。 GC_3 与 ENC 没有显著相关性,表明第 3 位碱基对密码子使用偏好性的影响较小。

表 2 茶树 *CsTLRs* 基因密码子各参数相关系数

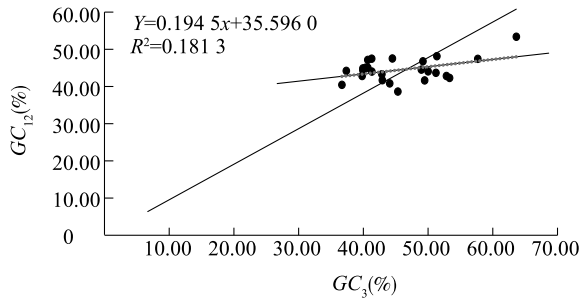
Table 2 Correlation coefficients between the codon usage parameters of *CsTLRs* genes in tea plant

项目	GC_1	GC_2	GC_3	GC_{all}	ENC	CAI
GC_1	1.000					
GC_2	0.254	1.000				
GC_3	0.177	0.483 *	1.000			
GC_{all}	0.544 **	0.759 **	0.859 **	1.000		
ENC	-0.081	-0.143	0.358	0.140	1.000	
CAI	0.312	-0.161	0.381	0.282	0.331	1.000

GC_1 : 密码子第 1 位碱基 G+C 含量; GC_2 : 密码子第 2 位碱基 G+C 含量; GC_3 : 密码子第 3 位碱基 G+C 含量; GC_{all} : 总 G+C 含量; ENC : 有效密码子数; CAI : 密码子适应指数。 * 表示显著相关($P < 0.05$); ** 表示极显著相关($P < 0.01$)。

2.2 中性绘图分析

图 1 显示, GC_{12} 含量为 38.67%~53.41%, GC_3 含量为 36.68%~63.64%, GC_{12} 和 GC_3 回归系数为 0.194 5, R^2 为 0.181 3。表明 GC_{12} 和 GC_3 之间无明显相关性, 碱基 G+C 组成高度保守, 基因突变贡献率占 19.45%, 自然选择贡献率占 80.55%, 则茶树 *CsTLRs* 基因的密码子使用偏好性更易受自然选择的影响。



GC_3 : 密码子第 3 位碱基 G+C 含量; GC_{12} : 密码子第 1 位碱基 G+C 含量 (GC_1) 和第 2 位碱基 G+C 含量 (GC_2) 的平均值。

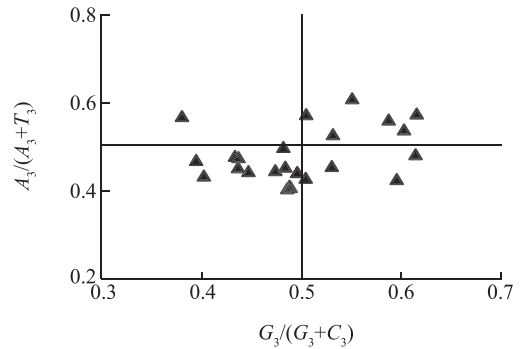
图 1 茶树 *CsTLRs* 基因密码子中性绘图分析

Fig.1 Neutrality plot analysis of codon usage in *CsTLRs* genes of tea plant

2.3 奇偶偏好性分析 (PR2-plot)

图 2 显示, 茶树 *CsTLRs* 基因散点分布不均匀, 表明茶树基因密码子偏好性受突变影响小, 进一步证实 *CsTLRs* 基因密码子偏好性主要受自然选择的影响。横坐标 $G_3/(G_3+C_3)$ 中比值大于 0.5 的散点

占比 40%, 表示碱基 C 使用频率略大; 纵坐标 $A_3/(A_3+T_3)$ 中比值大于 0.5 的散点占比 28%, 表示碱基 T 的使用频率明显高于碱基 A。



横坐标 $G_3/(G_3+C_3)$: 密码子第 3 位碱基 G 的含量与第 3 位碱基 G+C 含量的比值; 纵坐标 $A_3/(A_3+T_3)$: 密码子第 3 位碱基 A 的含量与第 3 位碱基 A+T 含量的比值。

图 2 茶树 *CsTLRs* 基因密码子奇偶偏好性分析 (PR2-plot)

Fig.2 Parity rule 2 plot (PR2-plot) analysis of codon usage bias in *CsTLRs* genes of tea plant

2.4 基于 GC_3 的有效密码子数分析 (ENC-plot)

如图 3 所示, 基因散点不同程度地偏离标准曲线, 表明茶树 *CsTLRs* 基因密码子偏好性主要受自然选择影响。96% 的基因散点偏离于标准曲线下方, 则基因偏好使用的特定的密码子表现出更低的 ENC 。为进一步确定中性突变和自然选择的影响程度, 结合 $ENC_{比值}$ 对差异进行量化分析(表 3), 发现只有 5 个基因 $ENC_{比值}$ 的绝对值 ≤ 0.05 , 80%

基因 $ENC_{\text{比值}}$ 的绝对值 > 0.05 , 表明茶树 *CsTLRs* 基因的密码子偏好性受自然选择影响更大, 证实了中性绘图和 PR2-plot 的分析结果。

2.5 相对同义密码子使用度 (RSCU) 分析

密码子 *RSCU* 大于 1.00 称为高频密码子, 图 4 显示, 茶树 *CsTLRs* 基因中有 29 个高频密码子, 其中密码子 AGA 的 *RSCU* 最大, 为 2.23。高频密码子中 U 结尾的密码子占 44.83%; A 结尾的密码子占 27.59%; G 结尾的密码子占 13.79%; C 结尾的密码子占 13.79%, 证明茶树 *CsTLRs* 基因偏好使用以 A/U 结尾的密码子。

表 3 茶树 *CsTLRs* 基因密码子 $ENC_{\text{比值}}$ 分布

Table 3 The distribution of the effective number of codons (ENC) ratio of *CsTLRs* genes of tea plant

范围	$ENC_{\text{比值}}$ (基因)	频数	频率 (%)
$[-0.05, 0.05]$	-0.01 (CSS0034420); 0.02 (CSS0049096, CSS0046429); 0.04 (CSS0041393); 0.05 (CSS0040328)	5	20.0
$(0.05, +\infty)$	0.06 (CSS0016632); 0.07 (CSS0023014, CSS0023113); 0.08 (CSS0026328, CSS0050347); 0.09 (CSS0007159); 0.10 (CSS0028457, CSS0050180); 0.11 (CSS0028845, CSS0027976, CSS0018453); 0.12 (CSS0027686, CSS0014780); 0.13 (CSS0005060, CSS0022757); 0.14 (CSS0032956); 0.16 (CSS0048639); 0.17 (CSS0026012); 0.20 (CSS0035631); 0.28 (CSS0033811)	20	80.0

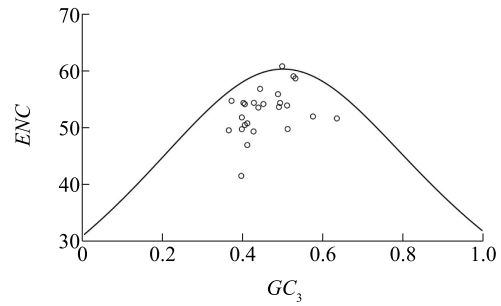
$ENC_{\text{比值}} = (ENC_{\text{理论}} - ENC_{\text{实际}}) / ENC_{\text{理论}}$ 。其中 $ENC_{\text{理论}}$ 为密码子偏好性仅由碱基组成决定的预期有效密码子数, $ENC_{\text{实际}}$ 为密码子实际使用频率得出的有效密码子数。

Ala	GCU(1.34)	GCC(1.03)	GCA(1.02)	GCG(0.61)	0	0
Cys	UGU(1.04)	UGC(0.96)	0	0	0	0
Asp	GAU(1.21)	GAC(0.79)	0	0	0	0
Glu	GAA(1.07)	GAG(0.93)	0	0	0	0
Phe	UUC(1.13)	UUU(0.87)	0	0	0	0
Gly	GGA(1.21)	GGU(1.08)	GGC(0.90)	GGG(0.81)	0	0
His	CAU(1.11)	CAC(0.89)	0	0	0	0
Ile	AUU(1.31)	AUC(0.97)	AUA(0.73)	0	0	0
Met	AUG(1.00)	0	0	0	0	0
Lys	AAG(1.01)	AAA(0.99)	0	0	0	0
Leu	UUG(1.51)	CUC(1.26)	CUU(1.18)	CUA(0.82)	CUG(0.68)	UUA(0.56)
Asn	AAU(1.02)	AAC(0.98)	0	0	0	0
Pro	CCA(1.60)	CCU(1.06)	CCG(0.89)	CCC(0.45)	0	0
Gln	CAA(1.25)	CAG(0.75)	0	0	0	0
Arg	AGA(2.23)	AGG(1.49)	CGG(0.66)	CGA(0.65)	CGC(0.62)	CGU(0.35)
Ser	UCU(1.19)	UCA(1.11)	AGU(1.01)	UCC(0.99)	AGC(0.87)	UCG(0.82)
Thr	ACC(1.35)	ACA(1.22)	ACU(0.99)	ACG(0.45)	0	0
Val	GUG(1.32)	GUU(1.29)	GUC(0.89)	GUA(0.50)	0	0
Trp	UGG(1.00)	0	0	0	0	0
Tyr	UAU(1.02)	UAC(0.98)	0	0	0	0

Ala: 丙氨酸; Cys: 半胱氨酸; Asp: 天冬氨酸; Glu: 谷氨酸; Phe: 苯丙氨酸; Gly: 甘氨酸; His: 组氨酸; Ile: 异亮氨酸; Met: 甲硫氨酸; Lys: 赖氨酸; Leu: 亮氨酸; Asn: 天冬酰胺; Pro: 脯氨酸; Gln: 谷氨酰胺; Arg: 精氨酸; Ser: 丝氨酸; Thr: 苏氨酸; Val: 缬氨酸; Trp: 色氨酸; Tyr: 酪氨酸。

图 4 茶树 *CsTLRs* 基因相对同义密码子使用度 (RSCU)

Fig.4 The relative synonymous codon usage (RSCU) of *CsTLRs* genes in tea plant



GC_3 : 密码子第 3 位碱基 G+C 含量; ENC : 有效密码子数。

图 3 茶树 *CsTLRs* 基因密码子 ENC -plot 绘图分析

Fig.3 ENC -plot analysis of codon usage bias in *CsTLRs* genes of tea plant

2.6 最优密码子分析

对茶树 *CsTLRs* 基因进行最优密码子分析,结果(表 4)显示,剔除终止密码子,通过 *ENC* 构建高偏性库和低偏性库,计算偏性库之间的 *RSCU* 差值 ($\Delta RSCU$),筛选出同时满足 $\Delta RSCU > 0.08$, $RSCU > 1.00$ 的密码子作为最优密码子,最终筛选出 20 个最

优密码子: *GCA*、*GCC*、*UGU*、*GAU*、*GAA*、*GGA*、*CAU*、*CUC*、*CUU*、*AAU*、*CCA*、*CCU*、*CAA*、*AGA*、*AGG*、*UCA*、*UCU*、*ACA*、*GUU*、*UAU*。其中 U 结尾的密码子占 45%,A 结尾的密码子占 40%,C 结尾的密码子占 10%,G 结尾的密码子占 5%,进一步证明了茶树 *CsTLRs* 基因偏好使用以 A/U 结尾的密码子。

表 4 茶树 *CsTLRs* 基因最优密码子分析

Table 4 Optimal codon analysis of *CsTLRs* genes in tea plant

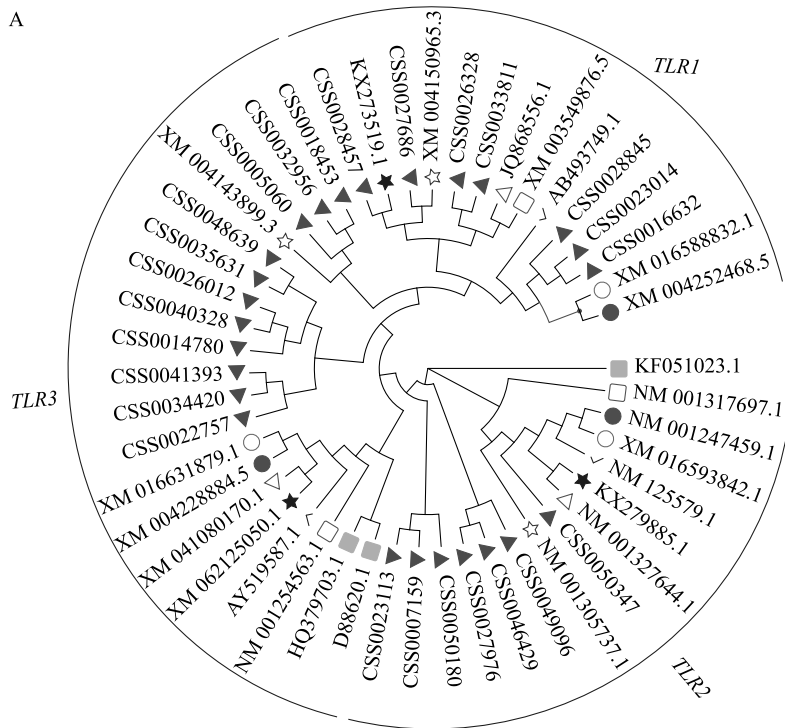
氨基酸	密码子	<i>RSCU</i>	高表达基因 <i>RSCU</i>	低表达基因 <i>RSCU</i>	$\Delta RSCU$	氨基酸	密码子	<i>RSCU</i>	高表达基因 <i>RSCU</i>	低表达基因 <i>RSCU</i>	$\Delta RSCU$
Ala	GCA	1.02	0.80	0.61	0.19	Pro	CCA	1.60	2.00	1.36	0.64
	GCC	1.03	1.49	1.15	0.34		CCC	0.45	0.40	0.47	-0.07
	GCG	0.61	0.23	0.68	-0.45		CCG	0.89	0.30	1.36	-1.06
	GCU	1.34	1.49	1.55	-0.06		CCU	1.06	1.30	0.81	0.49
Cys	UGC	0.96	0.57	0.97	-0.40	Gln	CAA	1.25	1.52	1.18	0.34
	UGU	1.04	1.43	1.03	0.40		CAG	0.75	0.48	0.82	-0.34
Asp	GAC	0.79	0.43	0.87	-0.44	Arg	AGA	2.23	2.79	1.06	1.73
	GAU	1.21	1.57	1.13	0.44		AGG	1.49	2.23	1.50	0.73
Glu	GAA	1.07	1.11	0.99	0.12	CGA	0.65	0.42	0.35	0.07	
	GAG	0.93	0.89	1.01	-0.12	CGC	0.62	0.42	0.88	-0.46	
Phe	UUC	1.13	0.60	1.18	-0.58	CGG	0.66	0.14	1.41	-1.27	
	UUU	0.87	1.40	0.82	0.58	CGU	0.35	0	0.79	-0.79	
Gly	GGA	1.21	1.31	1.03	0.28	Ser	AGC	0.87	0.71	0.68	0.03
	GGC	0.90	0.62	0.94	-0.32		AGU	1.01	1.12	1.09	0.03
	GGG	0.81	1.00	0.83	0.17		UCA	1.11	1.53	0.86	0.67
	GGU	1.08	1.08	1.20	-0.12		UCC	1.00	0.71	1.32	-0.61
His	CAC	0.89	0.57	1.18	-0.61	UCG	0.82	0.10	1.04	-0.94	
	CAU	1.11	1.43	0.82	0.61	UCU	1.19	1.83	1.01	0.82	
Ile	AUA	0.73	0.89	0.66	0.23	Thr	ACA	1.22	2.05	0.89	1.16
	AUC	0.97	0.97	0.92	0.05		ACC	1.35	1.12	1.58	-0.46
	AUU	1.31	1.14	1.42	-0.28		ACG	0.45	0.09	0.62	-0.53
Lys	AAA	0.99	0.86	0.87	-0.01	ACU	0.99	0.74	0.89	-0.15	
	AAG	1.01	1.14	1.13	0.01	Val	GUA	0.50	0.46	0.43	0.03
Leu	CUA	0.82	0.68	0.62	0.06		GUC	0.89	1.08	0.94	0.14
	CUC	1.26	1.44	1.27	0.17		GUG	1.32	0.77	1.57	-0.80
	CUG	0.68	0.76	0.65	0.11	GUU	1.29	1.69	1.06	0.63	
	CUU	1.18	1.27	1.12	0.15	Trp	UGG	1.00	1.00	1.00	0
	UUA	0.56	0.42	0.67	-0.25		Tyr	UAC	0.98	0.77	1.13
UUG	1.51	1.44	1.67	-0.23	UAU	1.02		1.23	0.87	0.36	
Met	AUG	1.00	1.00	1.00	0						
Asn	AAC	0.98	0.90	0.99	-0.09						
	AAU	1.02	1.10	1.01	0.09						

Ala: 丙氨酸; Cys: 半胱氨酸; Asp: 天冬氨酸; Glu: 谷氨酸; Phe: 苯丙氨酸; Gly: 甘氨酸; His: 组氨酸; Ile: 异亮氨酸; Met: 甲硫氨酸; Lys: 赖氨酸; Leu: 亮氨酸; Asn: 天冬酰胺; Pro: 脯氨酸; Gln: 谷氨酰胺; Arg: 精氨酸; Ser: 丝氨酸; Thr: 苏氨酸; Val: 缬氨酸; Trp: 色氨酸; Tyr: 酪氨酸。RSCU: 相对同义密码子使用度; 高表达基因 RSCU: 有效密码子数升序排列前 3 位基因的相对同义密码子使用度; 低表达基因 RSCU: 有效密码子数升序排列后 3 位基因的相对同义密码子使用度; $\Delta RSCU$: 高表达基因 RSCU-低表达基因 RSCU。粗体: 同时满足 $\Delta RSCU > 0.08$ 、 $RSCU > 1.00$ 的密码子。

2.7 基于 CDS 与 *RSCU* 的聚类分析

CDS 聚类是基于核苷酸序列的相似性 (如突变、选择压力) 构建进化关系,能够反映物种或基因的系统发育历史,而 *RSCU* 聚类是基于同义密码子使用频率的差异,揭示翻译选择压力、基因表达优化等,主要反映功能适应性^[27-29]。CDS 聚类分析结果表明茶树与杨树、黄瓜在遗传距离近分支上,说明茶树 *CsTLRs* 基因与杨树和黄瓜的亲缘关系较近 (图 5A)。基于 *TLRs* 基因 *RSCU* 的聚类分析结果显示,茶树同样与杨树和黄瓜聚类到一起,说明在 *TLRs* 基因上茶树与杨树、黄瓜有较相似的密码子使用偏好

(图 5B),同时番茄、烟草、水稻在 *TLRs* 基因上的 2 种聚类方式也高度相似,可见本研究大部分 *TLRs* 基因的 CDS 聚类与 *RSCU* 聚类结果相近,但在拟南芥、棉花和大豆上出现偏差,推测差异出现与基因的突变、转移等相关,导致基因密码子的使用偏好改变^[30],同时 *RSCU* 聚类分析可能未能全面捕捉所有影响密码子偏好性的要素,使结果无法准确反映物种进化关系^[31]。以上结果表明,密码子偏好性的聚类与物种的进化关系具有一定的相似性,但其分析结果还会受物种中基因的重组、倍增等突变的影响,*RSCU* 聚类可作为传统进化分析的良好补充。



▲ 茶树(*Camellia sinensis*); ★ 杨树(*Populus*); ☆ 黄瓜(*Cucumis sativus*); △ 棉花(*Gossypium hirsutum*); ● 番茄(*Solanum lycopersicum*); ○ 烟草(*Nicotiana tabacum*); ■ 水稻(*Oryza sativa*); □ 大豆(*Glycine max*); ✓ 拟南芥(*Arabidopsis thaliana*)

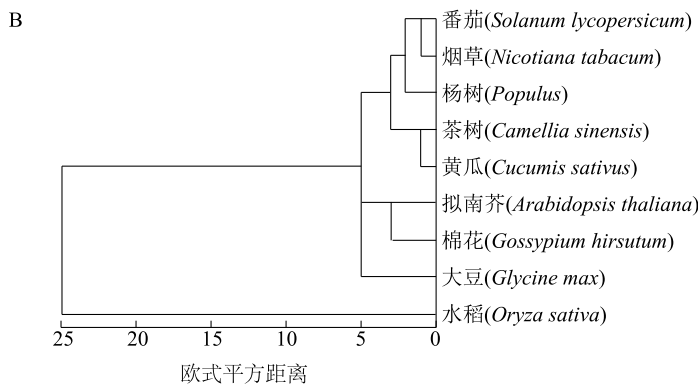


图 5 基于 *TLRs* 基因编码序列 (CDS) (A) 和相对同义密码子使用度 (*RSCU*) (B) 的聚类分析

Fig.5 Cluster analysis based on coding sequences (CDS) (A) and the relative synonymous codon usage (*RSCU*) (B) of *TLRs* genes

2.8 外源表达受体选择

将茶树 *CsTLRs* 基因密码子使用频率与烟草、番茄、拟南芥、大肠杆菌、酿酒酵母、根癌农杆菌基因组密码子使用频率相比较,比值在0.50~2.00 的密码子数量分别为 57 个、56 个、60 个、44 个、52 个、31

个(表 5),说明 3 种模式植物(拟南芥、烟草和番茄)均可作为茶树 *CsTLRs* 基因合适的遗传转化受体,若选择微生物异源表达 *CsTLRs* 基因,酿酒酵母优于大肠杆菌和根癌农杆菌表达系统。

表 5 茶树 *CsTLRs* 基因与 6 种模式生物基因组密码子使用频率比较

Table 5 Comparison of codon usage frequency of tea plant *CsTLRs* genes and the genomes of six model organisms

密码子	<i>CsTLRs</i> 使用频率 (%)	频率比值					
		<i>Cs/Nt</i>	<i>Cs/SI</i>	<i>Cs/At</i>	<i>Cs/Ec</i>	<i>Cs/Sc</i>	<i>Cs/Ag</i>
AAA	30.34	0.93	0.98	0.98	0.82	0.72	2.51
AAC	31.64	1.77	1.83	1.51	1.56	1.28	1.89
AAG	30.88	0.92	1.00	0.94	2.02	1.00	1.16
AAU	32.72	1.17	1.07	1.47	1.12	0.92	2.52
ACA	16.90	0.97	0.94	1.08	1.12	0.95	2.97
ACC	18.64	1.92	2.17	1.81	0.99	1.47	0.75
ACG	6.18	1.37	1.34	0.80	0.45	0.77	0.31
ACU	13.65	0.67	0.69	0.78	1.04	0.67	4.40
AGA	22.21	1.39	1.35	1.17	3.13	1.04	8.54
AGC	13.98	1.40	1.50	1.24	0.98	1.43	0.96
AGG	14.84	1.22	1.25	1.35	3.71	1.61	3.45
AGU	16.25	1.22	1.07	1.16	1.23	1.14	5.42
AUA	13.00	0.93	0.93	1.03	0.98	0.73	3.94
AUC	17.23	1.55	1.23	0.93	0.89	1.00	0.42
AUG	21.45	0.86	0.87	0.88	0.91	1.03	0.84
AUU	23.29	0.84	0.83	1.08	0.79	0.77	1.79
CAA	20.59	0.99	0.98	1.06	1.43	0.75	3.49
CAC	11.38	1.31	1.46	1.31	1.56	1.46	1.29
CAG	12.24	0.82	0.87	0.81	0.46	1.01	0.48
CAU	14.09	1.05	0.91	1.02	1.14	1.04	1.22
CCA	21.45	1.08	1.12	1.33	2.36	1.17	5.11
CCC	5.96	0.90	1.05	1.12	0.96	0.88	0.47
CCG	11.92	2.38	2.59	1.39	0.82	2.25	0.46
CCU	14.19	0.76	0.74	0.76	1.49	1.05	2.73
CGA	6.50	1.23	1.20	1.03	1.35	2.17	1.81
CGC	6.18	1.58	1.99	1.63	0.44	2.38	0.20
CGG	6.61	1.79	2.13	1.35	0.84	3.89	0.51
CGU	3.47	0.46	0.50	0.39	0.22	0.54	0.31
CUA	15.28	1.63	1.53	1.54	2.73	1.14	8.04
CUC	23.62	1.92	2.11	1.47	2.49	4.37	0.89
CUG	12.68	1.24	1.21	1.29	0.34	1.21	0.30
CUU	21.99	0.92	0.88	0.91	1.52	1.79	1.22

续表5 Continued5

密码子	<i>CsTLRs</i> 使用频率(%)	频率比值					
		<i>Cs/Nt</i>	<i>Cs/Sl</i>	<i>Cs/At</i>	<i>Cs/Ec</i>	<i>Cs/Sc</i>	<i>Cs/Ag</i>
GAA	27.19	0.76	0.78	0.79	0.77	0.60	0.84
GAC	18.85	1.12	1.26	1.10	1.05	0.93	0.69
GAG	23.84	0.81	0.90	0.74	1.23	1.24	0.95
GAU	28.71	0.78	0.73	0.78	0.85	0.76	1.03
GCA	12.24	0.53	0.55	0.70	0.53	0.76	0.81
GCC	12.35	0.99	1.22	1.20	0.57	0.98	0.24
GCG	7.37	1.27	1.42	0.82	0.35	1.19	0.20
GCU	16.04	0.51	0.52	0.57	0.85	0.76	1.38
GGA	18.64	0.80	0.73	0.77	1.37	1.71	2.59
GGC	13.87	1.24	1.43	1.51	0.67	1.42	0.28
GGG	12.46	1.19	1.15	1.22	1.01	2.08	1.45
GGU	16.58	0.74	0.69	0.75	0.70	0.69	0.95
GUA	5.63	0.49	0.50	0.57	0.43	0.48	1.66
GUC	10.08	0.91	1.00	0.79	0.77	0.85	0.33
GUG	14.84	0.89	0.93	0.85	0.75	1.37	0.58
GUU	14.52	0.54	0.52	0.53	0.67	0.66	1.08
UAC	9.97	0.74	0.80	0.73	0.85	0.67	1.15
UAU	10.40	0.58	0.56	0.71	0.48	0.55	0.73
UCA	17.88	1.02	0.86	0.98	1.36	0.96	4.47
UCC	16.04	1.57	1.62	1.43	1.65	1.13	0.94
UCG	13.11	2.47	2.34	1.41	1.60	1.52	0.76
UCU	19.07	0.95	0.90	0.76	1.46	0.81	5.15
UGC	11.05	1.53	1.65	1.53	2.01	2.30	1.70
UGG	20.48	1.68	1.52	1.64	1.53	1.97	1.64
UGU	12.03	1.23	1.11	1.15	2.04	1.48	6.68
UUA	10.40	0.78	0.72	0.82	0.60	0.40	8.67
UUC	20.26	1.13	1.16	0.98	1.46	1.10	0.70
UUG	28.28	1.27	1.17	1.35	2.19	1.04	3.14
UUU	15.60	0.62	0.60	0.72	0.64	0.60	1.34

Cs: 茶树 *TLRs* 基因密码子使用频率; *Nt*: 烟草基因组密码子使用频率; *Sl*: 番茄基因组密码子使用频率; *At*: 拟南芥基因组密码子使用频率; *Ec*: 大肠杆菌基因组密码子使用频率; *Sc*: 酿酒酵母基因组密码子使用频率; *Ag*: 根瘤农杆菌基因组密码子使用频率。

3 讨论

密码子偏好性是基因表达调控和进化的重要机制之一,它影响着转录和翻译的效率,进而影响蛋白质的表达水平^[32]。在生物信息学和基因组学研究中,密码子偏好性研究有助于基因功能的预测和基因组注释。目前,对茶树基因密码子偏好性的研究较少,但根据已有的研究结果可以得出茶树基因组

内部密码子使用偏好性是复杂和多样的^[33]。基因表达、mRNA 二级结构、基因长度、tRNA 丰度等多种因素都会影响密码子偏好性,即使是紧密相关的基因也可能根据自身的特性和环境适应性而采用不同的密码子使用模式^[28,34-36]。茶树 *CsTLRs* 基因 GC_3 平均值为 45.81%,筛选的 20 个最优密码子中,U 结尾的密码子占 45%,A 结尾的密码子占 40%,C 结尾的密码子占 10%,G 结尾的密码子占 5%,

CsTLRs 基因表现偏好使用以 A/U 碱基结尾的密码子。有研究发现,单子叶植物更偏好 G/C 结尾的密码子,双子叶植物更偏好 A/U 结尾的密码子^[28]。在茶树密码子末尾碱基使用偏好性研究中,前人发现茶树 *CsICE1*^[25]、*CsOMT*^[37]、*CsSAD*^[38]、*CsGPAT*^[39]、*CsLOX*^[40]、*CsWOX*^[41]、*CsSPDS*^[42]、*CsNRTI.I*^[43] 也偏好使用以 A/U 结尾的密码子,*CsActinI*^[44] 中以 A/U 或 G/C 结尾的密码子数量相近,无明显碱基偏好,而 *CsCBFI*^[45] 偏好 G/C 结尾的密码子。上述研究结果表明,大部分茶树基因偏好使用以 A/U 结尾的密码子,这与双子叶植物更偏好 A/U 结尾密码子的观点一致,并与赵洋等^[33] 的研究结果相符。茶树 *CsTLRs* 基因 *ENC*、*CAI* 和 *FOP* 分别为 53.06、0.18 和 0.39,说明茶树 *CsTLRs* 基因密码子使用偏好性较弱,表达水平较低。对茶树 *CsTLRs* 基因 *RSCU* 的分析确定了 29 个高频密码子,其中 AGA 的 *RSCU* 最高,为 2.23。在茶树基因 *CsLOX1*、*CsLOX6*、*CsLOX10*、*CsLOX12* 中使用频率最高的密码子也为 AGA^[40],由此推测茶树基因对 AGA 密码子的使用表现出高度保守性。

目前对密码子偏好性形成受自然选择还是突变压力影响的研究较多,例如自然选择是铁核桃叶绿体基因组密码子偏好性形成的主要原因^[46],而宽的 *AtrNiR* 基因密码子偏好性受突变压力影响更大^[47]。通过中性绘图、PR2-plot、*ENC*-plot 一系列分析发现,自然选择在塑造茶树 *CsTLRs* 基因的密码子偏好性方面发挥主导作用,与在 *CsOMT*^[37]、*CsLOX*^[40]、*CsWOX*^[41]、大理茶叶叶绿体基因组^[48]、黄药大头茶叶叶绿体基因组^[49] 中的研究结果一致,证明自然选择在密码子使用模式中的重要作用,同时基因突变、碱基组成等其他因素也会产生一定影响。通过对茶树 *CsTLRs* 进行 CDS 和 *RSCU* 聚类分析发现,亲缘关系更为紧密的物种之间,可能在密码子使用偏好上呈现出更高的相似度。但受不同进化因素的影响,如基因重组、倍增等突变,2 种聚类方式也存在不同程度的差异,对于一些基因(如 *CsICE1*^[25]、*CsCBFI*^[45]、*RuTTI2-I*^[50]) 的密码子 *RSCU* 聚类无法有效反映物种间的亲缘关系。并且 *RSCU* 聚类是通过分析不同物种间密码子偏好性的差异来构建关系结构,可能在参数计算过程中丢失关键信息而导致偏差^[31],一般认为基于 CDS 构建的聚类能更准确地反映物种间的关系。有研究结果表明,在较小的

分类单元,*RSCU* 聚类提供的信息更加精确,可能是因为在层次上物种间的遗传距离较短,密码子使用偏好受共同祖先的影响更大,因此更能反映出真实的亲缘关系^[51-52]。因此,综合使用 CDS 聚类和 *RSCU* 聚类更能显著地提高分类结果的精确度。对茶树 *CsTLRs* 基因外源表达受体进行分析发现,3 种模式植物(拟南芥、烟草和番茄)均可作为 *CsTLRs* 基因的遗传转化受体;在微生物异源表达中,与大肠杆菌和根瘤农杆菌相比,酿酒酵母是更好的外源表达载体。若要实现基因在外源物种中的高效表达,还需对部分偏好性差异较大的密码子进行改造,茶树 *CsTLRs* 基因的高效遗传转化仍有待进一步探索。综上所述,本研究结果将有助于理解茶树 *CsTLRs* 基因密码子的使用特征,为后续研究该基因调控茶树叶片毛状体的发育提供参考。

4 结论

本研究对茶树 *CsTLRs* 基因的密码子偏好性进行了分析,结果显示,其密码子偏好性较弱,且偏好以 A/U 结尾。在茶树 *CsTLRs* 基因中,共鉴定出 20 个最优密码子,自然选择是 *CsTLRs* 基因密码子偏好性形成的主要原因。茶树 *CsTLRs* 基因与杨树、黄瓜有较近的亲缘关系,且密码子使用偏好性相似。可选择拟南芥、烟草和番茄作为茶树 *CsTLRs* 基因异源表达的遗传转化的植物受体,在微生物中外源表达载体选择酿酒酵母更佳。

参考文献:

- [1] 梁菲菲. 密码子偏性的影响因素及研究意义[J]. 畜牧与饲料科学, 2010, 31(1): 118-119.
- [2] HERSHBERG R, PETROV D A. Selection on codon bias[J]. Annual Review of Genetics, 2008, 42: 287-299.
- [3] 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.
- [4] LIU X Y, LI Y, JI K K, et al. Genome-wide codon usage pattern analysis reveals the correlation between codon usage bias and gene expression in *Cuscuta australis* [J]. Genomics, 2020, 112(4): 2695-2702.
- [5] PARVATHY S T, UDAYASURIYAN V, BHADANA V. Codon usage bias[J]. Molecular Biology Reports, 2022, 49(1): 539-565.
- [6] 孙彬妹, 刘少群, 刘任坚, 等. 茶树茸毛的研究进展[J]. 茶叶通讯, 2018, 45(4): 3-6.
- [7] STRATMANN J W, BEQUETTE C J. Hairless but no longer clue-

- less; understanding glandular trichome development[J]. *Journal of Experimental Botany*, 2016, 67(18): 5285-5287.
- [8] HAUSER M T. Molecular basis of natural variation and environmental control of trichome patterning[J]. *Frontiers in Plant Science*, 2014, 5: 320.
- [9] BANDYOPADHYAY T, GOHAIN B, BHARALEE R, et al. Molecular landscape of *Helopeltis theivora* induced transcriptome and defense gene expression in tea[J]. *Plant Molecular Biology Reporter*, 2015, 33(4): 1042-1057.
- [10] 宋亚康, 张群峰, 张洁, 等. 茶毫氨基酸组成及矿质元素分析[J]. *茶叶科学*, 2017, 37(4): 339-346.
- [11] 郭桂义, 孙慕芳, 陈义, 等. 茶叶茸毛的化学成分测定[J]. *食品科学*, 2011, 32(8): 244-247.
- [12] LIU X Y, ZHOU F, WEN M C, et al. LC-MS and GC-MS based metabolomics analysis revealed the impact of tea trichomes on the chemical and flavor characteristics of white tea[J]. *Food Research International*, 2024, 191: 114740.
- [13] LONG P P, SU S X, WEN M C, et al. An insight into trichomes-deficiency and trichomes-rich black teas by comparative metabolomics; the impact of oxidized trichomes on metabolic profiles and infusion color[J]. *Food Research International*, 2024, 190: 114638.
- [14] LYU Z Y, LI J X, QIU S, et al. The transcription factors TLR1 and TLR2 negatively regulate trichome density and artemisinin levels in *Artemisia annua*[J]. *Journal of Integrative Plant Biology*, 2022, 64(6): 1212-1228.
- [15] DONG B R, XU Z H, WANG X X, et al. TrichomeLess regulator 3 is required for trichome initial and cuticle biosynthesis in *Artemisia annua*[J]. *Molecular Horticulture*, 2024, 4(1): 10.
- [16] LIU Y J, HOU H, JIANG X L, et al. A WD40 repeat protein from *Camellia sinensis* regulates anthocyanin and proanthocyanidin accumulation through the formation of MYB-bHLH-WD40 ternary complexes[J]. *International Journal of Molecular Sciences*, 2018, 19(6): 1686.
- [17] SUN B M, ZHU Z S, LIU R J, et al. TRANSPARENT TESTA GLABRA1 (TTG1) regulates leaf trichome density in tea *Camellia sinensis*[J]. *Nordic Journal of Botany*, 2020, 38(1): 38.
- [18] LI P H, FU J M, XU Y J, et al. CsMYB1 integrates the regulation of trichome development and catechins biosynthesis in tea plant domestication[J]. *New Phytologist*, 2022, 234(3): 902-917.
- [19] 刘任坚, 王玉源, 刘少群, 等. 茶树 *CsbHLH024* 和 *CsbHLH133* 转录因子功能鉴定[J]. *茶叶科学*, 2022, 42(3): 347-357.
- [20] LIU X E. A more accurate relationship between 'effective number of condons' and GC3s under assumptions of no selection[J]. *Computational Biology and Chemistry*, 2013, 42: 35-39.
- [21] 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.
- [22] WRIGHT F. The 'effective number of condons' used in a gene[J]. *Gene*, 1990, 87(1): 23-29.
- [23] 杨国锋, 苏昆龙, 赵怡然, 等. 蒺藜苜蓿叶绿体密码子偏好性分析[J]. *草业学报*, 2015, 24(12): 171-179.
- [24] 胡莎莎, 罗洪, 吴琦, 等. 苦荞叶绿体基因组密码子偏爱性分析[J]. *分子植物育种*, 2016, 14(2): 309-317.
- [25] 时慧, 王玉, 杨路成, 等. 茶树抗寒调控转录因子 ICE1 密码子偏性分析[J]. *园艺学报*, 2012, 39(7): 1341-1352.
- [26] 赵洋, 刘振, 杨培迪, 等. 密码子偏性分析及茶树中密码子偏性研究进展[J]. *茶叶通讯*, 2016, 43(2): 3-7.
- [27] SHARP P M, LI W H. The condon adaptation index; a measure of directional synonymous condon usage bias, and its potential applications[J]. *Nucleic Acids Research*, 1987, 15(3): 1281-1295.
- [28] DURET L. tRNA gene number and condon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes[J]. *Trends in Genetics*, 2000, 16(7): 287-289.
- [29] 张文娟. 基于密码子水平的生物信息学分析及进化研究[D]. 上海: 复旦大学, 2006.
- [30] 严子成, 蒋瑞平, 梁浩伟, 等. 川芎咖啡酸-O-甲基转移酶基因密码子偏好性与进化分析[J]. *应用与环境生物学报*, 2020, 26(4): 894-901.
- [31] 巫伟峰, 陈明杰, 陈发兴. '皇冠李' 苹果酸转运体基因 *ALMT4*、*ALMT9* 和 *iDT* 密码子偏好性分析[J]. *农业生物技术学报*, 2020, 28(1): 42-57.
- [32] BULMER M. The selection-mutation-drift theory of synonymous condon usage[J]. *Genetics*, 1991, 129(3): 897-907.
- [33] 赵洋, 杨阳, 刘振, 等. 茶树密码子用法分析[J]. *茶叶科学*, 2011, 31(4): 319-325.
- [34] MUKHOPADHYAY P, BASAK S, GHOSH T C. Differential selective constraints shaping condon usage pattern of housekeeping and tissue-specific homologous genes of rice and *Arabidopsis*[J]. *DNA Research*, 2008, 15(6): 347-356.
- [35] 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.
- [36] GU W J, ZHOU T, MA J M, et al. The relationship between synonymous condon usage and protein structure in *Escherichia coli* and *Homo sapiens*[J]. *Biosystems*, 2004, 73(2): 89-97.
- [37] 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.
- [38] PAN L L, WANG Y, HU J H, et al. Analysis of condon use features of stearyl-acyl carrier protein desaturase gene in *Camellia sinensis*[J]. *Journal of Theoretical Biology*, 2013, 334: 80-86.
- [39] LI C, PAN L L, WANG Y, et al. Codon bias of the gene for chloroplast glycerol-3-phosphate acyltransferase in *Camellia sinensis* (L.) O. Kuntze[J]. *Biochemical Systematics and Ecology*, 2014, 55: 212-218.
- [40] 周子维, 常笑君, 游芳宁, 等. 茶树脂肪氧合酶(*LOX*) 基因家族成员的分子进化及密码子偏好性分析[J]. *中国农业科技导报*, 2017, 19(12): 43-51.
- [41] 王占军, 吴子琦, 王朝霞, 等. 3个茶树品种 *WOX* 基因家族的进化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.
- [42] YOU E, WANG Y, DING Z T, et al. Codon usage bias analysis for the spermidine synthase gene from *Camellia sinensis* (L.) O. Kuntze[J]. *Genetics and Molecular Research*,2015,14(3):7368-7376.
- [43] 胡振民,万青,李欢,等. 茶树 *CsNRT1.1* 基因密码子使用特性分析[J]. *江苏农业学报*,2019,35(4):896-903.
- [44] 赵洋,刘振,杨培迪,等. 茶树 *CsActin1* 基因密码子偏性分析[J]. *茶叶通讯*,2014,41(4):13-17.
- [45] 郭秀丽,王玉,杨路成,等. 茶树 *CBF1* 基因密码子使用特性分析[J]. *遗传*,2012,34(12):1614-1623.
- [46] 原晓龙,郝佳波,王毅,等. 铁核桃叶绿体基因组密码子偏好性分析[J]. *分子植物育种*,2020,18(20):6671-6677.
- [47] 陈何,王晓,王乐,等. 苋菜 *AtrNiR* 基因密码子偏好性与进化分析[J]. *亚热带农业研究*,2021,17(1):48-56.
- [48] 佟岩,黄荟,王雨华. 森林茶园古茶树大理茶叶叶绿体基因组密码子偏好性及系统发育研究[J]. *茶叶科学*,2023,43(3):297-309.
- [49] 高灿,樊智丰,马长乐. 黄药大头茶叶叶绿体基因组密码子偏好性分析[J]. *西南林业大学学报(自然科学)*,2023,43(5):66-76.
- [50] 冯琛,张云婷,肖婕,等. 黑莓 *RuTT12-1* 基因密码子偏好性分析[J]. *基因组学与应用生物学*,2016,35(8):2133-2144.
- [51] CHRISTIANSON M L. Codon usage patterns distort phylogenies from or of DNA sequences[J]. *American Journal of Botany*,2005,92(8):1221-1233.
- [52] ZHOU H, WANG H, HUANG L F, et al. Heterogeneity in codon usages of sobemovirus genes[J]. *Archives of Virology*,2005,150(8):1591-1605.

(责任编辑:陈海霞)