

西瓜 *CIKNOX* 基因家族全基因组 鉴定及组织表达分析

陈雅心, 卢奥锋, 张力铭, 刘权权, 杨路明, 朱华玉, 刘东明

(河南农业大学园艺学院 郑州 450046)

摘要: *KNOX* 基因家族是编码同源异型盒蛋白的转录因子, 在植物生长发育和胁迫响应中发挥着重要的调控作用。为进一步挖掘西瓜 *KNOX* 转录因子家族成员信息, 探究分析其表达模式及基因功能, 通过生物信息学方法对西瓜 *KNOX* 基因家族成员进行了鉴定, 并对其理化性质、基因结构、系统进化及表达模式进行了分析。结果表明, 西瓜基因组中共有 12 个 *KNOX* 基因, 均定位于细胞核中, 符合转录因子属性特征; *KNOX1*、*KNOX2*、*ELK* 和 *Homeobox_KN* 这 4 类典型结构域均存在于大多数西瓜 *KNOX* 基因中, 表明这 4 类结构域在 *KNOX* 基因功能中具有重要作用; 系统发育分析将 *CIKNOX* 基因家族成员分为 Class I A、Class I B、Class II A 和 Class II B 4 个亚组, 启动子元件分析表明, *CIKNOX* 基因家族含有较多与生长发育和胁迫响应相关的顺式作用调控元件; 组织表达分析结果表明, *CIKNOX* 基因具有明显的组织特异表达特性, *CIKNOX6* 和 *CIKNOX11* 在根和茎中高表达, 而 *CIKNOX3* 和 *CIKNOX10* 在所有器官中表达量都相对较高, 但所有 *CIKNOX* 基因在果肉中的相对表达量都较低。研究结果为进一步深入探究 *CIKNOX* 基因在生长发育中的功能和组织表达特性奠定了基础。

关键词: 西瓜; *KNOX* 基因家族; 全基因组鉴定; 生物信息学分析; 组织表达

中图分类号: S651

文献标志码: A

文章编号: 1673-2871(2024)05-018-10

Genome-wide identification and expression analysis of *CIKNOX* gene family in watermelon

CHEN Yaxin, HU Aofeng, ZHANG Liming, LIU Quanquan, YANG Luming, ZHU Huayu, LIU Dongming

(Horticulture College of Henan Agricultural University, Zhengzhou 450046, Henan, China)

Abstract: KNOTTED-like homeodomain (*KNOX*) gene family is a class of transcription factors encoding homeobox protein, which plays an important role in plant growth and stress responses. To further investigate the watermelon *KNOX* transcription factor family members, the bioinformatics methods was utilized to identify and analyze their physical and chemical properties, gene structure, phylogeny, and expression patterns. The study revealed the presence of 12 *KNOX* genes in the watermelon genome, all of which were identified in the nucleus, the result was consistent with their characteristics as transcription factors. Most *KNOX* genes in watermelon exhibited the typical domains *KNOX1*, *KNOX2*, *ELK*, and *Homeobox_KN*, suggesting the significance of these domains in *KNOX* gene functionality. Phylogenetic analysis categorized *CIKNOX* gene family members into four subgroups: ClassIA, ClassIB, ClassIIA, and ClassIIB. Analysis of promoter elements indicated that the *CIKNOX* gene family harbored numerous cis-acting regulatory elements associated with growth, development, and stress response. The results of tissue expression analysis revealed distinct patterns for *CIKNOX* gene expression. Specifically, *CIKNOX6* and *CIKNOX11* exhibited high expression levels in roots and stems, whereas *CIKNOX3* and *CIKNOX10* showed relatively high expression across all organs. However, all *CIKNOX* genes had low relative expression levels in flesh. These findings provide a basis for future investigations into the functional roles and tissue-specific expression patterns of *CIKNOX* genes in growth and development.

Key words: Watermelon; *KNOX* gene family; Genome-wide identification; Bioinformatics analysis; Tissue expression

收稿日期: 2024-01-18; 修回日期: 2024-03-08

基金项目: 河南省科技攻关项目(222102110124); 河南省农业品种改良联合研究项目(2022010503)

作者简介: 陈雅心, 女, 在读硕士研究生, 研究方向为蔬菜遗传育种。E-mail: 2267453460@qq.com

通信作者: 刘东明, 男, 副教授, 研究方向为蔬菜遗传育种。E-mail: liudongming5668@163.com

西瓜 [*Citrullus lanatus* (Thunb.) Matsum. and Nakai] 为葫芦科西瓜属一年生蔓生草本植物,富含多种维生素和对人体有益的瓜氨酸、番茄红素等营养成分,无论鲜食还是制成果汁、糖果或蘸酱等都深受世界各地消费者的喜爱。同源异型盒基因家族(homeobox gene family)在动物和植物发育调控中具有重要作用,该基因家族包含高度保守的 180 个碱基对,编码一个含有 60 个氨基酸的三螺旋结构域,最早被克隆的 *homeobox* 基因来自果蝇^[1],随后植物中第一个 *homeobox* 基因——*Knotted-1* 在玉米中被分离获得^[2]。近年来研究发现,该基因家族广泛存在于陆生植物类群和特定的绿藻门植物中,但在红藻植物中尚未被发现^[3-4]。根据是否含有一个三氨基酸的环延伸基序,*homeobox* 类基因被分为 *TALE* 和 *non-TALE* 两类^[5],在动物中已经鉴定出 4 种类型的 *TALE* 基因,分别为 *MEIS*、*IRO*、*TGIF* 和 *PBC*,而在植物中仅有 *KNOX*(*KNOTTED-like homeobox*) 和 *BELL* (*BELL-like*) 两类基因被鉴定^[6]。*KNOX* 类基因通常含有 4 个典型结构域,分别为负责识别下游靶基因启动子序列的 C 端同源结构域(*Homeobox_KN*)、N 端的 *KNOX1* 和 *KNOX2* 以及参与转录抑制核定位信号、促进蛋白质-蛋白质相互作用的 *ELK* 结构域。由于 *MEIS* 和 *KNOX* 家族基因结构相似,*KNOX1* 和 *KNOX2* 结构域也称为 *MEINOX* 结构域,在他们的同源盒中,在第一和第二螺旋之间还有 3 个额外的氨基酸 (*P-Y-P*)^[6-7],该类结构域参与抑制靶基因表达和同二聚化。

根据序列相似性、内含子位置、系统发育关系和表达特点,*KNOX* 类基因常分为 I 类、II 类和 M 类^[4,8-9]。I 类亚家族包括 *SHOOT MERISTEMLESS* (*STM*)、*KNAT1*、*KNAT2* 和 *KNAT6* 等 4 个成员。作为转录激活因子或抑制因子,I 类亚家族基因在分生组织发育、叶形状的控制、激素稳态方面发挥重要作用,常在茎顶分生组织中表达^[9]。*STM* 基因在茎顶端分生组织中负责维持并调节花序结构,*KNAT1* 在拟南芥根中也显示出细胞特异性表达模式^[10],并与 *STM* 类基因具有冗余性^[11-13]。*KNAT2* 在茎顶端分生组织和根组织中表达^[14-15],而 *KNAT6* 在胚胎茎顶端分生组织、茎顶端分生组织边界^[14]和根韧皮部组织中表达^[16]。II 类亚家族基因中,*KNAT3*、*KNAT4* 和 *KNAT5* 在拟南芥根中显示出细胞特异性表达模式^[11]。*KNAT3* 在幼果、花序和根中表达量最高,而 *KNAT4* 在叶片和幼果中表达量最高^[17]。*KNAT5* 在幼芽和根新发育的伸长区表达,在表皮中

的表达可用来标记细胞分裂和伸长的边界^[11]。*KNAT7* 在拟南芥根的中央部位高度表达^[18],在拟南芥和杨树中与次生壁形成有关^[19]。M 类亚家族包含 *KNATM* 基因,在叶片近端-远端模式中起作用,并在成熟器官边界的器官原基近侧域中表达^[9]。

我国西瓜种植面积和产量常年居世界第一位^[20]。近年来,随着分子生物学技术的不断发展和西瓜参考基因组数据的不断优化,以精准定向聚合优异基因的分子育种已成为西瓜育种技术的主要发展方向。明确重要性状调控基因是实现从传统常规育种向分子标记辅助育种转型的基础。*KNOX* 家族基因具有组织特异性表达的特点,且在植物器官发育调控中具有重要作用。目前 *KNOX* 家族基因在葫芦科作物中的报道仅局限于黄瓜和瓠瓜,在黄瓜中有研究表明,*KNOX* 家族基因 *TKN2* 和 *TKN4* 与 *APRR2* 表达模式一致,推测其可能共同调控黄瓜叶绿体代谢机制^[21]。在瓠瓜上已经进行了 *KNOX* 全基因组水平的鉴定和组织表达分析^[22],而针对西瓜还未有系统的研究报道。笔者在全基因组水平上对西瓜 *KNOX* 基因家族成员进行了鉴定分析,共筛选获得了 12 个西瓜 *KNOX* 基因,对其理化性质、亚细胞定位、系统进化树、染色体定位和启动子元件进行了分析,并选用遗传性状稳定的西瓜材料 WT2 来探究其组织表达模式,以期为进一步了解 *KNOX* 的进化特征和功能提供参考。

1 材料与方法

1.1 植物材料

试验材料西瓜高代自交系 WT2,由河南农业大学园艺学院瓜类作物基因组与分子育种实验室收集保存,性状稳定遗传。试验于 2023 年 5—11 月在河南农业大学园艺学院瓜类作物基因组与分子育种实验室进行。

取籽粒饱满的 WT2 西瓜种子,55 °C 温汤浸种催芽后播种于穴盘中,放置在人工气候箱(28 °C, 16 h 光照/8 h 黑暗,相对湿度 70%)内。于 3 叶 1 心时定植在温室大棚中,并合理实施浇水施肥、病虫害防治等栽培管理措施,开花期严格自交授粉。授粉后 28 d 分别取根、茎、叶、花、果皮和果肉等组织部位,采用随机区组法进行取样,每个样本设置 3 次重复。所有样品用锡箔纸包裹后放于液氮中速冻,随后存放于-80 °C 超低温冰箱备用。

1.2 方法

1.2.1 *CIKNOX* 基因家族成员的鉴定和理化性质预

测 西瓜参考基因组、蛋白序列及注释文件均来自葫芦科基因组数据库 (<http://cucurbitgenomics.org/organism/21>), 采用 97103-V2 版本数据^[23]。拟南芥 (*Arabidopsis thaliana* L.) 参考基因组、蛋白序列、注释文件及 *KNOX* 基因家族成员数据均下载自 TAIR 网站 (<https://www.arabidopsis.org/index.jsp>)。 *KNOX* 基因家族结构域的隐马尔科夫模型文件下载于 pfam 数据库 (<http://pfam.xfam.org/>)^[24]。为获得准确的 *CIKNOX* 基因鉴定信息, 首先在 pfam 数据库下载 *KNOX* 基因家族结构域的隐马尔科夫模型文件, 使用 hmmer 软件对西瓜蛋白数据库进行比对搜索^[25]; 然后以拟南芥 *KNOX* 蛋白序列为探针, 在西瓜基因组数据库进行搜索。将以上两种方法获得的 *KNOX* 蛋白序列在 NCBI-CDD 网站 (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 进行保守结构域鉴定, 得到含有 *KNOX* 结构域的候选蛋白^[26]。利用在线工具 ProtParamtool (<https://web.expasy.org/protparam/>) 分析西瓜 *KNOX* 家族成员的氨基酸数目、蛋白分子质量、等电点等理化特性^[27]。利用 Cell-PLoc2.0 (<http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/>) 预测各基因的亚细胞定位^[28]。

1.2.2 *CIKNOX* 基因的染色体定位和基因结构分析 从葫芦科基因组数据库获得每条染色体的长度和 *CIKNOX* 基因的位置信息, 结合 TBtools 软件绘制 *CIKNOX* 的染色体定位图。基于葫芦科基因组数据库 (<http://cucurbitgenomics.org/>) 中提取的西瓜所有 *CIKNOX* 基因的基因组序列和 CDS 序列信息, 利用 GSDS 2.0 程序 (<http://gsds.gao-lab.org/>) 分析 *CIKNOX* 的外显子和内含子^[29], 利用 MEME (<http://meme-suite.org/tools/meme>) 分析 *CIKNOX* 蛋白的保守基序^[30], 并通过 TBtools 软件对基因结构和蛋白保守基序进行可视化分析^[31]。

1.2.3 *CIKNOX* 家族的系统进化和启动子顺式作用元件分析 从葫芦科基因组数据库 (<http://cucurbitgenomics.org/>)、拟南芥数据库 (<https://www.arabidopsis.org/>) 和水稻基因组数据库 (<http://rice.plantbiology.msu.edu/>) 获得西瓜、甜瓜、黄瓜、葫芦、瓠瓜、冬瓜、水稻和拟南芥 *KNOX* 家族成员编码的氨基酸序列, 利用 MUSCLE (<http://www.drive5.com/muscle/>) 软件对所有 *KNOX* 蛋白序列进行比对, 并利用 Jalview (<http://www.jalview.org/>) 软件对对比结果进行编辑^[32]。利用 IQ-TREE (<http://www.iqtree.org/>) 的最大似然方法构建 *KNOX* 的系统发育树, 并将 Bootstrap 参数设置为 1000^[33]。从葫芦科基因组数

据库中获得 *CIKNOX* 基因编码序列上游 2000 bp 的基因组 DNA 序列, 然后利用 PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) 对 *CIKNOX* 基因启动子区域的顺式作用元件进行分析^[34], 使用 TBtools 软件对元件的数量与分布进行可视化分析。

1.2.4 *CIKNOX* 基因的组织表达分析 取西瓜各组织样品置于含有液氮的研钵中迅速研磨成粉末, 在研磨过程中不断向研钵中加入液氮, 防止样品在空气中暴露时间过长造成降解。后续步骤参照北京华越洋 RNA 提取试剂盒对各组织部位总 RNA 进行提取, 通过凝胶电泳和超微量分光光度计 (P200+, Plextech, USA) 检测 RNA 质量, 用诺维赞反转录试剂盒通过两步法反转录成 cDNA。

根据葫芦科基因组数据库 (<http://cucurbitgenomics.org/>) 和 NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>)^[35] 网站设计特异性定量引物 (表 1)。以 *ClActin* 为内参基因, 按照荧光定量 PCR 试剂盒步骤, 以西瓜 cDNA 为模板检测基因的表达量, 每个检测样本设置 3 个生物学重复和 3 个技术重复, 采用 StepOnePlus™ Real-Time PCR 仪 (Thermo Scientific, Massachusetts, USA) 进行荧光定量分析, 并按照 $2^{-\Delta\Delta Ct}$ 的方法计算基因相对表达量^[36]。

2 结果与分析

2.1 *CIKNOX* 家族基因的鉴定

KNOX 蛋白包含 4 个保守结构域, 从 N 端到 C 端依次为 *KNOX1*、*KNOX2*、*ELK* 和 *Homeobox_KN*。西瓜中这 4 个结构域的基因数分别为 11、9、8 和 21 个, 但仅有 11 个成员同时包含 *Homeobox_KN* 结构域和其他类型的结构域。为了获得准确的数据, 以 9 个拟南芥 *AtKNOX* 蛋白序列在西瓜蛋白数据库进行 blast 检索, 通过序列同源比对, 最终从西瓜基因组数据库中获得了 12 个 *CIKNOX* 成员。综合上述两种分析方法, 鉴定到 12 个西瓜 *CIKNOX* 蛋白, 根据基因所在染色体位置分别命名为 *CIKNOX1* ~ *CIKNOX12* (表 2)。理化特性分析结果 (表 2) 表明, 西瓜 *KNOX* 家族基因编码蛋白分子质量为 13.59 (*CIKNOX9*) ~ 51.78 ku (*CIKNOX5*), 编码氨基酸为 116 (*CIKNOX9*) ~ 467 个 (*CIKNOX10* 和 *CIKNOX5*), 蛋白等电点为 4.44 (*CIKNOX8*) ~ 8.89 (*CIKNOX9*)。亚细胞定位预测结果表明, 所有 *CIKNOX* 基因都定位在细胞核中 (表 2), 这与 *KNOX* 基因的转录因子属性特征相符。

表 1 *CIKNOX* 定量引物序列
Table 1 Primer sequence for *CIKNOX* qRT-PCR

基因名称 Gene name	正向引物(5'-3') Forward primer sequence(5'- 3')	反向引物(5'-3') Reverse primer sequence(5'- 3')
<i>CIKNOX1</i>	TTCTTCTCAGCAGCACCACC	TCACATGTTGTTGAAGCTGCTC
<i>CIKNOX2</i>	CATGGCTCATCCCCACTACC	GCACACACTTCTTCAAGCCG
<i>CIKNOX3</i>	CTCGCTGCCTACGTCAACT	GATCATCCCCACGTCCCG
<i>CIKNOX4</i>	CCCCTTATCCAAGCTGGTT	GAGGGTGATTTTACAGGCCA
<i>CIKNOX5</i>	CATTGAGCAGTCCCTACAAGC	TTTGCATCGCTGTTGACTTGG
<i>CIKNOX6</i>	CCTCCTCGAAGCCTACATGG	GCTGCCGTGCTTCAAATTCT
<i>CIKNOX7</i>	TCCCCTTATCCTCGTTTGCTC	GTCGCTTTCTTGACGGATTCC
<i>CIKNOX8</i>	GCGTTCGTGAACTGCCAAAA	GCCGATATCACTTCCACGGT
<i>CIKNOX9</i>	CTTCACTGGTGGAACAGCCA	TTTTCGGGTTAGCCTGTCTG
<i>CIKNOX10</i>	AAATACTCGTTCGCTGGGCAA	CGAACATGCTGCTGGAGTTG
<i>CIKNOX11</i>	CCAAGCTTCTTGGATTGCCA	GGTGGACGCCATTGAAGACC
<i>CIKNOX12</i>	CCCAGAAGATGCTCTCACCC	ACAAGCAAGCTATCTCCGGAG
<i>ClActin</i>	CCATGTATGTTGCCATCCAG	GGATAGCATGGGGTAGAGCA

表 2 西瓜 *KNOX* 理化性质、亚细胞定位
Table 2 Physical and chemical characteristics, subcellular localization of *KNOX* gene in watermelon

基因名称 Gene name	基因 ID Gene ID	位置 Location	氨基酸数 Amino acid number	蛋白质质量 Molecular mass/ku	等电点 pI	亚细胞定位 Subcellular localization
<i>CIKNOX1</i>	Cla97C01G003320	3115592-3118800	301	33.89	6.16	细胞核 Nucleus
<i>CIKNOX2</i>	Cla97C01G021290	33581956-33585215	351	39.14	6.01	细胞核 Nucleus
<i>CIKNOX3</i>	Cla97C02G028600	2087790-2091705	342	38.59	6.54	细胞核 Nucleus
<i>CIKNOX4</i>	Cla97C02G040680	28605041-28608189	325	36.47	5.12	细胞核 Nucleus
<i>CIKNOX5</i>	Cla97C05G097580	26974963-26979195	467	51.78	6.10	细胞核 Nucleus
<i>CIKNOX6</i>	Cla97C05G098920	28102310-28106247	363	41.66	6.03	细胞核 Nucleus
<i>CIKNOX7</i>	Cla97C05G107680	34523357-34530079	324	36.86	5.12	细胞核 Nucleus
<i>CIKNOX8</i>	Cla97C06G126730	28458880-28459828	164	17.50	4.44	细胞核 Nucleus
<i>CIKNOX9</i>	Cla97C06G126740	28465450-28466835	116	13.59	8.89	细胞核 Nucleus
<i>CIKNOX10</i>	Cla97C08G156950	24581049-24585113	467	51.61	5.59	细胞核 Nucleus
<i>CIKNOX11</i>	Cla97C08G160490	27279508-27284950	378	42.88	5.89	细胞核 Nucleus
<i>CIKNOX12</i>	Cla97C08G161580	27996438-28009722	275	31.38	5.49	细胞核 Nucleus

2.2 *CIKNOX* 家族基因的染色体位置和基因结构

染色体定位结果表明,西瓜 *CIKNOX* 基因家族的 12 个家族成员集中分布于 5 条不同的染色体上(图 1-A)。其中,5 号染色体和 8 号染色体上分布有 3 个家族成员,成员数量最多,1、2、6 号染色体上则分别分布有 2 个家族成员,其他染色体上没有 *CIKNOX* 基因。西瓜 *CIKNOX* 基因家族成员中,*CIKNOX8* 和 *CIKNOX9* 分布于 6 号染色体上,且相邻物理距离仅为 4.62 kb,推断其为串联重复基因簇。由图 1-B 可以看出,西瓜 *CIKNOX* 家族基因均含有内含子,且内含子数目和长度差异较大。其中,*CIKNOX8* 和 *CIKNOX9* 只含有 1 个内含子,*CIKNOX2* 和 *CIKNOX3* 含有 3 个内含子,*CIKNOX1*、*CIKNOX4*、*CIKNOX6*、*CIKNOX7*、

CIKNOX11 和 *CIKNOX12* 均含有 4 个内含子,*CIKNOX5* 和 *CIKNOX10* 含有 5 个内含子。

2.3 *CIKNOX* 家族的系统进化分析

通过对甜瓜、黄瓜、冬瓜、瓠瓜和水稻数据库的检索,并对其保守结构域分析验证,最终分别在各作物上获得的 *KNOX* 基因数量为 12、11、11、12 和 13 个。将这些基因与获得的 12 个 *CIKNOX* 基因和 8 个 *ATKNOX* 基因编码的氨基酸序列结合,利用 MUSCLE、Jalview 和 IQ-TREE 软件构建 *KNOX* 系统发育树。由图 2 可以看出,所有 *KNOX* 基因家族分为 Class I 和 Class II 两个分支,其中 *KNOX* Class I 又进一步分为 IA 和 IB 2 个亚组;*KNOX* Class II 分为 IIA 和 IIB 2 个亚组。12 个 *CIKNOX* 基因中属于 Class IA 的有 5 个,其中

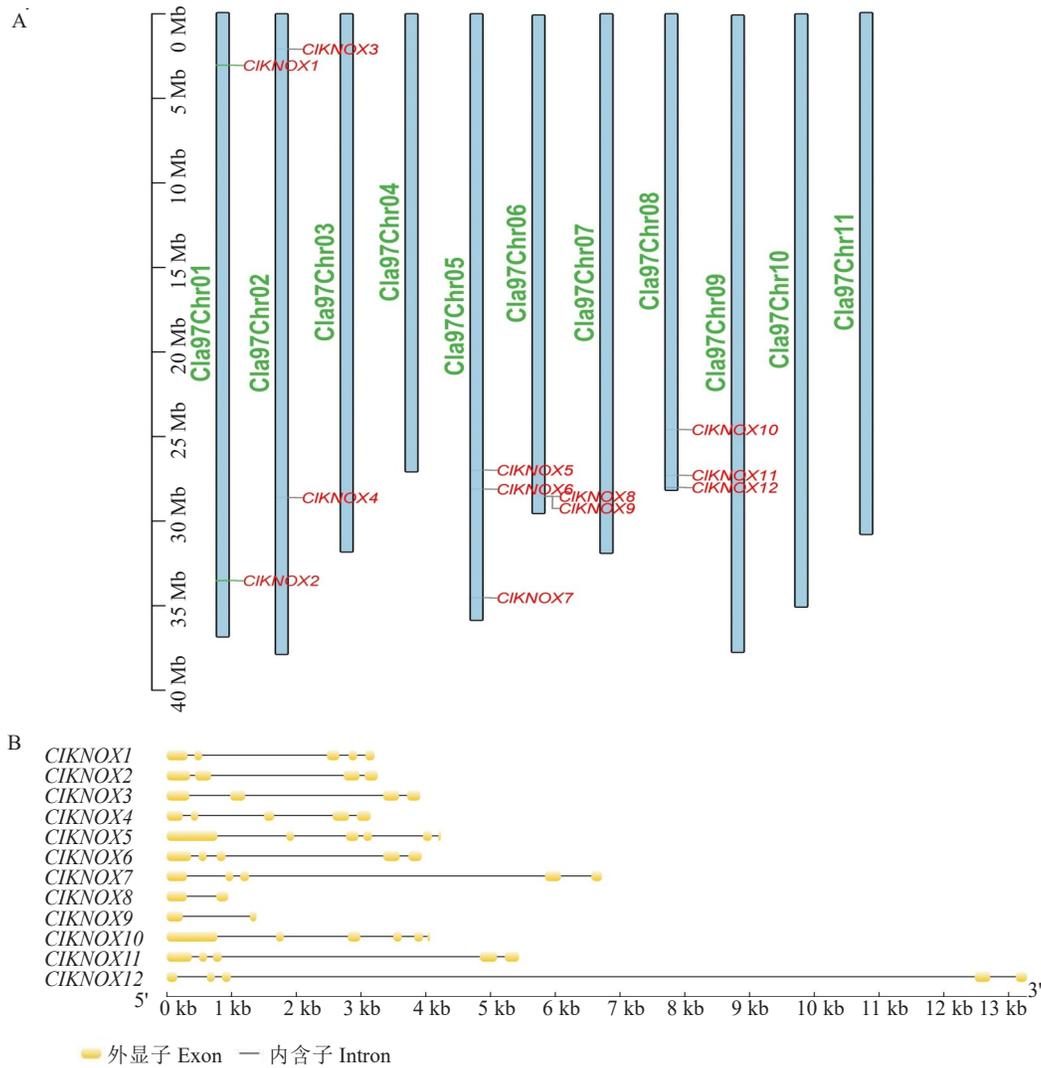


图1 西瓜 *KNOX* 基因在染色体上的分布(A)及基因结构(B)

Fig. 1 Chromosome distribution of *KNOX* gene(A) and gene structure(B) in watermelon

CIKNOX2 和 *CIKNOX3* 与 *AT1G62360.1* 属于同一分支,亲缘关系最近;*CIKNOX6* 和 *AT4G08150.1* 属于同一分支,*CIKNOX9* 和 *CIKNOX11* 同属于 Class IA 亚组。I B 亚组有 4 个 *CIKNOX* 成员,分别是 *CIKNOX4*、*CIKNOX7*、*CIKNOX8* 和 *CIKNOX12*,与 *AT1G70510.1* 和 *AT1G23380.1* 同属于 I B 亚组,亲缘关系较近。Class II A 亚组的成员最少,仅含有 *CIKNOX1*,该基因与 *AT1G62990.1* 亲缘关系最近,Class II B 亚组含有 2 个 *CIKNOX* 成员,分别是 *CIKNOX5* 和 *CIKNOX10*。

2.4 *CIKNOX* 基因家族保守结构域分析

CIKNOX 基因家族的保守结构域分析结果如图 3 所示,西瓜 *KNOX* 基因保守性较强的 8 类基序分别为 Motif 1 (RQVLLDWNSHYKW-PYPTEEEKVALAESTGLDQKQINNWFNQRKR - HWKP)、Motif 2 (IKAKILAHPLYPRLLSAYVDC-

QKVGAPPEVAPRLEEIRAES)、Motif 3 (DPRAE-DRELKDELLRKYSYJSSLKZEFLLKKKKGKLP-KDA)、Motif 4 (DPELDQ FMEAYCEMLVKY-KEELSRPFREA)、Motif 5 (AMEAVMACWEIEQS-LQSLTGVSPGEGTGAT MSDDDDDQ)、Motif 6 (SEDMQFVVMGDGAHPPYYA)、Motif 7 (TLFLN-RIESQLSTLCNGSF) 和 Motif 8 (DANLFDGSLEG-Homeobox_KNAMGFGPLIPTESERS)。其中,Motif 1 对应 Homeobox_KN 结构域,Motif 2 对应 KNOX1 结构域,Motif 3 对应 ELK 结构域,Motif 4 对应 KNOX2 结构域。*CIKNOX* 基因家族成员中,除了 *CIKNOX8* 和 *CIKNOX9* 基因,都含有 4 个典型的结构域,即 KNOX1、KNOX2、ELK 和 Homeobox_KN;*CIKNOX9* 具有 Homeobox_KN 和 ELK 结构域;*CIKNOX8* 只含有 KNOX1 结构域。4 个 Motif 基序中,Motif 1、Motif 2、Motif 3 出现频次为 11 次,Mo-

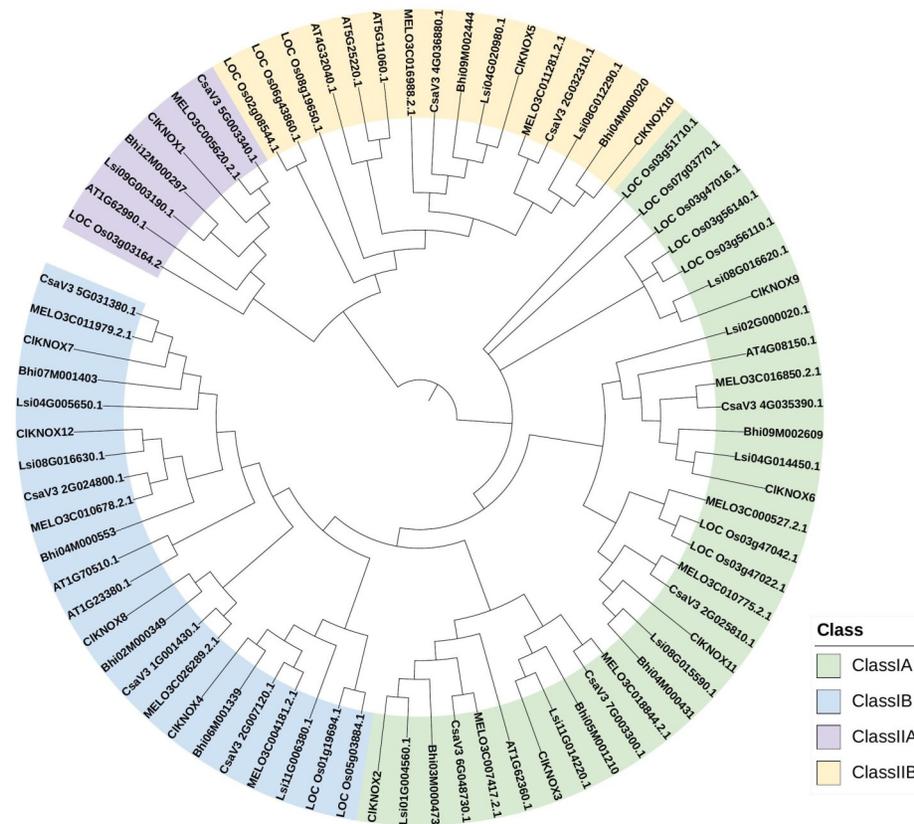


图2 西瓜与拟南芥、黄瓜、甜瓜、冬瓜、瓠瓜和水稻 *KNOX* 基因家族系统进化分析

Fig. 2 Phylogenetic analysis of *KNOX* gene family among watermelon, *Arabidopsis thaliana*, cucumber, melon, wax gourd, bottle gourd and rice

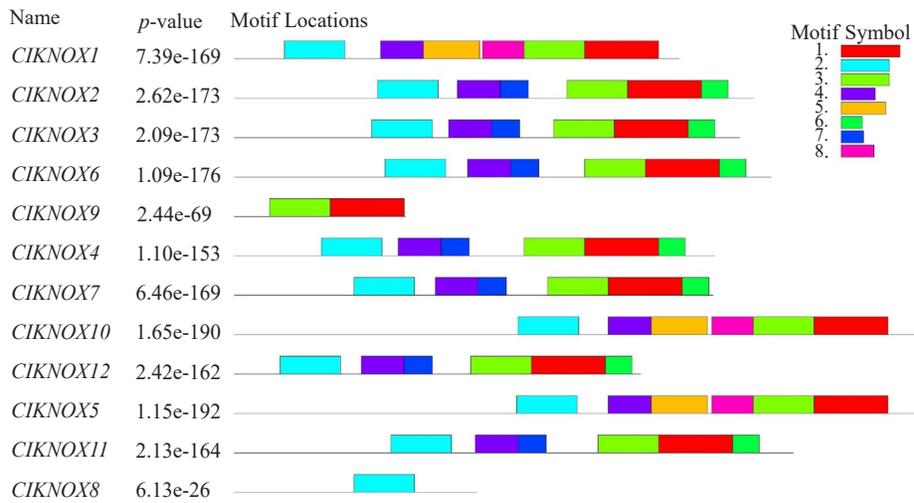


图3 西瓜 *KNOX* 基因家族保守基序与结构域

Fig. 3 Conserved motifs and domains of watermelon *KNOX* gene family

tif 4 的出现频次为 10 次,由此可见 *CIKNOX* 家族基因结构域高度保守。同时,从进化关系上看,亲缘关系较近的同亚家族的 *CIKNOX* 基因在基因结构组成和蛋白结构域分布上也具有较高保守性。

2.5 *CIKNOX* 基因的启动子顺式作用元件分析

顺式作用元件在基因表达调控中具有重要作用,相似功能的基因可能在启动子中具有类似的调控元件。为了更好地探究 *CIKNOX* 家族基因的表达调控机制,提取了 *CIKNOX* 家族成员的上游

2000 bp 的基因组序列,并结合 PlantCARE 数据库对其顺式作用元件进行预测。结果如图 4 所示,几乎所有 *CIKNOX* 基因的启动子区域都存在 4 种类型的顺式作用元件,除了基本启动子元件外,还包括植物生长发育、非生物胁迫反应、激素反应相关的启动子元件。作为与叶绿体发育紧密相关的基因家族,对其 motif 统计发现,启动子中多种元件与光反应相关,数量占到了总元件数量的 56.45%,如 Box 4、GA-motif、MRE、GT1-motif、chs-CMA1a、G-box、GATA-motif、TCT-motif、I-box、TCCC-motif、ATCT-motif、G-box、ATC-motif、3-AF1 binding site、LAMP- element、chs- CMA2a、AE- box、Sp1、TCCC-motif、AAAC-motif 和 ACE 等启动子元件均

与光反应有关,表明其基因表达调控受光的影响较大。与逆境胁迫反应相关元件约占 20.97%,主要包含 TC-rich repeats、ARE、LTR、MBS、WUN-motif 等与植物厌氧、低温、干旱、创伤等反应相关的元件;与激素响应相关的元件约占 14.92%,主要包含 ABRE、TATC-box、P-box、TGA-box、TCA-element 等与脱落酸、赤霉素、生长素、水杨酸等植物激素相关的元件;植物生长发育相关元件约占 7.26%,其中多为 MSA-like、GCN4_motif、CAT- box、O2-site、AACA_motif、AT-rich element 等响应细胞周期调控、胚乳发育、分生组织表达等生长发育过程的元件。综上所述,*CIKNOX* 基因可能在调节植物生长发育和非生物胁迫响应过程中发挥重要作用。

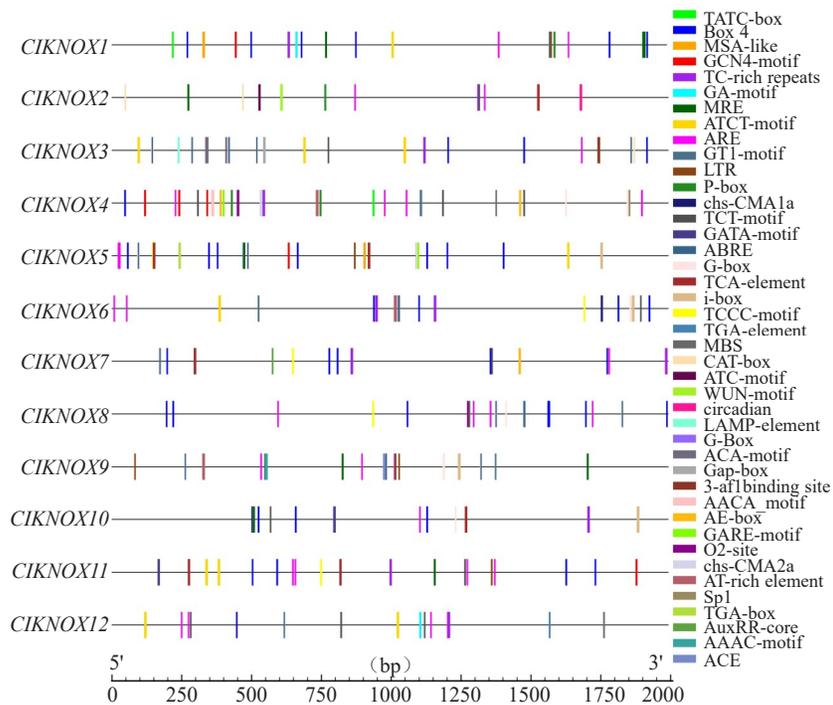


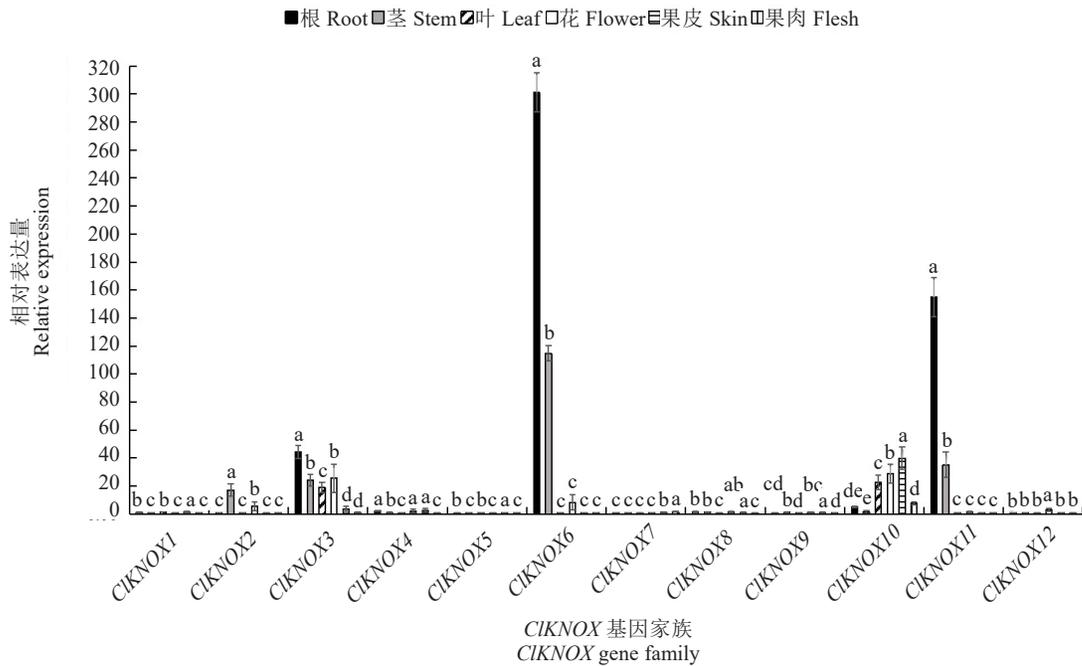
图 4 西瓜 *KNOX* 启动子顺式作用元件分析

Fig. 4 Analysis of cis-acting elements of watermelon *KNOX* promoter

2.6 *CIKNOX* 基因家族的组织表达分析

为探索 *CIKNOX* 家族基因在西瓜中的表达模式,采用荧光定量 qRT-PCR 的方法分析了其在西瓜 WT2 不同组织(根、茎、叶、花、果)中的表达量。由图 5 可以看出,*KNOX* 家族基因在西瓜不同组织中的表达差异较大。在根中,*CIKNOX6* 和 *CIKNOX11* 的表达量较高,其次为 *CIKNOX3* 和 *CIKNOX10*,其余基因的相对表达量都较低;在茎中,*CIKNOX6* 的相对表达量最高,*CIKNOX11*、*CIKNOX3*、*CIKNOX2* 的相对表达量次之;在叶片中,*CIKNOX3* 和 *CIKNOX10* 的相对表达量较高,其余基因都只有微

弱表达;在花中,*CIKNOX10* 和 *CIKNOX3* 的相对表达量较高,其次为 *CIKNOX6* 和 *CIKNOX2*;在果皮中,*CIKNOX10* 的相对表达量最高,其余基因的相对表达量都较低;在果肉中,大多 *CIKNOX* 基因的相对表达量都较低,*CIKNOX8* 和 *CIKNOX12* 不表达。总体而言,*CIKNOX6* 和 *CIKNOX11* 具有相似的表达模式,均在根和茎中具有相对较高的表达量,在其他器官中的表达量均较低。*CIKNOX3* 和 *CIKNOX10* 在所有检测组织中表达量均相对较高,而 *CIKNOX5* 和 *CIKNOX7* 在所有组织中几乎都不表达。



注：同一基因不同小写字母表示在 0.05 水平差异显著。

Note: Different small letters of the same gene indicate significant difference at 0.05 level.

图 5 西瓜 *KNOX* 基因在不同组织中的相对表达量

Fig. 5 The relative expression of *KNOX* gene in different tissues of watermelon

3 讨论与结论

作为一类在植物中广泛存在的转录因子家族，*KNOX* 基因家族在植物的发育和形态建成中起着重要的调控作用，其家族成员在不同的组织和发育阶段中表达具有差异性。随着高通量测序技术的发展，*KNOX* 家族基因已在多个物种中广泛研究，但目前关于西瓜 *CIKNOX* 基因家族尚未有研究报道。笔者结合 *KNOX* 基因家族的典型结构域特征和西瓜基因组数据库信息，首次对 *CIKNOX* 基因家族进行了鉴定，并对其基因结构、蛋白质保守结构域、系统进化和基因表达模式进行分析，为进一步研究 *CIKNOX* 基因的功能奠定了基础。

KNOX 基因广泛存在于植物基因组中，且在不同物种中的数量差异较大，拟南芥中有 8 个^[37]，瓠瓜中有 12 个^[22]，水稻中有 13 个^[38]，杨树中有 15 个^[39]，玉米^[40]和苹果^[41]中各有 22 个，笔者共鉴定出 12 个 *CIKNOX* 基因，数量与同为葫芦科植物的瓠瓜相同，与拟南芥、苹果、玉米的数量差异较大，这种现象通常是由基因进化过程中的部分重复和串联重复等造成的^[42]。进化分析结果表明，*CIKNOX* 中存在多个序列相似度较高的基因，推断其为同源基因。但染色体位置分析结果显示仅有 2 个 *CIKNOX* 基因

为串联重复，推断 *CIKNOX* 基因家族数目的扩增主要由部分重复导致。亚细胞定位预测结果表明，12 个 *CIKNOX* 基因全部定位于细胞核，这与瓠瓜等作物相似^[22]，并符合其转录因子的属性，表明其在生长发育中可能发挥着调控功能。

拟南芥 *KNOX* 基因 Class I 亚家族有 4 个成员 (*KNAT1*、*KNAT2*、*KNAT6* 和 *STM*)，西瓜 Class I 亚家族有 9 个成员基因；拟南芥 *KNOX* 基因 Class II 亚家族也有 4 个成员 (*KNAT3*、*KNAT4*、*KNAT5* 和 *KNAT7*)，西瓜 Class II 亚家族基因成员为 3 个，说明在进化过程中西瓜 *KNOX* 的 Class I 亚家族成员发生数量的扩增，而 Class II 亚家族成员没有发生大量扩增，由此推断 Class I 亚家族基因可能更容易发生基因重复。前人研究表明，基因结构与基因表达方式有密切的对应关系^[8]。Class I 类成员 (*CIKNOX3*、*CIKNOX6*、*CIKNOX11*) 主要在根和茎中表达水平较高，在其他组织的表达水平均较低，这与其他植物 *KNOX* 类基因的表达模式相符合，推测这些基因在西瓜分生组织的发育和调控中起着重要作用。西瓜 12 个 *KNOX* 基因家族成员中，9 个都含有 *KNOX1*、*KNOX2*、*ELK* 和 *Homeobox_KN* 这 4 个典型的结构域，但 *CIKNOX9* 只含有 *Homeobox_KN* 和 *ELK* 结构域，而 *CIKNOX8* 只含有 *KNOX1* 结构

域。缺失部分保守结构域的 *CIKNOX9* 和 *CIKNOX8* 可能是假基因,假基因的存在可能是基因复制不完全的结果,通常不具有完整的基因功能,但这些假基因的存在又为将来新基因的进化提供了可能^[43-44]。

WT2 作为野生型亲本,通过基因定位技术配置遗传群体,已经获得了调控西瓜黄化^[45]、绿花^[46]、短下胚轴和果皮腹纹^[47]等基因,是研究基因组织表达特性的理想材料。前期研究表明,*KNOX* 基因在植物不同组织器官中的表达具有差异性^[22,40,48]。笔者在本研究中发现,*CIKNOX* 基因家族成员在西瓜不同组织中的相对表达量同样具有多样性和较大的差异性。*CIKNOX3* 和 *CIKNOX10* 在西瓜所有器官中表达量均相对较高,推测这两个基因可能参与西瓜各主要器官的生长发育进程。而其他少数 *CIKNOX* 基因(如 *CIKNOX6* 和 *CIKNOX11*) 在根和茎中的表达水平相对较高,在叶、花和果实中的表达量普遍较低,这一表达水平与小麦^[38]、瓠瓜^[22]、大豆^[49]等相似。*CIKNOX6* 在根和茎中表达量最高,系统发育分析表明,*CIKNOX6* 与瓠瓜 *Lsi04G014450* 位于同一进化支,而在瓠瓜中的研究表明,*Lsi04G014450* 在根和茎中的表达量同样最高^[22],两基因具有相似的表达模式,推测该基因可能参与了西瓜根系的生长发育和茎顶端分生组织的分化。*CIKNOX10* 是 *CIKNOX* 家族成员中唯一在果皮中表达量较高的基因,推测其可能参与西瓜果皮的颜色调控。*CIKNOX5* 和 *CIKNOX7* 在所有组织中几乎都不表达,推测其可能存在功能缺失。*CIKNOX* 基因在不同组织中表达模式的差异性暗示了其功能表达的差异性,这为进一步解析 *CIKNOX* 基因在西瓜生长发育和形态建成中的作用奠定了重要基础。

综上所述,*CIKNOX* 基因家族的组织表达模式和启动子元件分析结果表明,该基因家族成员在进化过程中可能发生了功能分化,虽然大多数 *CIKNOX* 基因都含有 KNOX1、KNOX2、ELK 和 Homeobox_KN 这 4 类典型结构域,但不同 *CIKNOX* 基因在不同组织中表达模式的多样性暗示了其功能的差异性,*CIKNOX3*、*CIKNOX6* 和 *CIKNOX11* 可能与根和茎的发育相关,而 *CIKNOX3* 和 *CIKNOX10* 则可能与叶片和花的发育更加密切,不同 *CIKNOX* 基因在不同组织的特异性表达和模式为进一步深入探究 *CIKNOX* 基因在生长发育中的功能奠定了基础。

参考文献

- [1] GEHRING W J, AFFOLTER M, BURGLIN T. Homeodomain proteins[J]. Annual Review of Biochemistry, 1994, 63: 487-526.
- [2] VOLLBRECHT E, VEIT B, SINHA N, et al. The developmental gene *Knotted-1* is a member of a maize homeobox gene family[J]. Nature, 1991, 350(6315): 241-243.
- [3] HAKE S, SMITH H M S, HOLTAN H, et al. The role of *KNOX* genes in plant development[J]. Annual Review of Cell and Developmental Biology, 2004, 20: 125-151.
- [4] GAO J, YANG X, ZHAO W, et al. Evolution, diversification, and expression of *KNOX* proteins in plants[J]. Frontiers in Plant Science, 2015, 6: 882.
- [5] LEE J H, LIN H W, JOO S, et al. Early sexual origins of homeoprotein heterodimerization and evolution of the plant *KNOX/BELL* family[J]. Cell, 2008, 133(5): 829-840.
- [6] BURGLIN T R. Analysis of TALE superclass homeobox genes (*MEIS*, *PBC*, *KNOX*, *Iroquois*, *TGIF*) reveals a novel domain conserved between plants and animals[J]. Nucleic Acids Research, 1997, 25(21): 4173-4180.
- [7] CHEN H, ROSIN F M, PRAT S, et al. Interacting transcription factors from the three-amino acid loop extension superclass regulate tuber formation[J]. Plant Physiology, 2003, 132(3): 1391-1404.
- [8] KERSTETTER R, VOLLBRECHT E, LOWE B, et al. Sequence analysis and expression patterns divide the maize *knotted1*-like homeobox genes into two classes[J]. The Plant Cell, 1994, 6(12): 1877-1887.
- [9] MAGNANI E, HAKE S. *KNOX* lost the OX: The *Arabidopsis KNATM* gene defines a novel class of *KNOX* transcriptional regulators missing the homeodomain[J]. Plant Cell, 2008, 20(4): 875-887.
- [10] LONG J A, MOAN E I, MEDFORD J I, et al. A member of the *KNOTTED* class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*[J]. Nature, 1996, 379(6560): 66-69.
- [11] TRUERNIT E, SIEMERING K R, HODGE S, et al. A map of *KNAT* gene expression in the *Arabidopsis* root[J]. Plant Molecular Biology, 2006, 60(1): 1-20.
- [12] DOUGLAS S J, CHUCK G, DENGLER R E, et al. *KNAT1* and *ERECTA* regulate inflorescence architecture in *Arabidopsis*[J]. Plant Cell, 2002, 14(3): 547-558.
- [13] VENGLAT S P, DUMONCEAUX T, ROZWADOWSKI K, et al. The homeobox gene *BREVIPEDICELLUS* is a key regulator of inflorescence architecture in *Arabidopsis*[J]. Proceedings of the National Academy of Sciences, 2002, 99(7): 4730-4735.
- [14] BELLES-BOIX E, HAMANT O, WITIAK S M, et al. *KNAT6*: An *Arabidopsis* homeobox gene involved in meristem activity and organ separation[J]. Plant Cell, 2006, 18(8): 1900-1907.
- [15] HAMANT O, NOGUE F, BELLES-BOIX E, et al. The *KNAT2* homeodomain protein interacts with ethylene and cytokinin signaling[J]. Plant Physiology, 2002, 130(2): 657-665.
- [16] DEAN G, CASSON S, LINDSEY K. *KNAT6* gene of *Arabidopsis* is expressed in roots and is required for correct lateral root

- formation[J]. *Plant Molecular Biology*, 2004, 54(1): 71-84.
- [17] SERIKAWA K A, MARTINEZ-LABORDA A, ZAMBRYSKI P. Three *knotted1*-like homeobox genes in *Arabidopsis*[J]. *Plant Molecular Biology*, 1996, 32: 673-683.
- [18] BIRNBAUM K, SHASHA D E, WANG J Y, et al. A gene expression map of the *Arabidopsis* root[J]. *Science*, 2003, 302(5652): 1956-1960.
- [19] LI E Y, BHARGAVA A, QIANG W Y, et al. The Class II KNOX gene *KNAT7* negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in *Populus*[J]. *New Phytologist*, 2012, 194(1): 102-115.
- [20] 张磊, 武竞春, 张海英, 等. 西瓜重要性状功能基因研究进展[J]. *园艺学报*, 2023, 50(12): 2748-2764.
- [21] LIU H, JIAO J Q, LIANG X J, et al. Map-based cloning, identification and characterization of the *w* gene controlling white immature fruit color in cucumber (*Cucumis sativus* L.) [J]. *Theoretical and Applied Genetics*. 2016, 129(7): 1247-1256.
- [22] 徐晓燕, 黄小忠, 刘万仙, 等. 瓠瓜 KNOX 基因家族全基因组鉴定及组织表达分析[J]. *西北农林科技大学学报(自然科学版)*, 2024, 52(3): 113-122.
- [23] GUO S G, ZHANG J G, SUN H H, et al. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions[J]. *Nature Genetics*, 2013, 45(1): 51-58.
- [24] MISTRY J, CHUGURANSKY S, WILLIAMS L, et al. Pfam: The protein families database in 2021[J]. *Nucleic Acids Research*, 2021, 49(D1): 412-419.
- [25] FINN R D, CLEMENTS J, EDDY S R. HMMER web server: Interactive sequence similarity searching[J]. *Nucleic Acids Research*, 2011, 39(2): 29-37.
- [26] MARCHLER-BAUER A, LU S, ANDERSON J B, et al. CDD: A conserved domain database for the functional annotation of proteins[J]. *Nucleic Acids Research*, 2010, 39(1): 225-229.
- [27] GASTEIGER E, GATTIKER A, HOOGLAND C, et al. ExPASy: The proteomics server for in-depth protein knowledge and analysis[J]. *Nucleic Acids Research*, 2003, 31(13): 3784-3788.
- [28] 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.
- [29] HU B, JIN J, GUO A Y, et al. GSDB 2.0: An upgraded gene feature visualization server[J]. *Bioinformatics*, 2015, 31(8): 1296-1297.
- [30] BAILEY T L, JOHNSON J, GRANT C E, et al. The MEME suite[J]. *Nucleic Acids Research*, 2015, 43(W1): 39-49.
- [31] CHEN C J, WU Y, LI J W, et al. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining[J]. *Molecular Plant*, 2023, 16(11): 1733-1742.
- [32] EDGAR R C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput[J]. *Nucleic Acids Research*, 2004, 32(5): 1792-1797.
- [33] MINH B Q, SCHMIDT H A, CHERNOMOR O, et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era[J]. *Molecular Biology and Evolution*, 2020, 37(8): 2461.
- [34] LESCOT M, DEHAIS P, THIJS G, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences[J]. *Nucleic Acids Research*, 2002, 30(1): 325-327.
- [35] WOFFELMEN C. DNAMAN for Windows, Version 5. 2. 10: Lynon Biosoft[D]. Leiden, The Netherlands: Institute of Molecular Plant Sciences, Leiden University, 2004.
- [36] LIVAK K J, SCHMITTGEN T D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method[J]. *Methods*, 2001, 25(4): 402-408.
- [37] 宋晓峰, 鲁成凯, 董晓娜, 等. 花生 KNOX 基因家族的全基因组鉴定及组织表达分析[J]. *山东农业科学*, 2023, 55(8): 1-10.
- [38] 宋宁宁, 梁慧慧, 安义伟, 等. 小麦 KNOX 基因家族鉴定与生物信息学分析[J]. *分子植物育种*, 2022, 20(7): 2105-2117.
- [39] XIONG H, SHI A, WU D, et al. Genome-Wide identification, classification and evolutionary expansion of KNOX gene family in rice (*Oryza sativa*) and populus (*Populus trichocarpa*) [J]. *American Journal of Plant Sciences*, 2018, 9(6): 1071-1092.
- [40] 赵长江, 宋巨奇, 都梦翔, 等. 玉米 KNOX 基因家族鉴定及组织和逆境表达分析[J]. *西北植物学报*, 2021, 41(7): 1109-1119.
- [41] JIA P, XING L B, ZHANG C G, et al. MdKNOX15, a class I knotted-like transcription factor of apple, controls flowering and plant height by regulating GA levels through promoting the Md-GA2ox7 transcription[J]. *Environmental and Experimental Botany*, 2021, 185: 104411.
- [42] FLAGEL L E, WENDEL J F. Gene duplication and evolutionary novelty in plants[J]. *New Phytologist*, 2009, 183(3): 557-564.
- [43] LIU J J, EKRAMODDOULLAH A K M. Isolation, genetic variation and expression of TIR-NBS-LRR resistance gene analogs from western white pine (*Pinus monticola* Dougl. ex. D. Don.) [J]. *Molecular Genetics and Genomics*, 2004, 270(5): 432-441.
- [44] NAIR R A, THOMAS G. Isolation, characterization and expression studies of resistance gene candidates (RGCs) from *Zingiber* spp. [J]. *Theoretical and Applied Genetics*, 2007, 116(1): 123-134.
- [45] 段世享. 西瓜植株黄化基因 *Clyp* 的图位克隆与功能验证[D]. 郑州: 河南农业大学, 2023.
- [46] 孙东玲. 西瓜绿花基因 *CIGF* 和“星月西瓜”基因 *MS* 的定位研究[D]. 郑州: 河南农业大学, 2022.
- [47] 王登科. 西瓜果皮覆纹基因 (*CIGS*) 和短下胚轴基因 (*Clsh*) 的精细定位[D]. 郑州: 河南农业大学, 2022.
- [48] 叶明辉, 赵朋, 牛洋, 等. 马铃薯同源异形框基因家族的鉴定和表达分析[J]. *农业生物技术学报*, 2021, 29(2): 224-239.
- [49] 李方正, 杨素欣, 吴春霞, 等. 大豆 KNOX 基因家族的结构和表达分析[J]. *植物学报*, 2012, 47(3): 236-247.