

杜仲 TCP 转录因子鉴定及生物信息学分析

刘俊^{1,3,4}, 李龙², 吴耀松¹, 刘燕¹, 任闪闪¹, 陈玉龙^{1*}

1. 河南中医药大学中医药科学院 河南省中医方证信号传导重点实验室, 河南 郑州 450046
2. 西北农林科技大学林学院, 陕西 杨凌 712100
3. 国际竹藤中心 国家林业和草原局竹藤科学与技术重点开放实验室, 北京 100102
4. 国际竹藤中心 安徽太平试验中心, 安徽 黄山 245716

摘要: 目的 对杜仲 TCP (Teosinte branched 1/Cycloidea/Proliferating) 基因家族进行筛选分析, 以期对 *EuTCPs* 基因功能深入研究奠定基础。方法 以杜仲基因组数据库为基础, 通过生物信息学系统分析 TCP 基因家族的理化性质、系统进化、基因结构、启动子顺式作用元件及其在杜仲叶片发育及胶形成中的表达水平。结果 从杜仲基因组中共鉴定到 14 个 *EuTCPs*, 氨基酸数目介于 139~492, 理论等电点分布在 5.54~9.72, 相对分子质量区域为 18 880~53 620, 亚细胞定位预测在细胞核中, 均为亲水性蛋白。系统进化分为 PCF、CIN 和 CYC/TB1 3 个亚家族, 分别包含 5、6 和 3 个 *EuTCPs* 蛋白。表达模式分析显示, *EuTCPs* 在杜仲叶片不同发育时期存在显著差异, 每个发育时期 *EuTCPs* 表达量各不相同, *EuTCPs* 正调控杜仲胶的形成。结论 杜仲中 TCP 家族包含 14 个成员, 各成员的分子特征和组织表达特异性存在一定的差异, *EuTCPs* 可能参与杜仲叶片发育及杜仲胶的形成。

关键词: 杜仲; TCP 转录因子; 系统进化; 表达分析; 叶片发育

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2022)12-3755-11

DOI: 10.7501/j.issn.0253-2670.2022.12.022

Identification and bioinformatics analysis of TCP transcription factors in *Eucommia ulmoides*

LIU Jun^{1,3,4}, LI Long², WU Yao-song¹, LIU Yan¹, REN Shan-shan¹, CHEN Yu-long¹

1. Henan key Laboratory of TCM Prescription and Syndrome Signaling, Academy of Chinese Medicine, Henan University of Chinese Medicine, Zhengzhou Henna 450046, China
2. College of Forestry, Northwest Agriculture and Forestry University, Yangling 712100, China
3. International Center for Bamboo and Rattan, Key Open Laboratory of Bamboo and Rattan Science and Technology, China National Forestry and Grassland Administration, Beijing 100102, China
4. Anhui Taiping Experimental Station of International Center for Bamboo and Rattan, Huangshan 245716, China

Abstract: Objective In order to provide a basis for *EuTCPs* gene function research, *Eucommia ulmoides* TCP (Teosinte branched 1/Cycloidea/Proliferating) gene family was screened and analyzed in this study. **Methods** Based on *E. ulmoides* genome database, bioinformatics methods was used to systematically analyze the physical and chemical properties, phylogenetic evolution, gene structure, promoter cis-acting element and their expression in leaf development and gum formation of LBD gene family. **Results** In this study, a total of 14 *EuTCPs* were identified from *E. ulmoides* genome, *EuTCPs* encodes 139—492 amino acids, and theoretical isoelectric points distribution were 5.54—9.72, molecular weight between 18 880 to 53 620 by physicochemical analysis, all of *EuTCPs* proteins were hydrophilic proteins and expressed in nucleus. *EuTCP* gene family was divided into three subfamilies: PCF, CIN and CYC/TB1, which contain five, six and three *EuTCPs* proteins respectively. Expression pattern analysis showed that *EuTCPs* played different roles in leaf development of *E. ulmoides* and positively regulated the formation of gutta-percha. **Conclusion** In this study, the members of the TCP family were comprehensively identified and bioinformatics analysis from the genomic level of *E. ulmoides*, which provides basic data for further functional research of *EuTCP* genes.

Key words: *Eucommia ulmoides* Oliver; TCP gene family; phylogeny evolution; expression analysis; leaf development

收稿日期: 2021-11-06

基金项目: 河南省高等学校重点科研项目 (22A360005); 国家林业和草原局/北京市共建竹藤科学与技术重点实验室开放基金 (ICBR-2020-05); 国际竹藤中心安徽太平试验中心开放课题基金 (1632021006-4); 河南中医药大学博士科研基金资助项目 (RSBSJJ2019-04)

作者简介: 刘俊 (1990—), 女, 助理研究员, 博士, 主要从事药用植物分子生物学研究。E-mail: liujun_0325@163.com

*通信作者: 陈玉龙, 男, 教授, 博士, 主要从事中医药防治炎症与免疫研究。E-mail: cyl72621@163.com

TCP (Teosinte branched 1/Cycloidea/ Proliferating) 蛋白一般由 60 个氨基酸组成, 含有一个非典型的基序螺旋 (bHLH) 结构和 R 结构域^[1], 分为 Class I 和 Class II 2 个亚家族, Class I 又称为 PCF, Class II 进一步划分为 CIN 和 CYC/TB1 亚家族^[2]。TCP 转录因子参与植物胚胎发育^[3]、侧枝形成^[4-5]、开花调控^[6]、胁迫响应^[7]以及激素信号^[8]等生物学过程。

TB1 和 TB1-like 蛋白属于 CYC/TB1 亚家族, 抑制拟南芥分枝发育^[8]。AtTCP4 通过诱导脂肪氧化酶 (*LIPOXYGENASE2*) 基因表达, 促进茉莉酸 (JA) 生物合成和成熟叶片衰老^[9], 在花器官中 AtTCP4 抑制 *LIPOXYGENASE2* 基因表达^[10]。AtTCP4 通过靶向生长素生物合成酶基因 *YUCCA5* 调控生长素生物合成^[11], AtTCP14 是 PCF 亚家族成员, 在种子萌发过程中发挥负调控作用^[12]。毛竹 *PeTCP10* 通过脱落酸 (ABA) 依赖信号在抗旱中发挥正调控作用, 以茉莉酮酸甲酯 (Me-JA) 介导途径负调控侧根生长^[13]。*OsPCF2* 通过正调控 *NHX1* 影响水稻耐盐性^[14], *OsCP21* 和 *OsPCF6* 是 microRNA319 的靶基因, 在冷胁迫中发挥正调控作用^[15]。在拟南芥中超表达 *OsTCP19* 诱导 *IAA3*、*ABI3*、*ABI4* 上调表达, *LOX2* 下调表达, 导致侧根数目减少, 提高转基因植株胁迫耐受性^[16]。

杜仲 *Eucommia ulmoides* Oliver 是我国特有单科、单属、单种的多年生第三纪孑遗树种, 雌雄异株, 具有橡胶和药材的双重性质^[17]。广泛分布于温带及亚热带地区, 适应能力强, 在酸性或盐碱性的贫瘠土壤中均能生长良好^[18]。其树皮、叶片及种子均可入药, 富含多种活性成分, 对多种疾病的预防及提高人体免疫力具有显著的效果, 是我国二级珍稀保护植物。目前, 对杜仲的研究与利用多局限于资源收集、药理学及形态学研究, 分子遗传学方面的报道相对较少, 严重制约了杜仲分子育种及相关生物学特性研究。目前 TCP 转录因子已在多个物种中进行了研究, 然而杜仲 TCP 基因家族还未报道, 本研究以杜仲基因组数据^[19]为基础, 从基因组水平对 TCP 基因家族进行鉴定、理化性质、进化关系、保守基序以及表达模式分析, 以期为进一步探索杜仲 TCPs 基因功能奠定基础。

1 材料与方法

1.1 杜仲 TCP 基因家族鉴定及理化性质分析

从杜仲基因组数据库 Genome Warehouse (<https://bigd.big.ac.cn/gwh/Assembly/13/show>) 中下

载 TCP 蛋白候选序列, 利用 NCBI 的 (Conserved Domain Search Service (CD Search)) 分析蛋白结构域, 保留含有完整 bHLH 结构域的序列。通过在线软件 ProtParam (<http://web.expasy.org/protparam/>) 分析蛋白理化性质, 使用 Plant-mPLOC (<http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/>) 预测 EuTCPs 蛋白亚细胞定位, 利用在线工具 ExPASy (<https://www.expasy.org/tools>) 预测 EuTCPs 氨基酸数量、相对分子质量、理论等电点, 通过 Expasy (<https://web.expasy.org/protscale/>) 软件分析蛋白的亲疏水性, 利用 SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.plpage=npsa_sopma.html) 和 Phyre 2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) 软件分别预测蛋白的二级结构和三级结构。

1.2 杜仲 TCPs 蛋白系统发育树构建

通过 Clustal X1.83 软件对杜仲、水稻、拟南芥和毛竹 TCPs 氨基酸序列进行多序列比对, 利用 MEGA 6.0 的邻接法 (neighbor-joining) 构建系统发育树, 重复次数设置为 1000 次, 其他参数为默认值^[20]。

1.3 保守基序及启动子分析

通过 MEME (<http://meme-suite.org/>) 在线软件对 EuTCP 基因家族成员进行基序分析 (参数是: any number of Repetitions (anr), maximum number of Motifs=10, minimum width $\geq$ 6, and maximum width $\leq$ 50)。为了解析 EuTCPs 启动子区域顺势作用元件, 对 EuTCPs 基因 ATG 上游 2000 bp 序列进行查找分离, 利用 Plant CARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) 软件进行启动子顺势作用元件分析。

1.4 杜仲 EuTCP 基因家族表达模式分析

从 NCBI 的 Short Read Archive (SRA) 数据库中下载杜仲叶片不同发育时期 (叶芽、初生叶、幼叶、老叶) (版本号: SRP218063)^[21]及不同胶含量 (低含量、高含量) (版本号: SRP158357)^[22]转录组数据, EuTCPs 基因相对表达丰度用 FPKM 值表示, 对该数值取对数进行统计分析, 通过 MeV4-9-0 工具绘制基因表达图谱, 表达量最高的用绿色方框表示, 表达量最低的用黄色方框表示。

1.5 杜仲 TCPs 蛋白互作网络预测

利用 STRING 软件 (<https://string-db.org/>) 上传 EuTCPs 蛋白序列, 选择拟南芥数据库进行序列比

对, 根据已知拟南芥蛋白互作关系, 通过 Blast 比对杜仲同源蛋白, 通过 Cytoscape 3.7.0 软件对 EuTCP 基因家族蛋白互作信息进行评估和预测^[23]。

2 结果与分析

2.1 杜仲 TCP 基因家族鉴定及理化性质分析

通过 Genome Warehouse 数据库, 从杜仲基因组中共鉴定到 14 个 *EuTCPs* 基因(表 1), 利用 Pfam 和 NCBI 的 Conserved Domain Search 在线分析软件对 EuTCPs 蛋白进行结构域验证。结果显示, 14 个

EuTCPs 蛋白均含有非典型的 bHLH (BASIC、HELIX I、LOOP 和 HELIX II) 特征结构域, 分别命名为 EuTCP1~EuTCP14。使用 ExPASy 工具, 对 EuTCP 家族成员进行蛋白质理化性质分析, EuTCP14 蛋白最长, 编码 492 个氨基酸, EuTCP13 蛋白最短, 编码 139 个氨基酸, 相对分子质量分布区域为 18 880~53 620, 等电点范围是 5.54 (EuTCP14) 到 9.72 (EuTCP6)。亚细胞定位预测结果显示, EuTCPs 均定位在细胞核中(表 1)。

表 1 杜仲 TCP 蛋白序列特征及亚细胞定位预测

Table 1 Amino acid sequence characteristics and predicted subcellular location of *E. ulmoides* TCP proteins

基因 ID	基因名	拟南芥同源基因	染色体定位	CDS 长度/bp	蛋白质大小/aa	相对分子质量	等电点	亚细胞定位
EUC12473-RA	<i>EuTCP1</i>	AT2G37000	Super-Scaffold_57	627	208	22 441.51	8.85	细胞核
EUC20760-RA	<i>EuTCP2</i>	AT2G31070	scaffold865_obj	780	259	29 028.36	8.48	细胞核
EUC13147-RA	<i>EuTCP3</i>	AT4G18390	Super-Scaffold_67	1311	436	47 303.13	8.82	细胞核
EUC03320-RA	<i>EuTCP4</i>	AT3G27010	Super-Scaffold_177	930	309	32 689.32	8.99	细胞核
EUC22153-RA	<i>EuTCP5</i>	AT3G02150	scaffold1183_obj	1029	342	38 298.69	6.90	细胞核
EUC10434-RA	<i>EuTCP6</i>	AT5G23280	Super-Scaffold_46	798	265	27 785.13	9.72	细胞核
EUC05665-RA	<i>EuTCP7</i>	AT1G68800	Super-Scaffold_28	1083	360	41 030.62	7.75	细胞核
EUC03513-RA	<i>EuTCP8</i>	AT3G02150	Super-Scaffold_174	876	291	32 694.71	9.44	细胞核
EUC15623-RA	<i>EuTCP9</i>	AT1G67260	scaffold1436_obj	747	248	27 907.40	8.86	细胞核
EUC12970-RA	<i>EuTCP10</i>	AT3G47620	Super-Scaffold_113	1125	374	40 069.24	7.26	细胞核
EUC06461-RA	<i>EuTCP11</i>	AT1G53230	scaffold166_obj	879	292	32 398.81	5.65	细胞核
EUC10433-RA	<i>EuTCP12</i>	AT5G23280	Super-Scaffold_46	798	265	27 785.13	9.72	细胞核
EUC15622-RA	<i>EuTCP13</i>	AT1G67260	scaffold1436_obj	507	169	18 878.14	6.41	细胞核
EUC07713-RA	<i>EuTCP14</i>	AT3G15030	Super-Scaffold_600	1479	492	53 623.46	5.54	细胞核

2.2 杜仲 TCPs 蛋白结构分析

二级结构分析显示, EuTCPs 蛋白主要由 α 螺旋、无规则卷曲、延伸链和 β 转角组成, 顺序均是由无规则卷曲> α 螺旋>延伸链> β 转角, 具体数值见表 2。 α 螺

旋在 DNA 结合基序中发挥重要作用, 无规则卷曲易受侧链相互作用, 构成活性部位和功能部位。EuTCPs 蛋白三级结构非常相似, 结果如图 1 所示, 无规则卷曲所占比例最大, 蛋白空间结构不同, 决定功能的差异。

表 2 杜仲 TCP 基因家族二级结构分析

Table 2 Secondary structures analysis of *E. ulmoides* TCP gene family

基因名	最低值	位点	最高值	位点	α 螺旋 (占比/%)	延伸链 (占比/%)	β 转角 (占比/%)	无规则卷曲 (占比/%)
<i>EuTCP1</i>	-2.733	12	2.056	150	37 (17.79)	27 (12.98)	15 (7.21)	129 (62.02)
<i>EuTCP2</i>	-2.889	36	1.278	232	39 (15.06)	43 (16.60)	13 (5.02)	164 (63.32)
<i>EuTCP3</i>	-3.400	35	1.700	340	71 (16.28)	49 (11.24)	9 (2.06)	307 (70.41)
<i>EuTCP4</i>	-3.300	13	2.022	135	79 (25.57)	31 (10.03)	14 (4.53)	185 (59.87)
<i>EuTCP5</i>	-3.144	333	0.811	48	27 (7.89)	27 (7.89)	0 (0.00)	278 (81.29)
<i>EuTCP6</i>	-2.944	19	1.778	57	72 (27.17)	42 (15.85)	13 (4.91)	138 (52.08)
<i>EuTCP7</i>	-3.789	235	1.722	314	129 (35.83)	25 (6.94)	9 (2.50)	197 (54.72)
<i>EuTCP8</i>	-2.522	127	1.411	229	28 (9.62)	34 (11.68)	7 (2.41)	222 (76.29)
<i>EuTCP9</i>	-2.756	84	1.200	166	84 (33.87)	30 (12.10)	10 (4.03)	124 (50.00)
<i>EuTCP10</i>	-3.533	225	1.444	54	53 (14.17)	46 (12.30)	17 (4.55)	258 (68.98)
<i>EuTCP11</i>	-2.978	31	1.167	108	61 (20.89)	40 (13.70)	10 (3.42)	181 (61.99)
<i>EuTCP12</i>	-2.944	19	1.778	57	72 (27.17)	42 (15.85)	13 (4.91)	138 (52.08)
<i>EuTCP13</i>	-3.056	87	1.022	30	55 (32.54)	17 (10.06)	2 (1.18)	95 (56.21)
<i>EuTCP14</i>	-3.800	226	1.300	452	87 (17.68)	59 (11.99)	20 (4.07)	326 (66.26)

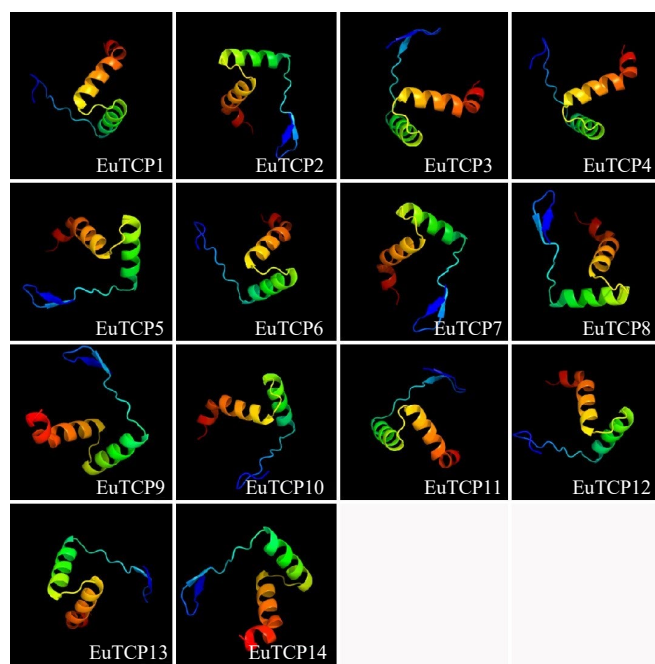


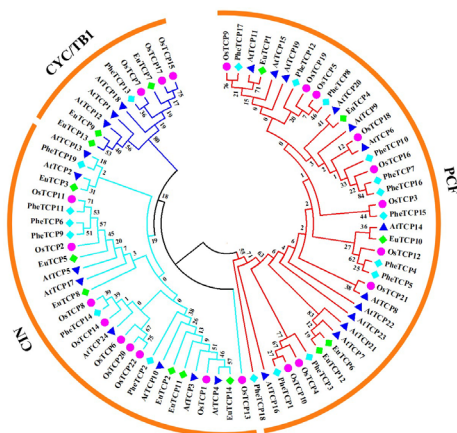
图 1 杜仲 TCPs 三级结构预测

Fig. 1 Tertiary structure prediction of *E. ulmoides* TCPs

2.3 杜仲 TCPs 蛋白系统进化分析

利用 MEGA 6.0 分析软件, 对 14 个杜仲 EuTCPs, 22 个水稻 OsTCPs^[24], 24 个拟南芥 AtTCPs^[25]和 19 个毛竹 PheTCPs^[26]蛋白进行 1000 次重复搜索, 构建系统进化树。结果如图 2 所示: 79 个 TCPs 蛋白分为 Class I 和 Class II 2 个大亚家族, 其中 Class I 又称为 PCF 亚家族; Class II 亚

族进一步划分为 CYC/TB1 和 CIN 亚家族。PCF 亚家族所包含 TCPs 成员数量最多, 含有 40 个 TCPs 蛋白, 5 个 EuTCPs, 13 个 AtTCPs, 10 个 OsTCPs 和 12 个 PtTCPs, 其次是 CIN 亚家族, 由 29 个 TCPs 蛋白组成, CYC/TB1 亚家族所含蛋白数量最少, 仅包含 10 个 TCPs 蛋白, 3 个 AtTCPs, 3 个 EuTCPs, 3 个 OsTCPs, 1 个 PheTCPs。



利用 MEGA 6.0 的邻接法构建系统进化树, 不同物种 TCPs 蛋白用不同颜色标记, 蓝色三角形代表拟南芥, 红色圆形代表水稻, 蓝绿色菱形代表毛竹, 杜仲 TCPs 蛋白由绿色菱形标记

The phylogenetic tree was generated with MEGA 6.0 software using the neighbour-joining method, TCP proteins from the different species were marked with different colored dots, blue triangles represented Arabidopsis, red cycles represented rice, blue-green diamond represented moso bamboo, and *E. ulmoides* TCPs proteins were marked by green rhombuses

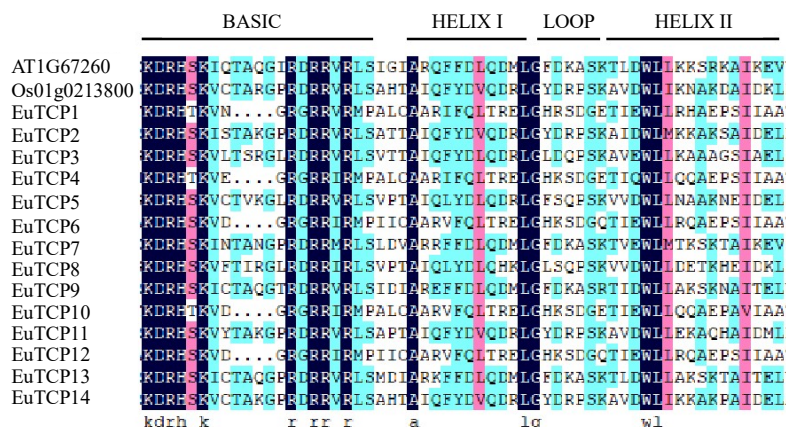
图 2 杜仲、水稻、拟南芥和毛竹 TCPs 蛋白系统进化树

Fig. 2 Phylogenetic tree of TCP proteins from *E. ulmoides*, rice, Arabidopsis and moso bamboo

2.4 杜仲 TCPs 结构域序列分析

为了分析杜仲 TCP 转录因子家族保守结构域序列特征, 利用 DNAMAN7.0 工具对杜仲、拟南芥和水稻 TCP 家族成员氨基酸序列进行比对, 结果如图 3 所示, 杜仲 TCP 结构域氨基酸分布与拟南芥和水稻十分相似, 14 个杜仲 TCP 家族成员均含有完整的非典型性 bHLH 结构域, 暗示杜仲 TCP 结构域高度保守性。BASIC 区域由 19 个氨

基酸组成, 其中有 9 个氨基酸是高度保守的, HLH 区域含有 5 个高度保守的氨基酸。杜仲 TCP 转录因子家族有 5 个蛋白(EuTCP1、EuTCP4、EuTCP6、EuTCP10 和 EuTCP12)BASIC 区域缺少 4 个氨基酸, 图 2 系统进化树结果显示, 他们均属于 Class I 亚家族, 其余蛋白均属于 Class II 亚家族, 这与人研究结果相一致^[2], 表明杜仲 TCP 转录因子家族进化的保守性。



结构域名称在序列上方标记, 黑色横线代表该区域的范围。黑色表示氨基酸是保守的, 红色和蓝绿色表示氨基酸是相似的, 无颜色表示氨基酸保守性较低

The structural domain names were marked above the sequences, with the black line representing the range of the region. Conserved amino acids were in black, similar amino acids were in red and turquoise, absence of color indicated amino acids were less conserved.

图 3 拟南芥、水稻和杜仲 TCP 蛋白保守结构域氨基酸序列比对

Fig. 3 Alignment of amino acid sequences in conserved domains from members of TCP proteins of Arabidopsis, rice, and *E. ulmoides*

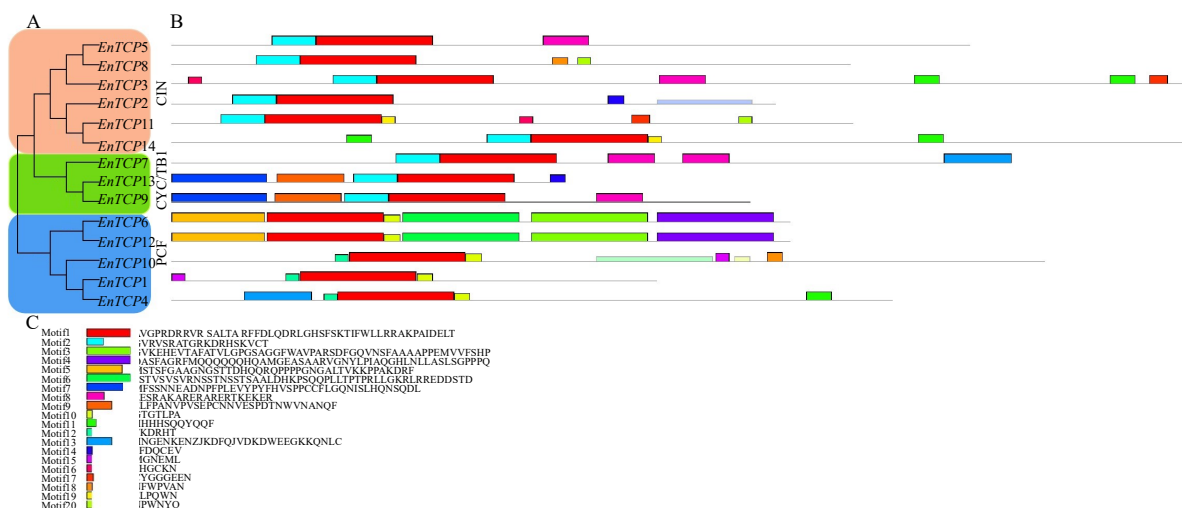
2.5 杜仲 TCP 转录因子家族基序分析

为了解析 *EuTCPs* 基因结构多样性和进化关系, 利用 MEME 在线软件对 *EuTCPs* 转录因子家族进行基序分析。结果如图 4 所示, *EuTCPs* 转录因子分为 3 个亚家族 (PCF、CYC/TB1 和 CIN), 进化关系与图 2 结果相一致, 20 个保守基序, 分别命名为 Motif 1~Motif 20, 14 条 *EuTCPs* 蛋白均含有 Motif 1 基序, 表明均属于 TCP 转录因子家族。同一亚家族 *EuTCPs* 蛋白基序具有高度相似性, CIN 亚家族均含有 Motif 1 和 Motif 2, PCF 亚家族成员均含有 Motif 1 和 Motif 19, 表明同一亚家族 *EuTCPs* 可能具有相似的功能。不同亚家族基序存在显著差异, 有些 Motifs 只在特定亚家族中存在, 例如, Motif 3~Motif 6 只存在于 PCF 亚家族, 只有 CYC/TB1 亚家族含有 Motif 7 和 Motif 9, 暗示某些基序在亚家族功能特异性方面发挥重要作用。

2.6 杜仲 TCPs 启动子顺式作用元件分析

为了探索 *EuTCPs* 基因功能和表达调控模式,

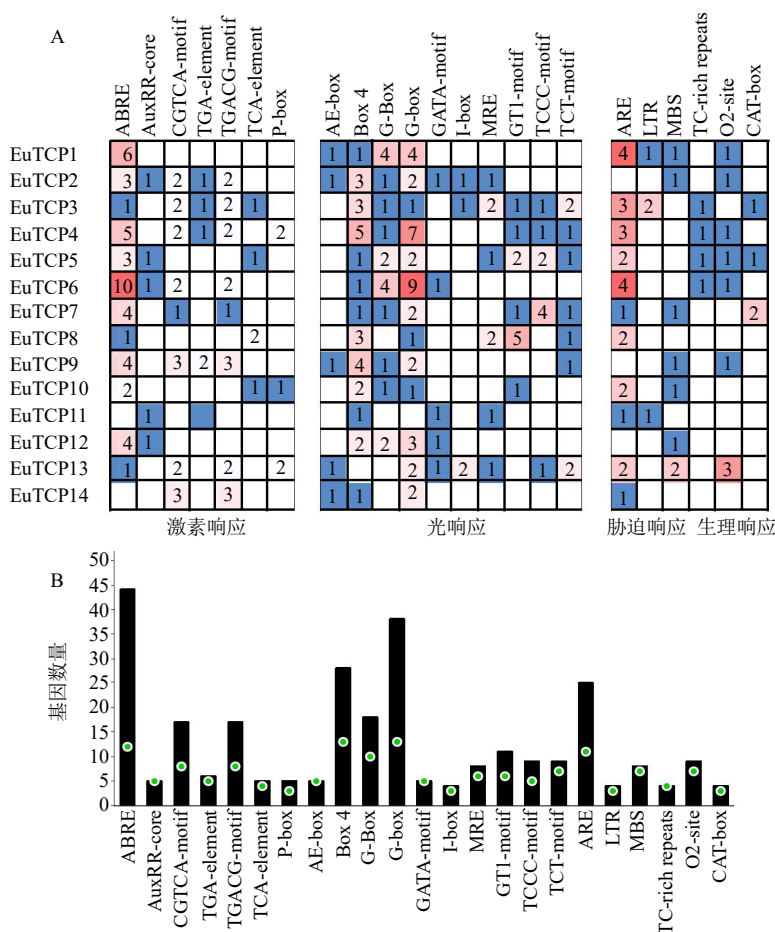
利用 Plant CARE 在线分析软件对 *EuTCPs* 起始密码子 (ATG) 上游 2000 bp 序列进行顺式作用元件分析, 结果如图 5 所示, *EuTCPs* 启动子中不仅含有基本顺式作用元件, 还存在 4 种类型元件: (1) 激素调控元件, 如赤霉素响应元件 ABRE; 生长素响应元件 AuxRR-core; 水杨酸响应元件 CGTCA-motif; (2) 光响应元件, 如 Box 4、G-box、G-Box、GT1-motif、GATA-motif 等; (3) 胁迫响应元件, 如干旱胁迫响应元件 MBS, 低温响应元件: LTR 和厌氧胁迫相关元件 ARE 等; (4) 生理响应元件, 如 O₂-site、CAT-box 等, 推测 *EuTCPs* 可能在杜仲生长发育、胁迫响应、激素调节以及光周期调控中发挥重要作用。*EuTCPs* 基因中光响应元件数量最多, 共有 135 个, 其中 G-box 元件有 38 个, 含有 28 个 Box 4 元件, 暗示 *EuTCPs* 基因的转录可能受光周期调控。*EuTCPs* 启动子区域含有 44 个 ABRE 元件和 25 个 ARE 元件(图 5-B), 表明 *EuTCPs* 可能参与 ABA 调节和厌氧调控。



A-利用 MEGA 6.0 的邻接法构建 EuTCP 基因家族系统进化树 B-EuTCPs 基因结构分析 C-保守基序氨基酸分布
A-the phylogenetic tree of EuTCP gene family was generated with MEGA 6.0 software using the neighbour-joining method B-the gene structure analysis of EuTCPs C-conserved motif amino acid distribution.

图 4 EuTCPs 蛋白保守基序分析

Fig. 4 Conservative motif analysis of EuTCP proteins



A-每个 EuTCP 基因启动子区域顺式作用元件的数目 B-顺式作用元件数量和含有顺式作用元件 EuTCPs 基因数量
A-the number of cis-acting elements in the promoter region of each EuTCP gene B-the number of cis-elements and the EuTCP genes containing cis-acting elements

图 5 EuTCPs 启动子顺式作用元件分布

Fig. 5 Cis-elements distributed in promoters of EuTCPs

2.7 杜仲 TCP 基因家族表达模式分析

为了解析 *EuTCPs* 基因在杜仲叶片不同发育时期及胶形成中的功能, 根据转录组数据, 检测 *EuTCPs* 基因的表达模式(图中显示基因编号与文章中基因命名一致)。结果如图 6、7 所示, 不同 *EuTCPs* 基因在杜仲叶片不同发育时期及胶含量中表达丰度存在显著差异, 大部分 *EuTCPs* 基因在杜仲叶片发育时期表达丰度较低, 有 2 个 *EuTCPs* 基因(*EuTCP11* 和 *EuTCP14*) 在杜仲叶片中未检测到, 推测 *EuTCP11* 和 *EuTCP14* 在杜仲叶片中不发挥作用。*EuTCPs* 基因在杜仲叶片不

同发育时期表达趋势不尽相同, 暗示 *EuTCPs* 在杜仲叶片发育中具有不同的功能, 5 个基因(*EuTCP1*、*EuTCP7*、*EuTCP8*、*EuTCP9*、*EuTCP13*) 在杜仲叶芽中表达量最高, 4 个(*EuTCP2*、*EuTCP3*、*EuTCP6*、*EuTCP12*) 基因在初生叶中表达丰度最高, 3 个基因(*EuTCP4*、*EuTCP5* 和 *EuTCP10*) 在老叶中转录水平最高。

为了预测 *EuTCPs* 在杜仲胶形成中的功能, 利用不同胶含量的转录组数据, 检测 *EuTCPs* 基因的转录水平。结果如图 7 所示, 大部分 *EuTCPs* 在杜仲不同胶含量中表达水平较低, 并且低胶含量叶片

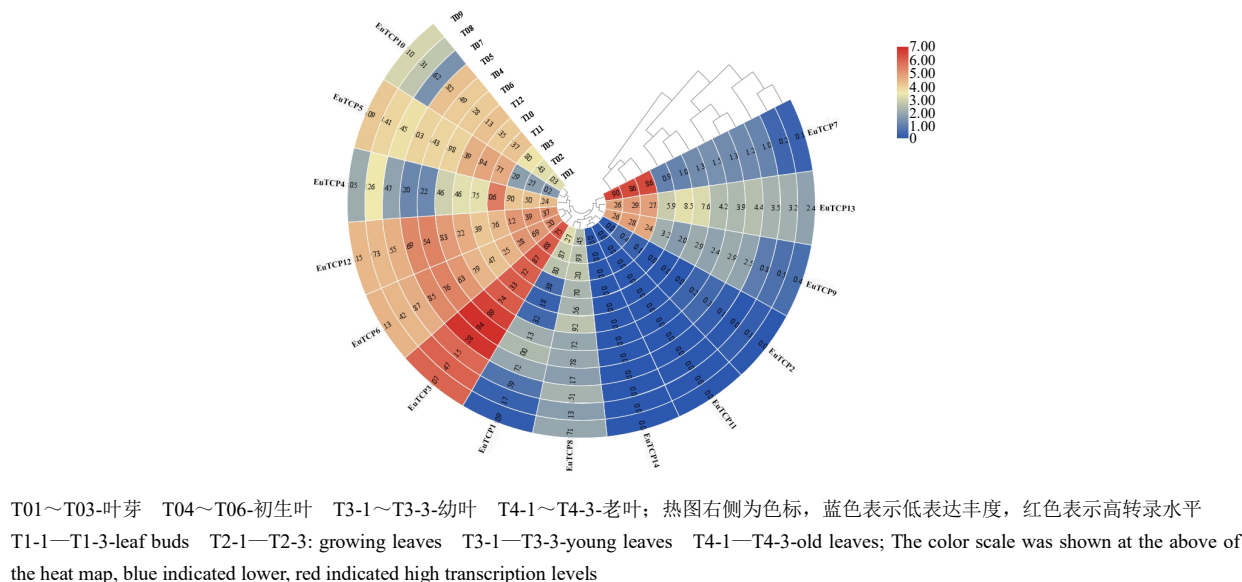


图 6 *EuTCPs* 基因在杜仲叶片不同发育时期的表达模式

Fig. 6 Expression patterns of *EuTCPs* genes at different development stages of *E. ulmoides* leaves.

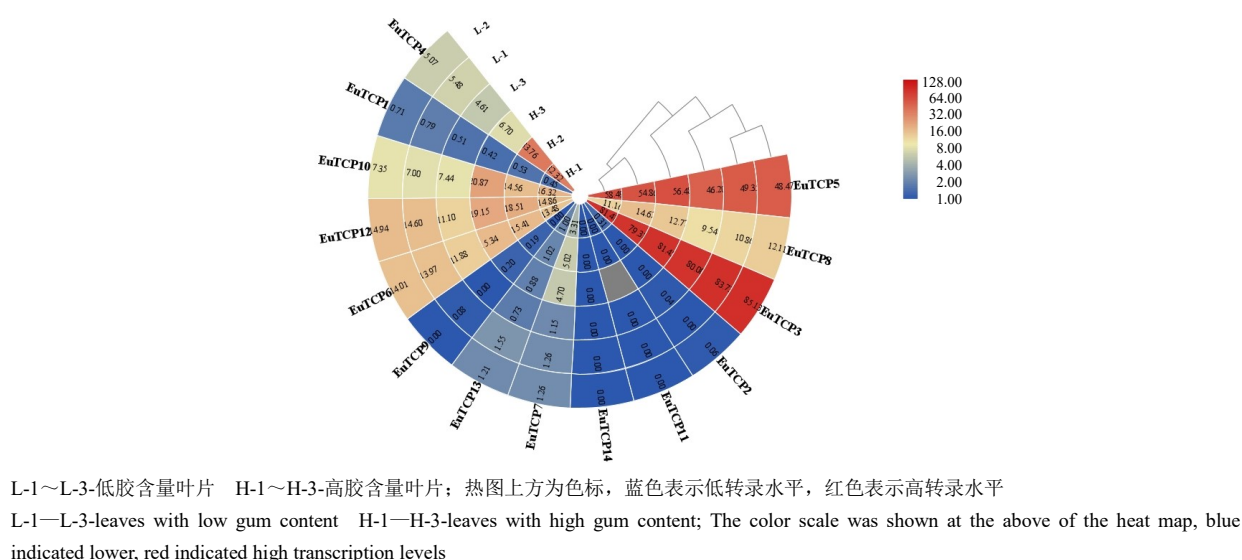


图 7 *EuTCPs* 基因在杜仲不同胶含量叶片中的表达模式

Fig. 7 Expression pattern of *EuTCPs* genes in leaves with different gum content

中的转录水平低于高胶含量叶片,例如: *EuTCP10* 在高胶含量叶片中的 FPKM 平均值是 17.25, 在低胶含量叶片中为 7.26, *EuTCP4* 在高胶含量中的转录水平是 24.26, 在低胶中为 5.05, 推测 *EuTCPs* 在杜仲胶形成过程中发挥正调控作用。

2.8 杜仲 TCPs 蛋白互作网络预测

表达模式结果显示, *EuTCP3* 在杜仲叶片不同发育阶段和胶含量中表达量均较高, 特别在初生叶中, FPKM 值为 111, 暗示 *EuTCP3* 在杜仲叶片发育和胶形成过程中发挥重要作用。为了检测 *EuTCP3* 与其他蛋白的互作关系, 利用 STRING 数据库, 预测 *EuTCP3* 的互作蛋白。结果如图 8 所示, *EuTCP3* 可以与 10 个蛋白发生相互作用, 其中 7 个属于 *EuTCP* 基因家族, 参与种子萌发、生物钟调节、细胞周期以及蛋白转录调控。MYB-like 转录因子, 在发育种子的原花青素积累中发挥关键的的决定性作用, LOB 转录因子 AS2, 是叶片正面细胞增殖负调控因子, 调节叶片对称层的形成和脉络的建立, AN1-like (SAP11) 锌指蛋白, 参与植物非生物胁迫响应^[27], 推测 *EuTCP3* 可能与其他蛋白相互作用共同参与杜仲叶片生长发育以及胁迫响应。

利用 Cytoscape 3.7.0 软件, 构建 *EuTCP* 转录因子家族共表达调控网络。结果如图 9 所示, 调控网络中共有 11 个节点 (代表 *EuTCPs* 蛋白) 和 35 条边 (代表蛋白质之间的相互作用), 表明 14 个 *EuTCPs* 中有 11 个蛋白存在互作关系, 其中 *EuTCP3*、*EuTCP5* 和 *EuTCP8* 蛋白互作数量最多, 均与 8 个 *EuTCPs* 蛋白存在互作关系, 暗示 *EuTCP3*、*EuTCP5*、*EuTCP8* 可能处于 *EuTCP* 基因家族核心位置, *EuTCP1*、*EuTCP10*、*EuTCP11* 和 *EuTCP14* 与 7 个 *EuTCPs* 蛋白互作, *EuTCP4* 和 *EuTCP6* 预测与 5 个 *EuTCPs* 蛋白互作, *EuTCP7* 与 4 个 *EuTCPs* 蛋白互作, *EuTCP9* 与 2 个 *EuTCPs* 蛋白发生互作关系。

3 讨论

杜仲是绞木目 (Garryales) 杜仲科 (Eucommiaceae) 杜仲属 *Eucommia* L. 中唯一现存的一个种, 是我国主要天然橡胶树种, 也是重要的木本油料和名贵中药材^[28], 药理作用广泛^[29], 主要分布于我国的陕西、湖南、贵州、云南、四川、江西等地区^[30]。TCP 蛋白是植物特有的参与生长发育、非生物胁迫、叶片衰老、侧枝形成、花粉发育等多种生物学过程的一类转录因子^[13,31]。目前为止, TCP

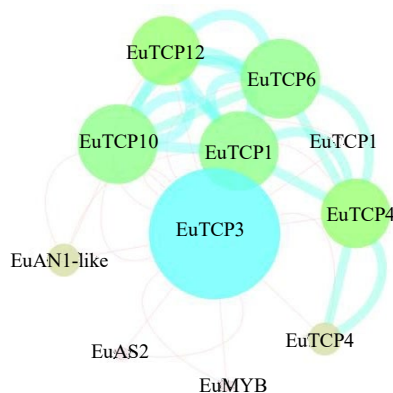


图 8 *EuTCP3* 蛋白互作网络

Fig. 8 Interaction network of *EuTCP3*

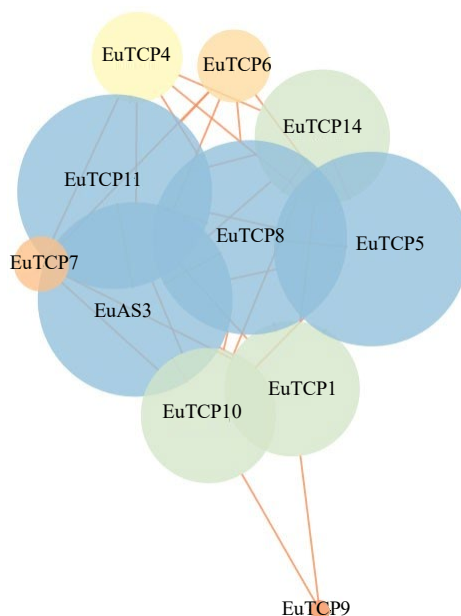


图 9 *EuTCP* 基因家族蛋白互作预测

Fig. 9 Prediction of protein interactions between *EuTCP* gene family

基因家族已在多个物种中进行了报道, 拟南芥中有 24 个 *AtTCPs*^[25], 水稻含有 22 个 *OsTCPs*^[24], 玉米中存在 29 个 *ZmTCPs* 成员^[32], 棉花中包含 38 个 *GrTCPs*^[33], 番茄中有 30 个 *SITCPs*^[34], 黄瓜中存在 22 个 *CsTCPs*^[35], 柑橘中含有 20 个 *CsTCPs*^[36], 西瓜中有 27 个 *CITCPs*^[37], 葡萄中包含 18 个 *VvTCPs*^[38], 苹果中 52 个 *MdTCPs*^[39], 菠萝中有 9 个^[40]、毛竹中存在 19 个 *PheTCPs*^[26]等。本研究以杜仲基因组数据为基础, 利用生物信息学方法, 从杜仲基因组中共鉴定到 14 个 *EuTCPs* 基因, 分为 3 个亚家族, 这与人研究结果相一致。PCF 亚家族包含 5 个 *EuTCPs* 蛋白, CYC/TB1 亚家族含有 3 个,

6 个 *EuTCPs* 蛋白属于 CIN 亚家族。序列比对结果显示, 14 个 *EuTCPs* 与拟南芥、水稻 TCP 蛋白结构域具有高度的相似性, N 末端均含有非典型的 basic-helix-loop-helix 结构域(图 3), 基序分析发现, Motif 1 基序存在于所有 *EuTCPs* 转录因子(图 4), 表明他们均属于 TCP 转录因子家族, 在基因进化过程中具有较保守性。根据保守结构域是否含有 4 种氨基酸残基, 将 TCP 基因家族分为 2 个亚家族, Class I 亚家族保守结构域含有 16 个氨基酸残基, Class II 亚家族 TCP 结构域包含 20 个氨基酸, 保守结构域外含有促进蛋白质相互作用的精氨酸富集 R 结构域^[1], *EuTCP3*、*EuTCP5*、*EuTCP7* 和 *EuTCP9* 均含有 motif 8 (R 结构域), 与 *AtTCPs*、*OsTCPs*^[24]、*BpTCPs*^[41]、*VvTCPs*^[37]、*PmTCPs*^[42] 和 *FvTCPs*^[43] 等具有相似性, 推测 *EuTCP3*、*EuTCP5*、*EuTCP7* 和 *EuTCP9* 可能具有相似的功能。

聚类分析结果显示, *EuTCP4* 与 *AtTCP20* 聚为同一分支, *AtTCP20* 参与调节拟南芥细胞扩张、分裂和分化, 在硝酸盐诱导的系统信号通路中发挥重要作用^[44], 菊花 *CmTCP20* 为 *AtTCP20* 的同源基因, 超表达 *CmTCP20* 增加转基因菊花和拟南芥的侧根数量和长度^[45]。*EuTCP10* 是 *AtTCP14* 的同源基因, *AtTCP14* 主要在胚胎维管组织中表达, 促进幼嫩节间细胞增殖, 抑制叶片细胞扩增, *attcp14* 突变体提高种子萌发对外源脱落酸和多效唑的敏感性, *AtTCP14* 参与激活拟南芥胚胎生长^[46], *CmTCP14* 异源转化拟南芥导致细胞周期相关基因下调表达, 抑制器官大小, 叶绿素和类胡萝卜素含量升高, 延长花期, 延缓叶片衰老^[47]。*EuTCP6* 和 *EuTCP12* 聚为同一分支的 *AtTCP7*、*AtTCP8*、*AtTCP21*、*AtTCP22* 和 *AtTCP23* 均在拟南芥幼嫩叶片中表达, 与野生型相比, *attcp23* 突变体增加叶片面积, *tcp8-1*、*tcp15-1*、*tcp22-1*、*tcp23-1* 四重突变体和 *tcp8-1*、*tcp15-1*、*tcp22-1*、*tcp23-1*、*tcp21-1* 五重突变体莲座叶数目减少, 叶片增大^[48]。*EuTCP14* 为 *AtTCP4* 的同源基因, *AtTCP4* 在幼苗子叶中表达, 激活 *SAUR16* 和 *SAUR50* 的光响应表达, 通过光敏色素相互作用因子 (PIFs) 调节 *AtTCP4* 与 *SAUR14*、*SAUR16* 和 *SAUR50* 启动子的结合, 黑暗条件下 *PIF3* 抑制 *AtTCP4* 与上述启动子的结合^[49]。*AtTCP4* 与 *CONSTANS (CO)* 启动子结合, 黄昏前后诱导 *CO* 基因表达, *AtTCP4* 与细胞核中另一个开花调节因子 *GIGANTEA (GI)* 结合, 诱导 *CO* 以 *GI* 依赖方式表达, 调节发育和开花调控^[50]。

EuTCP3 与 *AtTCP2* 互为同源基因, 超表达 *AtTCP2* 导致转基因拟南芥在光依赖条件下下胚轴矮小, 特别是蓝光条件下, *AtTCP2* 与隐花色素蛋白 (CRY1) 互作, 正调控 *HYH* 和 *HY5* 基因表达^[51]。*AtTCP17* 与 *EuTCP3* 亲缘关系最近, *AtTCP17* 在阴暗条件下显著积累, *AtTCP17* 及其同源基因 *AtTCP5* 和 *AtTCP13* 通过 PIF 依赖和独立途径诱导生长素生物合成, 在介导阴暗诱导下胚轴伸长中发挥重要作用。白光下, *tcp5*、*tcp13*、*tcp17* 三重突变体 (*3tcp*) 表现出下胚轴缺陷; 阴暗条件显著降低 *3tcp* 下胚轴伸长, *TCP17* 与 *PIFs* 和 *YUCCAs* 的启动子结合, 诱导生长素表达^[52], 推测 *EuTCPs* 转录因子在杜仲生长发育、激素调节和开花调控中发挥重要作用。

综上所述, 本研究以杜仲全基因组数据为背景, 对杜仲 TCP 转录因子家族进行了全面的生物信息学分析, 共鉴定出 14 个 *EuTCPs*, 分为 Class I 与 Class II 2 大类, 进一步划分为 3 个亚家族 (PCF、CYC/TB1 和 CIN)。理化性质分析显示, *EuTCPs* 编码 139~492 个氨基酸, 理论等电点分布于 5.54~9.72, 相对分子质量为 18 880~53 620, 亚细胞定位于细胞核中, 均为亲水性蛋白, 二级结构以 α -螺旋和无规则卷曲为主, 均含非典型 bHLH 保守结构域。表达模式分析显示, *EuTCPs* 基因家族在杜仲叶片发育中具有不同的表达模式, *EuTCPs* 在高胶含量中的转录水平高于低胶含量, 推测 *EuTCPs* 在杜仲胶形成中发挥正向调控作用, 本研究为深入探索 *EuTCPs* 基因功能奠定理论基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Cubas P, Lauter N, Doebley J, et al. The TCP domain: A motif found in proteins regulating plant growth and development [J]. *Plant J*, 1999, 18(2): 215-222.
- [2] Martín-Trillo M, Cubas P. TCP genes: A family snapshot ten years later [J]. *Trends Plant Sci*, 2010, 15(1): 31-39.
- [3] Takeda T, Suwa Y, Suzuki M, et al. The OsTB1 gene negatively regulates lateral branching in rice [J]. *Plant J*, 2003, 33(3): 513-520.
- [4] Maurya J P, Singh R K, Miskolczi P C, et al. Branching regulator BRC1 mediates photoperiodic control of seasonal growth in hybrid aspen [J]. *Curr Biol*, 2020, 30(1): 122-126.
- [5] Wang M, le Moigne M A, Bertheloot J, et al. BRANCHED1: A key hub of shoot branching [J]. *Front Plant Sci*, 2019, 10: 76.

- [6] Yang X, Zhao X G, Li C Q, *et al.* Distinct regulatory changes underlying differential expression of teosinte branched1-cycloidea-proliferating cell factor genes associated with petal variations in zygomorphic flowers of *Petrocosmea* spp. of the family Gesneriaceae [J]. *Plant Physiol*, 2015, 169(3): 2138-2151.
- [7] Guan P Z, Ripoll J J, Wang R H, *et al.* Interacting TCP and NLP transcription factors control plant responses to nitrate availability [J]. *Proc Natl Acad Sci USA*, 2017, 114(9): 2419-2424.
- [8] Finlayson S A. *Arabidopsis* Teosinte Branched1-like 1 regulates axillary bud outgrowth and is homologous to monocot Teosinte Branched1 [J]. *Plant Cell Physiol*, 2007, 48(5): 667-677.
- [9] Danisman S, van der Wal F, Dhondt S, *et al.* *Arabidopsis* class I and class II TCP transcription factors regulate jasmonic acid metabolism and leaf development antagonistically [J]. *Plant Physiol*, 2012, 159(4): 1511-1523.
- [10] Rubio-Somoza I, Weigel D. Coordination of flower maturation by a regulatory circuit of three microRNAs [J]. *PLoS Genet*, 2013, 9(3): e1003374.
- [11] Challa K R, Aggarwal P, Nath U. Activation of YUCCA5 by the transcription factor TCP4 integrates developmental and environmental signals to promote hypocotyl elongation in *Arabidopsis* [J]. *Plant Cell*, 2016, 28(9): 2117-2130.
- [12] Rueda-Romero P, Barrero-Sicilia C, Gómez-Cadenas A, *et al.* *Arabidopsis thaliana* DOF₆ negatively affects germination in non-after-ripened seeds and interacts with TCP14 [J]. *J Exp Bot*, 2012, 63(5): 1937-1949.
- [13] Liu H L, Gao Y M, Wu M, *et al.* TCP10, a TCP transcription factor in moso bamboo (*Phyllostachys edulis*), confers drought tolerance to transgenic plants [J]. *Environ Exp Bot*, 2020, 172: 104002.
- [14] Almeida D M, Gregorio G B, Oliveira M M, *et al.* Five novel transcription factors as potential regulators of OsNHX1 gene expression in a salt tolerant rice genotype [J]. *Plant Mol Biol*, 2017, 93(1/2): 61-77.
- [15] Wang S T, Sun X L, Hoshino Y, *et al.* microRNA319 positively regulates cold tolerance by targeting OsPCF₆ and OsTCP21 in rice (*Oryza sativa* L.) [J]. *PLoS One*, 2014, 9(3): e91357.
- [16] Mukhopadhyay P, Tyagi A K. Erratum: OsTCP19 influences developmental and abiotic stress signaling by modulating ABI4-mediated pathways [J]. *Sci Rep*, 2015, 5: 12381.
- [17] 杜红岩. 中国杜仲图志 [M]. 北京: 中国林业出版社, 2014.
- [18] 刘建斌. 杜仲种子的化学成分研究 [D]. 西安: 西北大学, 2006.
- [19] Wuyun T N, Wang L, Liu H M, *et al.* The hardy rubber tree genome provides insights into the evolution of polyisoprene biosynthesis [J]. *Mol Plant*, 2018, 11(3): 429-442.
- [20] Tamura K, Stecher G, Peterson D, *et al.* MEGA6: molecular evolutionary genetics analysis version 6.0 [J]. *Mol Biol Evol*, 2013, 30(12): 2725-2729.
- [21] Li L, Liu M H, Shi K, *et al.* Dynamic changes in metabolite accumulation and the transcriptome during leaf growth and development in *Eucommia ulmoides* [J]. *Int J Mol Sci*, 2019, 20(16): 4030.
- [22] Ye J, Han W J, Fan R S, *et al.* Integration of transcriptomes, small RNAs, and degradome sequencing to identify putative miRNAs and their targets related to Eu-rubber biosynthesis in *Eucommia ulmoides* [J]. *Genes*, 2019, 10(8): 623.
- [23] Liu J, Cheng Z C, Li X Y, *et al.* Expression analysis and regulation network identification of the *CONSTANS*-like gene family in moso bamboo (*Phyllostachys edulis*) under photoperiod treatments [J]. *DNA Cell Biol*, 2019, 38(7): 607-626.
- [24] Yao X, Ma H, Wang J, *et al.* Genome-wide comparative analysis and expression pattern of TCP gene families in *Arabidopsis thaliana* and *Oryza sativa* [J]. *J Integr Plant Biol*, 2007, 49(6): 885-897.
- [25] Cubas P. *Role of TCP genes in the evolution of morphological characters in angiosperms* [M] CRC Press: Systematics Association Special Volumes, 2002: 247-266.
- [26] 刘俊, 黄容, 程占超, 等. 毛竹 TCP 基因家族全基因组鉴定与分析 [J]. 基因组学与应用生物学, 2018, 37(12): 5388-5397.
- [27] Giri J, Vij S, Dansana P K, *et al.* Rice A20/AN₁ zinc-finger containing stress-associated proteins (SAP1/11) and a receptor-like cytoplasmic kinase (OsRLCK253) interact via A20 zinc-finger and confer abiotic stress tolerance in transgenic *Arabidopsis* plants [J]. *New Phytol*, 2011, 191(3): 721-732.
- [28] Wang L, Jing T, Li T Z, *et al.* Identification and expression analysis of the *Eucommia ulmoides* farnesyl diphosphate synthase gene family to reveal the key gene involved in rubber biosynthesis [J]. *Acta Physiol Plant*, 2017, 40(1): 1-5.
- [29] 柴士伟, 马静, 庄朋伟, 等. 杜仲有效成分在帕金森病治疗中的药理作用研究进展 [J]. 药物评价研究, 2019, 42(5): 1003-1009.

- [30] 张荣荣. 土壤干旱下杜仲截干苗生长和光合生理响应及引种栽培研究 [D]. 太谷: 山西农业大学, 2016.
- [31] Manassero N G, Viola I L, Welchen E, *et al.* TCP transcription factors: Architectures of plant form [J]. *Biomol Concepts*, 2013, 4(2): 111-127.
- [32] Chai W B, Jiang P F, Huang G Y, *et al.* Identification and expression profiling analysis of TCP family genes involved in growth and development in maize [J]. *Physiol Mol Biol Plants*, 2017, 23(4): 779-791.
- [33] Ma J, Wang Q L, Sun R R, *et al.* Genome-wide identification and expression analysis of TCP transcription factors in *Gossypium raimondii* [J]. *Sci Rep*, 2014, 4: 6645.
- [34] Parapunova V, Busscher M, Busscher-Lange J, *et al.* Identification, cloning and characterization of the tomato TCP transcription factor family [J]. *BMC Plant Biol*, 2014, 14: 157.
- [35] 袁伟, 顾掌根, 褚伟雄, 等. 黄瓜转录因子 TCP 基因家族的鉴定及结构分析 [J]. 分子植物育种, 2014, 12(2): 287-295.
- [36] 周延培, 张雅剑, 伊华林. 柑橘 TCP 家族生物信息学及表达谱分析 [J]. 果树学报, 2016, 33(5): 513-522.
- [37] Shi P B, Guy K M, Wu W F, *et al.* Genome-wide identification and expression analysis of the CITCP transcription factors in *Citrullus lanatus* [J]. *BMC Plant Biol*, 2016, 16: 85.
- [38] Leng X P, Wei H R, Xu X Z, *et al.* Genome-wide identification and transcript analysis of TCP transcription factors in grapevine [J]. *BMC Genomics*, 2019, 20(1): 786.
- [39] 张士刚. 苹果 TCP 转录因子家族生物信息学分析 [J]. 山东农业科学, 2014, 46(5): 12-17.
- [40] 阮城城, 胡福初, 罗志文, 等. 菠萝 TCP 基因家族的鉴定及成花诱导阶段的表达谱分析 [J]. 果树学报, 2020, 37(11): 1623-1635.
- [41] Zhao M L, Peng X J, Chen N Z, *et al.* Genome-wide identification of the TCP gene family in *Broussonetia papyrifera* and functional analysis of BpTCP8, 14 and 19 in shoot branching [J]. *Plants (Basel)*, 2020, 9(10): 1301.
- [42] Zhou Y Z, Xu Z D, Zhao K, *et al.* Genome-wide identification, characterization and expression analysis of the TCP gene family in *Prunus mume* [J]. *Front Plant Sci*, 2016, 7: 1301.
- [43] Wei W, Hu Y, Cui M Y, *et al.* Identification and transcript analysis of the TCP transcription factors in the diploid woodland strawberry *Fragaria vesca* [J]. *Front Plant Sci*, 2016, 7: 1937.
- [44] Guan P Z, Wang R C, Nacry P, *et al.* Nitrate foraging by *Arabidopsis* roots is mediated by the transcription factor TCP20 through the systemic signaling pathway [J]. *Proc Natl Acad Sci USA*, 2014, 111(42): 15267-15272.
- [45] Fan H M, Sun C H, Wen L Z, *et al.* CmTCP20 plays a key role in nitrate and auxin signaling-regulated lateral root development in *Chrysanthemum* [J]. *Plant Cell Physiol*, 2019, 60(7): 1581-1594.
- [46] Tatematsu K, Nakabayashi K, Kamiya Y, *et al.* Transcription factor AtTCP14 regulates embryonic growth potential during seed germination in *Arabidopsis thaliana* [J]. *Plant J*, 2008, 53(1): 42-52.
- [47] Zhang T, Qu Y X, Wang H B, *et al.* The heterologous expression of a *Chrysanthemum* TCP-P transcription factor CmTCP14 suppresses organ size and delays senescence in *Arabidopsis thaliana* [J]. *Plant Physiol Biochem*, 2017, 115: 239-248.
- [48] Aguilar-Martínez J A, Sinha N. Analysis of the role of *Arabidopsis* class I TCP genes AtTCP7, AtTCP8, AtTCP22, and AtTCP23 in leaf development [J]. *Front Plant Sci*, 2013, 4: 406.
- [49] Dong J, Sun N, Yang J, *et al.* The transcription factors TCP4 and PIF3 antagonistically regulate organ-specific light induction of SAUR genes to modulate *Cotyledon* opening during de-etiolation in *Arabidopsis* [J]. *Plant Cell*, 2019, 31(5): 1155-1170.
- [50] Kubota A, Ito S, Shim J S, *et al.* TCP4-dependent induction of CONSTANS transcription requires *GIGANTEA* in photoperiodic flowering in *Arabidopsis* [J]. *PLoS Genet*, 2017, 13(6): e1006856.
- [51] He Z M, Zhao X Y, Kong F N, *et al.* TCP2 positively regulates HY5/HYH and photomorphogenesis in *Arabidopsis* [J]. *J Exp Bot*, 2016, 67(3): 775-785.
- [52] Zhou Y, Zhang D Z, An J X, *et al.* TCP transcription factors regulate shade avoidance via directly mediating the expression of both *PHYTOCHROME INTERACTING FACTORS* and auxin biosynthetic genes [J]. *Plant Physiology*, 2018, 176(2): 1850-1861.

[责任编辑 时圣明]