

徐珍珍, 季 为, 陈祥龙, 等. 异常棉 *NF-YA* 家族基因鉴定及其在干旱胁迫下的表达分析[J]. 江苏农业学报, 2023, 39(6): 1265-1274.
doi: 10.3969/j.issn.1000-4440.2023.06.001

异常棉 *NF-YA* 家族基因鉴定及其在干旱胁迫下的表达分析

徐珍珍, 季 为, 陈祥龙, 徐 鹏, 郭 琪, 沈新莲

(江苏省农业科学院经济作物研究所/农业农村部长江下游棉花与油菜重点实验室, 江苏 南京 210014)

摘要: *NF-YA* 转录因子在植物生长发育和逆境响应中扮演重要角色。本研究通过生物信息学手段, 分析棉花二倍体野生种异常棉基因组中 *NF-YA* 家族基因的数目、染色体分布、系统进化、保守基序、结构、顺式作用元件、表达模式及编码蛋白质的亚细胞定位、理化性质等。结果表明, 异常棉 *NF-YA* 基因家族 (*GoanoNF-YA*) 包含 17 个成员, 分布在 10 条染色体上, 编码的蛋白质长度为 190~603 个氨基酸, 理论等电点为 7.87~10.07, 16 个成员包含 3~5 个外显子且编码的蛋白质定位于细胞核。拟南芥和异常棉 *NF-YA* 基因在系统进化树上分为 6 类, 每类基因在系统进化树上具有相同或相似类型和排列顺序的基序。5 个 *NF-YA* 基因在异常棉根、茎、叶和花中广泛表达, 15 个 *NF-YA* 基因对于干旱持续时间有不同程度的响应, 11 个 *NF-YA* 基因含有干旱响应顺式作用元件 MBS。本研究结果为进一步解析异常棉 *NF-YA* 家族基因的功能提供了理论依据。

关键词: 异常棉; *NF-YA* 基因家族; 生物信息; 干旱; 表达模式

中图分类号: S562.035.3 文献标识码: A 文章编号: 1000-4440(2023)06-1265-10

Identification of *NF-YA* family genes in *Gossypium anomalum* and their expression analysis under drought stress

XU Zhen-zhen, JI Wei, CHEN Xiang-long, XU Peng, GUO Qi, SHEN Xin-lian

(*Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences/Key Laboratory of Cotton and Rapeseed, Ministry of Agriculture and Rural Affairs, Nanjing 210014, China*)

Abstract: *NF-YA* transcription factors play an important role in plant growth and development and stress response. In this study, the number, chromosome distribution, phylogenetic evolution, conserved motifs, gene structure, *cis*-acting elements and expression patterns of *NF-YA* family genes in the genome of *Gossypium anomalum* and the physical and chemical properties and subcellular localization of encoded proteins were analyzed by bioinformatics. The results showed that there were 17 members of *NF-YA* gene family in the *G. anomalum* genome (*GoanoNF-YA*), which were distributed on 10 chromosomes. The length of the encoded proteins was 190–603 amino acids, and the theoretical isoelectric point was 7.87–10.07. Sixteen *GoanoNF-YA* members contained three to five exons and encoded proteins located in the nucleus. The *NF-YA* genes of *Arabidopsis thaliana* and *G. anomalum* were divided into six groups, and each group had the same or similar type and order of motifs. Five *NF-YA* genes were widely expressed in roots, stems, leaves and flowers in *G. anomalum*, 15 *NF-YA* genes had different degrees of response to drought stress, and 11 *NF-YA* genes contained drought-response *cis*-acting element MBS. The results of this study provide a theoretical basis for further exploring the function of *GoanoNF-YA* family genes in *G. anomalum*.

Key words: *Gossypium anomalum*; *NF-YA* gene family; bioinformatics; drought stress; expression pattern

收稿日期: 2023-03-14

基金项目: 国家重点研发计划项目(2022YFD1200300); 江苏省重点研发计划项目(BE2022364)

作者简介: 徐珍珍(1984-), 女, 山东泰安人, 博士, 副研究员, 主要从事植物分子育种研究。(E-mail) le20032@163.com

通讯作者: 沈新莲, (E-mail) xlshen68@126.com

核因子 Y (Nuclear factor Y, NF-Y) 又称为血红素激活蛋白 (Heme activator protein, HAP) 或者 CCAAT 结合因子 (CCAAT-binding factor, CBF), 是

一类能够结合基因启动子特异序列 CCAAT 的转录因子,普遍存在于酵母、植物和动物中^[1]。NF-Y 由 NF-YA (CBF-A 或 HAP2)、NF-YB (CBF-B 或 HAP3) 和 NF-YC (CBF-C 或 HAP5) 3 个亚基蛋白组成^[2]。各亚基都具有独特的高度保守核心区,其中 NF-YA 蛋白具有 53 个氨基酸组成的保守核心区,包括保守结构域 A1 和 A2。其中 A1 是 NF-YA 与 NF-YB/NF-YC 异源二聚体相互作用的 N 端,而结构域 A2 是 NF-YA 参与 DNA 结合的 C 端,在特异识别靶基因启动子 CCAAT 序列中发挥重要作用^[3]。一般情况下,NF-Y 调控下游基因表达的机制有 3 种。第一,NF-YA、NF-YB 和 NF-YC 构成活性异源三聚体,与下游目标基因的启动子 CCAAT-box 区域结合^[4];第二,NF-YB 和 NF-YC 形成异源二聚体 NF-YB/NF-YC,然后与特定转录因子结合形成复合物来调节下游基因表达,而 NF-YA 能够阻止 NF-YB/NF-YC 二聚体与特定转录因子结合,从而抑制其功能^[5-6];第三,NF-Y 亚基通过与其他蛋白复合物相互作用来调控下游基因的转录水平^[4]。目前 NF-YA 基因家族在多个物种中已有报道。水稻 (*Oryza sativa*)^[7] 和拟南芥 (*Arabidopsis thaliana*)^[8] 基因组中各包含 10 个 NF-YA 成员。谷子 (*Setaria italica*)^[9] 和玉米 (*Zea mays*)^[10] 中分别鉴定到 10 个和 14 个成员;油菜 (*Brassica napus*)^[11]、大豆 (*Glycine max*)^[12]、鹰嘴豆 (*Cicer arietinum*)^[13]、茶树 (*Camellia sinensis*)^[14] 和桃 (*Prunus persica*)^[15] 基因组中分别鉴定出 14 个、21 个、8 个、10 个和 6 个成员。越来越多的研究表明,NF-YA 基因广泛参与植物胚胎发育、种子形成和萌发^[16-17]、根的形态建成^[18]、开花^[19] 和光合作用^[20] 等生长发育过程,响应干旱^[9]、高温^[21] 和高盐^[9, 21] 等逆境胁迫。

棉花 (*Gossypium* spp.) 是世界上重要的纺织工业原料,是饲料蛋白和食用油的重要来源之一,在中国国民经济中具有重要地位。近年来,粮棉争地矛盾的加剧加速了中国棉花生产的区域化,年降水量少、干旱问题十分突出的新疆地区已成为中国棉花主产区。目前广泛种植的四倍体栽培种陆地棉 (*G. hirsutum*) 是在正常环境条件下选育出的,抗旱性弱。二倍体野生种异常棉 (*G. anomalum*) 生长于撒哈拉沙漠边缘和非洲西南部^[22],植株矮小且密被茸毛,叶色浅、叶片小,具有较强的抗旱性^[23]。因此,利用异常棉抗旱基因改良栽培种陆地棉资源是中国棉花产业发展的重

要方向。目前,异常棉基因组数据已在公共数据库公布^[24],但对异常棉抗旱性状相关的 NF-YA 家族基因分析未见报道。本研究基于异常棉参考基因组信息,利用生物信息学手段对 NF-YA 家族基因进行全基因组鉴定,分析该家族基因的系统进化、保守基序、基因结构、顺式作用元件、组织特异性表达、干旱胁迫下的表达模式以及该家族基因编码蛋白质的亚细胞定位、理化性质等,旨在为后续深入研究异常棉 NF-YA 基因家族在干旱胁迫下的功能提供理论依据。

1 材料与方法

1.1 试验材料

拟南芥 NF-YA 家族基因编码的氨基酸序列来自 TAIR (<https://www.arabidopsis.org/>) 网站;异常棉基因组注释的氨基酸和编码序列 (CDS) 来自 CottonGen 网站 (<https://www.cottongen.org/>)^[21];异常棉根、茎、叶和花的转录组数据以及干旱胁迫下异常棉叶片的转录组数据来自 NCBI 网站 (<https://www.ncbi.nlm.nih.gov/>)^[24-25]。

1.2 异常棉 NF-YA 家族基因的全基因组鉴定及编码的蛋白质理化性质分析和亚细胞定位

首先在 SMART 网站 (<http://smart.embl-heidelberg.de/>) 预测拟南芥 NF-YA 蛋白家族的保守结构域,均含有 CBF_B_NFYA 结构域 (PF02045),然后在 Pfam 网站 (<https://pfam-legacy.xfam.org/>) 下载种子文件 PF02045.18,利用搜索工具 HMMER3.1b1 检索异常棉氨基酸序列,并将获得的 NF-YA 蛋白成员在 NCBI CDD 和 SMART 蛋白数据库进行验证,最终将通过验证的基因作为异常棉 NF-YA 家族基因 (*GoanoNF-YA*) 并进行命名;GoanoNF-YA 蛋白家族成员的等电点和相对分子质量信息利用在线网站 (https://web.expasy.org/compute_pi/) 进行预测,利用在线软件 Cell-PLoc 2.0 (<http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/>) 进行异常棉 NF-YA 蛋白家族成员的亚细胞定位预测;利用 TBtools 软件提取 *GoanoNF-YA* 家族基因在染色体的位置信息。

1.3 异常棉 NF-YA 家族基因的系统进化、基因结构和顺式作用元件分析

利用 MEGA11 软件的邻接法 (Neighbor-Joining, NJ) 构建拟南芥和异常棉 NF-YA 家族基因的系统进化树,并进行 1 000 次 Boot strap 抽样自检;利用 GS-DS 在线软件 (<http://gsds.cbi.pku.edu.cn>) 分析 *Go-*

anoNF-YA 基因的外显子和内含子结构;利用 MEME 软件 (<http://meme-suite.org/tools/meme>) 对 *GoanoNF-YA* 基因进行保守基序 (motif) 分析;利用 PlantCARE 软件 (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html>) 分析异常棉 *NF-YA* 家族基因启动子顺式作用元件。

1.4 异常棉 *NF-YA* 家族基因的表达分析

利用 Tophat^[26] 和 Cufflink^[27] 软件包分析异常棉根、茎、叶和花及异常棉叶片在干旱胁迫下的转录组测序数据 Clean data 并进行表达量计算,用 FPKM (Fragments per kilobase of transcript permillion fragments mapped) 表示。提取 *NF-YA* 家族基因的表达数据,进行标准化后,利用 TBtools 绘制 *GoanoNF-YA*

基因表达量热图。

2 结果与分析

2.1 异常棉 *NF-YA* 家族基因的全基因组鉴定及编码蛋白质的理化性质分析和亚细胞定位

根据异常棉氨基酸数据库共鉴定出 17 个异常棉 *NF-YA* 家族基因,根据基因所在的染色体及位置分别命名为 *GoanoNF-YA1* ~ *GoanoNF-YA17*。*NF-YA* 家族基因编码的氨基酸序列长度为 190~603 aa,平均 315 aa。蛋白质的相对分子质量为 21 890.3~66 522.6,平均 34 774.3。理论等电点为 7.87~10.07,平均 8.94,均为碱性。除 *GoanoNF-YA4* 定位在叶绿体,其余 16 个成员均定位于细胞核(表 1)。

表 1 异常棉 *NF-YA* 家族基因的基本信息

Table 1 Information of the *NF-YA* family genes in *Gossypium anomalum*

基因编号	基因名称	基因编码的蛋白质			
		氨基酸序列长度	相对分子质量	理论等电点	位置
<i>Goano01G0359</i>	<i>GoanoNF-YA1</i>	330	36 193.9	9.04	细胞核
<i>Goano01G1339</i>	<i>GoanoNF-YA2</i>	353	39 125.0	8.93	细胞核
<i>Goano02G0945</i>	<i>GoanoNF-YA3</i>	325	35 911.5	8.67	细胞核
<i>Goano04G1265</i>	<i>GoanoNF-YA4</i>	603	66 522.6	8.43	叶绿体
<i>Goano04G2414</i>	<i>GoanoNF-YA5</i>	335	36 419.0	8.74	细胞核
<i>Goano05G1226</i>	<i>GoanoNF-YA6</i>	203	22 483.9	8.87	细胞核
<i>Goano05G1530</i>	<i>GoanoNF-YA7</i>	327	36 012.3	8.97	细胞核
<i>Goano06G1011</i>	<i>GoanoNF-YA8</i>	190	21 890.3	10.07	细胞核
<i>Goano07G0154</i>	<i>GoanoNF-YA9</i>	302	33 821.9	9.13	细胞核
<i>Goano08G0686</i>	<i>GoanoNF-YA10</i>	304	33 486.5	8.68	细胞核
<i>Goano08G2057</i>	<i>GoanoNF-YA11</i>	286	31 252.8	8.66	细胞核
<i>Goano10G0040</i>	<i>GoanoNF-YA12</i>	203	22 331.7	9.52	细胞核
<i>Goano10G0560</i>	<i>GoanoNF-YA13</i>	332	36 470.7	8.47	细胞核
<i>Goano10G2650</i>	<i>GoanoNF-YA14</i>	309	34 754.3	9.76	细胞核
<i>Goano11G4166</i>	<i>GoanoNF-YA15</i>	314	34 376.7	9.49	细胞核
<i>Goano13G0199</i>	<i>GoanoNF-YA16</i>	315	34 011.5	7.87	细胞核
<i>Goano13G0445</i>	<i>GoanoNF-YA17</i>	323	36 098.2	8.62	细胞核

异常棉 *NF-YA* 家族基因分布在 10 条染色体上。10 号染色体上有 3 个 *NF-YA* 基因;1 号、4 号、5 号、8 号和 13 号染色体上各有 2 个 *NF-YA* 基因;2 号、6 号、7 号和 11 号染色体上各有 1 个 *NF-YA* 基因;3 号、9 号和 12 号染色体上没有 *NF-YA* 基因(图 1)。

2.2 异常棉 *NF-YA* 家族基因的系统进化树和保守基序

异常棉和拟南芥 *NF-YA* 家族基因编码的氨基酸

序列比对后构建的系统进化树如图 2 所示。从系统进化树可以看出,异常棉和拟南芥 *NF-YA* 家族基因可分为 6 类(图 2)。I 类中有 5 个异常棉 *NF-YA* 基因(*GoanoNF-YA3*、*GoanoNF-YA2*、*GoanoNF-YA10*、*GoanoNF-YA11* 和 *GoanoNF-YA16*)和 2 个拟南芥 *NF-YA* 基因(*AtNF-YA1* 和 *AtNF-YA9*)。II 类中有 3 个异常棉 *NF-YA* 基因(*GoanoNF-YA15*、*GoanoNF-YA4* 和 *GoanoNF-YA1*)和 2 个拟南芥 *NF-YA* 基因(*AtNF-YA2* 和 *AtNF-YA10*)。III 类

中有 2 个异常棉 *NF-YA* 基因 (*GoanoNF-YA6* 和 *GoanoNF-YA12*) 和 2 个拟南芥 *NF-YA* 基因 (*AtNF-YA7* 和 *AtNF-YA4*)。IV 类中仅有 3 个异常棉 *NF-YA* 基因 (*GoanoNF-YA8*、*GoanoNF-YA17* 和 *GoanoNF-YA9*)。V 类中

有 4 个异常棉 *NF-YA* 基因 (*GoanoNF-YA5*、*GoanoNF-YA13*、*GoanoNF-YA7* 和 *GoanoNF-YA14*) 和 2 个拟南芥 *NF-YA* 基因 (*AtNF-YA5* 和 *AtNF-YA6*)。VI 类中仅有 2 个拟南芥 *NF-YA* 基因 (*AtNF-YA8* 和 *AtNF-YA3*)。

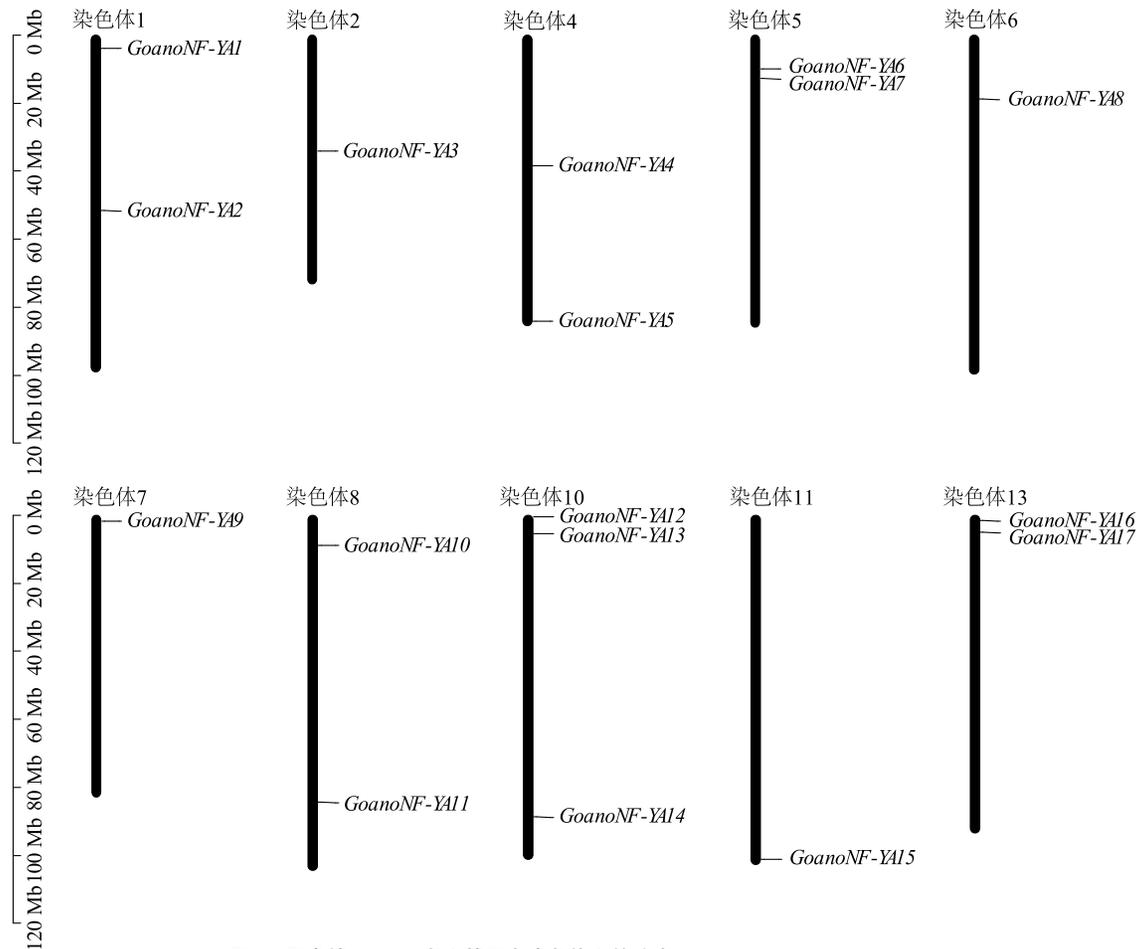
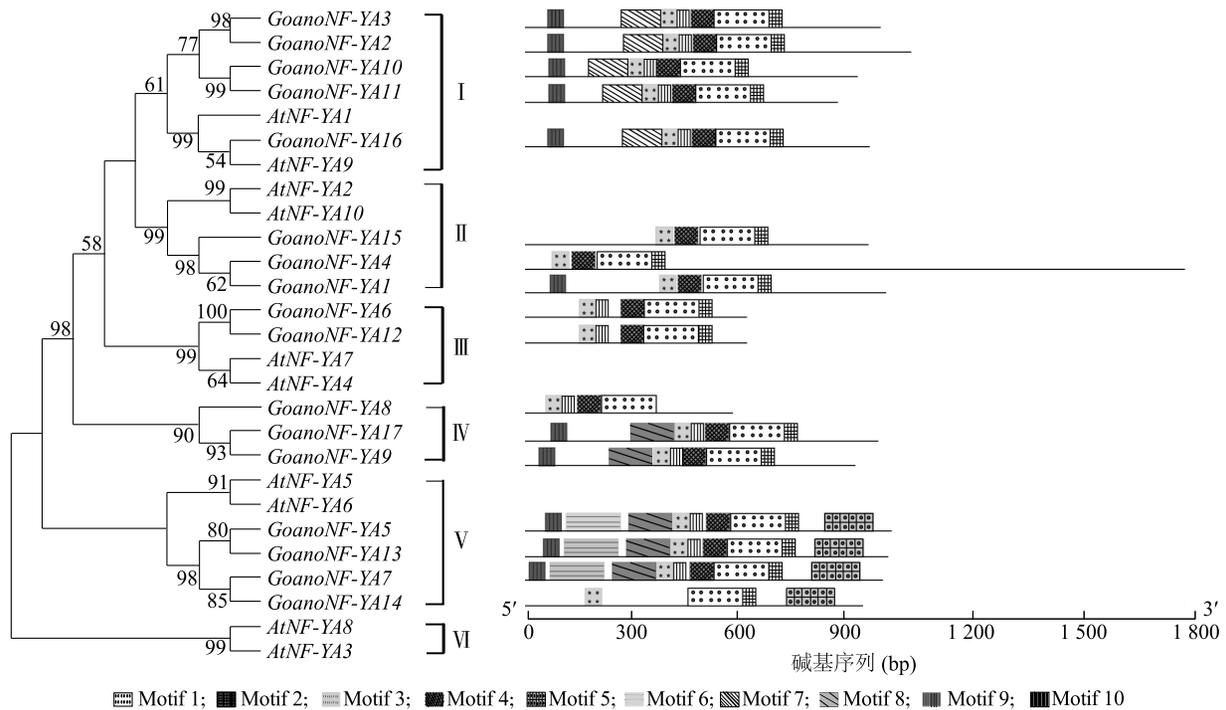


图 1 异常棉 *NF-YA* 家族基因在染色体上的分布

Fig.1 Chromosomal distribution of the *NF-YA* family genes in *G. anomalum*

系统进化树上分支近的 *NF-YA* 家族成员具有相同或类似的基序类型、数目和排列顺序 (图 2)。I 类中, 5 个 *GoanoNF-YA* 成员 (*GoanoNF-YA3*、*GoanoNF-YA2*、*GoanoNF-YA10*、*GoanoNF-YA11* 和 *GoanoNF-YA16*) 均具有 7 个相同类型的保守基序 (motif1、motif2、motif3、motif4、motif7、motif9 和 motif10) 且排列顺序相同; II 类中, *GoanoNF-YA15* 和 *GoanoNF-YA4* 基因均具有 4 个相同类型的保守基序 (motif1、motif2、motif3 和 motif7) 且排列顺序相同, 而 *GoanoNF-YA1* 基因中多了一个 motif9, 但其余 4 个保守基序的类型和排列顺序与 *GoanoNF-YA15* 和 *GoanoNF-YA4* 基因一致; III 类中, *GoanoNF-YA6* 和 *GoanoNF-YA12* 基因均包含 5 个相同类型的保守基序

(motif1、motif2、motif3、motif4 和 motif10) 且排序顺序完全一致; IV 类中, *GoanoNF-YA17* 和 *GoanoNF-YA9* 基因含有 7 个相同类型的保守基序 (motif1、motif2、motif3、motif4、motif8、motif9 和 motif10) 且排列顺序相同, 而 *GoanoNF-YA8* 基因仅包含与 *GoanoNF-YA17* 和 *GoanoNF-YA9* 相同类型的 motif1、motif3、motif4 和 motif10 4 个保守基序; V 类中, *GoanoNF-YA5*、*GoanoNF-YA13* 和 *GoanoNF-YA7* 基因含有 9 个相同类型的保守基序 (motif1、motif2、motif3、motif4、motif5、motif6、motif8、motif9 和 motif10) 且排列顺序相同, 而 *GoanoNF-YA14* 仅含有与 *GoanoNF-YA5*、*GoanoNF-YA13* 和 *GoanoNF-YA7* 相同类型的 4 个保守基序 (motif1、motif2、motif3 和 motif5)。



Motif 1~Motif 10 分别为保守基序 1~保守基序 10。

图 2 异常棉和拟南芥 *NF-YA* 基因系统进化树及异常棉 *NF-YA* 基因的保守基序

Fig.2 Phylogenetic tree of *NF-YA* genes in *Gossypium anomalum* and *Arabidopsis thaliana* and conserved motifs of *NF-YA* genes in *G. anomalum*

2.3 异常棉 *NF-YA* 家族基因的结构

异常棉 *NF-YA* 家族基因的结构分析结果表明,除了 *GoanoNF-YA4* 含有 9 个外显子外,其余 16 个成员含有 3~5 个外显子,系统进化树上分支近的 *NF-YA* 家族成员具有相同或类似的基因结构(图 3)。I 类中,5 个基因 *GoanoNF-YA3*、*GoanoNF-YA2*、*GoanoNF-YA10*、*GoanoNF-YA11* 和 *GoanoNF-YA16* 均含有 5 个外显子,其中 *GoanoNF-YA2* 基因的内含子较长。II 类中,*GoanoNF-YA15* 和 *GoanoNF-YA1* 基因均含有 5 个外显子,*GoanoNF-YA4* 基因含有 9 个大小不等的外显子;*GoanoNF-YA15* 基因具有较长的内含子。III 类中,*GoanoNF-YA6* 和 *GoanoNF-YA12* 基因均含有 5 个外显子,第 1 个外显子均较短;*GoanoNF-YA6* 内含子较长。IV 类中,*GoanoNF-YA17* 和 *GoanoNF-YA9* 基因含有 5 个外显子,*GoanoNF-YA8* 仅含有 4 个外显子;3 个基因的内含子均较短。V 类中,*GoanoNF-YA5*、*GoanoNF-YA13* 和 *GoanoNF-YA7* 等基因均含有 5 个外显子,而 *GoanoNF-YA14* 基因仅含有 3 个外显子;4 个基因成员的内含子均较短(图 3)。

2.4 异常棉 *NF-YA* 家族基因启动子顺式作用元件

异常棉 *NF-YA* 家族基因的相关顺式作用元件分为 5 类,包括胁迫、激素响应、生长发育、光反应和其他相关

元件(图 4)。胁迫响应顺式作用元件包括 ARE、DRE-core、GC-motif、LTR、MBS、TC-rich repeats 和 WUN-motif。*GoanoNF-YA2*、*GoanoNF-YA11*、*GoanoNF-YA16*、*GoanoNF-YA15*、*GoanoNF-YA12*、*GoanoNF-YA17*、*GoanoNF-YA9*、*GoanoNF-YA8* 和 *GoanoNF-YA14* 等 9 个基因含有 1~2 个干旱响应顺式作用元件 MBS。*GoanoNF-YA15*、*GoanoNF-YA1*、*GoanoNF-YA8* 和 *GoanoNF-YA7* 等 4 个基因含有 1~2 个干旱、低温和盐胁迫响应顺式作用元件 DRE-core。17 个 *GoanoNF-YA* 基因都含有抗氧化顺式作用元件 ARE, 1~7 个不等;仅 *GoanoNF-YA17* 基因含有 GC-motif 无氧诱导响应元件。激素响应顺式作用元件包括 ABRE、AuxRR-core、CGTCA-motif、GARE-motif、P-box、plant_AP-2-like、TATC-box、TCA-element、TGACC-motif 和 TGA-element。*GoanoNF-YA3*、*GoanoNF-YA12*、*GoanoNF-YA8* 和 *GoanoNF-YA13* 等 4 个基因不含有脱落酸响应顺式作用元件 ABRE,其余 13 个 *GoanoNF-YA* 基因都含有 ABRE, 1~5 个不等。生长发育相关的顺式作用元件包括 CAT-box、GCN4_motif、HD-Zip、MBSI、Myb-related、O2-site 和 RY-element。异常棉 *NF-YA* 基因启动子分别包含 7 个 CAT-box、1 个 GCN4_motif、4 个 HD-Zip、1 个 MBSI、127 个 Myb-related、13 个 O2-site 和 4 个 RY-element,涉及分生组织发育、胚乳发育、栅栏组织分化和种子发育的特异调

控。参与光反应的顺式作用元件包括 AAAC-motif、AE-box、AT1-motif、ATCT-motif、AT-rich sequence、Box 4、CCAAT-box、chs-CMA、circadian、G-Box、GA-motif、GATA-motif、GT1-motif、I-box、MRE、TCCC-motif、TCT-motif 和 W box。其他顺式作用元件包括 TATA-box、CAAT-box、Myc-related、AT~TATA-box、AAGAA-motif 和 ERE 等。

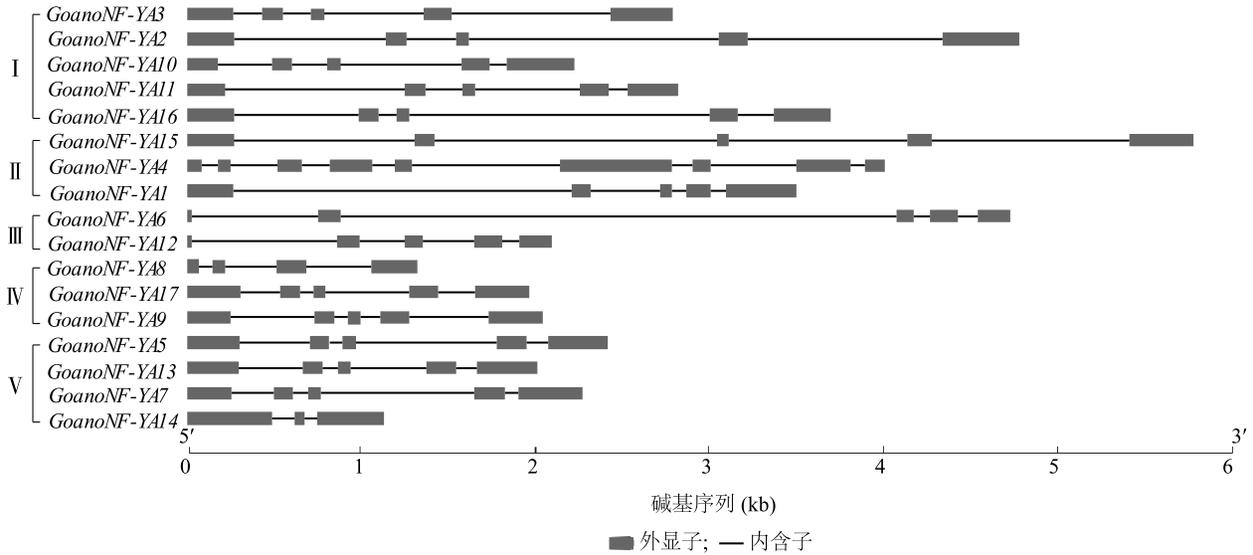


图 3 异常棉 NF-YA 基因结构特征

Fig.3 Gene structure of NF-YA family in *G. anomalum*

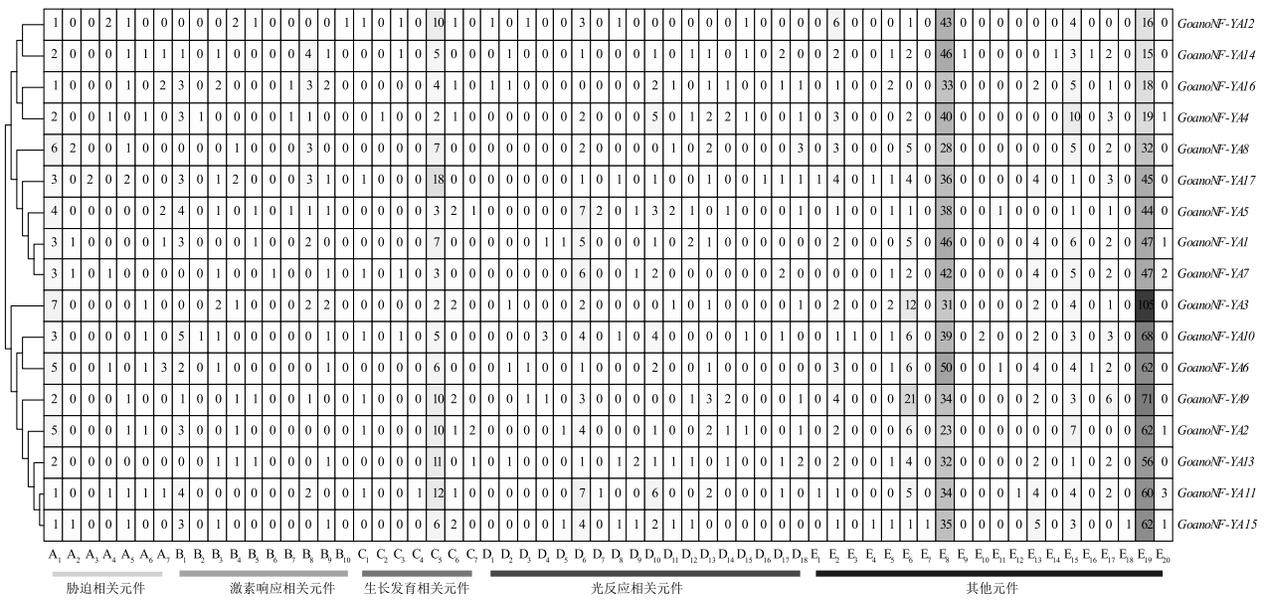


图 4 异常棉 NF-YA 家族基因的顺式作用元件

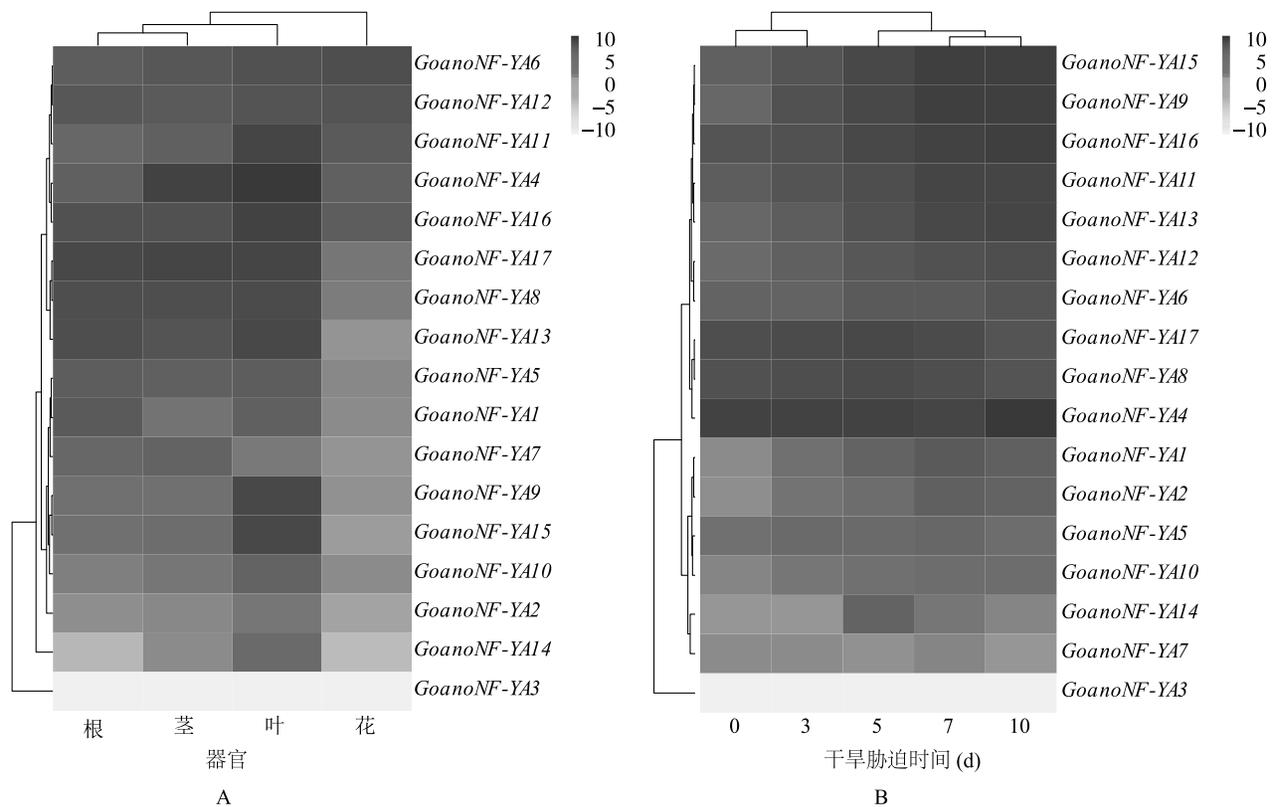
Fig.4 The cis-acting elements NF-YA family genes in *G. anomalum*

2.5 异常棉 NF-YA 家族基因的表达模式

异常棉 NF-YA 家族基因的组织表达模式如图 5A 所示。从图中可以看出, *GoanoNF-YA6*、*GoanoNF-YA12*、*GoanoNF-YA11*、*GoanoNF-YA4* 和 *GoanoNF-YA16* 等基因在根、茎、叶和花中均有较高表达; *GoanoNF-YA17*、*GoanoNF-YA8*、*GoanoNF-YA13*、*GoanoNF-YA5*、*GoanoNF-YA1* 和 *GoanoNF-YA7* 等基因在根、茎和叶中表达量较高,而在花中表达量较低; *GoanoNF-YA9*、*GoanoNF-YA15*、*GoanoNF-YA10*、*GoanoNF-YA2* 和 *GoanoNF-YA14* 等基因在叶中表达量最高,根、茎中表达量次之,花中表达量最低; *GoanoNF-YA3* 基因在根、茎、叶和花中几乎不表达。

不同干旱持续时间(0 d、3 d、5 d、7 d 和 10 d)下,异常棉叶片 *GoanoNF-YA* 家族基因的表达模式如图 5B 所示。从图中可以看出, *GoanoNF-YA15*、

GoanoNF-YA9、*GoanoNF-YA16*、*GoanoNF-YA11*、*GoanoNF-YA13*、*GoanoNF-YA12* 和 *GoanoNF-YA6* 共 7 个基因在干旱胁迫处理下表达量较高,且随着干旱胁迫时间增加表达量呈现上升的趋势; *GoanoNF-YA17* 和 *GoanoNF-YA8* 等基因在干旱胁迫处理后,表达量先升高后降低; *GoanoNF-YA4* 基因在干旱胁迫处理后表达量没有明显的变化; *GoanoNF-YA1*、*GoanoNF-YA2*、*GoanoNF-YA5* 和 *GoanoNF-YA10* 等基因在干旱胁迫处理下表达量较低,但随着干旱胁迫时间增加表达量呈现上升的趋势; *GoanoNF-YA14* 基因在干旱 5 d 时表达量最高,而在其他时期表达量较低; *GoanoNF-YA7* 基因在干旱胁迫期间表达量均较低, *GoanoNF-YA3* 基因在干旱胁迫期间均不表达。



A:异常棉不同器官 NF-YA 家族基因的表达模式;B:异常棉 NF-YA 家族基因在干旱胁迫过程中的表达变化特征。

图 5 异常棉 NF-YA 家族基因的表达

Fig.5 The expression of NF-YA family genes in *G. anomalum*

3 讨论

NF-YA 为 NF-Y 复合体的一个亚基蛋白,虽进

化上较为保守,但具有多样性的生物功能^[28-31]。本研究在二倍体野生种异常棉基因组中鉴定到 17 个 *GoanoNF-YA* 成员,同其他单子叶植物水稻、谷子和

玉米以及双子叶植物拟南芥、油菜、大豆、鹰嘴豆、桃、白菜和茶树等^[7-15]一样, *NF-YA* 基因在异常棉基因组中也发生了扩张。这在一定程度上表明, *NF-YA* 家族成员在进化过程中可能存在功能分化和遗传冗余, 从而帮助植物在长期的进化过程中形成灵活的、复杂的和多功能的转录调控系统以适应复杂多变的环境。系统进化分析结果表明, 拟南芥和异常棉 *NF-YA* 家族成员可分为 6 类, 每类 2 个物种的基因数目不同, 且 2 个物种都存在单独聚类的情况。如 IV 类中仅有异常棉 *NF-YA* 成员, 没有拟南芥相关成员, 而 VI 类中仅有拟南芥 2 个成员 *AtNF-YA8* 和 *AtNF-YA3*, 没有异常棉 *NF-YA* 成员。这说明 *NF-YA* 基因在物种进化中发生了适应性进化, 这与以往报道^[7-15]是一致的。通过保守基序和基因结构预测, 系统进化树上分支近的 *NF-YA* 家族成员具有相同或相似类型和排列顺序的基序, 推测分支较近的 *NF-YA* 家族成员间的功能可能存在着相似性。异常棉 *GoanoNF-YA* 家族基因中除 *GoanoNF-YA4* 外, 其余 16 个成员均包含 3~5 个外显子, 与其他作物具有相似的外显子数目^[9-11]。因此, 可以认为 *NF-YA* 外显子数目在进化上较为保守。

Zhao 等^[4] 研究结果表明, 拟南芥 *NF-YA* 基因能调控胚胎发生、根系生长、种子萌发、叶片发育、花期和逆境胁迫等生物学过程, 这种基因的多样性功能与其启动子的顺式作用元件有关。启动子顺式作用元件通过调控目的基因的转录进而影响植物的生长发育、非生物胁迫响应和光反应等^[28-30]。本研究发现异常棉 *GoanoNF-YA* 基因 5' 上游 2 000 bp 的序列中存在胁迫、植物激素、生长发育和光响应等相关顺式作用元件。与胁迫相关的元件包括干旱胁迫响应元件 MBS、脱落酸响应元件 ABRE、干旱、低温和盐胁迫响应元件 DRE 和抗氧化作用响应元件 GC-motif 等; 与生长发育相关的元件有 CAT-box、GCN4-motif、HD-Zip、RY-element、O2-site 和 MBSI 等, 分别响应分生组织发育、胚乳发育、栅栏组织分化、种子发育、玉米醇溶蛋白代谢调节和类黄酮生物合成等过程; 与光反应相关的顺式作用元件有 ATC-motif、ATCT-motif、Box4 和 MRE 等。这些结果表明异常棉 *GoanoNF-YA* 基因具有多样的生物功能。

NF-YA 基因在植物体中呈现多样化的组织表达模式。拟南芥 10 个 *NF-YA* 基因中, 有 8 个基因在胚胎中表达, 5 个基因 (*AtNF-YA1*、*AtNF-YA2*、*AtNF-*

YA3、*AtNF-YA7* 和 *AtNF-YA9*) 在胚乳中表达^[17]。亚洲棉 (*G. arboretum*) 的 *GaNF-YA4* 基因在种子和纤维发育过程中表达量较高^[32]。陆地棉 (*G. hirsutum*) 的 *GhNF-YA18* 基因在茎、叶、花瓣和雄蕊中表达量较高, 而 *GhNF-YA27* 基因在茎中表达量较高^[33]。小麦中 *TaNF-YA3*、*TaNF-YA4*、*TaNF-YA7* 和 *TaNF-YA9* 等基因优先在胚乳组织中表达^[34]。本研究中, 异常棉 17 个 *GoanoNF-YA* 家族基因, 有 5 个基因在根、茎、叶和花中均高表达, 11 个基因在根、茎、叶均较高表达, 但在花中表达量较低。这种基因的多样化组织表达模式说明 *NF-YA* 基因参与了较多的生物学过程, 具有多样化的生物功能, 在植物的生长发育和逆境响应中扮演着非常重要的角色。拟南芥 *AtNF-YA5* 基因在干旱胁迫 14 d 后表达量上调, 其异位表达增强了植株抗旱性^[35]。在拟南芥中过表达大豆 *GmNF-YA3* 基因可提高植株的抗旱性^[36]。本研究中, 异常棉 *GoanoNF-YA* 基因在干旱胁迫后, 除了 *GoanoNF-YA7* 和 *GoanoNF-YA3* 外, 其余 15 个基因均对干旱胁迫产生了响应。进化树上与 *AtNF-YA5* 同源性较近的 *GoanoNF-YA13* 在异常棉干旱胁迫处理下表达量均较高, 且随着干旱胁迫时间增加表达量呈现上升的趋势。

4 结论

本研究通过生物信息学手段从棉花野生种异常棉基因组中鉴定出了 17 个 *NF-YA* 家族基因, 分布在 10 条染色体上, 编码的蛋白质长度为 190~603 个氨基酸, 理论等电点为 7.87~10.07, 16 个基因包含 3~5 个外显子。异常棉 *NF-YA* 基因与拟南芥 *NF-YA* 基因在系统进化树上可分为 6 类, 每类成员具有相似或者相同类型和排列顺序的基序。异常棉 *NF-YA* 家族基因具有组织特异性表达模式且参与了干旱胁迫响应, 与 *AtNF-YA5* 同源性较近的 *GoanoNF-YA13* 在干旱胁迫处理下表达量较高, 推测其在调控干旱胁迫响应中发挥重要作用。本研究结果为进一步开展异常棉 *NF-YA* 基因家族功能研究奠定了基础。

参考文献:

- [1] TESTA A, DONATI G, YAN P, et al. Chromatin immunoprecipitation (ChIP) on chip experiments uncover a widespread distribution of NF-Y binding CCAAT sites outside of core promoters [J].

- The Journal of Biological Chemistry, 2005, 280(14): 13606-13615.
- [2] ROMIER C, COCCHIARELLA F, MANTOVANI R, et al. The NFYB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y [J]. Journal of Biological Chemistry, 2003, 278(2): 1336-1345.
- [3] XING Y Y, ZHANG S, OLESEN J T, et al. Subunit interaction in the CCAAT-binding heteromeric complex is mediated by a very short alpha-helix in HAP2 [J]. Proceedings of the National Academy of Sciences of the United States of America, 1994, 91(8): 3009-3013.
- [4] ZHAO H, WU D, KONG F, et al. The *Arabidopsis thaliana* nuclear factor Y transcription factors [J]. Front Plant Sci, 2017, 7: 2045.
- [5] SATO H, MIZOI J, TANAKA H, et al. *Arabidopsis* DPB3-1, a DREB2A interactor, specifically enhances heat stress-induced gene expression by forming a heat stress-specific transcriptional complex with NF-Y subunits [J]. Plant Cell, 2014, 26(12): 4954-4973.
- [6] HACKENBERG D, WU Y, VOIGT A, et al. Studies on differential nuclear translocation mechanism and assembly of the three subunits of the *Arabidopsis thaliana* transcription factor NF-Y [J]. Mol Plant, 2012, 5(4): 876-888.
- [7] THIRUMURUGAN T, ITO Y, KUBO T A, et al. Identification, characterization and interaction of *HAP* family genes in rice [J]. Molecular Genetics and Genomics, 2008, 279(3): 279-289.
- [8] SIEFERS N, DANG K K, KUMIMOTO R W, et al. Tissue-specific expression patterns of *Arabidopsis* NF-Y transcription factors suggest potential for extensive combinatorial complexity [J]. Plant Physiologist, 2009, 149(2): 625-641.
- [9] FENG Z J, HE G H, ZHENG W J, et al. Foxtail millet NF-Y families: genome-wide survey and evolution analyses identified two functional genes important in abiotic stresses [J]. Frontiers in Plant Science, 2015, 6: 1142.
- [10] ZHANG Z B, LI X L, ZHANG C, et al. Isolation, structural analysis, and expression characteristics of the maize nuclear factor Y gene families [J]. Biochemical Biophysical Research Communications, 2016, 478(2): 752-758.
- [11] LIANG M X, YIN X Z, LIN Z Y, et al. Identification and characterization of NF-Y transcription factor families in Canola (*Brassica napus* L.) [J]. Planta, 2014, 239(1): 107-126.
- [12] QUACH T, NGUYEN H, VALLIYODAN B, et al. Genome-wide expression analysis of soybean NF-Y genes reveals potential function in development and drought response [J]. Molecular Genetic Genomics, 2015, 290(3): 1095-1115.
- [13] CHU H D, NGUYEN K H, WATANABE Y, et al. Identification, structural characterization and gene expression analysis of members of the nuclear factor-Y family in Chickpea (*Cicer arietinum* L.) under dehydration and abscisic acid treatments [J]. International Journal of Molecular Sciences, 2018, 19(11): 3290.
- [14] WANG P J, ZHENG Y C, GUO Y C, et al. Identification, expression, and putative target gene analysis of nuclear factor-Y (NF-Y) transcription factors in tea plant (*Camellia sinensis*) [J]. Planta, 2019, 250(1): 1671-1686.
- [15] LI M, LI G X, LIU W, et al. Genome-wide analysis of the NF-Y gene family in peach (*Prunus persica* L.) [J]. BMC Genomics, 2019, 20: 612.
- [16] MU J Y, TAN H L, HONG S L, et al. *Arabidopsis* transcription factor genes *NF-YA1*, 5, 6, and 9 play redundant roles in male gametogenesis, embryogenesis, and seed development [J]. Molecular Plant, 2013, 6(1): 188-201.
- [17] SIRIWARDANA C L, KUMIMOTO R W, JONES D S, et al. Gene family analysis of the *Arabidopsis* NF-Y transcription factors reveals opposing abscisic acid responses during seed germination [J]. Plant Molecular Biology Reporter, 2014, 32(5): 971-986.
- [18] SORIN C, DECLERCK M, CHRIST A, et al. A miR169 isoform regulates specific NF-Y targets and root architecture in *Arabidopsis* [J]. New Phytologist, 2014, 202(4): 1197-1211.
- [19] SIRIWARDANA C L, KUMIMOTO R W, JONES D S, et al. Nuclear factor Y, subunit A (NF-YA) proteins positively regulate flowering and act through *FLOWERING LOCUS T* [J]. PLoS Genetics, 2016, 12(12): e1006496.
- [20] ALAM M M, TANAKA T, NAKAMURA H, et al. Overexpression of a rice heme activator protein gene (*OsHAP2E*) confers resistance to pathogens, salinity and drought, and increases photosynthesis and tiller number [J]. Plant Biotechnology Journal, 2014, 13(1): 85-96.
- [21] LI Y J, FANG Y, FU Y R, et al. NFYA1 is involved in regulation of postgermination growth arrest under salt stress in *Arabidopsis* [J]. PLoS ONE, 2013, 8(4): e61289.
- [22] FRYXELL P A. A revised taxonomic interpretation of *Gossypium* L. (Malvaceae) [J]. Rheedea, 1992, 2(2): 108-165.
- [23] GANESH S N, VIVEK P C, SUBHASH S M, et al. Interspecific hybridization in *Gossypium* L.: characterization of progenies with different ploidy-confirmed multigenomic backgrounds [J]. Plant Breeding, 2013, 132: 211-216.
- [24] XU Z Z, CHEN J D, MENG S, et al. Genome sequence of *Gossypium anomalum* facilitates interspecific introgression breeding [J]. Plant Communications, 2022, 3(5): 100350.
- [25] JI W, YU H, SHANGGUAN Y, et al. Transcriptome profiling of *Gossypium anomalum* seedlings reveals key regulators and metabolic pathways in response to drought stress [J]. Plants (Basel), 2023, 12(2): 312.
- [26] TRAPNELL C, PACHTER L, SALZBERG S L. TopHat: discovering splice junctions with RNA-Seq [J]. Bioinformatics, 2009, 25(9): 1105-1111.
- [27] TRAPNELL C, ROBERTS A, GOFF L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks [J]. Nature Protocols, 2012, 7(3): 562-578.
- [28] 宋秋明,李大勇,张慧娟,等. 植物 NF-Y 转录因子的生物学功

- 能及其研究进展 [J]. 植物生理学报, 2015, 51: 623-632.
- [29] 丁慧霞, 刘 凤, 张利娟, 等. 植物中 NF-Y 转录因子的结构和功能研究进展 [J]. 分子植物育种, 2017, 15: 1691-1701.
- [30] 李 娟, 高 凯, 安新民. 转录因子 NF-Y 在植物生长发育和逆境胁迫响应中的作用 [J]. 中国细胞生物学学报, 2019, 41: 2434-2442.
- [31] 刘雪梅, 王 悦, 李玮琦, 等. 辣椒 miR169 及其靶基因的生物信息学分析 [J]. 江苏农业学报, 2021, 37(6): 1510-1515.
- [32] 王园园, 赵春月, 孙润润, 等. 亚洲棉 *NF-YA* 基因家族的全基因组鉴定及表达分析 [J]. 分子植物育种, 2021, 19(14): 4564-4573.
- [33] 潘 奥, 王静静, 孙福来, 等. 陆地棉 *NF-YA* 基因家族的全基因组鉴定与功能分析 [J]. 棉花学报, 2020, 32(4): 316-328.
- [34] STEPHENSON T J, MCINTYRE C L, COLLET C, et al. Genome-wide identification and expression analysis of the NF-Y family of transcription factors in *Triticum aestivum* [J]. Plant Molecular Biology, 2007, 65(1/2): 77-92.
- [35] LI W X, OONO Y, ZHU J, et al. The *Arabidopsis* NFYA5 transcription factor is regulated transcriptionally and post transcriptionally to promote drought resistance [J]. Plant Cell, 2008, 20(8): 2238-2251.
- [36] NI Z Y, HU Z, JIANG Q Y, et al. *GmNFYA3*, a target gene of miR169, is a positive regulator of plant tolerance to drought stress [J]. Plant Molecular Biology, 2013, 82(1/2): 113-129.

(责任编辑:石春林)