

任昂彦, 陈 涛, 李晓莎, 等. 向日葵 *COBRA* 家族基因的全基因组鉴定与分析[J]. 江苏农业学报, 2026, 42(7): 1338-1348.
doi: 10.3969/j.issn.1000-4440.2026.07.005

向日葵 *COBRA* 家族基因的全基因组鉴定与分析

任昂彦¹, 陈 涛¹, 李晓莎¹, 贺 江¹, 杨 华¹, 高 进¹, 刘爱萍², 陈环宇¹, 戚永奎¹
(1. 江苏沿海地区农业科学研究所, 江苏 盐城 224002; 2. 建湖县芦沟镇农业农村办公室, 江苏 建湖 224711)

摘要: *COBRA* 蛋白是一类糖基磷脂酰肌醇(GPI)锚定蛋白, 在植物细胞壁生物合成、细胞分裂、细胞扩张和防御反应等过程中发挥重要作用。本研究在向日葵全基因组数据中鉴定出不均匀分布在 11 条染色体上的 16 个 *COBRA* 基因家族成员。基于生物信息学分析, 向日葵 *COBRA* 家族基因可划分为 2 个亚家族, 二者在结构特征上存在显著差异, 包括基因结构以及编码的 *COBRA* 蛋白关键保守结构域(半胱氨酸-半胱氨酸-缬氨酸-丝氨酸, CCVS)变异、亲疏水模式分化、GPI 锚定修饰位点分化等。启动子顺式作用元件分析发现, 该家族基因均含 2 种以上参与逆境胁迫响应、激素信号传导或发育调控相关的顺式作用元件。蛋白质互作预测结果显示, HaCOBRA10 蛋白和 HaCOBRA14 蛋白均与 2 个纤维素合成酶蛋白(CESA4、CESA8)及 2 个糖基转移酶蛋白(IRX3、IRX9)互作。组织差异表达和胁迫响应分析发现, 该家族基因在生长发育调控和逆境胁迫响应中可能具有功能多样性。本研究结果为向日葵细胞壁形成机制研究及抗逆育种提供了新靶点, 并为进一步探究 *COBRA* 蛋白在植物中的生物学功能奠定了理论基础。

关键词: 向日葵; *COBRA* 蛋白; 纤维素; 细胞壁

中图分类号: S565.5 **文献标识码:** A **文章编号:** 1000-4440(2026)07-1338-11

Genome-wide identification and analysis of *COBRA* family genes in sunflower

REN Angyan¹, CHEN Tao¹, LI Xiaosha¹, HE Jiang¹, YANG Hua¹, GAO Jin¹, LIU Aiping²,
CHEN Huanyu¹, QI Yongkui¹

(1. Institute of Agricultural Sciences in the Coastal District of Jiangsu Province, Yancheng 224002, China; 2. Agriculture and Rural Affairs Office of Lugou Town, Jianhu County, Jianhu 224711, China)

Abstract: *COBRA* proteins are a class of glycosyl-phosphatidylinositol (GPI)-anchored proteins that play important roles in plant cell wall biosynthesis, cell division, cell expansion, and defense responses. In this study, 16 *COBRA* genes unevenly distributed across 11 chromosomes were identified in the sunflower whole-genome. Based on bioinformatics analysis, this family was classified into two subfamilies, with significant differences in structural features, including gene structure, variations in the key conserved domain (cysteine-cysteine-valine-serine, CCVS), divergence in hydrophobicity/hydrophilicity patterns, and divergence in GPI-anchoring modification sites. Analysis of promoter *cis*-acting elements revealed that all mem-

bers contained at least two types of *cis*-acting elements involved in stress response, hormone signaling, or developmental regulation. The protein-protein interaction prediction results showed that both HaCOBRA10 and HaCOBRA14 interacted with two cellulose synthases (CESA4, CESA8) and two glycosyltransferases (IRX3, IRX9). Tissue-specific expression and stress response analyses suggested that genes in this family may exhibit functional diversity in growth, development, and stress adaptation. This

收稿日期: 2025-07-30

基金项目: 江苏省盐城市应用基础研究计划项目(YCBK2025063、YCBK2025071); 农业农村部沿海盐碱地农业科学观测实验站开放课题(YHS202105、YHS202306); 江苏省农业科技自主创新基金项目[CX(23)1040]; 江苏省基础研究计划青年基金项目(BK20240495)

作者简介: 任昂彦(1989-), 女, 山东菏泽人, 博士, 助理研究员, 主要从事经济作物遗传育种研究。(E-mail) ray0918@163.com

通讯作者: 戚永奎, (E-mail) qiyongkuiworld@sina.cn

study provides new targets for research on cell wall formation mechanism and stress resistance breeding in sunflower, and lays a theoretical foundation for further investigation into the biological functions of COBRA proteins in plants.

Key words: sunflower; COBRA protein; cellulose; cell wall

植物细胞壁作为动态基质,不仅能为细胞提供机械支撑,还可通过调控细胞扩张速率与方向,直接影响细胞形态建成。纤维素微纤丝在细胞壁中的沉积模式与排列方向决定细胞的形态特征及扩张取向^[1-2]。COBRA 是一类定位于细胞膜的 GPI 锚定蛋白,核心功能是调控植物细胞壁纤维素微纤丝的沉积方向与细胞定向伸长。高等植物 COBRA 蛋白结构较为保守,其 C 端包含 GPI 修饰位点(ω -site),该位点在纤维素的组装与结晶过程中至关重要^[3]。此外,该类蛋白质还具有 N 端分泌信号肽、中间亲水区、潜在 N-糖基化位点以及保守的半胱氨酸-半胱氨酸-缬氨酸-丝氨酸(CCVS)结构域等典型特征^[4]。目前,COBRA 蛋白已在拟南芥^[5]、玉米^[6]、水稻^[7]、谷子^[8]、番茄^[9]等多种植物中被鉴定和报道。

大量研究结果表明,COBRA 蛋白在植物纤维素合成过程中发挥关键作用,同时可间接影响植株的机械强度。拟南芥 COBRA 蛋白缺失会导致细胞扩张受阻、纤维素结晶度下降及细胞壁沉积异常^[5],并引发花粉管伸长缺陷^[10]。水稻 *bcl* 突变体茎秆纤维素含量降幅达 80%,导致茎秆机械强度显著下降^[7];玉米 *Bk2* 通过调控茎秆纤维素含量,影响植株机械支撑强度^[11];高粱 COBRA 家族基因 *BC1* 基因也被证实参与纤维素合成,其突变体同样表现出茎秆脆性增加、易折断的表型^[12]。近期研究还发现,该家族基因编码的蛋白质还参与植物逆境防御响应^[5,13]。

向日葵(*Helianthus annuus* L.)作为重要的油料作物,其茎秆纤维素含量丰富,占秸秆干质量的32%~39%^[14-15],占茎髓干质量达50%^[16],是生产燃料乙醇、吸附材料、纳米纤维素等物质的优质原料。目前关于向日葵 COBRA 蛋白的鉴定及功能研究尚未见报道,本研究针对向日葵 COBRA 家族基因进行全基因组鉴定与系统分析,为深入解析向日葵纤维素合成与沉积机制以及抗逆育种提供理论基础。

1 材料与方法

1.1 向日葵 COBRA 家族基因的全基因组鉴定及染色体定位

以 12 条拟南芥 COBRA 家族蛋白质氨基酸序

列为检索序列,在向日葵基因组数据库 IARN (<https://www.heliogene.org/>) 执行 PlastP 获取同源蛋白质氨基酸序列。经去除重复后,利用 Pfam (<http://pfam-legacy.xfam.org/>)^[17]、SMART (<http://smart.embl-heidelberg.de/>)^[18] 和 NCBI CDD (<https://www.ncbi.nlm.nih.gov/cdd/>) 等在线蛋白质结构域数据库与分析平台,对候选序列进行保守结构域预测,筛选包含 COBRA 保守结构域(PF04833)的蛋白质氨基酸序列,最终鉴定为向日葵 COBRA 家族蛋白成员(HaCOBRA)。

从向日葵基因组数据库获取 17 条染色体长度及候选基因染色体定位信息,利用 Map Gene 2 Chromosome v2 (http://mg2c.iask.in/mg2c_v2.0/) 在线绘制 HaCOBRA 基因的染色体定位图谱,并根据染色体定位信息将筛选出的该家族基因命名为 HaCOBRA1~HaCOBRA16。利用 EMBOSS (<https://www.ebi.ac.uk/Tools/psa/>) 进行核苷酸序列相似性分析。将同一条染色体上 100 kb 区域内的同源基因定义为基因簇;若序列相似性超过 70%,则判定为串联重复基因簇^[19-20]。

1.2 COBRA 蛋白理化性质、亚细胞定位及 CCVS 结构保守性分析

利用 Expasy 在线平台 (<https://www.expasy.org/>) 预测 COBRA 蛋白的等电点、相对分子量等理化性质^[21],利用 SignalP-5.0 (<http://www.cbs.dtu.dk/services/SignalP/>) 在线预测信号肽位点,利用 Plant-mPLoc (<http://www.csbio.sjtu.edu.cn/bioinf/plant/>) 预测蛋白质亚细胞定位^[22]。利用 MULTALIN (<http://multalin.toulouse.inra.fr/multalin/>) 对 HaCOBRA 蛋白氨基酸序列进行多重序列比对,并分析 CCVS 保守结构域的序列保守性。

1.3 进化关系分析、保守基序和基因结构分析

分别从水稻 (<http://rice.plantbiology.msu.edu/>)、玉米 (<https://maizegdb.org/>)、谷子 (<https://phytozome.jgi.doe.gov/pz/portal.html>) 基因组数据库获取 3 个物种 COBRA 蛋白氨基酸序列,利用 ClustalW 将其与拟南芥和向日葵 COBRA 蛋白氨基酸序列进行多序列比对,并通过 MEGA 6.0 软件构建不同物种 COBRA 蛋白的系统

发育树;对除 HaCOBRA6 外的 15 条 HaCOBRA 蛋白氨基酸序列进行 ClustalW 比对,利用 MEGA 6.0 构建向日葵 COBRA 家族系统发育树。

基于 16 条 HaCOBRA 蛋白氨基酸序列,利用 MEME Suite 5.3.3 在线平台 (<https://meme-suite.org/meme/tools/meme>) 分析其保守基序,检索数量设为 6,其余参数为默认值。

在向日葵基因组数据库获取 HaCOBRA 基因组序列及编码氨基酸的核苷酸序列,利用 Gene Structure Display Server 2.0 (<https://gsds.gao-lab.org/>) 绘制向日葵外显子-内含子基因结构图^[23]。

1.4 蛋白质亲/疏水性拓扑和 GPI 修饰位点的预测

利用 Expasy 在线平台 ProtScale 工具 (<https://web.expasy.org/protscale/>), 选用 Kyte & Doolittle (Hydropath.) 方法,分析向日葵 16 条 COBRA 蛋白氨基酸序列的亲/疏水性。利用 GPI Lipid Anchor Project (https://mendel.imp.ac.at/gpi/gpi_server.html) 预测各蛋白质氨基酸序列中 GPI 修饰位点^[24]。

1.5 启动子顺式作用元件及蛋白质-蛋白质互作网络分析

基于向日葵基因组数据库,分别获取 16 个 HaCOBRA 基因起始密码子上游 1.5 kb 启动子核苷酸序列,利用 Plant-CARE 在线数据库 (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) 预测启动子区域顺式作用元件分布信息。

对向日葵 COBRA 家族的 16 个成员的氨基酸序列,通过共同和单独检索的方式,在 STRING 数据库 (<https://cn.string-db.org/>) 预测互作蛋白质^[25],预测蛋白质的功能注释来自 UniProtKB 数据库 (<https://www.uniprot.org/>)。

1.6 向日葵 COBRA 基因的表达模式分析

基于 SunExpress 公共数据库 (<https://www.heliogene.org/sunexpress.html>) 获取基因 HaCOBRA 组织表达模式及非生物胁迫响应模式下表达数据。组织表达谱包括茎、舌状花、管状花、雄蕊、雌蕊、成熟叶、种子、根;胁迫响应包括 NaCl、聚乙二醇 (PEG) 处理下的根和叶片组织。利用 MeV4.9.0 软件对基因表达量热图进行可视化。

2 结果与分析

2.1 向日葵 COBRA 家族基因鉴定及其基本特征

基于生物信息学技术和向日葵全基因组数据,

本研究共鉴定出 16 个向日葵 COBRA 基因家族成员,这些基因不均匀地分布在 11 条染色体上,根据染色体定位命名为 HaCOBRA1~HaCOBRA16 (表 1、图 1)。第 7 号染色体上有 5 个基因,其中 HaCOBRA7~HaCOBRA9 形成基因簇,但未形成串联重复基因簇。物理化学特性分析发现成员间差异较大,HaCOBRA5 蛋白氨基酸序列长度最短、相对分子量最小,HaCOBRA12 蛋白氨基酸序列长度最长、相对分子量最大。等电点变化范围相对较小,HaCOBRA16 蛋白等电点最低 (5.26),HaCOBRA12 蛋白等电点最高 (9.49)。经预测,12 个 HaCOBRA 蛋白成员氨基酸序列含有 N 端信号肽,4 个 HaCOBRA 蛋白 (HaCOBRA3、HaCOBRA5、HaCOBRA6、HaCOBRA11) 未检测到 N 端信号肽。脂肪族氨基酸指数差异较小,范围为 64.50~78.35,说明该家族蛋白质成员热稳定性较为一致。经蛋白质亚细胞定位预测,发现有 15 个成员定位于细胞膜,HaCOBRA6 同时定位于细胞膜和细胞核。上述结果表明,尽管向日葵 COBRA 家族蛋白质成员之间存在显著分化,但仍具备 COBRA 蛋白的典型特征,体现了该家族蛋白质在进化过程中的保守性。

2.2 向日葵 COBRA 蛋白家族系统发育及基因结构

基因家族起源于祖先基因的复制与分化,跨物种进化分析可揭示生命的起源与演化规律。为明确向日葵 COBRA 蛋白家族的系统发育地位及演化特征,本研究选取单子叶植物谷子、水稻、玉米和双子叶模式植物拟南芥 COBRA 蛋白家族成员氨基酸序列,与 15 个 HaCOBRA 蛋白氨基酸序列进行多系列比对 (因 HaCOBRA6 蛋白不能和其他蛋白质聚类,故在本研究中排除),并采用最大似然法构建系统发育树 (图 2)。从图 2 可见,54 个 COBRA 蛋白可明显划分为 Subfamily I 和 Subfamily II 两个亚家族,前者含 38 个成员,后者含 16 个成员;15 个 HaCOBRA 蛋白中有 11 个属于 Subfamily I,4 个属于 Subfamily II。从物种间进化关系来看,HaCOBRA 蛋白与双子叶植物拟南芥亲缘关系更近,而单子叶植物的 COBRA 蛋白则聚类为独立分支。此外,向日葵 COBRA 家族存在多个旁系同源基因对,暗示其在进化过程中可能发生基因复制事件。综上,供试 5 个物种中向日葵 COBRA 蛋白进化分支集中,为进一步解析该蛋白质家族的功能保守性与物种特异性提供了依据。

表 1 向日葵 COBRA 蛋白理化特性及亚细胞定位

Table 1 Physicochemical characteristics and subcellular localization of COBRA proteins in sunflower

蛋白质名称	基因号	COBRA 结构域(aa)	脂肪族氨 基酸指数	相对 分子量	等电点	氨基酸长度 (aa)	信号肽 (aa)	亚细胞定位
HaCOBRA1	HanXRQr2_Ch01g0044441	228~406	72.08	71 278.83	5.33	654	23~24	细胞膜
HaCOBRA2	HanXRQr2_Ch02g0061321	45~209	75.28	49 202.74	8.96	436	21~22	细胞膜
HaCOBRA3	HanXRQr2_Ch04g0156001	1~158	77.36	42 063.15	8.61	379	—	细胞膜
HaCOBRA4	HanXRQr2_Ch06g0243321	217~396	76.99	70 572.29	5.49	642	30~31	细胞膜
HaCOBRA5	HanXRQr2_Ch07g0309031	31~75	70.13	8 878.14	7.93	75	—	细胞膜
HaCOBRA6	HanXRQr2_Ch07g0309041	242~352	64.50	40 541.61	8.51	358	—	细胞膜和细胞核
HaCOBRA7	HanXRQr2_Ch07g0309691	56~219	69.38	50 503.95	8.89	451	32~33	细胞膜
HaCOBRA8	HanXRQr2_Ch07g0309701	56~219	78.35	27 707.99	8.84	255	32~33	细胞膜
HaCOBRA9	HanXRQr2_Ch07g0309711	56~222	67.56	35 862.93	8.29	320	32~33	细胞膜
HaCOBRA10	HanXRQr2_Ch09g0400221	43~206	69.49	48 260.74	8.98	430	19~20	细胞膜
HaCOBRA11	HanXRQr2_Ch10g0439551	47~209	73.45	47 859.09	6.71	426	—	细胞膜
HaCOBRA12	HanXRQr2_Ch11g0485261	224~403	72.53	74 342.37	9.49	656	21~22	细胞膜
HaCOBRA13	HanXRQr2_Ch14g0652471	56~219	69.64	50 486.96	8.94	449	32~33	细胞膜
HaCOBRA14	HanXRQr2_Ch15g0711681	43~206	69.56	48 213.76	9.21	430	19~20	细胞膜
HaCOBRA15	HanXRQr2_Ch15g0711691	55~218	64.80	50 551.83	8.92	448	31~32	细胞膜
HaCOBRA16	HanXRQr2_Ch16g0746491	223~401	76.57	70 690.42	5.26	647	17~18	细胞膜

-表示未检测到信号肽。

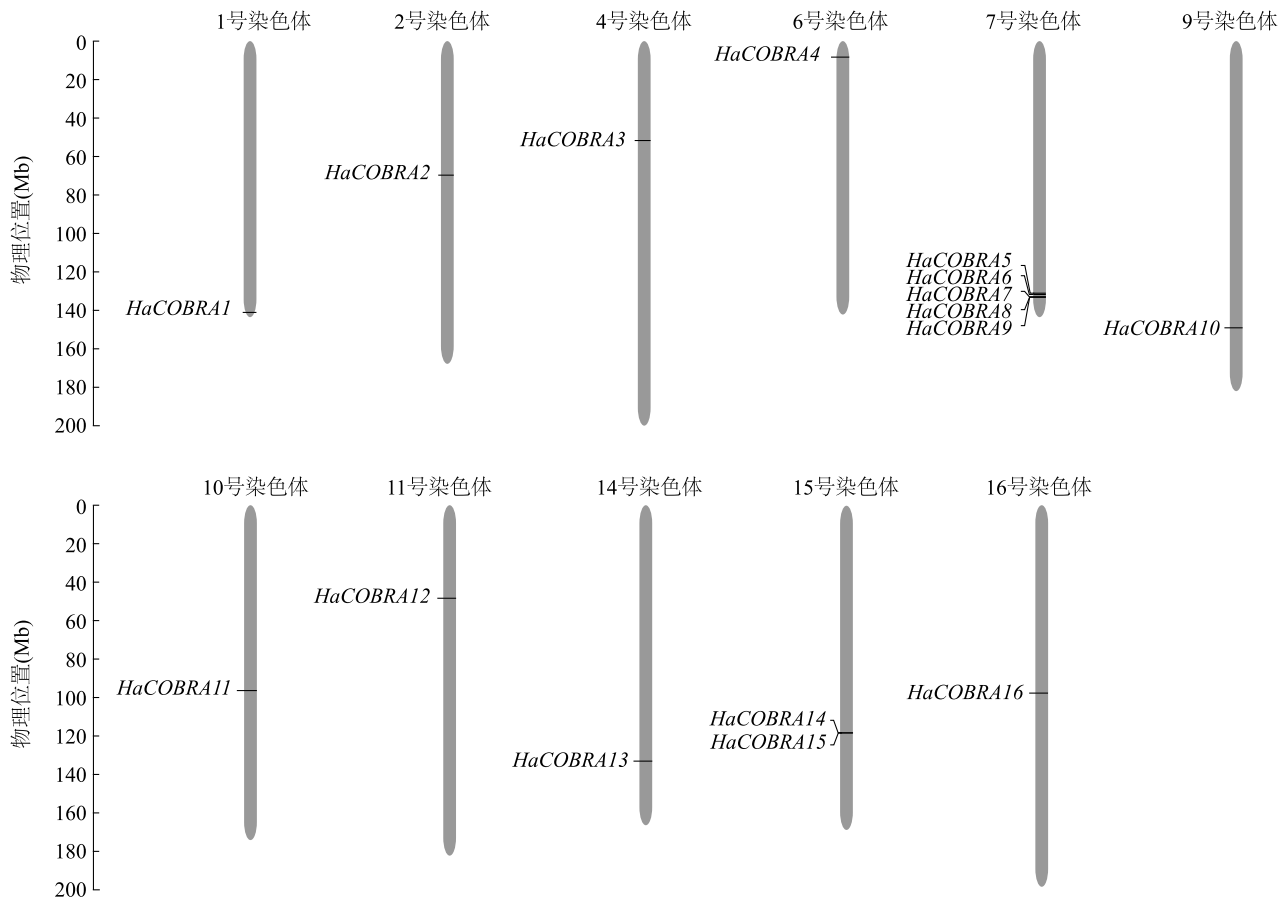
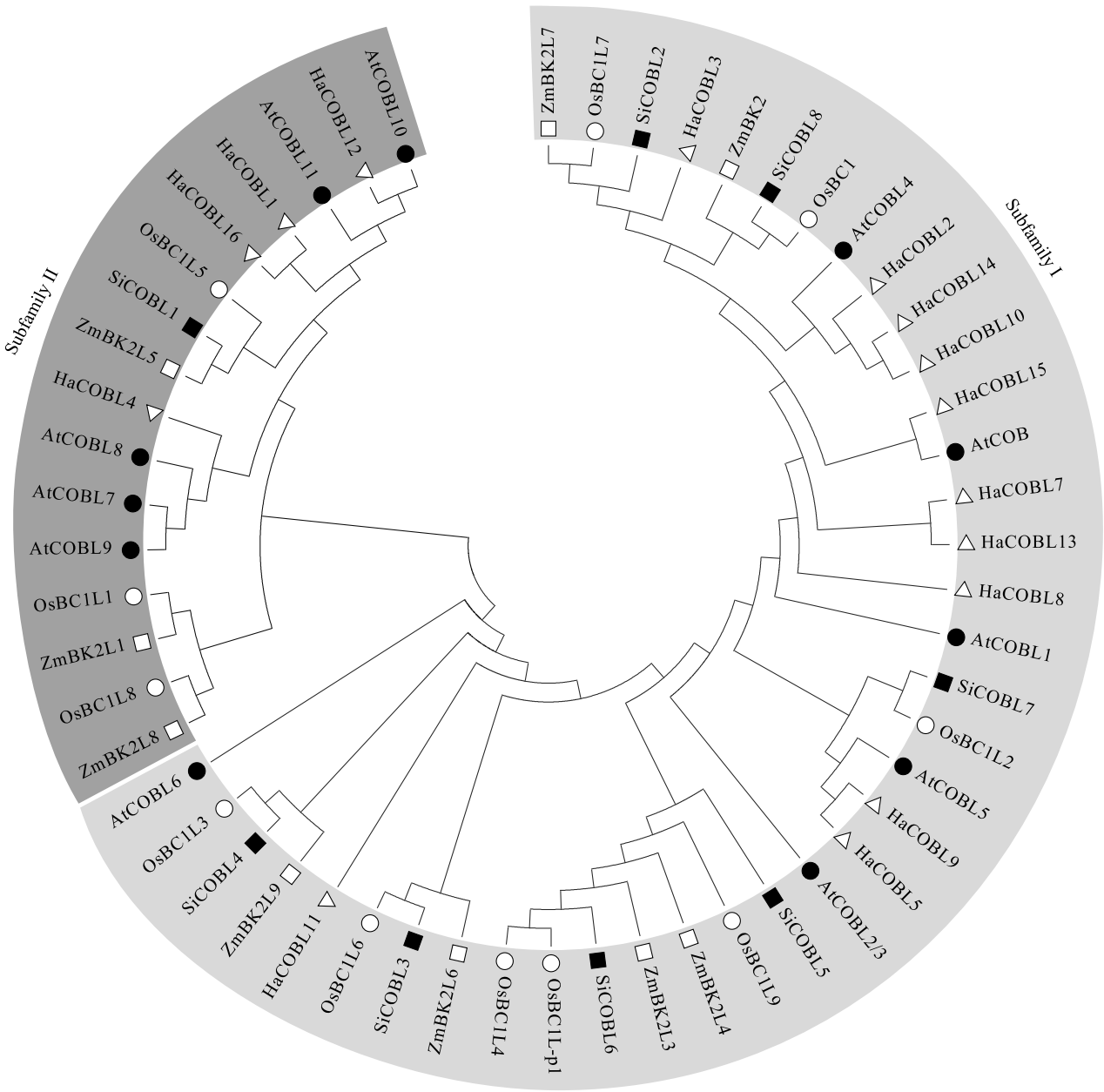


图 1 向日葵 COBRA 基因的染色体定位图谱

Fig.1 Chromosomal localization of COBRA genes in sunflower

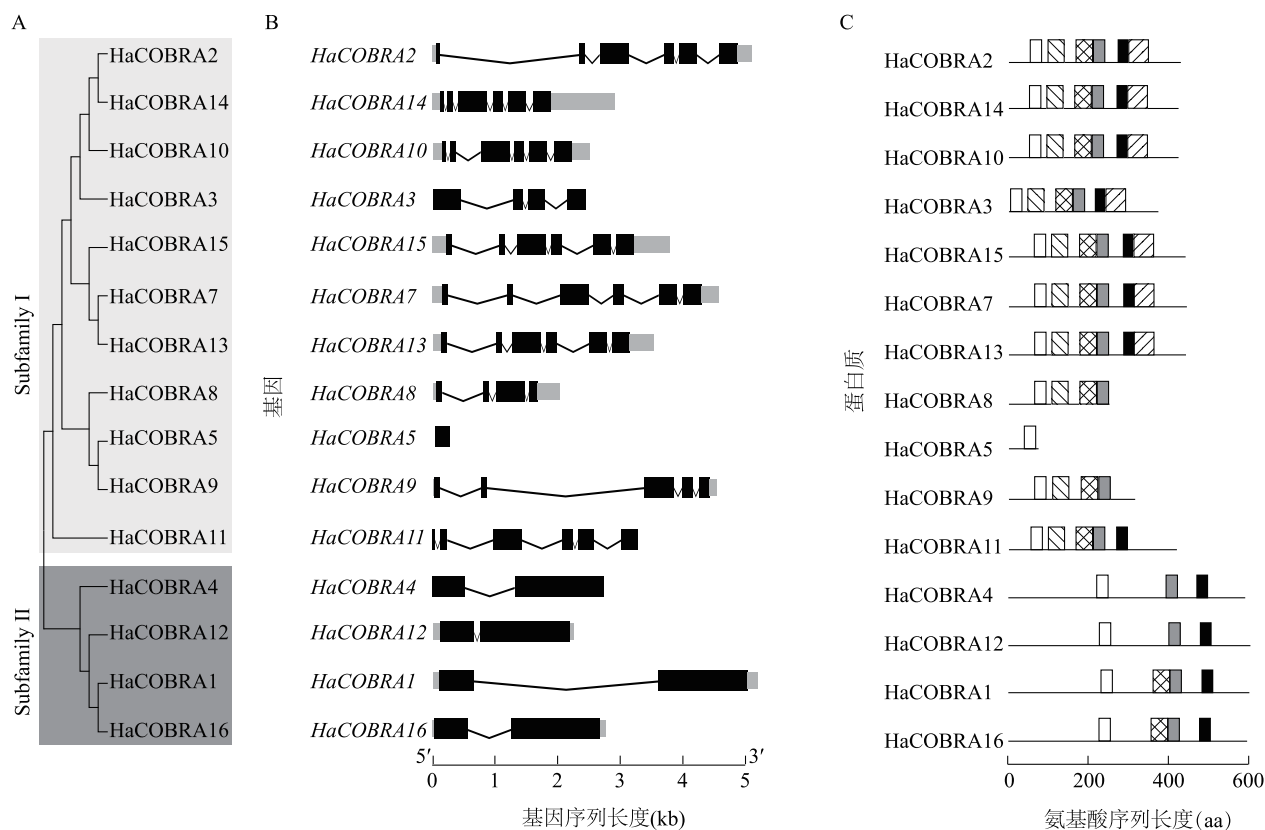


△ 向日葵(*Helianthus annuus*); □ 玉米(*Zea mays*); ○ 水稻(*Oryza sativa*); ■ 谷子(*Setaria italica*); ● 拟南芥(*Arabidopsis thaliana*)

图 2 5 个物种 COBRA 蛋白系统发育树
Fig.2 Phylogenetic tree of COBRA proteins in five species

基于 15 个向日葵 COBRA 蛋白成员构建系统发育树,结果(图 3A)发现,15 个蛋白质仍可划分为 2 个亚家族,且存在 HaCOBRA2、HaCOBRA14 等 4 个旁系同源蛋白对,提示该物种内可能发生了特定的基因复制事件。从图 3B 可知,亚家族间基因结构存在显著差异。Subfamily I 中除 HaCOBRA5 基因仅含有 1 个外显子外,其余均含 4~6 个外显子;Subfamily II 中基因均含有 2 个外显子。这进一步证实

了亚家族分类的可靠性。蛋白质的保守基序与蛋白质潜在功能密切相关。从图 3C 可见,15 个 HaCOBRA 蛋白中共鉴定到 6 个保守基序,仅 motif1 为全家族共有。Subfamily I 中,7 个成员包含检测到的 6 个基序,HaCOBRA5 蛋白仅含 motif1;Subfamily II 中 4 个成员的基序相对较少,HaCOBRA4 蛋白和 HaCOBRA12 蛋白包含 3 个 motif,HaCOBRA1 蛋白和 HaCOBRA16 蛋白包含 4 个 motif。



A:系统发育树; B:基因结构; C:保守基序。图B中, ■ 外显子; ▒ 非编码区; ∨ 内含子。图C中, □ motif 1(基序序列:5'-QQYRHISPPG WTLGWTWAKKEVIWSMVGA-3'); ▒ motif 2(基序序列:5'-KTPTCCVSFSFYNETIVPCPTCAGCEN-3'); ■ motif 3(基序序列:5'-CPIR VHWHVKTBYKDYWRKVTITNFN-3'); ▒ motif 4(基序序列:5'-GPGYTGPVAVVVKPTKFJTPDGRRTVQALMTWNVTCTYSQFL-3'); ▒ motif 5(基序序列:5'-PHCCCKTPTVVDLLPGVPYNQIANCKGKGVISSWGQDPNN-3'); ▒ motif 6(基序序列:5'-YRMNYTQWTLVV QHPNLBNVTQVFSFNYKPLTPYGSINDTAMFYGVKFFYN-3')。

图3 向日葵 *COBRA* 蛋白家族进化树、保守基序及基因结构分析

Fig.3 Phylogenetic tree, conserved motifs, and gene structure analysis of the *COBRA* protein family in sunflower (*Helianthus annuus* L.)

2.3 向日葵 *COBRA* 蛋白保守性及理化特性

已有研究表明,植物 *COBRA* 蛋白的 CCVS 结构高度保守,可能参与金属离子结合或二硫键的形成^[26-27]。本研究通过多重序列比对分析,结果(图4)显示,12条蛋白质氨基酸序列含有保守的 CCVS 结构,其余4个蛋白质发生缺失或突变。其中,HaCOBRA5 蛋白、HaCOBRA6 蛋白表现为缺失;HaCOBRA8 蛋白的 CCVS 结构突变为 CCSS 结构;HaCOBRA9 蛋白的 CCVS 结构突变为 CCTS 结构,这种变异可能影响二硫键形成效率,赋予蛋白质新的功能特性。

亲水性/疏水性分析可揭示蛋白质的结构特征与进化适应性。本研究分析发现,在向日葵中11个 HaCOBRA 蛋白含典型的中间亲水区。HaCOBRA3、HaCOBRA11 蛋白仅 C 端包含疏水区;HaCOBRA5

蛋白无疏水区,可能是可溶性蛋白;HaCOBRA6 蛋白的疏水区位于序列中部,可能形成单次跨膜螺旋;HaCOBRA8 蛋白仅 N 端有疏水区,说明亲水/疏水拓扑结构发生适应性进化变异。同时,GPI 锚定蛋白 C 端疏水区常含 GPI 修饰位点(ω -site)。向日葵 *COBRA* 家族蛋白均定位出该位点,说明均具备 GPI 锚定修饰潜能。此外, ω -site 氨基酸类型存在分化,11个为天冬酰胺,HaCOBRA1、HaCOBRA4、HaCOBRA12、HaCOBRA16 为丝氨酸,HaCOBRA5 为甘氨酸,这种变异可能影响锚定效率或蛋白质膜定位特性,为其功能特异性提供结构基础。

2.4 *HaCOBRA* 基因启动子顺式作用元件

顺式作用元件可通过与转录因子结合,介导基因对环境胁迫、激素信号及生长发育的响应。本研究对16个 *HaCOBRA* 基因启动子序列进行预测分

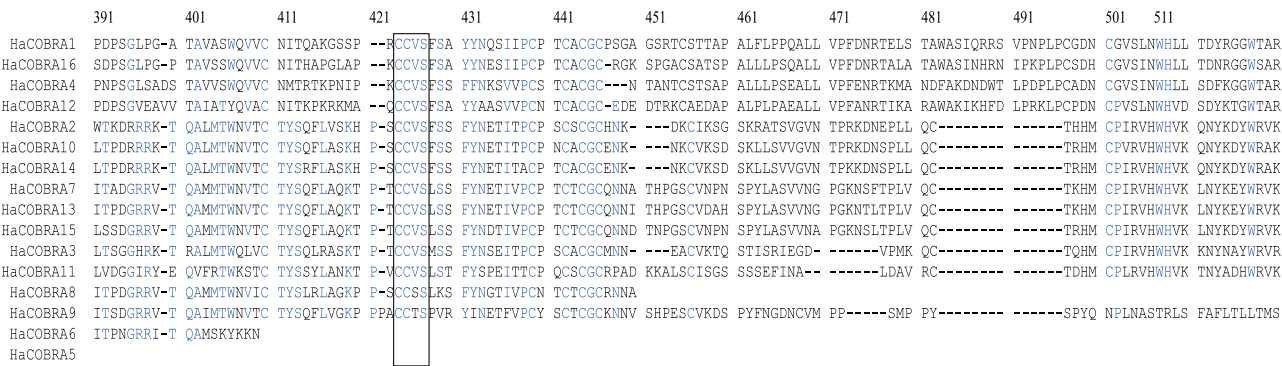
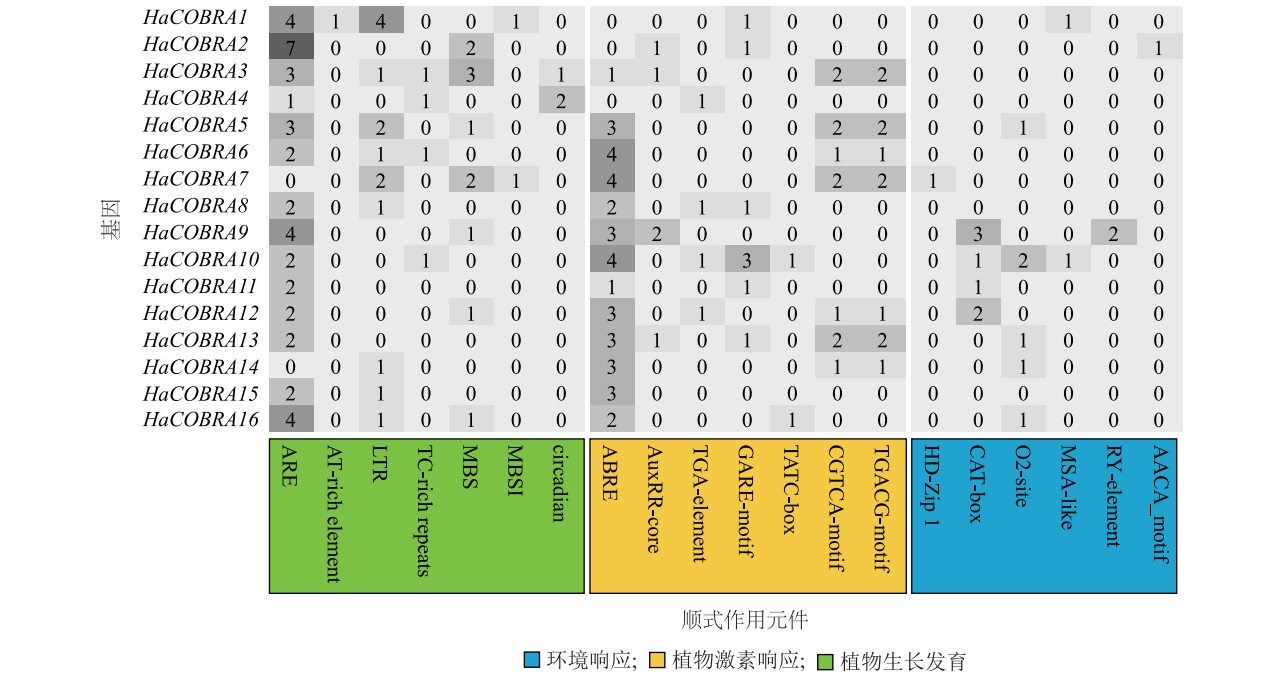


图 4 HaCOBRA 蛋白氨基酸序列中 CCVS 结构域的保守性
Fig.4 Conservation of CCVS domain in the amino acid sequences of HaCOBRA proteins

析,鉴定出 3 类核心顺式作用元件,分别参与环境响应、植物激素响应和植物生长发育调控(图 5)。从图 5 可见,11 个 *HaCOBRA* 基因启动子同时包含 3 种类型的顺式作用元件,另外 5 个 *HaCOBRA* 基因未检测到与生长发育相关的元件,表明 *HaCOBRA* 基因家族成员广泛参与环境胁迫应答、激素信号传

导及生长发育调控,但成员间可能存在功能分化。值得注意的是,*HaCOBRA1* 基因启动子中富集 10 个环境胁迫响应元件,包括 6 个厌氧胁迫响应元件 (ARE、AT-rich element)/缺氧胁迫响应元件 (MBSI) 和 4 个低温胁迫响应元件 (LTR),说明该基因在应对非生物胁迫时可能发挥关键作用。



ARE:参与厌氧胁迫响应元件;AT-rich element:参与厌氧胁迫响应元件;LTR:参与低温响应元件;TC-rich repeats:参与防御和应激反应元件;MBS:参与干旱胁迫响应元件;MBSI:参与环境缺氧响应元件;circadian:参与昼夜节律调控元件;ABRE:参与脱落酸响应元件;AuxRR-core:参与生长素响应元件;TGA-element:参与生长素响应元件;GARE-motif:参与赤霉素响应元件;TATC-box:参与赤霉素响应元件;CGTCA-motif:参与茉莉酸甲酯响应元件;TGACG-motif:参与茉莉酸甲酯响应元件;HD-Zip 1:参与栅栏叶肉细胞分化元件;CAT-box:与分生组织表达相关元件;O2-site:米肌蛋白代谢调控元件;MSA-like:细胞周期调控元件;RY-element:种子特异性调控元件;AACA_motif:胚乳特异性调控元件。

图 5 向日葵 *COBRA* 基因启动子的顺式作用元件
Fig.5 The cis-acting elements in the promoters of sunflower *COBRA* genes

2.5 向日葵 *COBRA* 蛋白互作

为挖掘 *HaCOBRA* 蛋白在纤维素合成途径中潜在的互作蛋白质及共表达基因,本研究对 16 个家族成员的蛋白质氨基酸序列进行了互作预测分析。结果显示,未检测到该家族成员之间存在直接互作关系;在单独检索时,仅 4 个成员 (*HaCOBRA10*、*HaCOBRA12*、*HaCOBRA14*、*HaCOBRA16*) 存在潜在的互作网络(图 6)。其中 *HaCOBRA10* 和 *HaCOBRA14* 预测的直接共表达蛋白质高度一致,这些共表达的蛋白质多与纤维素合成直接相关,包含 2 个纤维素合成酶蛋白(*CESA4*、*CESA8*)、3 个假定纤维素合成酶蛋白(*A0A251TTY1*、*A0A251RPA9* 和 *A0A251UJT5*)、2 个参与细胞壁木聚糖合成的糖基转移酶蛋白(*IRX9*、*IRX3*)及 2 个功能未知的蛋白(*A0A251V7R1* 和 *A0A251UD97*),说明这两个基因可能直接参与纤维素合成过程。*HaCOBRA12* 与 *HaCOBRA16* 预测的直接互作蛋白质完全一致,但未鉴定出已知的与纤维素合成或相关细胞壁代谢功能的蛋白质,说明二者可能有未知新功能。

2.6 向日葵 *COBRA* 基因表达模式

表达谱分析结果(图 7)显示,*HaCOBRA* 基因存在显著的组织表达与胁迫响应差异。从图 7A 可见,*HaCOBRA1*、*HaCOBRA7*、*HaCOBRA13*、*HaCOBRA14*、*HaCOBRA15* 在茎、舌状花、管状花等 8 种检测器官(组织)中均高表达,推测这些基因参与调控向日葵生长发育。*HaCOBRA16* 在根中高表达,暗示其调控根系发育。从图 7B 可见,与对照相比,叶片中,盐和干旱胁迫下 *HaCOBRA3*、*HaCOBRA4*、*HaCOBRA5*、*HaCOBRA6*、*HaCOBRA9*、*HaCOBRA11*、*HaCOBRA12* 基因上调表达,*HaCOBRA7*、*HaCOBRA8*、*HaCOBRA10*、*HaCOBRA13*、*HaCOBRA15*、*HaCOBRA16* 基因下调表达,且响应趋势一致;*HaCOBRA1* 为差异化响应。根中,*HaCOBRA8*、*HaCOBRA10*、*HaCOBRA14*、*HaCOBRA15* 基因在两种胁迫下均上调表达,*HaCOBRA6*、*HaCOBRA9* 基因下调表达;*HaCOBRA1* 基因在干旱胁迫下上调表达。综上,*HaCOBRA* 基因通过组织特异性表达参与胁迫响应,根系中上调表达基因可能是核心胁迫应答调控因子。

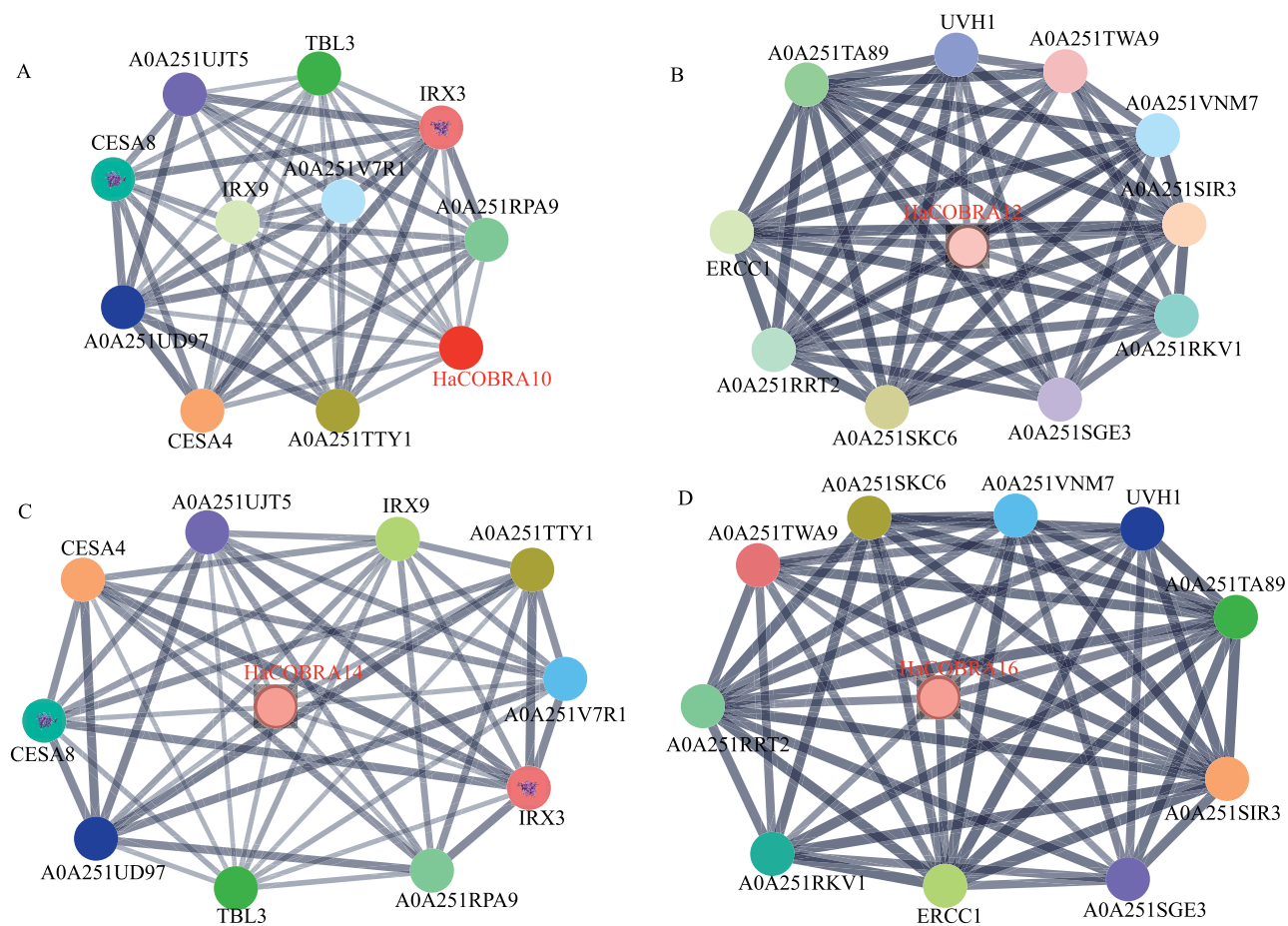
3 讨论

COBRA 家族蛋白是植物特有的一类规模小且

结构保守的蛋白质,对植物细胞定向分裂、细胞扩张、纤维素合成及根毛顶端细胞发育等具有关键调控作用^[26-27]。目前,针对拟南芥^[5,27]、水稻^[7,28]和玉米^[6,11]*COBRA* 家族蛋白的研究比较深入,但有关向日葵 *COBRA* 家族蛋白的报道相对较少。本研究通过鉴定分析,共获得 16 个向日葵 *COBRA* 蛋白,其成员数量略多于拟南芥(12 个)^[5]、谷子(9 个)^[8]、玉米(9 个)^[6]、水稻(11 个)^[28]等物种。在家族分类上与上述物种相似,均表现为 Subfamily I 成员多于 Subfamily II,说明该家族在不同物种中的进化具有相似性;在基因结构上,2 个亚家族成员的外显子数量与拟南芥^[5]和谷子^[8]相似。

植物 *COBRA* 蛋白在结构上高度保守,包含 1 个亲水区、CCVS 结构、N 端糖基化位点及信号肽等^[15]。与拟南芥^[5]、水稻^[28]、玉米^[6]、谷子^[8]、冬枣^[29]等物种的 *COBRA* 家族蛋白质相比,向日葵 *COBRA* 家族蛋白在序列结构上表现出更高的多样性。据报道,大豆中有 2 个 *COBL* 蛋白缺失 CCVS 结构域^[30],本研究发现 4 个 *HaCOBRA* 蛋白存在结构域缺失或位点突变,且其编码基因均定位于 7 号染色体,这表明 *HaCOBRA* 基因在进化过程中可能发生了缺失或变异事件。

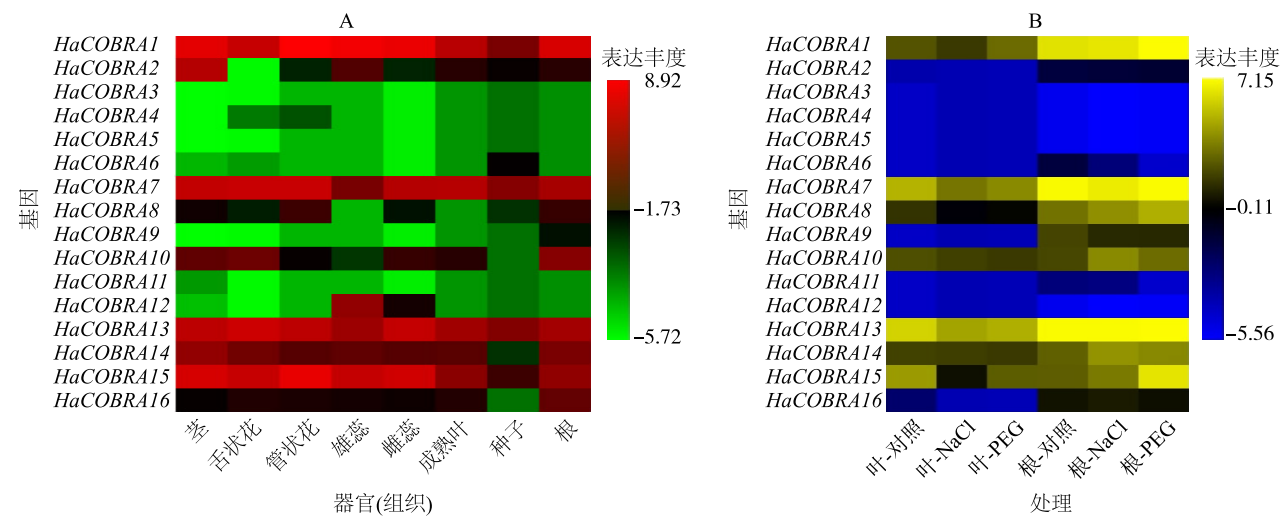
编码糖基磷脂酰基锚定蛋白的植物 *COBRA* 基因,常与纤维素合成酶(*CESA*)基因家族成员共表达,参与调控植物细胞壁扩张、纤维素沉积等^[5,31]。2013 年,Liu 等^[3]研究证实水稻 *COBRA* 蛋白 OsBC1 可与 *CESA* 蛋白互作,共同参与调控纤维素的正常排列。拟南芥 *AtCOBL4* 基因与 *AtCESA4-7-8* 基因高度共表达,参与调控纤维素结晶^[32]。与 *AtCOBL4* 基因同源的玉米 *ZmBK2* 基因在各组织中高表达,并参与次生细胞壁纤维素的沉积^[11,33]。本研究发现,*HaCOBRA2*、*HaCOBRA10*、*HaCOBRA14* 基因和 *AtCOBL4* 基因高度同源,预测 *HaCOBRA10* 蛋白及 *HaCOBRA14* 蛋白与多个纤维素合成酶互作,包含 2 个纤维素合成酶(*CESA4*、*CESA8*)及 2 个糖基转移酶(*IRX9*、*IRX3*)。除此之外,*HaCOBRA10* 基因的启动子区域包含多个调控生长发育及环境响应的顺式作用元件,结合组织中高表达及在胁迫下的表达模式,表明 *HaCOBRA10* 基因可能与 *CESA* 基因互作共同调控纤维素合成,参与生长发育必须过程以及胁迫响应过程,是潜在的关键基因。



线条粗细度表示数据支持的强度。

图 6 HaCOBRA10、HaCOBRA12、HaCOBRA14、HaCOBRA16 蛋白互作网络

Fig.6 Protein-protein interaction networks of HaCOBRA10, HaCOBRA12, HaCOBRA14 and HaCOBRA16



A:组织表达;B:胁迫响应表达。叶-对照:未经胁迫处理的叶片,叶-NaCl:盐胁迫处理的叶片,叶-PEG:干旱胁迫处理的叶片;根-对照:未经胁迫处理的根,根-NaCl:盐胁迫处理的根,根-PEG:干旱胁迫处理的根。

图 7 HaCOBRA 基因的表达模式

Fig.7 Expression profiles of HaCOBRA genes

4 结 论

本研究在向日葵全基因组中鉴定出 16 个 *HaCOBRA* 家族基因,这些基因在染色体上呈现不均匀分布,多数基因保持植物 *COBRA* 特有的典型特征。启动子分析、表达模式和共表达分析共同揭示,大多数 *HaCOBRA* 基因参与生长发育中纤维素形成和胁迫响应,其中 *HaCOBRA10* 与 *HaCOBRA14* 是潜在的纤维素合成关键基因。本研究结果初步明确了向日葵 *HaCOBRA* 家族基因参与纤维素生物合成的调控作用,为加快向日葵细胞壁性状形成分子机制的研究奠定了基础。

参考文献:

- [1] Cosgrove D J. Plant cell wall extensibility: connecting plant cell growth with cell wall structure, mechanics, and the action of wall-modifying enzymes[J]. *Journal of Experimental Botany*, 2016, 67(2): 463-476.
- [2] Cosgrove D J. Plant expansins: diversity and interactions with plant cell walls[J]. *Current Opinion in Plant Biology*, 2015, 25: 162-172.
- [3] Liu L F, Shang-Guan K K, Zhang B C, et al. Brittle culm1, a COBRA-like protein, functions in cellulose assembly through binding cellulose microfibrils[J]. *PLoS Genetics*, 2013, 9(8): 1003704.
- [4] Sorek N, Ryan A, Carroll A, et al. Functional analysis of the COBRA-like family[EB/OL]. (2016-02-10) [2025-06-27] <https://europepmc.org/article/PPR/PPR38561>.
- [5] Schindelman G, Morikami A, Jung J, et al. *COBRA* encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in *Arabidopsis*[J]. *Genes & Development*, 2001, 15(9): 1115-1127.
- [6] Cruz-García F, Gómez A, Zúñiga J J, et al. Cloning and characterization of a *COBRA*-like gene expressed *de novo* during maize germination[J]. *Seed Science Research*, 2003, 13(3): 209-217.
- [7] Li Y H, Qian Q, Zhou Y H, et al. *BRITTLE CULM1*, which encodes a COBRA-like protein, affects the mechanical properties of rice plants[J]. *The Plant Cell*, 2003, 15(9): 2020-2031.
- [8] 任昂彦,陈环宇,王根平,等. 谷子 *COBRA* 家族成员的全基因组鉴定及表达模式分析[J]. *分子植物育种*, 2024, 22(23): 7669-7677.
- [9] 曹颖,唐晓凤,刘永胜. 番茄 *COBRA* 基因克隆、表达模式及生物信息学分析[J]. *植物研究*, 2012, 32(3): 304-310.
- [10] Li S, Ge F R, Xu M, et al. *Arabidopsis* COBRA-like 10, a GPI-anchored protein, mediates directional growth of pollen tubes[J]. *The Plant Journal*, 2013, 74(3): 486-497.
- [11] Sindhu A, Langewisch T, Olek A, et al. Maize Brittle stalk2 encodes a COBRA-like protein expressed in early organ development but required for tissue flexibility at maturity[J]. *Plant Physiology*, 2007, 145(4): 1444-1459.
- [12] Li P, Liu Y R, Tan W Q, et al. Brittle culm 1 encodes a COBRA-like protein involved in secondary cell wall cellulose biosynthesis in sorghum[J]. *Plant & Cell Physiology*, 2019, 60(4): 788-801.
- [13] Niu E L, Fang S, Shang X G, et al. Ectopic expression of Gh-COBL9A, a cotton glycosyl-phosphatidyl inositol-anchored protein encoding gene, promotes cell elongation, thickening and increased plant biomass in transgenic *Arabidopsis*[J]. *Molecular Genetics and Genomics*, 2018, 293(5): 1191-1204.
- [14] Antonopoulou G, Dimitrellos G, Beohide A S, et al. Chemical pretreatment of sunflower straw biomass: the effect on chemical composition and structural changes[J]. *Waste and Biomass Valorization*, 2015, 6(5): 733-746.
- [15] De Souza J B, Michelin M, Amâncio F L R, et al. Sunflower stalk as a carbon source inductive for fungal xylanase production[J]. *Industrial Crops and Products*, 2020, 153: 112368.
- [16] Yin Z D, Pan Z L, Wang C, et al. Composition, structure and mechanical properties of several natural cellular materials[J]. *Chinese Science Bulletin*, 2007, 52(21): 2903-2908.
- [17] Bateman A, Birney E, Durbin R, et al. The Pfam protein families database[J]. *Nucleic Acids Research*, 2000, 28(1): 263-266.
- [18] Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020[J]. *Nucleic Acids Research*, 2021, 49(1): 458-460.
- [19] Wang M, Xu Z C, Ding A M, et al. Genome-wide identification and expression profiling analysis of the xyloglucan endotransglucosylase/hydrolase gene family in tobacco (*Nicotiana tabacum* L.)[J]. *Genes*, 2018, 9(6): 273.
- [20] Ren A Y, Ahmed R I, Chen H Y, et al. Genome-wide identification, characterization and expression patterns of the pectin methyl-esterase inhibitor genes in sorghum bicolor[J]. *Genes*, 2019, 10(10): 755.
- [21] Wilkins M R, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools in the ExPASy server[J]. *Methods in Molecular Biology*, 1999, 112: 531-552.
- [22] Chou K C, Shen H B. Cell-PLOC: a package of Web servers for predicting subcellular localization of proteins in various organisms[J]. *Nature Protocols*, 2008, 3(2): 153-162.
- [23] Hu B, Jin J P, Guo A Y, et al. GSDB 2.0: an upgraded gene feature visualization server[J]. *Bioinformatics*, 2015, 31(8): 1296-1297.
- [24] Eisenhaber B, Bork P, Eisenhaber F. Prediction of potential GPI-modification sites in preprotein sequences[J]. *Journal of Molecular Biology*, 1999, 292(3): 741-758.
- [25] Szklarczyk D, Franceschini A, Kuhn M, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored[J]. *Nucleic Acids Research*, 2011, 39: 561-568.
- [26] Brady S M, Song S, Dhugga K S, et al. Combining expression and

- comparative evolutionary analysis of the *COBRA* gene family[J]. *Plant Physiology*, 2007, 143(1): 172-187.
- [27] Roudier F, Schindelman G, Desalle R, et al. The *COBRA* family of putative GPI-anchored proteins in *Arabidopsis*. a new fellowship in expansion[J]. *Plant Physiology*, 2002, 130(2): 538-548.
- [28] Hu L F, He H H, Zhu C L, et al. Genome-wide identification and phylogenetic analysis of the *Chalcone* synthase gene family in rice [J]. *Journal of Plant Research*, 2017, 130(1): 95-105.
- [29] 盛松柏, 田 菊, 庞晓明. 冬枣 *COBRA* 基因家族全基因组鉴定及表达分析[J]. *分子植物育种*, 2018, 16(1): 61-68.
- [30] Sangi S, Araújo P M, Coelho F S, et al. Genome-wide analysis of the *COBRA*-like gene family supports gene expansion through whole-genome duplication in soybean (*Glycine max*) [J]. *Plants*, 2021, 10(1): 167.
- [31] Persson S, Wei H R, Milne J, et al. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102(24): 8633-8638.
- [32] Sorek N, Sorek H, Kijac A, et al. The *Arabidopsis* *COBRA* protein facilitates cellulose crystallization at the plasma membrane [J]. *Journal of Biological Chemistry*, 2014, 289(50): 34911-34920.
- [33] Ching A, Dhugga K S, Appenzeller L, et al. Brittle stalk 2 encodes a putative glycosylphosphatidylinositol-anchored protein that affects mechanical strength of maize tissues by altering the composition and structure of secondary cell walls[J]. *Planta*, 2006, 224(5): 1174-1184.

(责任编辑:黄克玲)