

DOI:10.3969/j.issn.1672-5565.201702001

# 亚洲棉 *ARF* 基因家族全基因组鉴定与分析

冯庭\*, 司国阳, 李国辉, 郭宁, 蔡永萍

(安徽农业大学 生命科学学院, 合肥 230036)

**摘要:**生长素响应因子(*ARF*)是植物中可以特异性结合生长素应答基因启动子的转录因子,在植物的生长发育中起着重要的调控作用。为了更好的了解亚洲棉中*ARF*基因家族的种类和数量,本文利用生物信息学的方法,分析了亚洲棉基因组中*ARF*基因的种类、进化关系、物理定位以及基因的结构和保守基序。结果表明在亚洲棉基因组中发现并初步确定了29个*ARF*基因,命名为*GaARF*,通过基因定位发现除在2、3、5、11号染色体上没有*ARF*基因外,其他染色体上都有*ARF*基因存在;而且在10号染色体上还存在着基因簇;进化树分析表明,亚洲棉部分*ARF*蛋白和已知功能的拟南芥聚类到一起;表达模式分析显示,不同组织不同处理条件下*ARF*基因的表达存在差异。

**关键词:**亚洲棉; *ARF*; 进化关系

中图分类号: S759.95 文献标志码: A 文章编号: 1672-5565(2018)01-022-07

## Gome-wide identification and analysis of *ARF* gene family in *Gossypium arboretum* L.

FENG Ting\*, SI Guoyang, LI Guohui, GUO Ning, CAI Yongping

(School of Life Sciences, Anhui Agriculture University, Hefei 230036, China)

**Abstract:** Auxin response factor (*ARF*) is an important transcription factor that can specifically bind auxin responsive gene promoter in plants, and influences the plant growth and development. To better understand the type and quantity of *ARF* gene families in *Gossypium arboretum*, we used the method of bioinformatics to analyze the structure types physical mapping of genes and conserved motifs of genome in *Gossypium arboretum*. The results show that 29 *ARF* genes, named *GaARF* were identified in the genome of *Gossypium arboretum*. All chromosomes contain the *ARF* gene except the 2, 3, 5 and 11 by analyzing the position of genes. In addition, there is a gene cluster on chromosome 10. Phylogenetic analysis showed that some *ARF* proteins and proteins with known function in *Arabidopsis* were clustered together. Expression profile analysis indicated the expression of *ARF* genes were diverse in different tissues.

**Keywords:** *Gossypium arboretum* L.; *ARF*; Evolution relationship

转录因子对植物的生长、发育和响应外界环境的变化起着重要的作用<sup>[1]</sup>,生长素响应因子(*ARF*)就是其中的一类,生长素与植物种子的萌发,器官的形成和向性反映等诸多生理过程有密切联系<sup>[2]</sup>。生长素响应因子能够与生长素响应原件TGTCTC序列特异性结合并调控生长素响应基因的表达<sup>[3]</sup>,这些响应基因包括*AUX/IAA*家族、*CH3*家族和受生长素调控的RNA调控的*SAUR*基因家族<sup>[4]</sup>。*ARF*家

族成员在N-端有一个能与DNA特异性结合的B3结构域(DBD);在中间有一个与转录活性有关的结构域(MR),MR区域的氨基酸构成和序列长度决定了其分子量和转录活性<sup>[5]</sup>;一般认为,中间区域含丝氨酸(S)、苏氨酸(T)和脯氨酸(P)是抑制子(RD),而富含谷氨酰胺(Q)的多为激活子(AD)<sup>[6]</sup>。在拟南芥中,*AtARF1*、2、3、4、9是转录抑制子;*AtARF5*、6、7、8、9是转录激活子。大多数

收稿日期:2017-02-17;修回日期:2017-03-14.

基金项目:安徽省生物学重点学科(2014SWQJ009).

\*通信作者:冯庭,男,在读硕士,研究方向:作物遗传学;E-mail:793114910@qq.com.

*ARF* 基因还含有 2 个 C 端的结构域(功能域 III 和 IV),能够与 *ARF* 蛋白本身或 *AUX/IAA* 结合形成同源或异源二聚体<sup>[7]</sup>。生长素响应因子蛋白是由一个大的基因家族编码的<sup>[7]</sup>,目前,拟南芥<sup>[8]</sup>、玉米<sup>[6]</sup>、雷蒙德氏棉<sup>[9]</sup>等物种的 *ARF* 基因家族均被鉴定出来,且关于模式植物 *ARF* 基因家族的功能也有很多研究。亚洲棉与雷蒙德氏棉同为二倍体棉花,有关 *ARF* 基因家族的报道还较少。

亚洲棉(*Gossypium arboreum* L.)是一种原产于亚洲的棉花种类的合称,是人类最早种植的农作物之一,同时也是目前世界最重要的经济棉种陆地棉的祖先。2014 年 4 月,亚洲棉基因组(A 基因组)测序的完成<sup>[10]</sup>,为研究 *ARF* 基因家族成员的生物学功能提供了丰富的资源。本研究利用生物信息学的方法,鉴定了亚洲棉 *ARF* 基因家族,命名为 *GaARF* 从基因水平上分析了 *ARF* 基因在亚洲棉中的数目、序列特征、进化关系,旨在为进一步研究生长素相关基因提供参考。

## 1 材料和方法

### 1.1 *GaARF* 基因家族的筛选及氨基酸序列属性分析

以锦葵科亚洲棉为研究对象,亚洲棉全基因组数据来源于中国农业科学院棉花研究所 CGP 数据库(<http://cgp.genomics.org.cn>)。使用“BioEdit”软件将获得的亚洲棉全基因组氨基酸序列建立一个本地数据库,然后以拟南芥 *ARF* 蛋白序列作为 Query 序列与建立的亚洲棉氨基酸序列进行 BlastP (E-value=0.001) 序列比对,将所得结果在 Pfam<sup>[11]</sup> 数据库(<http://pfam.xfam.org/>)和 CDD<sup>[12]</sup>(<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)进行检测,筛选出含 *ARF* 基因特征结构域(B3、*AUX/IAA*)的序列。将所得到的氨基酸序列使用 ExPASy<sup>[13]</sup>(<http://www.expasy.org/>)进行在线分析,以确定氨基酸的等电点(PI)、蛋白质的分子量(MW)和不稳定系数。

### 1.2 *GaARF* 基因家族进化树的构建与分类

利用 MEGA7.0<sup>[14]</sup> 软件里提供的 ClustalW 工具对亚洲棉 *ARF* 基因氨基酸序列进行多序列比对,然后通过邻接法(NJ, Neighbor-Joining),校验参数(Bootstrap)设置为 1 000,构建系统进化树。不同物种 *ARF* 蛋白进化树的绘制:采用拟南芥、玉米共计 83 个 *ARF* 蛋白序列进行综合分析,拟南芥、玉米 *ARF* 蛋白序列均来自 NCBI 数据库(<https://www.ncbi.nlm.nih.gov/>)。用上述方法进行综合分析,用上述方法绘制 *ARF* 蛋白家族的复合进化树。

### 1.3 染色体定位、基因结构及序列保守性分析

亚洲棉染色体长度及起始位点信息从棉花基因数据库(<https://www.cottongen.org>)中得到,使用 MapInspect 软件将得到的 *GaARF* 基因在染色体上的物理位置绘制成图。按照 *GaARF* 基因在染色体上的位置进行命名。采用 GSDB<sup>[15]</sup>(<http://gsds.cbi.pku.edu.cn/>)在线工具获得亚洲棉 *ARF* 家族外显子-内含子结构图。用已得到的蛋白序列,使用 MEME<sup>[16]</sup> 在线分析工具(<http://meme.sdsc.edu/>)分析得到亚洲棉 *ARF* 家族蛋白的 motif 模式。

### 1.4 信号肽检测及亚细胞定位

利用 TargetP<sup>[17]</sup> 信号肽预测在线网站(<http://www.cbs.dtu.dk/services/TargetP/>)信号肽预测程序,对 29 条亚洲棉 *ARF* 蛋白进行信号肽预测。

### 1.5 组织表达模式分析

在棉花转录组数据库(<http://www.plexdb.org/plex.php?database=Cotton>)中将亚洲棉基因数据通过 Blast 比对,得到 *GaARF* 基因在根和叶中的表达情况。

## 2 结果与分析

### 2.1 *GaARF* 基因家族成员鉴定

利用生物信息学方法,从亚洲棉中共获得 29 个 *ARF* 家族成员(见表 1),同时通过 EsPAsy 在线网站查询所有的 *GaARF* 基因的基本信息,并对蛋白质的长度、分子量、等电点、不稳定系数等理化性质进行分析。结果显示,亚洲棉中 29 个蛋白序列差异很大,编码 *ARF* 蛋白的氨基酸长度范围从 576 aa (*GaARF11*)到 1 141 aa (*GaARF14*),分子量范围从 63.44 kD (*GaARF11*)到 127.54 kD (*GaARF14*),等电点范围从 5.21 (*GaARF21*)到 8.93 (*GaARF24*)。从蛋白的基本特征可以看出该基因家族无论从基因长度到蛋白特性变化都很大,表明该基因家族具有不同的特性,发挥着不同的生物学作用。一般认为,蛋白的不稳定系数大于 40 时,则蛋白属于不稳定蛋白,本研究中亚洲棉蛋白的不稳定系数都大于 43,因此均属于不稳定蛋白。值得注意的是,亚洲棉大部分蛋白的等电点都小于 7,表明多数 *GaARF* 编码的蛋白呈弱酸性,可能在酸性的亚细胞环境中发挥相应的作用。

### 2.2 *GhARF* 基因结构和保守结构域分析

系统进化分析是比较同源序列和研究分子进化的方法,利用本研究中的 *GaARF* 蛋白构建进化树(见图 1),参照 SUN<sup>[9]</sup> 等人的分类方式,将 *GaARF* 基因家族分为: I、II、III (III a、III b、III c)、IV 四大

类,以此来更好的了解 *GaARF* 基因系统的进化关系。由图1可以看出,*GaARF* 基因中存在11个同源对,其中8对步长值(Bootstrap values)为100,说明这八对 *ARF* 基因的亲缘关系非常近;同一亚家族的基因序列和保守序列很相似,表明它们的功能可能很相似,体现了 *GaARF* 基因在进化过程的保守性;不同亚家族中的 *GARF* 基因在结构上有较大的差异,如第IV基因家族只有2~4个外显子,而其他三

类的外显子数量明显多于第IV基因家族。所有的 *GaARF* 均含 motif1、2、3、4,而不同亚家族间的保守区域不同,如亚家族IV中均含有 motif18,不含有 motif13、16,表明这些 motif 可能存在某些特定的功能。基因结构和保守基序的相似性和差异性从侧面反映了 *GaARF* 基因家族在漫长的进化过程中的相对保守性和为适应环境而产生的多样性。

表1 *GaARF* 基因家族的鉴定及特性Table 1 Characteristics of identified *GaARF* gene family

基因名称	基因ID	结构域	染色体	定位	蛋白特性				外显子数目	拟南芥同源基因
					氨基酸	分子量	等电点	不稳定系数		
<i>GaARF1</i>	Cotton_A_03644	B3, ARF, AUX/IAA	1	67 209 931-67 213 848	859	95 556.03	6.13	55.00	14	AT5G62000.3
<i>GaARF2</i>	Cotton_A_31049	B3, ARF, AUX/IAA	1	81 285 179-81 288 793	693	77 053.68	6.49	43.54	14	AT3G61830.1
<i>GaARF3</i>	Cotton_A_14407	B3, ARF, AUX/IAA	4	7 914 108-7 917 887	691	76 963.06	6.03	47.95	14	AT3G61830.1
<i>GaARF4</i>	Cotton_A_08273	B3, ARF	4	25 459 819-25 464 075	705	79 633.64	6.13	54.94	14	AT1G59750.2
<i>GaARF5</i>	Cotton_A_01955	B3, ARF, AUX/IAA	6	29 053 261-29 057 413	869	96 723.74	6.28	51.84	15	AT5G62000.3
<i>GaARF6</i>	Cotton_A_35443	B3, ARF, AUX/IAA	6	38 247 234-38 253 499	836	92 802.07	6.01	63.02	13	AT1G30330.1
<i>GaARF7</i>	Cotton_A_25871	B3, ARF, AUX/IAA	6	45 492 120-45 495 300	661	73 998.59	6.44	46.28	14	AT4G23980.1
<i>GaARF8</i>	Cotton_A_36154	B3, ARFAUX/IAA	7	73 625 377-73 630 142	734	82 026.47	6.26	58.87	15	AT4G23980.1
<i>GaARF9</i>	Cotton_A_04263	B3, ARF	7	79 440 487-79 443 060	686	75 503.71	8.37	50.89	4	AT2G28350.1
<i>GaARF10</i>	Cotton_A_40208	B3, ARF	8	10 980 998-10 985 644	740	80 779.82	5.70	56.98	10	AT2G33860.1
<i>GaARF11</i>	Cotton_A_18446	B3, ARF	8	44 741 459-44 744 067	576	63 418.77	6.58	45.01	2	AT1G77850.1
<i>GaARF12</i>	Cotton_A_16408	B3, ARF, AUX/IAA	8	84 349 362-84 354 587	942	104 299.21	5.66	53.19	14	AT1G19850.1
<i>GaARF13</i>	Cotton_A_10518	B3, ARF	8	103 387 522-103 390 330	654	72 505.21	8.67	46.37	4	AT2G28350.1
<i>GaARF14</i>	Cotton_A_27843	B3, ARF, AUX/IAA	9	57 390 631-57 396 654	1 141	127 542.08	6.36	68.53	13	AT1G19220.1
<i>GaARF15</i>	Cotton_A_03933	B3, ARF	9	61 960 171-61 963 626	719	78 895.73	6.81	51.49	11	AT2G33860.1
<i>GaARF16</i>	Cotton_A_02933	B3, ARF, AUX/IAA	9	79 277 687-73 277 687	907	100 340.59	5.95	69.56	14	AT1G30330.2
<i>GaARF17</i>	Cotton_A_23397	B3, ARF	9	91 004 948-91 007 211	685	75 643.03	6.33	46.25	3	AT4G30080.1
<i>GaARF18</i>	Cotton_A_22543	B3, ARF, AUX/IAA	10	13 533 060-13 537 247	677	76 121.91	8.18	51.39	14	AT5G62000.3
<i>GaARF19</i>	Cotton_A_18937	B3, ARF, AUX/IAA	10	15 597 527-15 606 378	1 029	114 988.05	7.31	46.92	20	AT4G23980.1
<i>GaARF20</i>	Cotton_A_26156	B3, ARF, AUX/IAA	10	15 987 202-111 131 684	902	99 524.82	5.93	69.43	14	AT1G30330.1
<i>GaARF21</i>	Cotton_A_11311	B3, ARF	10	1 915 730-19 120 421	705	77 249.11	6.59	56.61	11	AT2G33860.1
<i>GaARF22</i>	Cotton_A_11048	B3, ARF, AUX/IAA	10	21 629 555-21 634 399	814	89 881.63	6.68	49.68	12	AT5G60450.1
<i>GaARF23</i>	Cotton_A_16138	B3, ARF	10	24 051 262-24 154 263	578	63 435.47	6.40	47.92	2	AT1G77850.1
<i>GaARF24</i>	Cotton_A_07064	B3, ARF	10	88 003 872-88 006 976	695	77 315.76	8.93	47.37	4	AT2G28350.1
<i>GaARF25</i>	Cotton_A_24047	B3, ARF	10	109 631 933-109 634 411	695	76 393.3	7.25	50.11	3	AT4G30080.1
<i>GaARF26</i>	Cotton_A_06071	B3, ARF	12	94 383 684-94 389 563	1 099	122 028.3	6.20	58.76	13	AT1G19220.1
<i>GaARF27</i>	Cotton_A_06107	B3, ARF	12	94 722 686-94 724 907	674	75 004.46	6.59	50.87	3	AT4G30080.1
<i>GaARF28</i>	Cotton_A_27669	B3, ARF, AUX/IAA	13	59 279 756-59 284 569	950	104 464.34	5.21	51.36	14	AT1G19850.1
<i>GaARF29</i>	Cotton_A_16760	B3, ARF	13	111 131 684-15 993 089	737	81 824.14	6.03	65.15	13	AT5G37020.2

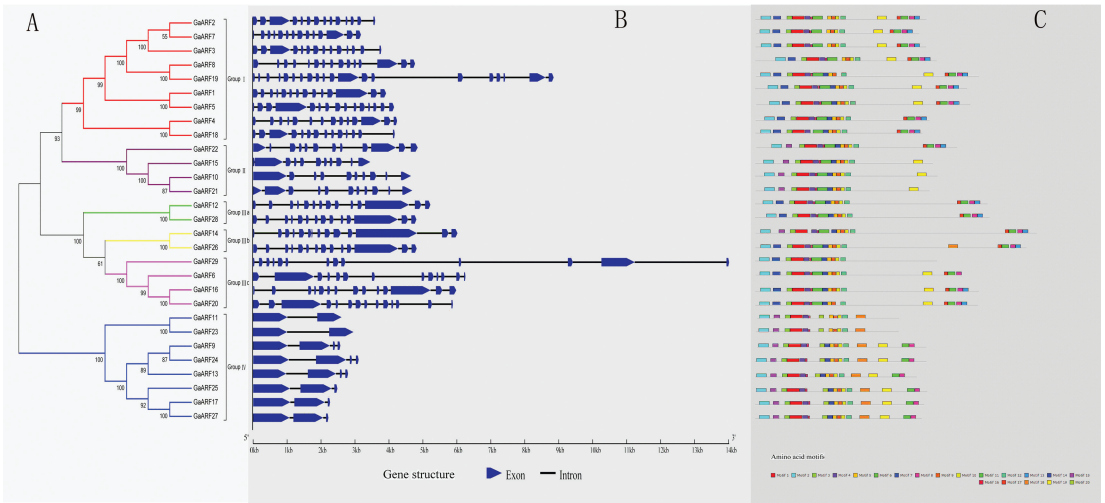


图 1 *GaARF* 基因家族的系统发育、基因结构及保守区域

Fig.1 Phylogenetic tree, gene structure and conserved region of *GaARF* gene family

### 2.3 *GaARF* 基因家族系统进化分析

为了更好的明确亚洲棉与单、双子叶物种间 *ARF* 基因亲缘关系的远近,将已公布的序列信息的单子叶植物玉米 (31 个); 双子叶植物亚洲棉 (29 个)、拟南芥 (23 个) 共计 83 个 *ARF* 蛋白构建了系统进化树 (见图 2)。从图中可以看出,83 个 *ARF* 家族成员被明显的分成四大类,由进化关系可以发现,

亚洲棉和拟南芥较多的聚类到一起,由此可以看出它们之间的进化关系较近。从亚洲棉与拟南芥的聚类关系可以看出,Class I、II 中的 *ARF* 蛋白是转录抑制子,可能在转录过程中抑制转录的发生;Class III 中的 *ARF* 蛋白是转录激活子,可能在转录过程中促进转录的进行。

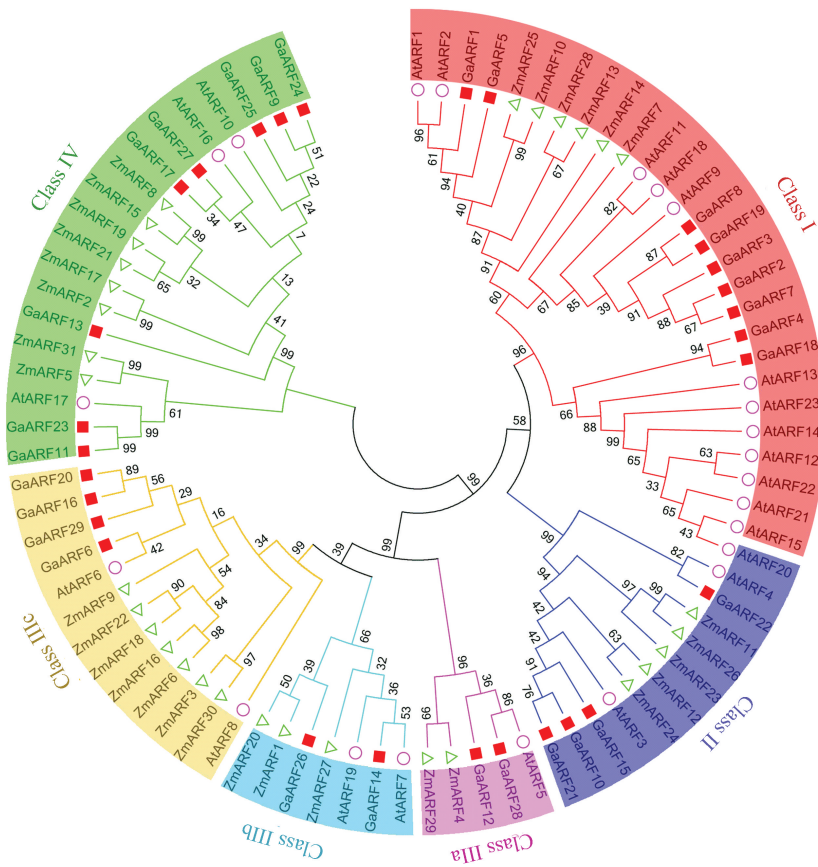


图 2 *ARF* 蛋白多物种系统进化树

Fig.2 Phylogenetic tree of *ARF* protein in multi-species

## 2.4 *GaARF* 基因家族染色体物理定位

物种间基因数量的增加被认为是促进物种进化的方式,基因家族数量增加的主要方式是基因复制,基因复制又分为基因内复制和基因间复制<sup>[18]</sup>。亚洲棉基因组在复制的过程中经历了几次的基因数量加倍(Gene duplication)事件,这些基因的复制品通常可幸免于选择压力<sup>[10]</sup>。这不仅为亚洲棉的进化提供了保障,也丰富了亚洲棉基因家族的多样性。为了进一步研究基因分化对 *GaARF* 基因家族的影响,我们利用 Mapinspect 软件将 29 个 *GARF* 基因定

位在 9 个染色体上(见图 3),其中 2、3、5、11 号染色体不含 *ARF* 基因。由图中可以看出,亚洲棉 29 个 *ARF* 基因是随机分布的,其中 10 号染色体分布最多,有 9 个 *ARF* 基因。一般认为 200 kb 核苷酸有 3 个以上基因的基因群,就认为是一个基因簇,亚洲棉中在 10 号染色体上存在一个基因簇,它们可能编码催化同一个新陈代谢途径中不同步骤的酶的结构基因。根据基因串联复制的定义, *GaARF2*、3、7; *GaARF6*、16、20; *GaARF17*、27; 还存在基因间的串联复制。

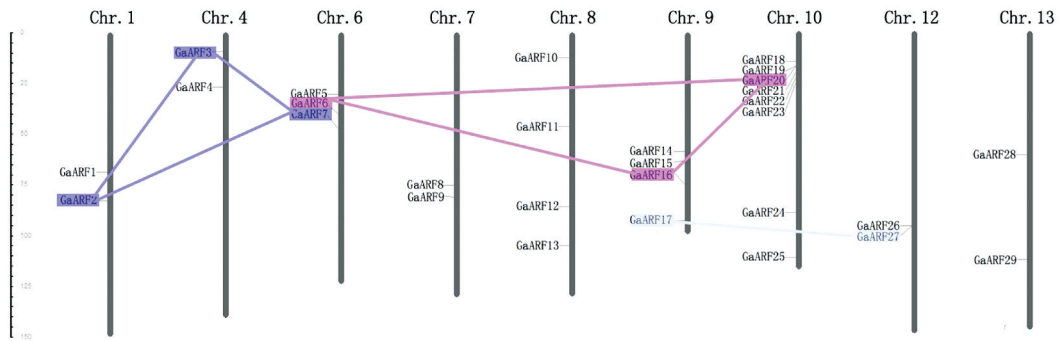


图 3 *GaARF* 基因染色体定位图

Fig.3 Chromosomal locations of *GaARF* gene family

## 2.5 亚洲棉信号肽检测及亚细胞定位预测

利用 TargetP 软件检测,发现有 7 条(*GaARF8*、10、15、20、21、22、23)定位到叶绿体上, *GaARF11* 蛋白定位到线粒体上,其余 21 条蛋白未检测到信号肽信息。

## 2.6 亚洲棉 *ARF* 基因家族表达模式分析

为了进一步了解 *ARF* 基因的功能和特性,对基因的组织特异性表达和诱导表达模式进行分析(见

图 4)。分析结果表明: *GaARF1*、2、3、7、12、26、28 在根中表达较高,推测它们可能参与根的形成; *GaARF1*、15、17、22、26 在叶中的表达较高,表明它们可能在叶的发育过程起到重要的作用。淹水处理后, *GaARF11*、13、24 在根和叶都存在差异表达,暗示着它们可能是参与响应涝害的关键基因,但 *ARF* 基因在不同部位的表达差异及作用还有待于进一步研究。

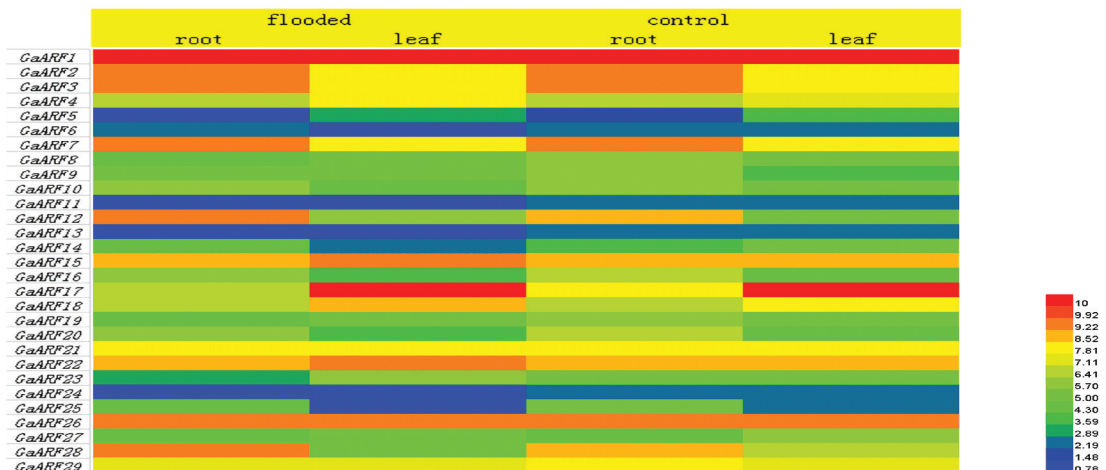


图 4 *GaARF* 基因家族表达热图

Fig.4 Heat map of *GaARF* gene expression

### 3 讨论

*ARF* 基因家族是参与生长素信号调节的重要转录因子,关于生长素的功能及作用机制的研究正在持续进行。Ulmasov 等于 1997 年利用酵母单技术以人工合成的 AuxREP3(4x)作为分子诱饵,筛选拟南芥 cDNA 表达文库,鉴定了第一个 *ARF* 基因,即 *AtARF1*<sup>[3]</sup>。到目前为止,在拟南芥中共发现 23 个 *ARF* 基因<sup>[8]</sup>,番茄中发现 21 个 *ARF* 基因<sup>[19]</sup>,玉米中发现 31 个 *ARF* 基因<sup>[6]</sup>,水稻中共发现 25 个 *ARF* 基因<sup>[20]</sup>,苹果中共发现 29 个 *ARF* 基因<sup>[21]</sup>,雷蒙德氏棉共发现 35 个<sup>[9]</sup>,目前在亚洲棉中还没有 *ARF* 基因家族的系统分析和研究,基于此,本文进行了亚洲棉 *ARF* 家族的生物信息学分析。

本研究从基因组水平上初步筛选并分析了 29 个亚洲棉 *ARF* 基因,并将其分为四大类,系统的阐述了其进化关系。前人研究发现,*AtARF1* 和 *AtARF2* 基因参与花器官衰老和脱落的过程<sup>[22]</sup>,通过系统进化树分析发现,*GaARF1* 和 *GaARF5* 与 *AtARF1*、2 同源性最高,可以推测亚洲棉中 *GaARF1* 和 *GaARF5* 也参与花器官衰老和脱落过程;*AtARF3* 控制 KANAD 蛋白合成,参与叶的极性反应,可以推测亚洲棉中同源性最高的 *GaARF15* 有着相似的作用;*AtARF7* 和 *AtARF19* 作为转录激活子,介导拟南芥侧根形成<sup>[8]</sup>,暗示着 *GaARF14* 可能有着相似的作用;*AtARF6* 和 *AtARF8* 基因参与调控逆境相关的茉莉酸合成<sup>[23]</sup>,可以推测亚洲棉中 *GaARF16*、20、29 可能参与逆境调控相关激素的合成;*AtARF10*、*AtARF16* 基因的单突变体没有明显的表型,而双突变型根冠和根系向地性生长发生异常<sup>[24]</sup>,亚洲棉中有 5 个基因和这两个基因属于一个类别,但具体作用还有待考究。

最近的研究还表明 *ARF* 基因的表达受到 microRNA 的调控, microRNA160 调控拟南芥 *AtARF10* 和 *AtARF16* 基因的表达<sup>[24]</sup>;玉米中, microRNA167 高表达的突变株能够导致 *ZmARF3*、9、16、18、22、30 转录本的降解,暗示 *ZmARF* 基因起到转录后的调控作用<sup>[6]</sup>。亚洲棉由于研究较少,目前还没有 *ARF* 基因诱导表达的报道。

前人研究表明,基因家族数量的增加主要是基因复制,基因复制又分为基因内和基因间复制<sup>[25]</sup>,从进化树中可以看出,亚洲棉中存在基因复制的现象。亚洲棉基因组(A)和雷蒙德氏棉基因组(D)在距今约 5 百万年(2~13 百万年)之前从同一祖先分化而来,二者的基因数目和基因序列都极为相近,染

色体水平上也保留了高度的共线性,但由于 A 基因组发生过多次大规模的反转座子插入事件,导致其基因组膨胀至超过 D 基因组的两倍。但亚洲棉中 *ARF* 基因家族数量多于雷蒙德氏棉,可能是因为亚洲棉含有的 41 330 个编码基因中,68.5% 的基因组序列是由重复序列组成的造成的<sup>[10]</sup>。

*ARF* 基因家族调节植物的激素代谢,在植物的生长发育中起到至关重要的作用。本研究以亚洲棉全基因组序列为研究背景,为研究 *ARF* 基因的功能奠定了基础。亚洲棉中 *ARF* 基因家族是如何调控生长素发挥作用还需要后人进一步研究。

### 参考文献(References)

- [1] HORI S, NOMURA T, SAKAGUCHI S. Control of regulatory cell development by the transcription factor foxp3[J]. *Science*, 2003, 299(5609): 1057–1061. DOI: 10.1126/science.1079490.
- [2] 吴蓓, 吴建勇, 蔡刘体, 等. 生长素反应因子[J]. *植物生理学通讯*, 2005, 41(3): 273–278. DOI: 10.13592/j.cnki.pj.2005.03.001.
- [3] WU Bei, WU Jianyong, CAI Liuti, et al. Auxin response factor[J]. *Plant Physiology Communications*, 2005, 41(3): 273–278. DOI: 10.13592/j.cnki.pj.2005.03.001.
- [3] ULMASOV T, HAGEN G, GUILFOYLE T J. Arf1, a transcription factor that binds to auxin response elements[J]. *Science*, 1997, 276(5320): 1865–1868. DOI: 10.1126/science.276.5320.1865.
- [4] ABEL S, THEOLOGIS A. Early genes and auxin action[J]. *Plant Physiology*, 1996, 111(1): 9–17. DOI: 10.1104/pp.111.1.9.
- [5] RAMOS J A, ZENSER N, LEYSER O, et al. Rapid degradation of auxin/indoleacetic acid proteins requires conserved amino acids of domain ii and is proteasome dependent[J]. *Plant Cell*, 2001, 13(10): 2349–2360. DOI: 10.1105/tpc.13.10.2349.
- [6] XING H, PUDAKE R N, GUO G, et al. Genome-wide identification and expression profiling of auxin response factor (arf) gene family in maize[J]. *BMC Genomics*, 2011, 12(1): 178. DOI: 10.1186/1471-2164-12-178.
- [7] TIWARI S B. The roles of auxin response factor domains in auxin-responsive transcription[J]. *Plant Cell*, 2003, 15(2): 533–543. DOI: 10.1105/tpc.008417.
- [8] OKUSHIMA Y, OVERVOORDE P J, ARIMA K, et al. Functional genomic analysis of the auxin response factor gene family members in arabidopsis thaliana: Unique and overlapping functions of arf7 and arf19. [J]. *Plant Cell*, 2005, 17(2): 444–463. DOI: 10.1105/tpc.104.028316.
- [9] SUN R, WANG K, GUO T, et al. Genome-wide identification of auxin response factor (arf) genes and its tissue-spe-

- cific prominent expression in *Gossypium raimondii* [J]. *Functional & Integrative Genomics*, 2015, 15 (4): 1–13. DOI:10.1007/s10142-015-0437-0.
- [10] LI F, FAN G, WANG K, et al. Genome sequence of the cultivated cotton *Gossypium arboreum* [J]. *Nature Genetics*, 2014, 46 (6): 567–572. DOI:10.1016/j.pgene.2015.09.001.
- [11] PUNTA M, COGGILL P C, Eberhardt R Y, et al. The Pfam protein families database [J]. *Nucleic Acids Research*, 2008, 36 (Database issue): 263–266. DOI:10.1002/047001153x.g306303.
- [12] MARCHLERBAUER A, LU S, ANDERSON J B, et al. CDD: A conserved domain database for the functional annotation of proteins [J]. *Nucleic Acids Research*, 2011, 39 (Database issue): D225–229. DOI:10.1093/database/bar058.
- [13] GASTEIGER E, HOOGLAND C, GATTIKER A, et al. The proteomics protocols handbook-chapter 52: Protein identification and analysis tools on the ExPASy server [J]. 2005, 571–607. DOI:10.1385/1-59259-890-0:571.
- [14] KUMAR S, STECHER G, TAMURA K. Mega7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets [J]. *Molecular Biology & Evolution*, 2016, 33 (7): 1870. DOI:10.1093/molbev/msw054.
- [15] GUO A Y, ZHU Q H, CHEN X, et al. [gsds]: A gene structure display server [J]. *Hereditas*, 2007, 29 (8): 1023. DOI:10.1360/yc-007-1023.
- [16] BAILEY T L, BODEN M, BUSKE F A, et al. MEME suite: Tools for motif discovery and searching [J]. *Nucleic Acids Research*, 2009, 37 (Web Server issue): W202–W208. DOI:10.1093/nar/gkp335.
- [17] EMANUELSSON O, BRUNAK S, VON H G, et al. Locating proteins in the cell using TargetP, SignalP and related tools [J]. *Nature Protocols*, 2007, 2 (4): 953–971. DOI:10.1038/nprot.2007.131.
- [18] SANKOFF D. Gene and genome duplication [J]. *Current Opinion in Genetics & Development*, 2001, 11 (6): 681. DOI:10.1002/9780470619902.ch15.
- [19] WU J, WANG F, CHENG L, et al. Identification, isolation and expression analysis of auxin response factor (ARF) genes in *Solanum lycopersicum* [J]. *Plant Cell Reports*, 2011, 30 (11): 2059–2073. DOI:10.1242/dev.067918.
- [20] WANG D, PEI K, FU Y, et al. Genome-wide analysis of the auxin response factors (ARF) gene family in rice (*Oryza sativa*) [J]. *Gene*, 2007, 394 (1–2): 13–24. DOI:10.1016/j.gene.2007.01.006.
- [21] 李慧峰, 冉昆, 何平, 等. 苹果生长素响应因子 (ARF) 基因家族全基因组鉴定及表达分析 [J]. *植物生理学报*, 2015, 51 (7): 1045–1054. DOI:10.13592/j.cnki.pj.2015.0173.
- LI Huifeng, RAN Kun, HE Ping, et al. Genome-wide identification and expression analysis of auxin response factor (ARF) gene family in apple [J]. *Plant Physiology Journal*, 2015, 51 (7): 1045–1054. DOI:10.13592/j.cnki.pj.2015.0173.
- [22] REMINGTON D L, VISION T J, GUILFOYLE T J, et al. Contrasting modes of diversification in the *aux/iaa* and *arf* gene families [J]. *Plant Physiology*, 2004, 135 (3): 1738–1752. DOI:10.1104/pp.104.039669.
- [23] NAGPAL P, ELLIS C M, WEBER H, et al. Auxin response factors *arf6* and *arf8* promote jasmonic acid production and flower maturation [J]. *Development*, 2005, 132 (18): 4107–4118. DOI:10.1242/dev.01955.
- [24] WANG J W, WANG L J, MAO Y B, et al. Control of root cap formation by microRNA-targeted auxin response factors in Arabidopsis [J]. *Plant Cell*, 2005, 17 (8): 2204–2216. DOI:10.1105/tpc.105.033076.
- [25] GUAN X, XU T, GAO S, et al. Temporal and spatial distribution of auxin response factor genes during tomato flower abscission [J]. *Journal of Plant Growth Regulation*, 2014, 33 (2): 317–327. DOI:10.1007/s00344-013-9377-x.