王伟杰,雷亚柯,张建航,等.花生GATA基因家族鉴定和表达分析[J].山西农业大学学报(自然科学版),2024,44(2):10-22.
Wang W J, Lei Y K, Zhang J H, et al.Identification and expression analysis of peanut GATA gene family[J].Journal of Shanxi Agricultural University(Natural Science Edition), 2024,44(2):10-22.
doi: 10.13842/j.cnki.issn1671-8151.202312032

花生GATA基因家族鉴定和表达分析

王伟杰,雷亚柯,张建航,展世杰,邓陈威,贾朝阳*

(周口市农业科学院,河南 周口,466000)

摘　要:[目的]探究GATA基因家族在花生生长发育和抗旱调控中的作用。[方法]通过生物信息学方法鉴定花生GATA家族成员,分析其系统进化、基因结构、组织表达模式,最后结合转录组数据和实时荧光定量PCR技术筛选花生GATA家族响应干旱胁迫的关键基因。[结果]本文鉴定到47个GATA类转录因子编码基因,它们不均匀地分布在除2号、4号和14号染色体以外的17条染色体上。理化性质分析表明,该家族属于亲水性蛋白,大部分基因呈酸性,只有*Arahy. QG9XFC. 1*具有信号肽。根据拟南芥家族同源基因聚类分析结果,可将花生GATA家族蛋白分为4个亚家族。其中第Ⅰ、Ⅱ、Ⅲ亚家族成员的锌指结构域特征模体为$C_{X2}C_{X18}C_{X2}C$,而第Ⅳ亚家族为$C_{X2}C_{X20}C_{X2}C$类型。大部分成员在A亚基因组和B亚基因组中呈对称分布,组织表达模式分析显示,超过半数的*GATA*家族成员在22个组织中表现不同程度的特异性表达。其中第Ⅰ亚家族*Arahy. LJYJ4M*、*Arahy. VS8GG1*、*Arahy. 3S8P0L*和*Arahy. I6JZDT*在所有组织中均显示较高的表达水平。然而*Arahy. QY7BN7*和*Arahy. F7WS4I*在叶片,*Arahy. C6NV9N*在花被中表达较高,*Arahy. 4Y7J59*和*Arahy. D56WMI*在种子中呈现组织特异性高表达特性。转录组数据分析发现,34个*GATA*基因响应干旱、低温、ABA和BR激素处理,5个*GATA*基因在干旱和低温胁迫下表达量发生显著上调,1个基因发生显著下调。进一步RT-qPCR分析表明,*Arahy. C6NV9N*和*Arahy. 2SUK7S*特异性响应PEG胁迫的基因,尤其前者在叶组织中胁迫早期呈现持续上调表达。[结论]花生中有47个*GATA*基因,其中*Arahy. C6NV9N*和*Arahy. 2SUK7S*可能是花生GATA家族响应干旱胁迫关键基因。

关键词:花生;转录因子;*GATA*;组织表达模式;干旱胁迫
中图分类号:S565.2　　**文献标识码**:A　　**文章编号**:1671-8151(2024)02-0010-13

转录因子,又称反式作用因子,能够特异性识别和结合靶基因的启动子,进而调控下游靶基因的表达[1]。研究人员根据转录因子与DNA上特定序列的结合方式的不同,将其分为不同的家族,比如:bZIP(碱性亮氨酸拉链)、MYB、ZFP(锌指蛋白)、MADS、WRKY、NAC(NAM、ATA-F1/2和CUC1/2)、PHD(植物保守结构域)等[2]。锌指蛋白(zinc-finger protein,ZFP)是指具备指状结构特征并且能够与锌结合的一类蛋白质的总称,是植物中重要的转录因子之一。根据锌指蛋白中氨基酸残基的数量和位置对其进一步分类,比如:C_2H_2型、GATA型、C_3H_6型等。已有研究报道表明,GATA在植物种子萌发、器官发育、碳代谢、氮代谢、逆境胁迫、器官形成等生物学过程中发挥着重要的调控作用[3]。

1988年,研究人员首次在鸡的珠蛋白基因的启动子上发现并报道了GATA蛋白,证明了该转录因子参与鸡的造血过程[4-5]。之后,研究人员又鉴定出了GATA-1~GATA-6等多个GATA蛋白[6],1993年,研究人员首次在植物烟草中鉴定出了GATA转录因子,并且命名为*NTL1*,研究表明该转录因子参与烟草氮代谢生物学过程[7]。因此表明*GATA*基因不仅仅存在于高等动物中,在植物中同样存在。随后研究人员在拟南芥中发现了

收稿日期:2023-12-14　　**修回日期**:2024-01-05
基金项目:河南省科研平台22年产业基金(225101610063);河南省优质专用型花生工程技术研究中心(20221108038)
作者简介:王伟杰(ORCID:0009-0004-4042-2870),硕士,研究方向:花生新品种选育及栽培
*通信作者:贾朝阳(ORCID:0009-0001-0347-125X),副研究员,硕士,E-mail:jcy2309@163.com

调控拟南芥下胚轴和叶柄伸长的GATA转录因子*ZIM*[8]、调控拟南芥种子萌发的GATA转录因子*BME3*[9]、调控拟南芥光形态建成转录因子*GATA2*[10];在水稻中发现了调控水稻器官形成的GATA家族转录因子*NECKLEAF1*[11]和叶绿素合成及叶绿体发育的GATA家族转录因子*Cga1*[12];在小麦中发现了转录因子*TaGATA1*通过直接调控TaABI5的表达增强小麦种子休眠能力,进而提高穗发芽抗性[13];在玉米中GATA家族转录因子*GATA7*和*GATA33*参与调控玉米顶尖分生组织生长和发育[14];在构树中发现GATA家族转录因子可以参与调控冷胁迫[15];杨树(Populus)转录因子*GATA13*和*GATA19*通过调控气孔的大小或关闭,提高植株的水分利用效率,进一步提高抗旱性[16]。

目前,关于GATA基因家族方面的研究大多数都还是停留在拟南芥和水稻等模式作物中,而花生GATA基因家族还鲜有报道,关于该家族生物信息学分析更少。因此,本研究通过对花生全基因组中的GATA基因家族进行鉴定和组织表达模式进行分析,旨在为花生GATA转录因子对于花生生长发育和抗逆调控分析奠定基础。

1　材料与方法

1.1　材料

本研究所选用的种子(周花5号、远杂9102),聚乙二醇6000(科密欧化学试剂有限公司)。

1.2　方法

1.2.1　花生GATA基因家族成员鉴定

以"*GATA*"为关键词在花生基因组数据库(https://peanutbase.usda.iastate.edu/home)中Gene&Gene Family工具进行检索,下载检索到的基因氨基酸序列(存在多个转录本的取最长),将氨基酸序列上传至NBCI-CDD网站((https://www.ncbi.nlm.nih.gov/cdd/))和SMART网站(https://smart.embl.de/)进行结构域分析,去除不含"ZFP-GATA"结构域的序列最终得到47个花生GATA基因家族成员。将花生GATA基因家族成员氨基酸序列上传至Expasy网站(https://web.expasy.org/protparam/)进行氨基酸理化性质分析包括:氨基酸数量、分子量、等电点和总平均亲水性,将所有成员理化性质汇总,使用Excel 2016软件制表。

1.2.2　花生GATA基因家族成员染色体定位

将GATA家族成员在花生基因组数据库中用Gene&Gene Family工具进行检索,下载染色体编号、起始位置以及终止位置,使用MapGene2Chrom(http://mg2c.iask.in/mg2c_v2.1/)工具进行可视化。

1.2.3　花生GATA基因家族系统进化关系

从拟南芥数据库TAIR(https://www.arabidopsis.org/)中获得拟南芥GATA基因家族氨基酸序列,将花生和拟南芥GATA基因家族氨基酸序列上传至MEGA 7.0软件,使用Clustal工具进行序列比对,比对结果进行合理的删除,使用邻接法(Neighbor-Joining Method)构建系统进化树,校验参数Bootstrap设置为重复1000次。导出并保存.nwk文件,使用ITOL(https://itol.embl.de/)工具对进化树进行美化,最后根据拟南芥GATA基因家族同源基因聚类分析结果,对花生GATA基因家族进行分类。

1.2.4　花生GATA基因家族基因结构分析

将花生GATA基因家族氨基酸序列上传至MEME网站进行motif分析,下载mast.xml文件;花生GATA基因家族氨基酸上传至CDD网站进行结构域分析,下载结果;将1.3中MEGA 7.0比对之后花生的氨基酸序列单独提取出来使用邻接法(Neighbor-Joining Method)构建系统进化树,校验参数Bootstrap设置为重复1000次),导出并保存.nwk文件;最后使用TBrools中Gene Structure View工具对上述步骤结果进行可视化。

1.2.5　花生GATA基因家族序列比较

将花生GATA基因家族氨基酸序列上传至MEGA 7.0软件,使用Clustal工具进行比对,找到GATA结构域并截图保存。

1.2.6　花生GATA蛋白信号肽、亚细胞定位预测

将花生GATA基因家族氨基酸序列上传至SignalP 6.0工具网站(https://services.healthtech.dtu.dk/services/SignalP-6.0/)进行信号肽预测;将花生GATA基因家族氨基酸序列上传至WoLF PSORT工具网站(https://wolfpsort.hgc.jp/)进行亚细胞定位。

1.2.7 花生GATA基因家族共线性分析

使用TBtools-Ⅱ(Toolbox for Biologist) v 2.001软件中Text merge for MCScanX工具提取花生物种内*GATA*基因间的关系,使用Circle Gene View工具进行可视化。

1.2.8 花生GATA基因家族组织表达、不同激素、干旱和低温胁迫处理响应分析

22个花生不同发育时期、不同组织[17]和不同激素、干旱和低温胁迫处理转录组数据分别来自于花生基因组数据库(https://peanutbase.usda.iastate.edu/home)和PGR网站(http://peanutgr.fafu.edu.cn/index.php)。使用R软件对基因FPKM值进行log2标准化后,使用Excel 2016绘制热图。

1.2.9 花生GATA基因家族在干旱响应分析

以耐旱品种远杂9102和周花5号为试验对象,使用75%酒精消毒10 min后,放置清水中进行催芽,待种子露白,定植于水培盒中,使用清水于人工气候箱(16 h光照/8 h黑暗;28 ℃)中培养,待植株发育至三叶期时使用20%PEG6000进行开始处理,于0、3、6、9、12 h分别对叶片和根系进行取样,液氮速冻后用于总RNA的提取。RT-qPCR分析采用诺唯赞生物科技股份有限公司荧光定量酶试剂盒(ChamQ Universal SYBR qPCR master Mix),在Bio-Rad CFX 96平台上进行,以*Actin11*为内参,利用primer 5.0进行引物设计(表1),每个样品3次重复。

表1 RT-qPCR引物信息

Table 1 RT-qPCR Primer information

目的基因 Purpose fragment	正向引物 Forward primer sequences	反向引物 Reverse primer sequences	片段大小/bp Fragment size
Arahy.C6NV9N	ACAGGAAGAGGAGGAAAGCC	GCACAAGAAAGGGCCATCAA	196
Arahy.2SUK7S	GATGATGATGGATCTCTTCATGG	TGTCAGGTCTTCAAAACCAACC	189
Actin11	CAGGCCGTTCTCTCCCTTTAT	CAGGCCGTTCTCTCCCTTTAT	162

2 结果与分析

2.1 花生GATA基因家族的鉴定与理化性质分析

通过花生基因组数据库检索,NCBI-CDD和SMART数据库验证,最终鉴定得到47个花生GATA基因家族成员,经Expasy数据库理化性质进行分析,结果表明(表2):GATA基因家族成员氨基酸数量范围在108~415个,分子量在12 119.71~45 242.82 Da,其中Arahy.C6NV9N分子量最小,Arahy.V4I8CJ分子量最大;等电点在4.5~10.19,小于7的有27个,占总数目的57.45%,而大于7的有20个,占总数目的42.55%;总平均亲水性均小于零。由此说明该家族成员大部分呈酸性,且均为亲水性蛋白。

表2 GATA基因家族成员鉴定和理化性质分析

Table 2 Basic physical and chemical information of GATA gene family in peanut

蛋白质ID Protein ID	氨基酸数 Number of Amino Acid	分子量/Da Molecular Weight	等电点 pI	总平均亲水性 GRAVY
Arahy.0U49HS.1	285	30 839.98	6.07	−0.70
Arahy.129W39.1	303	33 350.89	6.00	−0.85
Arahy.18HYB7.1	379	41 754.29	6.34	−0.68
Arahy.1UE20Q.1	224	25 348.52	6.82	−1.14
Arahy.23I7N3.1	307	34 739.63	6.46	−0.84
Arahy.2L20XB.1	376	41 855.97	4.75	−0.84
Arahy.2SUK7S.1	345	38 773.25	8.98	−1.07
Arahy.36H91F.1	164	17 331.93	10.19	−0.50
Arahy.3E7WJ6.1	358	40 639.20	8.96	−1.15

续表

蛋白质ID Protein ID	氨基酸数 Number of Amino Acid	分子量/Da Molecular Weight	等电点 pI	总平均亲水性 GRAVY
Arahy.3J6WSK.1	249	27 685.51	8.56	−0.79
Arahy.3S8P0L.1	352	37 994.67	6.01	−0.50
Arahy.4Y7J59.1	200	22 493.01	9.19	−0.29
Arahy.54XSGH.1	362	39 784.73	6.02	−0.58
Arahy.5U2LX3.1	362	39 790.70	6.02	−0.60
Arahy.69GBM5.1	243	29 025.26	5.95	−1.14
Arahy.86LJVS.1	380	42 318.54	4.75	−0.83
Arahy.91YEB7.1	225	25 302.44	6.78	−1.09
Arahy.B6VTS2.1	296	32 708.73	8.74	−0.61
Arahy.C126L3.1	147	15 730.66	10.02	−0.85
Arahy.C6NV9N.1	108	12 119.71	9.00	−0.85
Arahy.D56WMI.1	133	14 754.15	9.33	−0.55
Arahy.DG0T3J.1	378	41 109.54	6.61	−0.73
Arahy.E1UHJ9.1	308	34 070.78	6.22	−0.64
Arahy.EI5605.1	285	30 839.98	6.07	−0.70
Arahy.FDN5YK.1	167	17 945.21	9.94	−0.76
Arahy.FRWT0V.1	365	40 728.87	7.59	−0.56
Arahy.GUPV5U.1	309	34 092.03	6.95	−0.78
Arahy.HHX0PC.1	309	34 119.05	6.95	−0.77
Arahy.I6JZDT.1	352	38 037.74	6.19	−0.50
Arahy.KA6YJ2.1	378	40 996.29	6.64	−0.73
Arahy.L83AAN.1	243	28 861.03	6.26	−1.19
Arahy.LJYJ4M.1	337	36 989.37	5.40	−0.61
Arahy.MTW1LM.1	379	41 924.55	6.34	−0.69
Arahy.MV37LX.1	308	34 852.79	6.60	−0.84
Arahy.N4A1C9.1	247	27 876.60	8.63	−0.62
Arahy.NCU4DF.1	248	27 988.73	8.42	−0.60
Arahy.NH2FYW.1	365	40 879.04	7.61	−0.59
Arahy.QG9XFC.1	326	36 002.16	6.00	−0.70
Arahy.QV2A2L.1	136	15 009.08	9.79	−0.78
Arahy.QY7BN7.1	294	33 214.55	9.06	−0.94
Arahy.TF5DD7.1	249	27 663.46	8.36	−0.79
Arahy.V4I8CJ.1	415	45 242.82	6.19	−0.42
Arahy.V7E7MG.1	408	45 052.88	7.19	−0.43
Arahy.VS8GG1.1	337	37 006.44	5.62	−0.60
Arahy.X8E4UZ.1	296	32 718.72	8.53	−0.62
Arahy.Z6CAHI.1	308	34 178.92	6.37	−0.68
Arahy.F7WS4I.1	215	24 225.57	9.04	−0.83

2.2　花生GATA基因家族成员染色体定位

根据花生基因组数据库中GATA基因家族染色体位置信息,绘制染色体定位图(图1),结果显示:GATA基因家族分布在17条染色体上,第2、4和14号染色体上未检测到*GATA*基因。第16号染色体分布最多有5个*GATA*基因;其次是第3和11号染色体分布4个*GATA*基因;第1、5、6、9、10、13、15、18、19号染色体均分布3个*GATA*基因;第6和20号染色体上分布2个*GATA*基因,最后是第7、12、17号染色体分布最少只有1个*GATA*基因。说明GATA家族基因不均匀的分布在花生染色体上。

2.3　花生GATA基因家族系统进化树

为了进一步探究花生GATA基因家族成员之

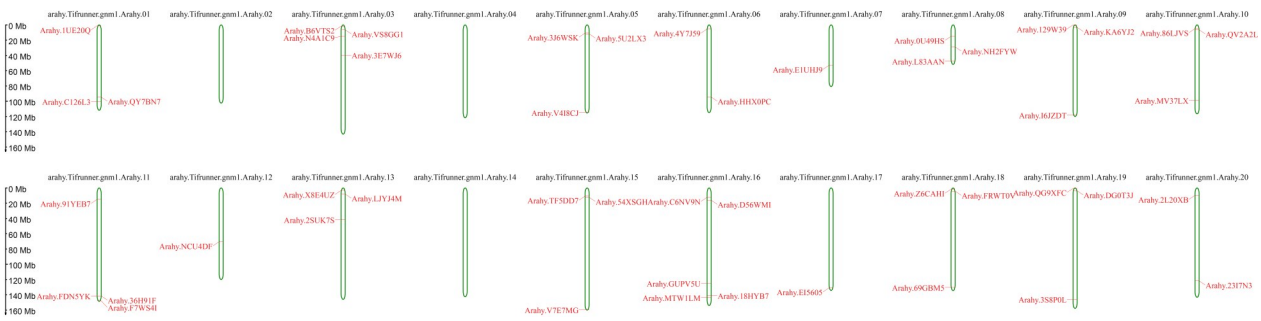


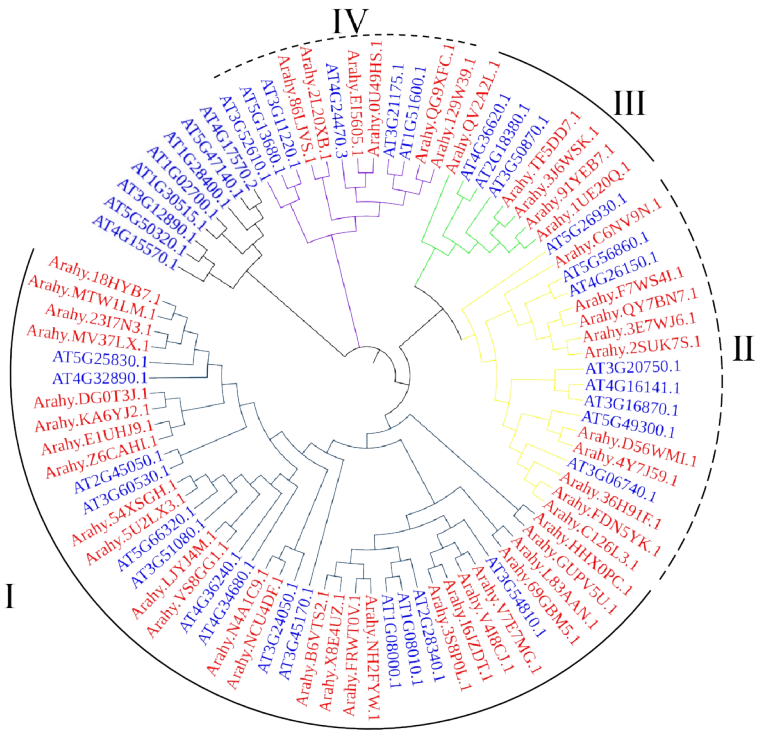
图1 花生GATA基因家族的染色体定位

Fig.1 Chromosome localization of GATA family genes in peanut

间进化关系,使用MAGA 7.0软件对花生GATA基因家族成员(红色)和拟南芥家族成员(蓝色)进行多重比较,构建系统进化树(图2)。结果表明:根据拟南芥家族同源基因聚类分析结果,将花生GATA基因家族分为4个亚家族,第Ⅰ亚家族成员有26个,第Ⅱ亚家族成员数量次之有10个,第Ⅲ亚家族成员最少有5个,第Ⅳ亚家族有6个成员。根据系统发育关系,可以推测基因亚家族成员之间具有较近进化距离,可能具有相同的生物学功能,因此可以根据家族成员进化关系推测GATA基因家族成员的功能。

2.4 花生GATA基因家族结构域和比对和分类

为了进一步揭示花生GATA基因功能特性,对该家族基因进行了基因结构分析,结果表明(图3、图4):改为"根据图3B可知,所有亚家族中第Ⅰ亚家族中检索到的基序最多含有8个不同基序,第Ⅲ亚家族含有基序最少,只有1个基序,进一步研究发现第Ⅰ亚家族均含有Motif2,第Ⅱ亚家族均含有Motif8,第Ⅲ亚家族均含有Motif1,第Ⅳ亚家族均含有Motif1和Motif5,不同的亚家族之间Motif数量和种类也不相同,这说明不同亚族中基序分布的不同导致其可能在进化过程中出现了功



AT:拟南芥;Arahy:花生

AT:*Arabidopsisthaliana*; Arahyp:*Arachis hypogaea L.*

图2 花生和拟南芥GATA基因家族进化分析

Fig.2 A phylogenetic tree of GATA gene family in peanut and Arabidopsis

能多样性演变;图3C可知,除了第Ⅱ亚家族部分*GATA*结构域在氨基酸序列的前端,其它基因*GATA*结构域皆在氨基酸序列后端;除*Arahy. 18HYB7*、*Arahy. MTW1LM*、*Arahy. 69GBM5*、*Arahy. V4I8CJ*以外所有基因均包含CDS区和UTR区,GATA基因外显子呈现1个(*Arahy. 69GBM5*)到11个(*Arahy. 2L20XB*、*Arahy. 86LJVS*)的结构变化,不同亚家族成员具有不同的CDS数量,而相同的亚家族中CDS数量相近(图3D)。同时,在相同亚族中GATA转录因子相似的结构特征表明其含有相似的功能。

通过MEGA 7.0中ClustalW工具比对发现,第Ⅰ、Ⅱ、Ⅲ亚家族成员锌指结构域类型为$C_{X2}C_{X18}C_{X2}C$,而第Ⅳ亚家族为$C_{X2}C_{X20}C_{X2}C$类型(图4)。其它氨基酸同一亚家族存在相似性,不同亚家族存在特异性,推测为物种进化的结果。

2.5　花生GATA基因家族信号肽预测

对花生GATA蛋白进行信号肽预测,发现除Arahy. QG9XFC. 1具有信号肽以外其它蛋白均不具备信号肽(图5)。

2.6　花生GATA基因家族亚细胞定位

为了进一步探索GATA基因家族蛋白的表达位置,我们使用WoLF PSORT工具对该家族蛋白进行亚细胞定位预测,结果显示(图6):有39个蛋白定位于细胞核中,占总数的83%,只有8个蛋白(Arahy. 0U49HS. 1、Arahy. 3J6WSK. 1、Arahy. 4Y7J59. 1、Arahy. EI5605. 1、Arahy. LJYJ4M. 1、Arahy. TF5DD7. 1、Arahy. V7E7MG. 1、Arahy. F7WS4I. 1)定位于叶绿体,占总数17%。说明大多数蛋白在细胞核中表达。

2.7　花生GATA基因家族共线性分析

为进一步探究花生GATA基因家族成员间的进化关系,使用TBtools软件对家族成员进行共线性分析。研究发现在47个成员中,共有19对家族成员在花生A亚基因组和B亚基因组相应的位置,占花生GATA基因家族成员总数的80%(图7)。进一步研究又发现这些成对出现的家族成员,在系统进化树也有较近的距离,因此可以推测为异源四倍体栽培种花生形成时,由2个二倍体野生种花生各自携带而来。剩下20%没有成对出现

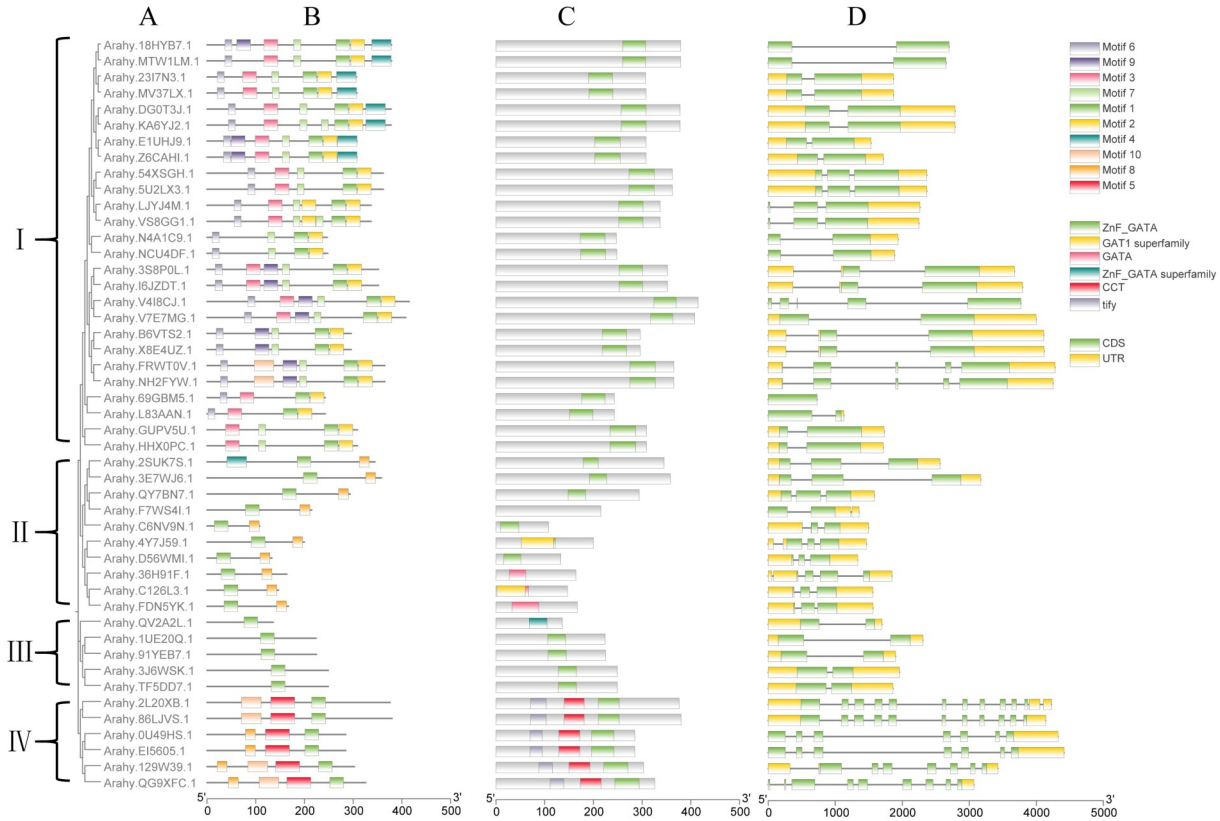


图3　花生GATA家族的系统发育关系、保守基序组成、结构域以及基因结构

Fig.3　Phylogenetic relationship, conserved motif, domain and exon-intron structures of peanut GATA gene family

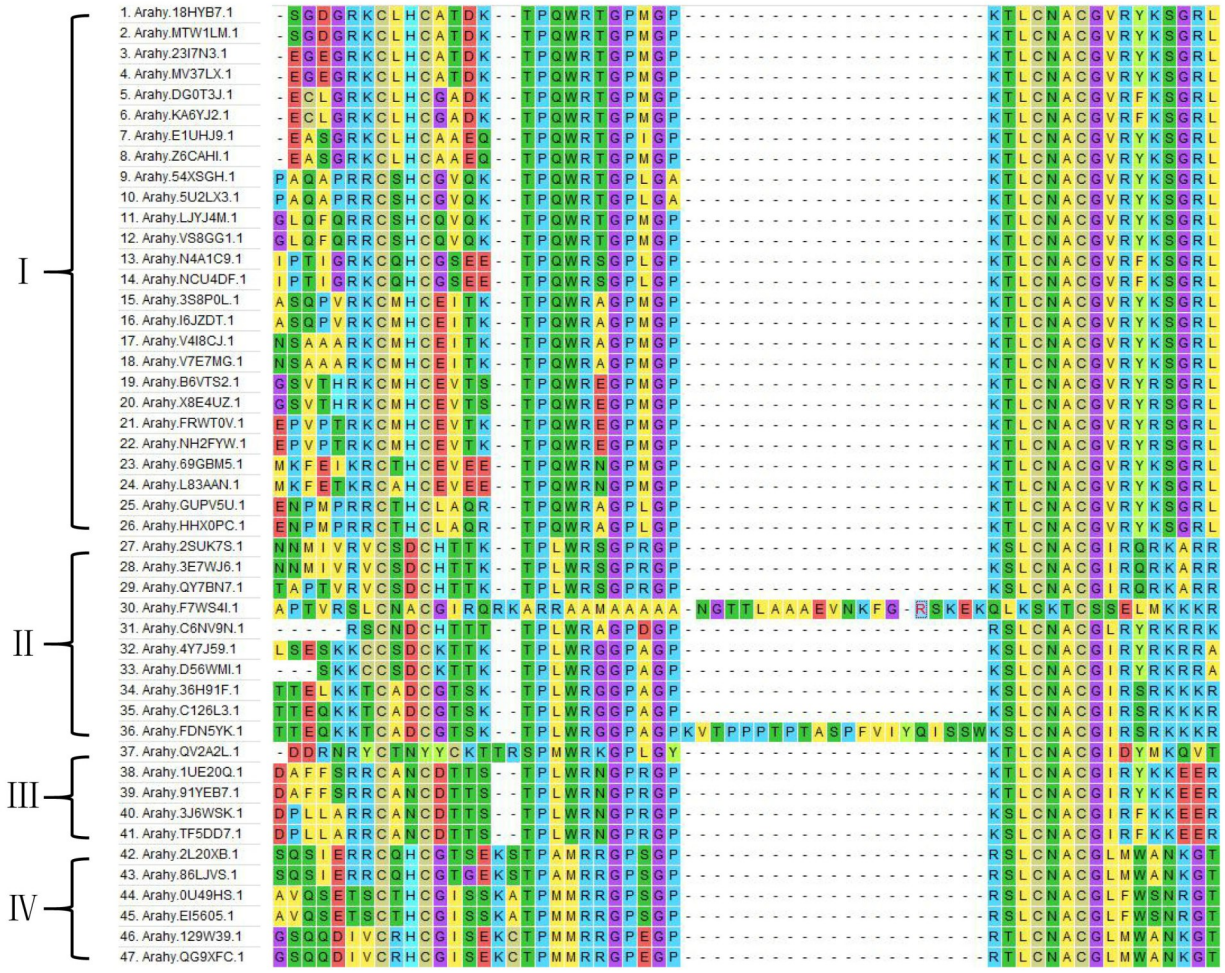
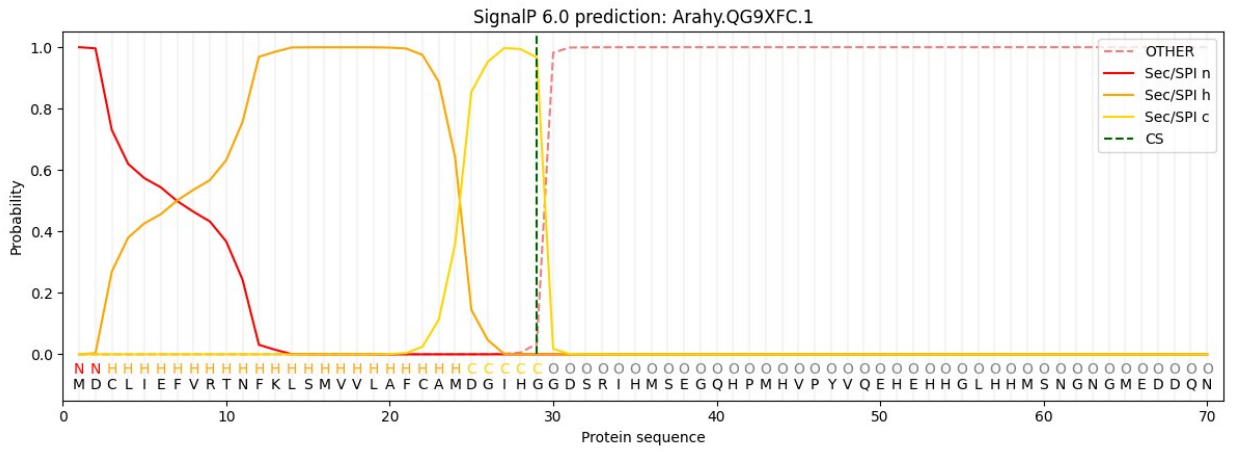


图4 花生GATA蛋白成员保守结构域

Fig.4 Conserved domains of peanut GATA family proteins


SignalP 6.0 prediction: Arahy.QG9XFC.1
OTHER
Sec/SPI n
Sec/SPI h
Sec/SPI c
CS
Probability
1.0
0.8
0.6
0.4
0.2
0.0
N N H H H H H H H H H H H H H H H H H H H H H C C C C O O O

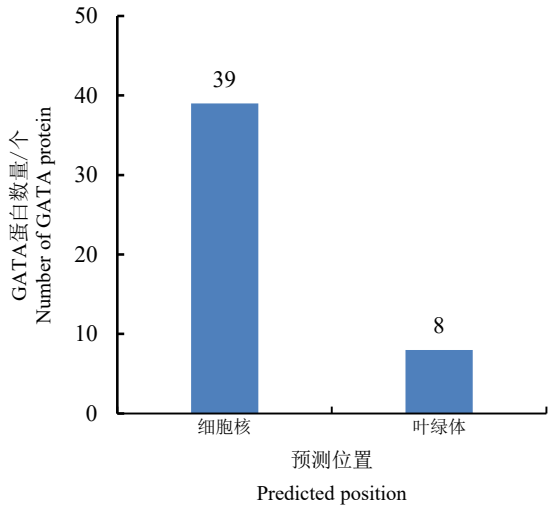


图6　花生GATA基因家族亚细胞定位预测

Fig.6　Prediction of subcellular localization of peanut GATA gene family

能,对基因组数据库不同组织转录组数据,绘制热图(图8),结果表明:在47个家族成员中有34个成员在22个组织中均有不同程度的表达(白色表示未检测到表达),其中*Arahy. LJYJ4M*、*Arahy. VS8GG1*、*Arahy. 3S8P0L*、*Arahy. I6JZDT*在所有组织中都具有较高的表达,而且他们均属第Ⅰ亚家族,而*Arahy. QY7BN7*和*Arahy. F7WS4I*在叶片中有较高的表达,*Arahy. C6NV9N*在花被中具有较高的表达,*Arahy. 4Y7J59*和*Arahy. D56WMI*在种子中具有较高的表达,进一步研究发现,该家族成员组织表达模式相似的基因其进化关系也比较近,因此可以推测其生物学功能也可能相似,根据其组织表达情况,推测GATA家族成员在花生生长发育过程中的功能,但是需要相关分子生物学试验加以论证。

2.9　花生GATA基因家族对不同激素、干旱和低温胁迫处理响应分析

为了进一步探究不同非生物胁迫和激素处理下*GATA*基因的表达模式,对干旱、低温、ABA和BR处理之后花生幼苗叶片的转录组数据进行分析,结果显示(图9):34个基因家族成员响应干旱、低温、ABA和BR处理(白色表示未检测到表达)。其中在干旱和低温处理下*Arahy. C6NV9N*、*Arahy. 4Y7J59*、*Arahy. 3E7WJ6*、*Arahy. D56WMI*和*Arahy. 2SUK7S*表达量均发生显著上调,而*Arahy. Z6CAHI*下调最为明显,进一步研究发现*Arahy. D56WMI*、*Arahy. 4Y7J59*、*Arahy. 2SUK7S*在受到干旱处理和ABA处理后基因表达变化出现了一致性,表明它们对于干旱胁迫的调控与ABA信号通路相关,通过对比干旱胁迫和低温胁迫基因表达变化情况可知,该家族成员对于干旱胁迫更为敏感且大多数基因表达量上调。

2.10　花生GATA基因家族PEG6000处理下表达分析

基于干旱处理的转录组数据,对干旱处理后表达量明显上升的5个基因(*Arahy. C6NV9N*、*Arahy. 4Y7J59*、*Arahy. 3E7WJ6*、*Arahy. D56WMI*、*Arahy. 2SUK7S*)进行PEG处理。采用20%的PEG模拟干旱进行处理,结果显示:只有*Arahy. C6NV9N*和*Arahy. 2SUK7S*响应PEG胁迫且表现出规律性变化,但响应模式不同(图10)。其中*Arahy. C6NV9N*在受到PEG胁迫后在根系中的表达量在3 h时显著增高,但在周花5号材料遗传背景下其表达差异仅为1.7倍,而在远杂9102材料遗传背景下表现为PEG胁迫后6和12 h双峰表达模式,表达量为未胁迫对照的2.6倍;叶片中*Arahy. C6NV9N*受到胁迫以后在2个品种中皆显著上调并且均在12 h时达到最高表达水平,最低上调幅度为4倍,最大上调幅度达16倍以上。*Arahy. 2SUK7S*在受到PEG胁迫后在根系中的表达量表现为先上升后下降,在周花5号中其表达量峰值在6 h达到显著水平,而远杂9102中峰值在9 h达到显著水平;在叶片中*Arahy. 2SUK7S*受到胁迫后,在周花5号中3 h时显著升高,胁迫9 h后降低到一般水平,而在远杂9102中其表达量在3 h无明显变化,在6 h时达到峰值,之后也降低到一般水平,在12 h时降到最低。

3　讨论与结论

能够特异性识别并结合靶基因的顺式元件的蛋白被称为转录因子。锌指转录因子是目前研究较为深入的转录因子家族之一,GATA型转录因子又是锌指蛋白转录因子中的一种重要类型。已有研究报道该家族影响着植物种子萌发、器官发育、碳代谢、氮代谢等[3]过程,本研究在栽培花生中鉴定得到47个GATA基因家族成员,成员数目比水稻(29个)[18]、玉米(37个)[19]、高粱(30个)[20]、辣椒(24个)[21]多,但是又比烟草(57个)[22]少,推测可

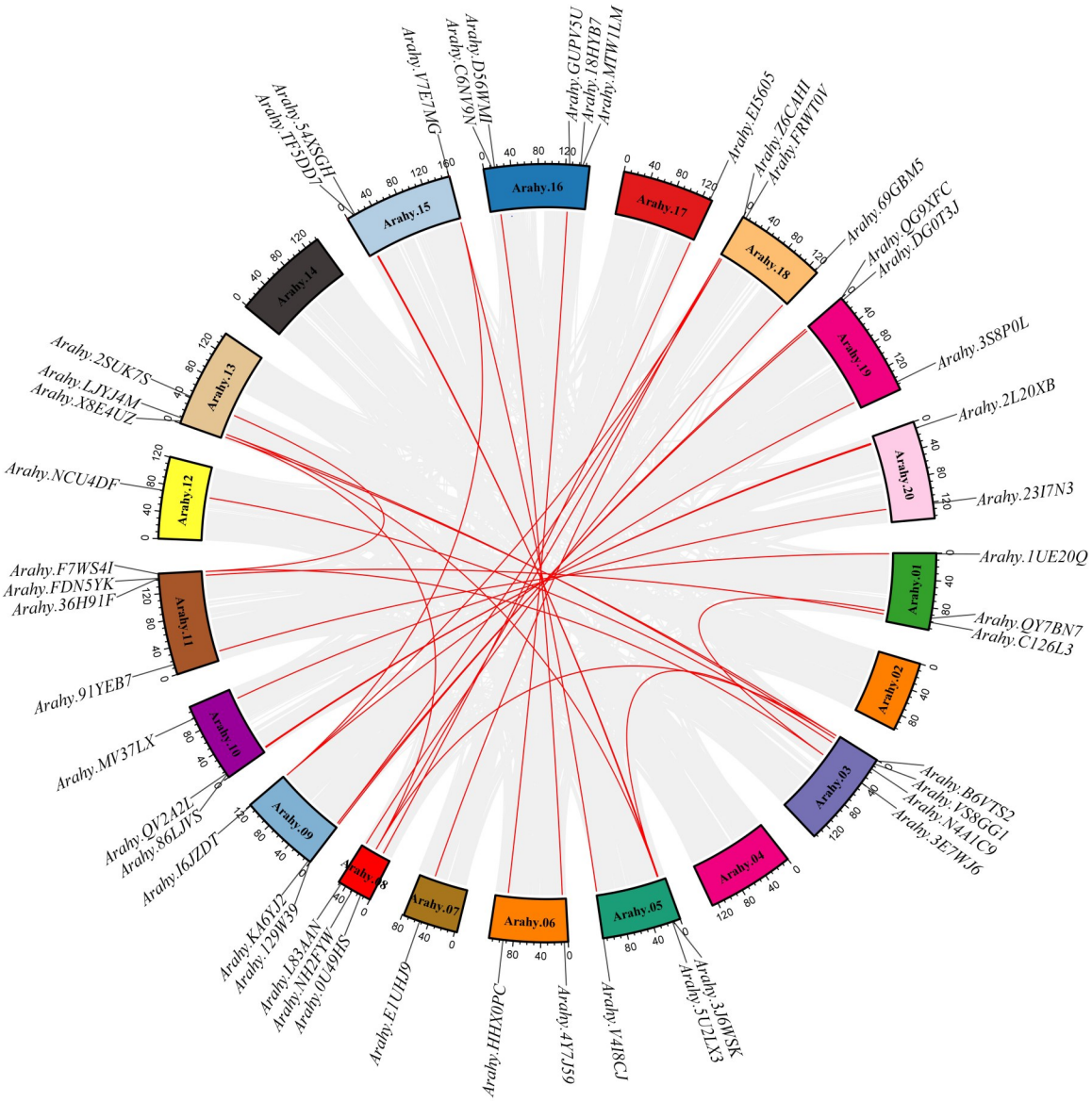


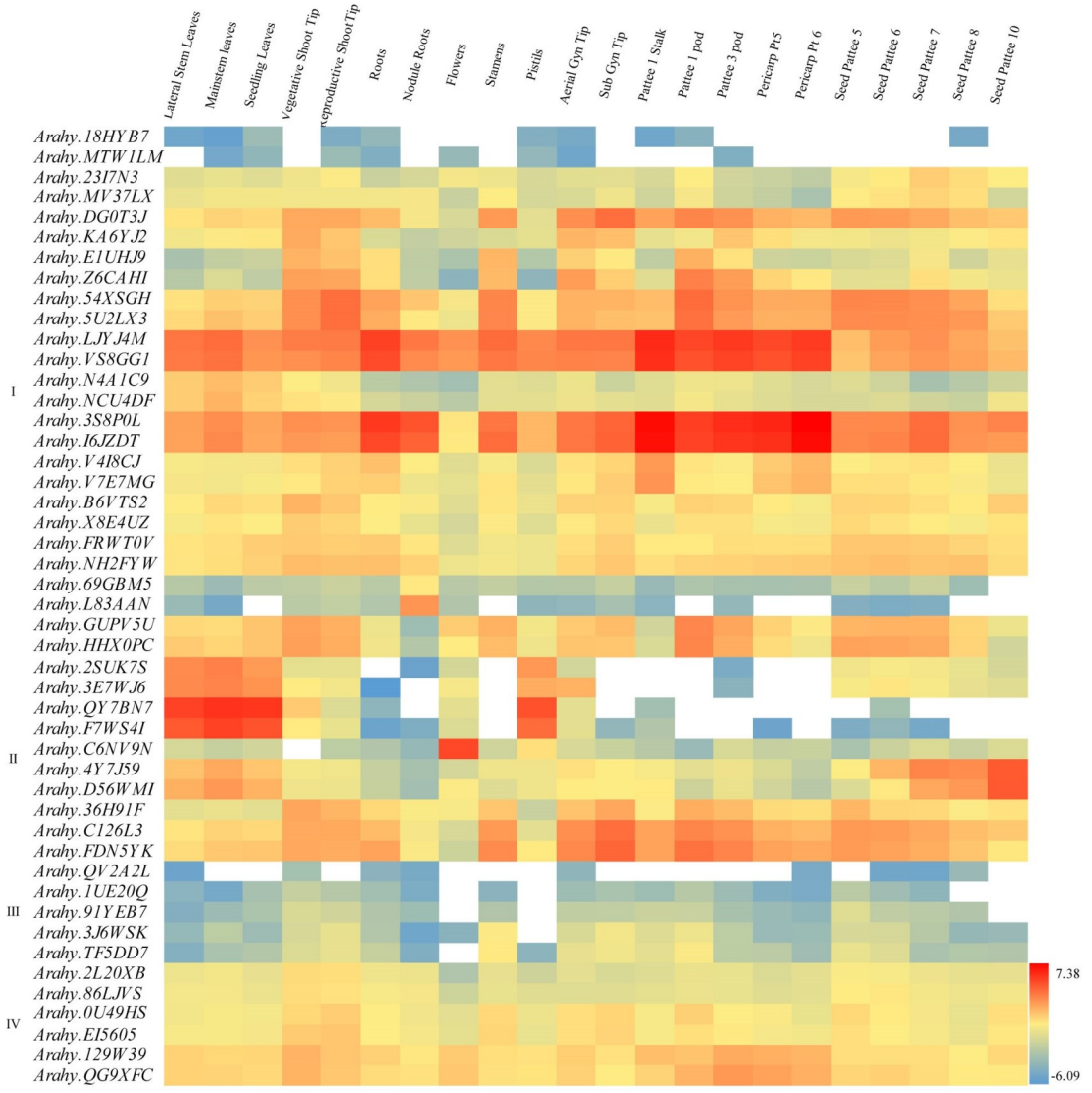
图7 花生*GATA*基因的共线性分析

Fig.7 The synteny of *GATA* genes in peanut

能与水稻、拟南芥和辣椒属于二倍体，而烟草和花生属于异源四倍体有关。

一级序列结构分析表明，花生GATA蛋白均为不稳定的亲水蛋白，有27个呈酸性，20个呈碱性；编码108～415个氨基酸，分子质量12 119.71～45 242.82 Da，其锌指结构域包括2种类型：$C_{X2}C_{X18}C_{X2}C$型和$C_{X2}C_{X20}C_{X2}C$型，与烟草GATA基因家族锌指结构域相似，且同一亚家族基因具有相同锌指结构域类型[22]。进一步对GATA家族成员基因结构研究发现，与已报道基因家族相似，同一亚家族成员之间其基因结构具有相似性，而不同亚家族成员之间具有特异性。因此可见或许正是由于这些家族成员之间的理化性质和基因结构的相似性和特异性，才使得GATA基因家族成员具有相似而又不同的生物学功能。

通过共线性分析发现在GATA基因家族中半数以上成员在A、B亚基因组之间对应位置存在共线性，这与花生中另外一个重要的转录因子家族bHLH转录因子共线性结果一致[23]，也从侧面印证了异源四倍体栽培花生基因组来自于二倍体野生花生*A. duranensis*和*A. ipaensis*祖先种。通过组织表达模式分析，发现第Ⅰ亚家族中表达量最高

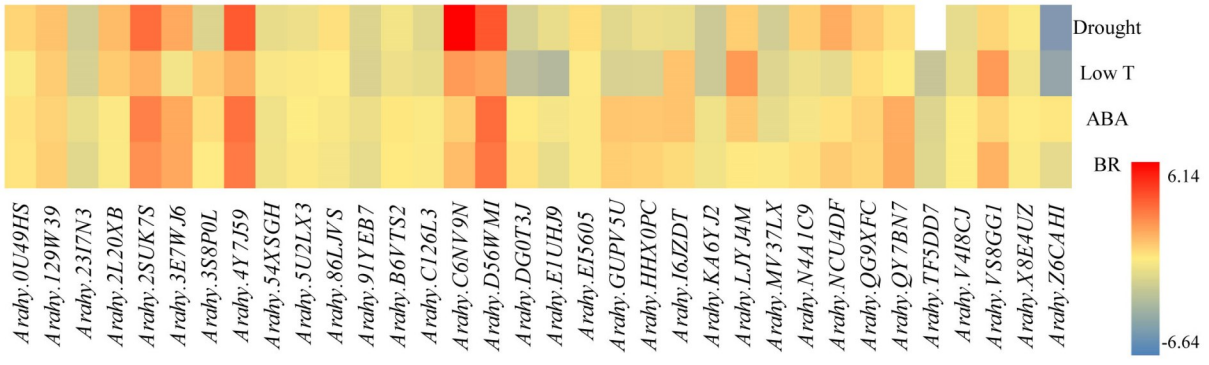


图中不同颜色代表基因表达水平,蓝色表示基因表达弱,红色表示基因表达强

Different colors are used to represent gene expression levels, blue indicates weak, red indicates strong

图8　花生*GATA*基因的表达模式分析

Fig.8　The expression pattern analysis of *GATA* genes in peanut

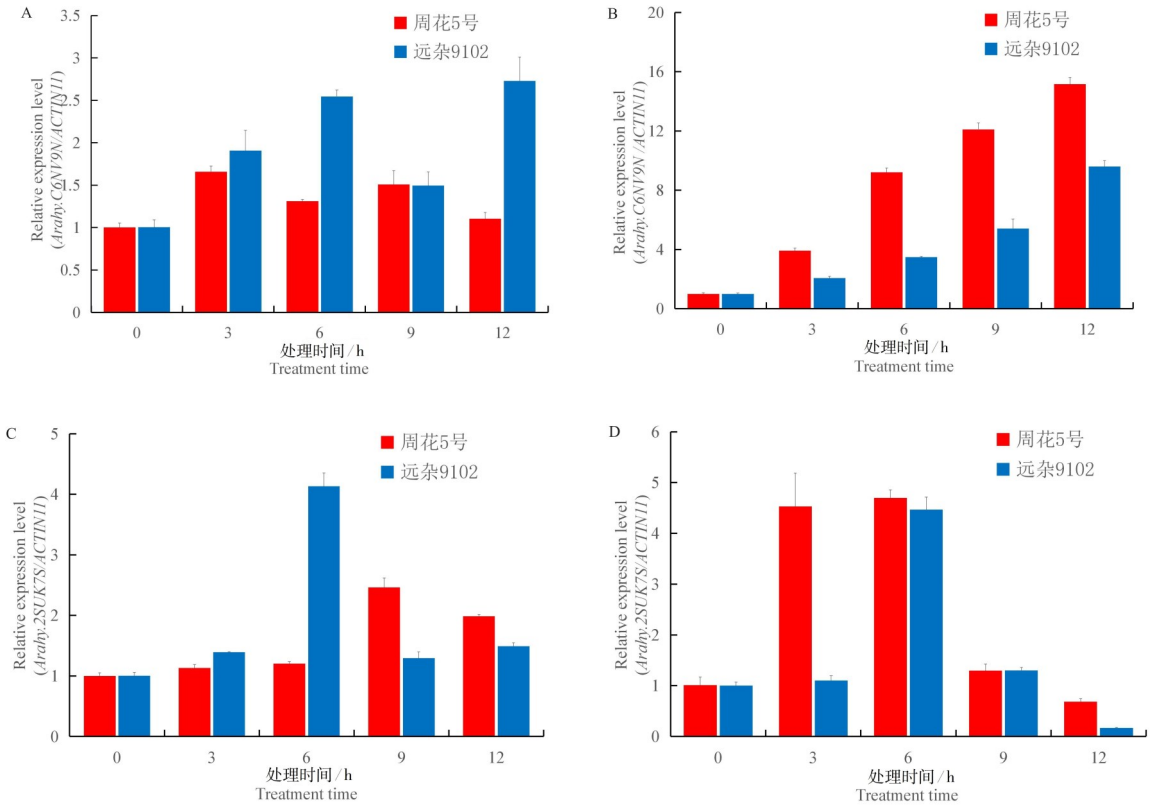


图中不同颜色代表基因表达水平,蓝色表示基因表达下调,红色表示基因表达上调

Different colors are used to represent gene expression levels, blue indicates down, red indicates up

图9　花生*GATA*基因在不同处理下的表达模式分析

Fig.9　The expression pattern analysis of *GATA* genes in peanut



A：*Arahy. C6NV9N*在根系中的表达量变化；B：*Arahy. C6NV9N*在叶片中的表达量变化；C：*Arahy. 2SUK7S*在根系中的表达量变化；D：*Arahy. 2SUK7S*在叶片中的表达量变化

A：Changes of *Arahy. C6NV9N* expression in root；B：Changes of *Arahy. C6NV9N* expression in leaf；C：Changes of *Arahy. 2SUK7S* expression in root；D：Changes of *Arahy. 2SUK7S* expression in leaf

图10 *Arahy.C6NV9N*和*Arahy.2SUK7S*在PEG6000处理下的表达分析

Fig.10 Expression analysis of *Arahy.C6NV9N* and *Arahy.2SUK7S* under PEG6000 treatment

组织为花生发育后期生殖生长阶段的器官中，因此推测第Ⅰ亚家族成员其生物学功能可能与后期种子发育成熟相关，而第Ⅱ亚家族则主要在叶片中突出表达，第Ⅲ、Ⅳ亚家族表达量都比较低，因此研究家族成员时要把重点放到第Ⅰ、Ⅱ亚家族。利用非生物胁迫和激素处理的花生转录组数据对GATA基因家族表达模式进行进一步分析，发现半数以上基因参与花生非生物胁迫应答，如在干旱胁迫和激素处理下*Arahy. D56WMI*、*Arahy. 4Y7J59*、*Arahy. 2SUK7S*基因表达变化出现了一致性。因此可以推测这3个基因对于干旱胁迫的调控可能是依赖ABA信号途径[24]。RT-qPCR验证结果显示*Arahy. C6NV9N*和*Arahy. 2SUK7S*在PEG处理后表达均发生显著变化，然而花生*GATA*基因具体抗旱调控功能有待进一步研究。

综上所述，本研究以花生GATA家族基因为对象，主要分析了GATA基因家族系统理化性质、进化关系、基因结构、共线性以及组织表达模式和非生物胁迫与激素响应，进一步加深了对于花生GATA基因家族成员的了解，为花生GATA基因家族成员进行下一步研究提供一定的基础，为抗逆、高产花生品种的选育提供新思路。

参 考 文 献

[1]Zhang J Z. Overexpression analysis of plant transcription factors [J]. Current Opinion in Plant Biology, 2003, 6(5): 430-440.

[2]侯思宇，孙朝霞，郭彬，等.大豆两个C2H2型转录因子基因序列特征及表达分析[J].植物生理学报，2014，50(5)：665-674.
Hou S Y, Sun C X, Guo B, et al. Cloning and expression analysis of two C_2H_2 transcription factors in soybean [J]. Plant Physiology Journal, 2014, 50(5): 665-674.

[3]王娟，兰海燕.GATA转录因子对植物发育和胁迫响应调控的研究进展[J].植物生理学报，2016，52(12)：1785-1794.

Wang J, Lan H Y. Advances in regulation of GATA transcription factor to plant development and stress responses [J]. Plant Physiology Journal, 2016, 52(12): 1785-1794.

[4]Evans T, Reitman M, Felsenfeld G. An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes [J]. Proceedings of the National Academy of Sciences of the United States of America, 1988, 85(16): 5976-5980.

[5]Omichinski J G, Clore G M, Schaad O, et al. NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1[J]. Science, 1993, 261(5120): 438-446.

[6]张未丽,刘智皓,吴风瑞,等.转录因子GATA-2在脊椎动物发育过程中的作用[J].四川动物,2009,28(6):945-948.
Zhang W L, Liu Z H, Wu F R, et al. Role of the GATA-2 transcription factor in vertebrates during development [J]. Sichuan Journal of Zoology, 2009, 28(6): 945-948

[7]Daniel-Vedele F, Caboche M. A tobacco cDNA clone encoding a GATA-1 zinc finger protein homologous to regulators of nitrogen metabolism in fungi [J]. Molecular Genetics and Genomics, 1993, 240(3): 365-373.

[8]Shikata M, Matsuda Y, Ando K, et al. Characterization of Arabidopsis ZIM, a member of a novel plant-specific GATA factor gene family [J]. Journal of Experimental Botany, 2004, 55(397): 631-639.

[9]Liu P P, Koizuka N, Martin R C, et al. The BME3 (Blue Micropylar End 3) GATA zinc finger transcription factor is a positive regulator of Arabidopsis seed germination [J]. The Plant Journal for Cell and Molecular Biology, 2005, 44(6): 960-971.

[10]Luo X M, Lin W H, Zhu S, et al. Integration of light- and brassinosteroid-signaling pathways by a GATA transcription factor in Arabidopsis [J]. Developmental Cell, 2010, 19(6): 872-883.

[11]Wang F, Zhu D M, Huang X, et al. Biochemical insights on degradation of Arabidopsis DELLA proteins gained from a cell-free assay system [J]. The Plant Cell, 2009, 21(8): 2378-2390.

[12]Hudson D, Guevara D R, Hand A J, et al. Rice cytokinin GATA transcription factor1 regulates chloroplast development and plant architecture [J]. Plant Physiology, 2013, 162(1): 132-144.

[13]Wei X N, Li Y Y, Zhu X L, et al. The GATA transcription factor TaGATA1 recruits demethylase TaELF6-A1 and enhances seed dormancy in wheat by directly regulating TaABI5 [J]. Journal of Integrative Plant Biology, 2023, 65(5): 1262-1276.

[14]Zhan J, Thakare D, Ma C, et al. RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation [J]. The Plant Cell, 2015, 27(3): 513-531.

[15]Peng X, Wu Q Q, Teng L H, et al. Transcriptional regulation of the paper mulberry under cold stress as revealed by a comprehensive analysis of transcription factors [J]. BMC Plant Biology, 2015, 15(1): 108.

[16]沈超.杨树PdGATA19和PdGATA13在干旱和低氮胁迫下的功能研究[D].北京:北京林业大学,2021.
Shen C. Functional analyses of poplar PdGATA19 and PdGATA13 in response to low-nitrogen and drought stress [D]. BeiJing: Beijing forestry university, 2021.

[17]Clevenger J, Chu Y, Scheffler B, et al. A developmental transcriptome map for allotetraploid *Arachis hypogaea* [J]. Frontiers Plant Science, 2016, 7: 1446.

[18]骆鹰,王有成,王伟平,等.水稻GATA基因家族生物信息学分析[J].分子植物育种,2018,16(17):5514-5522.
Luo Y, Wang Y C, Wang W P, et al. Bioinformatics analysis of GATA gene family in rice [J]. Molecular Plant Breeding, 2018, 16(17): 5514-5522.

[19]王延召,周波,韩小花,等.玉米GATA基因家族的全基因组鉴定及热胁迫下的表达分析[J].河南农业科学,2020,49(11):19-25.
Wang Y Z, Zhou B, Han X H, et al. Genome-wide identification of maize GATA gene family and expression analysis under heat stress [J]. Journal of Henan Agricultural Sciences, 2020, 49(11): 19-25.

[20]宋迎辉,朱灿灿,代书桃,等.高粱GATA基因家族的全基因组鉴定及表达分析[J].山东农业科学,2022,54(8):14-23.
Song Y H, Zhu C C, Dai S T, et al. Genome-wide identification and expression of GATA gene family in *Sorghum bicolor* [J]. Shandong Agricultural Sciences, 2022, 54(8): 14-23.

[21]袁岐,张春利,赵婷婷,等.辣椒GATA转录因子的生物信息学分析[J].中国农学通报,2017,33(17):24-31.
Yuan Q, Zhang C L, Zhao T T, et al. Bioinformatics analysis of GATA transcription factor in Pepper [J]. Chinese Agricultural Science Bulletin, 2017, 33(17): 24-31.

[22]杨玄松,谢雯榕,顾钢,等.普通烟草GATA转录因子家族鉴定及基因表达分析[J].烟草科技,2023,56(1):11-23.
Yang X S, Xie W R, Gu G, et al. Identification and gene expression analysis of GATA transcription factor family in *Nicotiana tabacum* [J]. Tobacco Science & Technology, 2023, 56(1): 11-23.

[23]王勇江,陈克平,姚勤.bHLH转录因子家族研究进展[J].遗传,2008,30(7):821-830.
Wang Y J, Chen K P, Yao Q. Progress of studies on bHLH transcription factor families [J]. Hereditas, 2008, 30(7): 821-830.

[24]Agurla S, Gahir S, Munemasa S, et al. Mechanism of stomatal closure in plants exposed to drought and cold stress [J]. Advances in Experimental Medicine and Biology, 2018, 1081: 215-232.

Identification and expression analysis of peanut GATA gene family

Wang Weijie, Lei Yake, Zhang Jianhang, Zhan Shijie, Deng Chenwei, Jia Chaoyang*

(*Zhoukou Academy of Agricutural Science, Zhoukou 466000, China*)

Abstract: [Objective] This study aimed to investigate the roles of the GATA gene family in the growth, development, and drought resistance regulation of peanuts. [Methods] The peanut GATA family members were identified using bioinformatics methods, and their systematic evolution, gene structure, and tissue expression patterns were analyzed. Finally, key genes responsive to drought stress in the peanut GATA family were screened by integrating transcriptome data and real-time fluorescence quantitative PCR technology. [Results] In this study, 47 GATA transcription factor encoding genes were identified, unevenly distributed on 17 chromosomes except for chromosomes 2, 4, and 14. Physicochemical property analysis showed that this family belongs to hydrophilic proteins, with most genes being acidic and only Arahy. QG9XFC. 1 possessing a signal peptide. Based on homologous gene clustering analysis with Arabidopsis family, peanut GATA family proteins were classified into four subfamilies. The zinc finger structure characteristic motif of members in subfamily I, II, and III were $C_{X2}C_{X18}C_{X2}C$, while subfamily IV was $C_{X2}C_{X20}C_{X2}C$ type. Most members were symmetrical distributed in A and B subgenomes. Tissue expression pattern analysis showed that over half of the GATA family members showed varying degrees of tissue-specific expression among 22 tissues. Among them, members, *Arahy. LJYJ4M*, *Arahy. VS8GG1*, *Arahy. 3S8P0L*, and *Arahy. I6JZDT* of subfamily I showed high expression levels in all tissues. However, *Arahy. QY7BN7* and *Arahy. F7WS4I* showed higher expression in leaves, *Arahy. C6NV9N* in petals, and *Arahy. 4Y7J59* and *Arahy. D56WMI* showed tissue-specific high expression in seeds. Transcriptome data analysis revealed that 34 GATA genes responded to drought, low temperature, ABA, and BR hormone treatments, with five GATA genes significantly up-regulated under drought and low-temperature stress, and one gene significantly down-regulated. Further RT-qPCR analysis showed that *Arahy. C6NV9N* and *Arahy. 2SUK7S* were specifically responsive to PEG stress, especially with the former showing sustained upregulation in leaf tissue during early stress. [Conclusion] There are 47 GATA genes in peanuts, among which *Arahy. C6NV9N* and *Arahy. 2SUK7S* may be key genes in the peanut GATA family's response to drought stress.

Keywords: Peanut, Transcription factor, GATA, Tissue expression pattern, Drought stress

(编辑:吕俊俐)