

◆特邀栏目◆

苦荞 CCT 基因家族生物信息学及表达分析*

石桃雄,赵佳利,田桃,陈庆富,邓娇**

(贵州师范大学荞麦产业技术研究中心,贵州贵阳 550001)

摘要: CCT 转录因子在调控植物花期、生长发育及抗非生物胁迫等方面发挥着重要的功能。本研究以拟南芥 *AtCCT* 基因家族为参考序列,利用本地 BLAST 并结合保守结构域等生物信息学工具,筛选出苦荞 *FtCCT* 基因家族成员,并对其理化性质、染色体分布、基因结构、系统进化及表达水平进行分析。结果显示:从苦荞中共鉴定出 35 个 *FtCCT* 基因,含 1-8 个内含子;编码蛋白有 117-753 个氨基酸残基,等电点为 4.96-9.51,均为亲水性蛋白。染色体定位分析表明,这些基因在 8 条染色体上均有分布。苦荞 *FtCCT* 基因家族含有 10 个保守基序和 5 个保守结构域,且都含有 CCT 保守结构域。系统进化分析表明,苦荞的 *FtCCT* 基因家族与拟南芥一样可分为 3 个亚家族,其中 CMF 亚家族的成员最多。35 个 *FtCCT* 基因在苦荞根、茎、叶和花中的表达水平具有差异性,在叶和花中具有高表达量的成员较多,只有少数的成员在根和茎中高表达。本研究为进一步解析 CCT 基因调控苦荞花期及生长发育奠定基础。

关键词: 苦荞 开花 *FtCCT* 系统进化 表达分析

中图分类号: S513 文献标识码: A 文章编号: 1005-9164(2021)03-0301-09

DOI: 10.13656/j.cnki.gxkx.20210830.007

0 引言

荞麦是一种蓼科(Polygonaceae)荞麦属(*Fagopyrum* Mill.)的双子叶小杂粮作物,甜荞(*Fagopyrum esculentum* Moench)和苦荞(*Fagopyrum tataricum* Gaertn.)是其常见的两个栽培种^[1]。苦荞含有丰富的多酚化合物、类黄酮和不饱和脂肪酸等生

物活性物质,能够降低胆固醇、抑制肿瘤、调节高血压、抗炎和治疗糖尿病等^[2,3]。此外,苦荞还含有丰富的蛋白质、氨基酸、多种维生素和纤维素等成分,是谷类作物中唯一具有七类营养物质的作物^[4]。因此,苦荞是一种很好的药食两用作物,具有非常高的药用价值和营养价值,且具有广泛的环境适应能力,能带来良好的经济效益^[3,5,6]。

收稿日期:2021-05-15

* 国家自然科学基金项目(31860408)、贵州省科技计划项目(黔科合基础[2020]1Y095)、国家现代农业产业体系荞麦育种岗位科学家专项(CARS-07-A5)和贵州省教育厅荞麦种质资源保育及创新重点实验室建设基金项目(黔教合 KY[2017]002)资助。

【作者简介】

石桃雄(1980-),女,副教授,硕士生导师,主要从事荞麦数量遗传学研究,E-mail:shitaoxiong@126.com。

【**通信作者】

邓娇(1986-),女,副教授,硕士生导师,主要从事植物生理和分子生物学研究,E-mail:ddj613@163.com。

【引用本文】

石桃雄,赵佳利,田桃,等. 苦荞 CCT 基因家族生物信息学及表达分析[J]. 广西科学,2021,28(3):301-309.

SHI TX,ZHAO J L,TIAN T, et al. Bioinformatics and Expression Analysis of CCT Gene Family in *Fagopyrum tataricum* [J]. Guangxi Sciences, 2021,28(3):301-309.

花期是植物生命周期中一个十分重要的发育阶段,对经济作物的生态适应性、产量和品质等有着非常重要的影响。苦荞花期的早晚与产量呈极显著负相关关系^[7],因此研究荞麦开花的调控机理对提高苦荞产量至关重要。据报道,植物 CCT(即 CO、COL 和 TOC1)转录因子家族具有调控花期、生物钟和响应光周期等功能^[8,9]。此外,CCT 基因还通过激素信号转导途径参与抵抗干旱、盐、高温、低温等非生物胁迫,减少这些非生物胁迫对植物造成的危害^[10,11]。因此 CCT 基因在植物的生长发育及抗逆境胁迫方面发挥着重要的作用。

CCT 基因编码的蛋白都含有 CCT 结构域,是由 CO-LIKE 蛋白、CONSTANS (CO)蛋白和 TIMING OF CAB1 (TOC1)蛋白的 C 末端的 43 个保守的氨基酸残基组成^[12]。根据保守结构域的不同,CCT 基因家族分为 3 个亚家族,分别为 CONSTANS-like (COL)亚家族、CCT motif (CMF)亚家族和 PRR 亚家族。COL 亚家族所编码蛋白均包含 1-2 个 B-box 结构域和 1 个 CCT 结构域,CMF 亚家族所编码蛋白仅含有 1 个 CCT 结构域,而 PRR 亚家族所编码蛋白均含有 1 个 CCT 结构域和 1 个 Receiver-like (REC) 结构域^[13]。目前,CCT 基因家族已在拟南芥^[14]、水稻^[8]、小麦^[11]、玉米^[15]和毛竹^[16]等多个物种中被报道。拟南芥的 CCT 基因家族有 40 个成员,其中 CO 基因是最先被发现的,它在长日照条件下能促进开花,突变后拟南芥表现出晚花性状^[17];TOC1 基因通过昼夜节律调控开花光周期^[18];COL1 基因在抵抗盐、干旱和温度等非生物胁迫中起到重要作用^[14]。水稻中有 41 个 CCT 基因,最新研究发现至少有 18 个成员参与了水稻的开花过程,其中有 4 个新发现的成员参与抽穗期的调控;3 个基因 *O_sCCT22*、*O_sCCT38* 和 *O_sCCT41* 表现出在长日照条件下抑制开花,在短日照条件下促进开花的作用,而 *O_sCCT03* 在长、短日照条件下均能促进水稻开花^[8]。玉米中的 CCT 基因家族成员比其他物种稍多,共有 57 个,其中有 15 个成员与花期显著相关^[15],而且 *ZmCOL3* 基因对开花、茎节数、总叶片株高和穗位高均有影响,呈现出多功能的特点^[19]。

在模式植物和大宗作物中 CCT 基因家族调控花期的研究报道已有不少,但在荞麦中还未见报道。本研究利用生物信息学工具对苦荞 CCT 基因家族进行全基因组范围的鉴定,并对其理化特性、染色体定位、基因结构、保守基序、保守结构域、系统进化和

表达模式等进行分析,拟为进一步探究苦荞 CCT 基因家族功能,并利用该基因家族改良荞麦品种奠定基础 and 提供理论依据。

1 材料与方法

1.1 苦荞基因组中 CCT 基因的筛选及理化性质分析

苦荞品种 'cv. Pinku1' 的基因组测序数据及其注释信息均从 NCBI 基因组数据库下载 (<http://www.mgbkbase.org/Pinku1/>)。从拟南芥信息资源数据库 TAIR (<https://www.arabidopsis.org/>) 下载 CCT 基因家族序列作为本地 BLAST 的参考序列 (query 序列)。运用 TBtools (v1.082) 软件进行 BLAST 比对,得到苦荞 CCT 基因家族候选序列。再运用 Conserved Domain Database (CDD) 数据库 (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb.cgi>) 对这些 CCT 基因的结构域进行分析,从候选序列中筛选出含有 CCT 结构域的序列,确认为苦荞的 CCT 基因家族成员。

运用 ExPaSy 提供的在线 ProtParam 软件 (<https://web.expasy.org/protparam/>), 分析苦荞 CCT 基因编码的蛋白质的理化性质,包括氨基酸残基的个数、相对分子量、等电点和亲水性等。

1.2 CCT 基因染色体分布

运用 TBtools (v1.082) 软件绘制苦荞 CCT 基因在染色体上的分布图并对其进行共线性分析。

1.3 CCT 基因结构、保守基序及结构域分析

首先运用 MEME 在线工具 (<http://meme-suite.org/>) 分析苦荞 CCT 蛋白的氨基酸序列的保守基序,基序个数设置为 10 个;然后用 TBtools (v1.082) 软件将预测结果 xml 文件进行可视化;最后利用 TBtools (v1.082) 软件分析苦荞 CCT 基因结构,并利用 CDD 数据库分析 CCT 基因保守结构域。

1.4 CCT 基因家族系统进化树构建

利用 MEGA 7.0 软件的 Cluster W 程序对苦荞和拟南芥的 CCT 基因家族进行多序列比对,采用 Neighbor Joining (NJ) 法构建系统进化树。

1.5 苦荞 CCT 基因的表达分析

采用 Zhang 等^[20] 发表的苦荞 'cv. Pinku1' 的转录组数据,根据 4 种组织(根、茎、叶和花)中的 FPKM 值,利用 TBtools (v1.082) 软件绘制 CCT 基因表达水平的热图。

2 结果与分析

2.1 苦荞 *CCT* 基因家族成员的鉴定及在染色体的定位

以拟南芥 40 个 *CCT* 基因家族成员为参考序列, 并结合 *CCT* 基因家族特有的 *CCT* 结构域, 在苦荞基因组数据库中共鉴定出 35 个 *CCT* 基因家族成员。染色体定位分析表明, 苦荞 35 个 *CCT* 基因在 8 条染色体上均有分布, 其中 7 号染色体上分布的 *CCT* 基因数目最多, 有 9 个; 其次是 8 号染色体上, 分布了 8 个成员(图 1)。分别有 5 个、4 个和 3 个 *CCT* 基因定位在 2 号、1 号和 5 号染色体上。3 号、4 号和 6 号染色体上各分布了 2 个 *CCT* 基因。共有 8 对 *CCT* 基因发生片段复制现象(图 1), 分别为 *FtPinG0006138300.01.T01* 和 *FtPinG0004072200.01.T01*、*FtPinG0004072200.01.T01* 和 *FtPinG0009564600.01.T01*、*FtPinG0009564600.01.T01* 和 *FtPinG0003531700.01.T01*、*FtPinG0001582600.01.T01* 和 *FtPinG0004578000.01.T01*、*FtPinG0008789300.01.T01* 和 *FtPinG0006802300.01.T01*、*FtPinG0008422500.01.T01* 和 *FtPinG0007624400.01.T01*、*FtPinG0007624400.01.T01* 和 *FtPinG0004084100.01.T01* 以及 *FtPinG0004084100.01.T01* 和 *FtPinG0008422500.01.T01*。

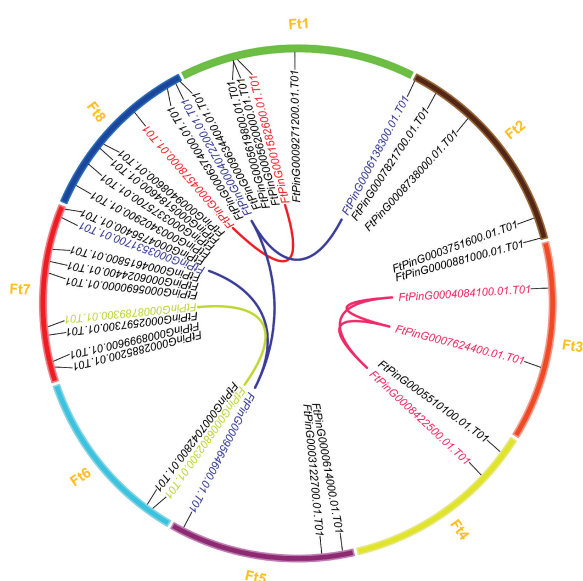


图 1 苦荞 *FtCCT* 基因在染色体上的分布及片段复制

Fig. 1 Distribution and fragment duplication of *FtCCT* genes of *F. tataricum* on chromosome

2.2 苦荞 *FtCCT* 系统发育、基因结构、保守基序及结构域

系统发育分析可将苦荞 35 个 *CCT* 基因分成三大类(图 2a), 共含有 10 个保守基序(图 2b)。第 I 类包含了大部分成员, 共 26 个, 又可细分为 4 小组: 第 1 小组有 10 个 *CCT* 基因, 除 *FtPinG0004084100.01.T01* 和 *FtPinG0007042800.01.T01* 只含有 Motif 1, 其他都含有 Motif 1 和 Motif 8; 第 2 小组含 4 个成员, 除 *FtPinG0005510100.01.T01* 缺少 Motif 10, 其他 3 个都具有 4 个 Motif (Motif 1、Motif 5、Motif 6、Motif 10); 第 3 小组由 4 个 *CCT* 基因组成, 其中 *FtPinG0005690000.01.T01* 只含有 Motif 1, 其他 3 个含有 Motif 1 和 Motif 2; 第 4 小组由 8 个成员组成, 其中 *FtPinG0006138300.01.T01* 仅含有 Motif 1, *FtPinG0006374000.01.T01* 含有 Motif 1、Motif 2, 其他 6 个 *CCT* 基因由 Motif 1、Motif 2、Motif 4 和 Motif 9 组成。第 II 类由 6 个 *CCT* 基因组成, 其中 2 个 *CCT* 基因仅含有 Motif 7, 2 个含有 Motif 1 和 Motif 7, 还有 2 个含有 Motif 1、Motif 2 和 Motif 7。第 III 类只有 3 个 *CCT* 基因成员, 都含有 Motif 1、Motif 5 和 Motif 10 (图 2a, b)。

35 个苦荞 *CCT* 基因中, 6 个 *CCT* 基因(*FtPinG0004084100.01.T01*、*FtPinG0007042800.01.T01*、*FtPinG0005690000.01.T01*、*FtPinG0006138300.01.T01*、*FtPinG0009634400.01.T01* 和 *FtPinG0004756400.01.T01*) 只含有 1 个基序, 其他 29 个 *CCT* 基因含有 2-4 个基序(图 2a, b)。除了两个成员(*FtPinG0009634400.01.T01* 和 *FtPinG0004756400.01.T01*) 不具有 Motif 1, 其他 33 个 *CCT* 基因成员都具有 Motif 1, 占比达到 94%; 12 个 *CCT* 基因成员具有 Motif 2 的基因, 比例为 34%。这说明 Motif 1 和 Motif 2 可能为苦荞 *CCT* 基因家族最保守的基序(图 2b)。在 Pfam 数据库中(<http://pfam.xfam.org/>), Motif 1 被注释为 *CCT* 结构域, Motif 2 和 Motif 4 被注释为 B-box 结构域, Motif 3 和 Motif 6 被注释为和调节信号响应相关的结构域, Motif 5 被注释为 GATA zinc finger, Motif 10 被注释为 tify 结构域, Motif 7、Motif 8 和 Motif 9 的注释未知。

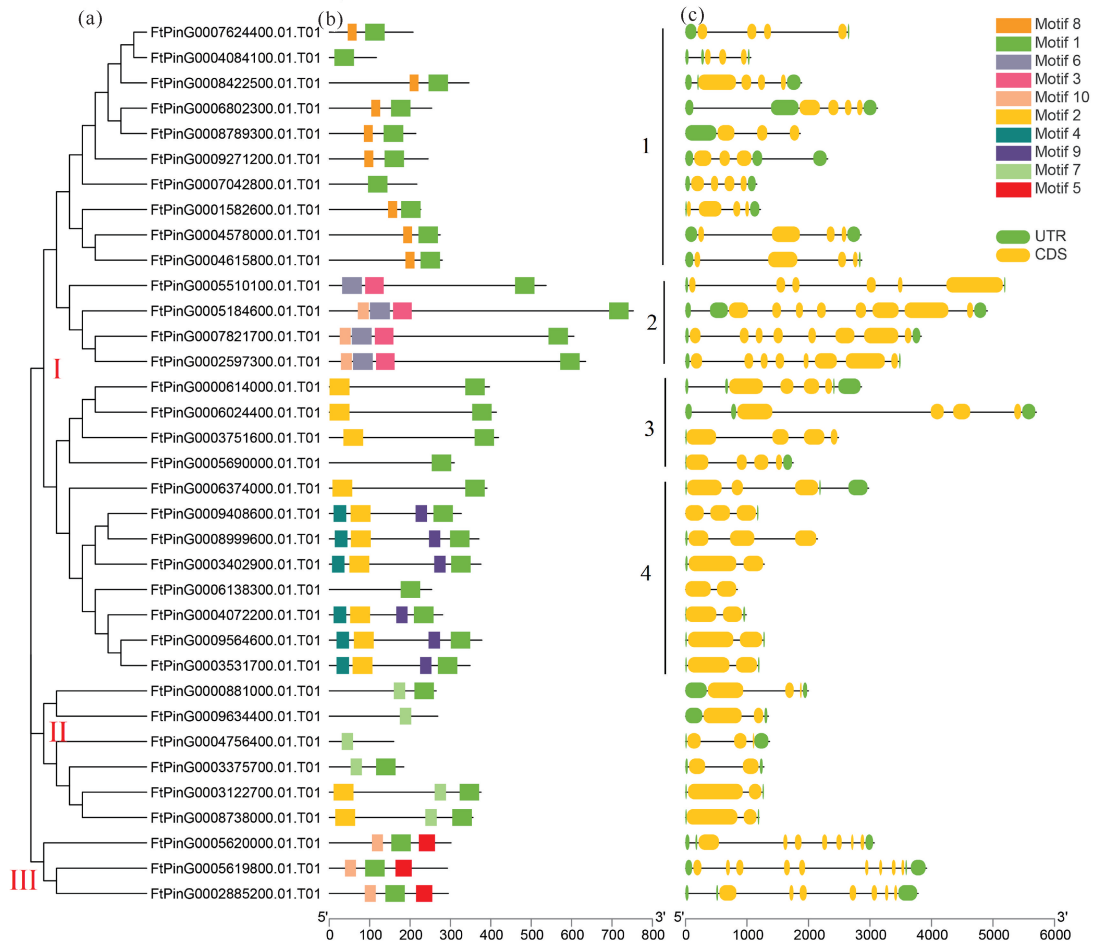


图2 苦荞 *FtCCT* 基因家族的系统发育树(a)、保守基序(b)及基因结构(c)

Fig. 2 Phylogenetic tree (a), conserved motif (b) and gene structure (c) of *FtCCT* gene family in *F. tataricum*

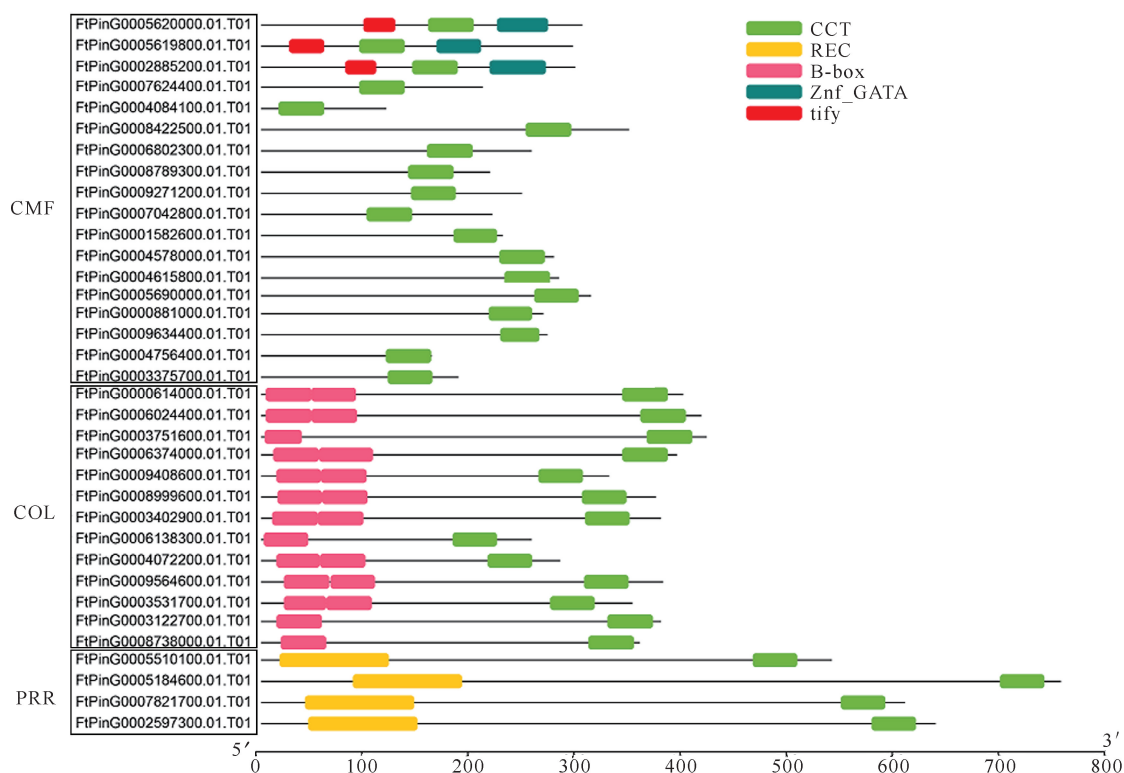
苦荞的 35 个 *CCT* 基因内含子数目为 1–8, 其中 *FtPinG0005619800.01.T01* 含有的内含子数目最多, 有 8 个; 其次是 *FtPinG0005184600.01.T01*、*FtPinG0007821700.01.T01* 和 *FtPinG0002597300.01.T01* (均位于第 I 类第 2 小组中), 均含有 7 个内含子(图 2c), 9 个 *CCT* 基因成员只有 1 个内含子。

苦荞 *CCT* 基因家族所编码蛋白中含有 5 种主要的保守结构域, 分别是 CCT、B-box、Znf-GATA、tify 和 REC (图 3), 其中 CTT 结构域在所有苦荞 *CCT* 基因编码蛋白中都有出现, 说明该结构域在植物进化过程中高度保守。苦荞 *CCT* 基因家族中的 CMF 亚家族的成员最多(18 个), 此亚族所编码的蛋白都含有 1 个 CCT 结构域, 其中还有 3 个含有 tify 结构和 Znf-GATA 结构域, 并且这 3 个蛋白的大小相似;

COL 亚家族由 13 个 *CCT* 基因组成, 所编码蛋白中除了 4 个成员只含有 1 个 B-box 和 1 个 CCT 结构域外, 其他成员都含有 2 个 B-box 和 1 个 CCT 结构域; PRR 亚家族中的 *CCT* 基因数目最少, 只有 4 个, 所编码蛋白由 1 个 CCT 和 1 个 REC 结构域组成。

2.3 苦荞 *FtCCT* 基因家族的理化性质

对 35 个 *FtCCT* 基因编码的蛋白质的理化性质分析表明, *CCT* 基因家族编码的蛋白质含有 117–753 个氨基酸残基, 相对分子质量为 4 082.90–82 741.52 Da, 预测的理论等电点为 4.96–9.51, 且以偏酸性和中性居多。预测的苦荞 *CCT* 基因家族蛋白质的亲水指数均为负值, 表明蛋白质均属于亲水性蛋白, 其中以 *FtPinG0007624400.01.T01* 的亲水性最强, 平均亲水指数为 -9.180 (表 1)。

图3 苦荞 *FtCCT* 基因家族保守结构域Fig. 3 Conserved domain of *FtCCT* gene family in *F. tataricum*表1 苦荞 *FtCCT* 基因家族的理化性质Table 1 Physicochemical properties of *FtCCT* gene family in *F. tataricum*

基因 ID Gene ID	氨基酸个数 Number of amino acids (aa)	分子量 Molecular weight (Da)	理论等电点 Theoretical pI	平均亲水指数 Grand average of hydrophaticity
<i>FtPinG0006802300.01.T01</i>	254	28 506.62	5.40	-0.070
<i>FtPinG0008789300.01.T01</i>	251	24 460.34	5.86	-0.675
<i>FtPinG0007624400.01.T01</i>	280	23 431.93	7.56	-9.180
<i>FtPinG0008422500.01.T01</i>	346	39 349.17	5.28	-0.776
<i>FtPinG0004084100.01.T01</i>	117	13 882.15	5.81	-1.475
<i>FtPinG0007042800.01.T01</i>	217	24 969.69	7.17	-0.925
<i>FtPinG0009271200.01.T01</i>	245	28 135.74	6.17	-1.049
<i>FtPinG0004578000.01.T01</i>	275	30 855.76	5.31	-0.468
<i>FtPinG0004615800.01.T01</i>	280	31 326.54	5.53	-0.380
<i>FtPinG0001582600.01.T01</i>	227	26 143.52	5.44	-0.530
<i>FtPinG0009564600.01.T01</i>	378	41 234.24	5.66	-0.334
<i>FtPinG0004072200.01.T01</i>	281	31 151.97	7.36	-0.371
<i>FtPinG0003402900.01.T01</i>	376	4 082.90	5.52	-0.432
<i>FtPinG0003531700.01.T01</i>	394	38 323.00	6.01	-0.255
<i>FtPinG0006138300.01.T01</i>	254	27 914.32	5.51	-0.345
<i>FtPinG0009408600.01.T01</i>	327	36 302.66	6.31	-0.522
<i>FtPinG0008999600.01.T01</i>	371	41 072.56	5.91	-0.631
<i>FtPinG0003122700.01.T01</i>	376	42 553.85	6.05	-0.668

续表 1

Continued table 1

基因 ID Gene ID	氨基酸个数 Number of amino acids (aa)	分子量 Molecular weight (Da)	理论等电点 Theoretical pI	平均亲水指数 Grand average of hydropathicity
<i>FtPinG0000881000.01.T01</i>	265	29 461.10	6.26	-0.752
<i>FtPinG0008738000.01.T01</i>	356	40 364.42	7.71	-0.790
<i>FtPinG0007821700.01.T01</i>	606	68 023.00	6.04	-0.691
<i>FtPinG0006374000.01.T01</i>	391	43 083.77	4.96	-0.556
<i>FtPinG0005510100.01.T01</i>	537	60 598.89	6.00	-0.745
<i>FtPinG0004756400.01.T01</i>	160	18 352.63	7.74	-0.827
<i>FtPinG0003375700.01.T01</i>	185	21 025.74	9.51	-0.908
<i>FtPinG0005620000.01.T01</i>	302	32 932.18	6.11	-0.768
<i>FtPinG0002597300.01.T01</i>	635	70 527.61	5.87	-0.783
<i>FtPinG0003751600.01.T01</i>	419	45 833.02	4.96	-0.455
<i>FtPinG0005184600.01.T01</i>	753	82 741.52	6.01	-0.821
<i>FtPinG0005690000.01.T01</i>	310	35 031.36	5.92	-1.083
<i>FtPinG0000614000.01.T01</i>	397	43 516.39	5.42	-0.620
<i>FtPinG0006024400.01.T01</i>	414	45 045.81	4.98	-0.519
<i>FtPinG0009634400.01.T01</i>	269	20 528.68	7.66	-0.710
<i>FtPinG0005619800.01.T01</i>	293	32 063.96	6.05	-0.545
<i>FtPinG0002885200.01.T01</i>	295	32 443.75	5.87	-0.889

2.4 苦荞 *FtCCT* 基因家族进化分析结果

系统进化分析发现苦荞和拟南芥 *CCT* 基因家族成员可分为三大类: I、II 和 III (图 4)。其中, I 类中的基因数目最丰富, 由 29 个拟南芥 *CCT* 基因和 22 个苦荞 *CCT* 基因组成, 可以进一步分为 I_A 和 I_B 两个亚类; II 类包含 8 个拟南芥和 10 个苦荞的 *CCT* 基因成员; III 类中拟南芥和苦荞各 3 个 *CCT* 基因成员。

根据姚新转等^[21]对拟南芥 *CCT* 基因家族的分类方法, 可以将苦荞的 35 个 *CCT* 基因家族分为 COL、CMF 和 PRR 3 个亚家族 (图 4), 与图 3 的结果基本一致。其中 CMF 亚家族包含了分布在进化树上的 II 类和 III 类中的全部 *CCT* 成员, 以及 I_A 和 I_B 两个亚类中的小分支; COL 亚家族主要分布在进化树中的 I_A 亚类以及 I_B 亚类中的一个分支上; PRR 亚家族由 I_B 的一个分支组成。

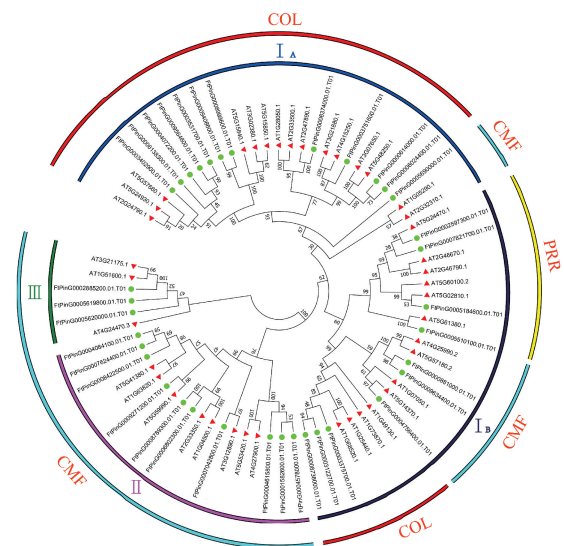
图 4 苦荞和拟南芥 *CCT* 基因家族系统进化

Fig. 4 Phylogenetic tree of *CCT* gene family from *F. tataricum* and *Arabidopsis*

2.5 苦荞 CCT 基因家族在组织中的表达模式

苦荞 35 个 CCT 基因在根、茎、叶和花 4 个组织中的表达模式如图 5 所示, 有 5 个 CCT 基因在根中表达水平高, 推测可能与根的生长发育调控有关; 只有 1 个 CCT 基因在茎中显著高表达, 推测其可能参与茎的生长发育过程。在叶中表达水平较高的 CCT 基因的数目较多, 而且大多数是 CMF 亚家族成员, 他们可能参与叶的生长发育与光形态建成。在花中表达水平较高的 CCT 基因的数目也较多, 主要是 COL 和 CMF 亚家族成员, 这些基因可能参与调控开花过程, 比如开花时间、花期长短等。

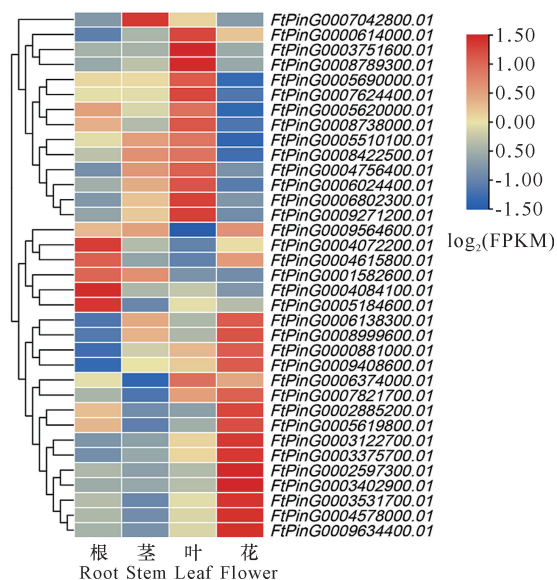


图 5 苦荞 CCT 基因家族在不同组织中的表达模式

Fig. 5 Expression pattern of CCT genes in different tissues of *F. tataricum*

3 讨论

植物 CCT 基因家族在调节植物的生长发育和抗逆境胁迫中起着至关重要的作用, 并且大多数成员参与开花调控, 与作物的产量也息息相关^[8,16,22,23]。本研究在全基因组水平上共鉴定到 35 个苦荞 *FtCCT* 基因, 数目与其他物种相差不大, 含有 1-8 个内含子(图 2)。这些苦荞 CCT 蛋白等电点为 4.96-9.51, 绝大多数以酸性为主, 均为亲水性蛋白(表 1)。35 个苦荞 *FtCCT* 基因在 8 条染色体上均有分布, 且 8 对 *FtCCT* 基因发生了片段复制事件(图 1)。片段重复在 CCT 基因家族成员的扩增中可能起着重要的作用, 推测这些同源性较高的复制基因具有相似的生物学功能。苦荞 CCT 蛋白与其他物种一样含 CCT、B-box、Znf-GATA、tify 和 REC 等 5 种主要的

保守结构域(图 3)。所有的苦荞 CCT 蛋白都含有 CCT 结构域, 说明该结构域在植物进化过程中高度保守。

为更好地了解苦荞 *FtCCT* 基因家族成员的分类和可能具有的功能, 本研究将 40 个拟南芥 *AtCCT* 家族成员与 35 个苦荞 *FtCCT* 家族成员进行蛋白序列比对并构建系统进化树(图 4)。结果表明, 两者的 CCT 基因具有较高的同源性, 说明 CCT 基因在物种间比较保守, 可能具有相似的功能。根据姚新转等^[21]对拟南芥 *AtCCT* 基因家族的分类方法, 苦荞的 35 个 *FtCCT* 基因家族也可以分为 COL、CMF 和 PRR 3 个亚家族, 且 CCT 基因家族的功能可能与 Motif 以及结构域相关。Putterill 等^[17]研究表明, 拟南芥的 *AtCO* 基因(*At15840.1*)属于 COL 亚家族, 具有促进开花的功能, 苦荞中两个 *AtCO* 同源基因 *FtPinG0008999600.01.T01* 和 *FtPinG0009408600.01.T01* 在花组织中具有较高的表达水平(图 5), 因此推测 *FtPinG0008999600.01.T01* 和 *FtPinG0009408600.01.T01* 可能具有相似的功能, 并且苦荞的 COL 亚家族的 12 个成员中有 7 个在花组织中高表达(图 5), 这与拟南芥的 COL 亚家族中的大部分成员参与拟南芥开花时间的调控一样^[18]。属于拟南芥 PRR 亚家族的 *TOC1* 被报道影响昼夜节律^[24], 苦荞中的同源基因 *FtPinG0005510100.01.T01* 可能也参与生物钟的调控。PRR 亚家族的 4 个成员中有 2 个在花中高表达, 在根和叶中各有 1 个表达水平较高(图 5), 因此, 不同的成员可能对不同组织的生长发育起着调控作用。CMF 亚家族对谷类作物的驯化和适应起着重要作用, 且对开花起着抑制作用^[25,26], 而苦荞中的 CMF 亚家族的成员最丰富, 其中有 8 个成员在叶中高表达, 有 6 个在花中高表达, 还有少数成员在根和茎中具有较高的表达水平, 这些基因可能调控不同组织的生长发育。因此, 苦荞 CCT 基因家族中的有些成员可能与拟南芥的同源基因具有相似的功能, 而特异的 Motif 和结构域可能是不同亚家族成员之间功能差异的主要原因。

4 结论

植物 CCT 基因家族在调节植物的生长发育、花期和抗逆境胁迫中起着至关重要的作用。本研究在全基因组水平上共鉴定出 35 个苦荞 CCT 基因, 在 8 条染色体上均有分布。苦荞 CCT 基因编码的蛋白的等电点为 4.96-9.51, 绝大多数以酸性为主, 均为

亲水性蛋白。基于 CCT、B-box、Znf-GATA、tify 和 REC 等 5 种主要的保守结构域, 将苦荞的 CCT 基因家族划分为 CMF、COL 和 PRR 3 个亚家族。苦荞和拟南芥 CCT 基因具有较高的同源性。在叶和花中高表达的 CCT 基因数目较多, 推测这些基因可能参与了生长发育、光形态建成或开花调控过程。以上研究结果为进一步探究苦荞 CCT 基因家族功能, 以及利用该基因家族改良荞麦花期奠定了基础。

参考文献

- [1] 盛晋华, 张雄杰, 陕方, 等. 内蒙古自治区荞麦生产开发现状与对策[J]. 作物杂志, 2009(6): 1-5.
- [2] 吴韬, 肖丽, 李伟丽. 苦荞的营养与功能成分研究进展[J]. 西华大学学报: 自然科学版, 2021, 40(2): 91-96.
- [3] JOSHI D C, ZHANG K, WANG C, et al. Strategic enhancement of genetic gain for nutraceutical development in buckwheat: A genomics-driven perspective [J]. *Biotechnology Advances*, 2020, 39: 107479.
- [4] 何伟俊, 曾荣, 白永亮, 等. 苦荞麦的营养价值及开发利用研究进展[J]. 农产品加工, 2019(23): 69-75.
- [5] TOLAINI V, DEL FIORE A, NOBILI C, et al. Exploitation of tartary buckwheat as sustainable ingredient for healthy foods production [J]. *Agriculture and Agricultural Science Procedia*, 2016, 8: 455-460.
- [6] 刘邻渭, 章华伟, 姜丽. 中国荞麦深加工的探索[J]. 西北农林科技大学学报: 自然科学版, 2002, 30(增刊): 83-85.
- [7] 吕丹, 黎瑞源, 郑冉, 等. 213 份苦荞种质资源主要农艺性状分析及高产种质筛选[J]. 南方农业学报, 2020, 51(10): 2429-2439.
- [8] ZHANG J, FAN X, HU Y, et al. Global analysis of CCT family knockout mutants identifies four genes involved in regulating heading date in rice [J]. *Journal of Integrative Plant Biology*, 2021, 63(5): 913-923.
- [9] 袁玺垒, 王振山, 贾小平, 等. 光周期调控植物开花分子机制以及 CCT 基因家族研究进展[J]. 浙江农业学报, 2020, 32(6): 1133-1140.
- [10] MA L, YI D, YANG J, et al. Genome-wide identification, expression analysis and functional study of CCT gene family in *Medicago truncatula* [J]. *Plants*, 2020, 9(4): 513.
- [11] 田书军. 小麦 TaCCT 基因家族的全基因组鉴定及非生物胁迫下的表达模式分析[D]. 杨凌: 西北农林科技大学, 2020.
- [12] ROBSON F, COSTA M M, HEPWORTH S R, et al. Functional importance of conserved domains in the flowering-time gene *CONSTANS* demonstrated by analysis of mutant alleles and transgenic plants [J]. *The Plant Journal*, 2001, 28(6): 619-631.
- [13] ZHANG L, LI Q, DONG H, et al. Three CCT domain-containing genes were identified to regulate heading date by candidate gene-based association mapping and transformation in rice [J]. *Scientific Reports*, 2015, 5(1): 7663.
- [14] 刘芳, 肖衡, 金玉环, 等. 短命植物小拟南芥光周期调控基因 *CONSTANS* 的克隆与表达特征分析[J]. 石河子大学学报: 自然科学版, 2019, 37(2): 190-199.
- [15] JIN M, LIU X, JIA W, et al. *ZmCOL3*, a CCT gene represses flowering in maize by interfering with the circadian clock and activating expression of *ZmCCT* [J]. *Journal of Integrative Plant Biology*, 2018, 60(6): 465-480.
- [16] 张文祥, 李敏, 王汇源, 等. 毛竹 CCT 基因家族成员的鉴定及表达分析[J/OL]. 分子植物育种: 1-12. [2021-05-09]. <http://kns.cnki.net/kcms/detail/46.1068.S.20210308.1708.028.html>.
- [17] PUTTERILL J, ROBSON F, LEE K, et al. The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors [J]. *Cell*, 1995, 80(6): 847-57.
- [18] STRAYER C, OYAMA T, SCHULTZ T F, et al. Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog [J]. *Science*, 2000, 289(5480): 768-771.
- [19] 贾伟. 玉米光周期敏感基因 *ZmCOL3* 的功能解析[D]. 沈阳: 吉林农业大学, 2018.
- [20] ZHANG L, LI X, MA B, et al. The Tartary buckwheat genome provides insights into rutin biosynthesis and abiotic stress tolerance [J]. *Molecular Plant*, 2017, 10(9): 1224-1237.
- [21] 姚新转, 张宝会, 赵德刚, 等. 茶树 CsCCT 基因家族的鉴定和表达模式分析[J/OL]. 分子植物育种: 1-27. [2021-05-13]. <http://kns.cnki.net/kcms/detail/46.1068.S.20210225.1749.011.html>.
- [22] 李晓华, 丁明亮, 乔玲, 等. 粗山羊草 CCT 基因家族进化及节律表达分析[J]. 植物遗传资源学报, 2017, 18(6): 1151-1158.
- [23] MENGARELLI D A, ZANOR M I. Genome-wide characterization and analysis of the CCT motif family genes in soybean (*Glycine max*) [J]. *Planta*, 2021, 253(1): 15.
- [24] TAKATA N, SAITO S, SAITO C T, et al. Phylogenetic footprint of the plant clock system in angiosperms:

- Evolutionary processes of pseudo-response regulators [J]. *BMC Evolutionary Biology*, 2010, 10(1):126.
- [25] LI Y, XU M. *CCT* family genes in cereal crops: A current overview [J]. *The Crop Journal*, 2017, 5(6): 449-458.
- [26] 章佳. 水稻 *CCT* 家族基因的功能研究和 *Hd1* 的重新克隆[D]. 武汉:华中农业大学, 2017.

Bioinformatics and Expression Analysis of *CCT* Gene Family in *Fagopyrum tataricum*

SHI Taoxiong, ZHAO Jiali, TIAN Tao, CHEN Qingfu, DENG Jiao

(Research Center of Buckwheat Industry Technology, Guizhou Normal University, Guiyang, Guizhou, 550001, China)

Abstract: *CCT* transcription factor plays an important role in the regulation of plant flowering, growth and development, as well as resistance to abiotic stress. The *CCT* gene family of *Arabidopsis thaliana* was used as the reference sequence in this study. The *FtCCT* gene family members of *Fagopyrum tataricum* were screened out by bioinformatics tools, including local BLAST and conserved domains. The physicochemical properties, chromosome distribution, gene structure, phylogenetic evolution and expression level were analyzed. The results showed that a total of 35 *FtCCT* genes were identified in *F. tataricum* where the introns ranged from 1 to 8. The encoded protein had 117–753 amino acid residues with isoelectric points of 4.96–9.51, all of which were hydrophilic proteins. Chromosomal mapping analysis showed that these genes were distributed on all 8 chromosomes. The *FtCCT* gene family contains 10 conserved motifs and 5 conserved domains, and all of them contained *CCT* conserved domains. Phylogenetic analysis showed that *CCT* gene family of *F. tataricum* could be divided into three subfamilies like *A. thaliana*, among which the CMF subfamily had the most members. The expression levels of 35 *FtCCT* genes in roots, stems, leaves and flowers of *F. tataricum* were different. There were more members with high expression levels in leaves and flowers, and only a few members were highly expressed in roots and stems. This study lays a foundation for further analysis of *CCT* gene regulation in flowering period and growth of *F. tataricum*.

Key words: *Fagopyrum tataricum*, flowering, *FtCCT*, phylogenetics, expression analysis

责任编辑:米慧芝



微信公众号投稿更便捷

联系电话:0771-2503923

邮箱:gxkx@gxas.cn

投稿系统网址: <http://gxkx.ijournal.cn/gxkx/ch>