BMC Genomics

RESEARCH

Open Access

Identification and analysis of the GATA gene family in onion (*Allium cepa* L.) in response to chromium and salt stress

Chirasmita Bose[1], Pratyush Kumar Das[2], Praveen Roylawar[3,4], Pravara Rupawate[5], Kiran Khandagale[4], Satyabrata Nanda[1*] and Suresh Gawande[4*]

Abstract

Background The GATA transcription factors play multifaceted roles in modulating vital physiological processes in plants. However, the GATA transcription factor family in onion (*Allium cepa* L.) has been explored to a limited extent. In the present study, a genome-wide survey of the GATA family and the subsequent characterization has been carried out in the onion genome.

Results In total, 24 *A. cepa* GATAs (*AcGATA1-AcGATA24*) have been identified in the onion genome. Chromosomal mapping revealed that all identified genes could be mapped onto different onion chromosomes or scaffolds. The gene duplication, synteny, and collinearity analysis of the *AcGATA*s suggested their divergence, expansion, and selection in onions. Phylogenetic analysis of the *AcGATA*s divided them into five groups along with other plant GATAs. Gene ontology and *cis*-regulatory element analysis results suggested that the *AcGATA*s could regulate crucial processes, such as growth and development, phytohormone signalling, and stress response. The tissue-specific expression study indicated that the *AcGATA*s expressed in multiple onion tissues. The expression analysis under subjected chromium and salt stress revealed that multiple *AcGATA*s get induced in response to the applied stresses. Lastly, the protein interaction network study predicted some key interacting partners of the *AcGATA*s that can regulate vital physiological processes in onions.

Conclusions The present study identified and characterized the GATA gene family in onions. Functional predictions and interaction network analysis suggested the roles of *AcGATA*s in modulating multiple onion physiological processes. The induced expression of *AcGATAs* under chromium and salt stress indicated their involvement in abiotic stress response in onions. Overall, the study provides newer insights into the GATA gene family and their possible roles in onions.

Keywords *Allium cepa*, Transcription factors, Stress response, Chromium, Salinity

*Correspondence:
Satyabrata Nanda
satyabrata.nanda@cutm.ac.in
Suresh Gawande
sureshgawande76@gmail.com
Full list of author information is available at the end of the article

Background

Plant transcription factors (TFs) regulate vital processes, including growth, development, metabolism, and stress response [1]. TFs regulate the gene expressions in plants by exploiting their two principal functions, including binding to appropriate DNA elements and facilitating the binding of other proteins at the promoter regions [2]. Several TF families, such as bHlH (basic helix-loop-helix), SPL (Squamosa-promoter binding protein-like), bZIP (Basic region-leucine zipper), DREB (Dehydration-responsive element binding protein), and GATA (GATA-binding factor) have been extensively studied for their important roles in plants [3]. The GATA is a well-distributed TF found in plants, animals, and microbes, such as fungi. The GATA name for this TF family is because of their peculiar capacity to bind to the T/AGATAA /G sequence on the promoter region [4]. The signature motif, C-X_2-C-X_{17-20}-C-X_2−C of the GATAs is evolutionarily conserved [4]. In plants, the GATA TF was first discovered in tobacco, and since then, the GATA TF family has been identified in several plants, including *Arabidopsis thaliana*, rice, wheat, soybean, chili, cucumber, poplar, and cotton [4–11]. On the other hand, the availability of whole genome sequences in the public domains has facilitated the genome-wide identification and characterization of the GATAs in several plants. Recently, GATA TF family genes have been identified in *Phoebe bournei*, *Sorghum bicolor*, *Dimocarpus longan*, and *Vitis vinifera* by employing a genome-wide identification strategy by accessing the next-generation sequencing data [12–15].

The GATAs play multifaceted roles in plants by regulating several crucial physiological processes, including growth, photosynthesis, seed germination, and stress response [12, 15–18]. For instance, *OsGATA7* and *OsGATA12* regulate rice growth and yield by controlling the grain architecture and grain shape, and tillering, respectively [19, 20]. Similarly, *OsGATA8* has been reported to enhance rice tolerance to abiotic stresses by controlling the expression of genes involved in the reactive oxygen species (ROS) pathway, photosynthesis, and stress tolerance [21]. In addition, *OsGATA16* has been reported to confer cold stress tolerance in rice. The overexpression of *OsGATA16* improved the cold tolerance of the overexpressing rice seedlings. Further study revealed that cold tolerance is associated with the *OsGATA16*-mediated repressing of *OsWRKY45-1* [22]. In *Phoebe bournei*, multiple GATAs were involved in various abiotic stress responses, including cold, heat, salinity, and drought [12]. Likewise, in *Ipomoea batatas*, *IbGATA24* positively modulates the plant tolerance against salinity and drought stress by interacting with IbCOP9-5a [23]. A recent study reported that *SlGATA17*, a GATA TF in tomato, interacts with SlHY5 protein to enhance salinity stress tolerance [24].

Onion (*Allium cepa* L.) is a widely used vegetable for its high culinary demands. Other than that, it has significant applications in traditional medicines. However, onion productivity has been adversely affected due to several environmental factors, including abiotic and biotic stresses. In abiotic stress, salinity, temperature, and heavy metal contamination are the major factors contributing to onion yield loss [25]. Odisha, a southeastern state of India, experiences several mining activities, including iron, aluminum, and chromium. Thus, agricultural lands near mining areas get contaminated, which affects crop production. The regions where onion is mainly produced in Odisha are predominantly affected by soil salinity and chromium contamination [26]. Therefore, studying onion physiology and gene expression dynamics under such stresses is essential to understand plant tolerance mechanisms. Recently, the onion genome data has been made available in the public domain, greatly facilitating the onion genomics research [27]. However, the GATA gene family in onion has not been well-explored. Considering the aforementioned facts, this study aims to identify and characterize the *GATA* family in onions and to evaluate their involvement in onion stress response against two of the major abiotic stresses, salinity and chromium. The study has been carried out by using a set of stringent bioinformatic analysis, including protein properties prediction, gene structure organization, conserved motifs and signature domain analysis, phylogenetic and gene duplication analysis, prediction of gene ontology (GO), *cis*-regulatory elements (CREs), chromosomal mapping, synteny, collinearity, and protein–protein interactions to characterize the onion GATAs (*AcGATA*s). In addition, the expression of the *AcGATA*s has been analyzed in different onion tissues and under salinity and chromium stress.

Results

Identification of characterization of *AcGATA*s in onion genome

In total, 24 GATAs were identified in the onion genome through a stringent bioinformatic analysis and named as *AcGATA1-24* (Table 1). All 24 onion GATAs were found to possess the signature GATA zinc finger Pfam domain (PF00320), which was confirmed by doing Simple Modular Architecture Research Tool (SMART) and Conserved Domain Database (CDD) searches (Fig. 1A). The gene structure organization of the *AcGATA*s was analyzed, and the results revealed that the number of introns/exons varied across the 24 onion GATAs (Fig. 1B). The number of exons per gene ranged from 1 (single exon, *AcGATA19*) to 9 exons (*AcGATA11*). Subsequently,

Table 1 The predicted physico-chemical properties of *AcGATAs* in onions

Name	Transcript ID	Chrom no	Start Position	End Position	Strand	Exons	Amino acids	MW (KDa)	pI	GRAVY Score	Localization
AcGATA1	g17314.t1	Scaffold_11379	13,506	15,828	-ve	2	311	33.72	6.02	−0.765	Nucleus
AcGATA2	g46105.t1	Scaffold_27692	61,684	62,335	+ ve	2	193	21.74	7.67	−0.715	Nucleus
AcGATA3	g79291.t1	Scaffold_35258	330,634	331,508	+ ve	2	266	29.66	7.77	−0.766	Nucleus
AcGATA4	g79608.t1	Scaffold_35317	639,049	639,552	+ ve	2	140	16.01	9.86	−0.924	Nucleus
AcGATA5	g113541.t1	2	275,625,548	275,626,753	-ve	3	221	24.95	9.49	−0.794	Nucleus
AcGATA6	g175187.t1	Scaffold_47551	750,642	757,938	-ve	6	478	53.64	5.57	−0.45	Nucleus
AcGATA7	g175189.t1	Scaffold_47551	798,911	800,014	+ ve	2	303	34.05	5.23	−0.903	Nucleus
AcGATA8	g177201.t1	2	432,527,986	432,529,354	-ve	2	322	35.4	8.06	−0.419	Nucleus
AcGATA9	g251888.t1	1	1,073,159	1,086,841	-ve	7	277	30.13	5.9	−0.644	Nucleus
AcGATA10	g291540.t1	Scaffold_63078	17,354	17,853	-ve	2	139	15.3	9.67	−0.532	Nucleus
AcGATA11	g361876.t1	Scaffold_72038	373,489	379,159	+ ve	9	543	61.12	6.43	−0.659	Nucleus
AcGATA12	g374580.t1	Scaffold_73658	527,752	528,645	-ve	3	243	27.57	9.43	−0.494	Nucleus
AcGATA13	g379039.t1	Scaffold_74191	548,927	554,759	+ ve	6	283	32.02	9.54	−0.726	Nucleus
AcGATA14	g380862.t1	Scaffold_74448	66,638	68,215	-ve	4	320	35.55	6.35	−0.308	Nucleus
AcGATA15	g393953.t1	Scaffold_76030	113,272	113,875	+ ve	2	173	18.99	8.43	−0.523	Nucleus
AcGATA16	g403354.t1	Scaffold_77156	22,312	27,645	-ve	6	441	49.16	8.13	−0.61	Nucleus
AcGATA17	g424699.t1	Scaffold_79626	746,386	752,238	-ve	2	275	31.61	9.17	−1.13	Nucleus
AcGATA18	g440419.t1	Scaffold_81367	787,247	788,795	-ve	4	417	47.85	9.22	−0.599	Nucleus
AcGATA19	g483965.t1	Scaffold_86072	104,391	104,978	-ve	1	195	21.66	9.28	−0.79	Nucleus
AcGATA20	g523096.t1	Scaffold_90934	75,152	76,228	+ ve	2	300	33.65	5.72	−0.885	Nucleus
AcGATA21	g499934.t1	3	145,856,037	145,857,673	-ve	3	209	23.85	8.79	−0.855	Nucleus
AcGATA22	g207213.t1	Scaffold_51830	819,482	820,333	-ve	2	255	29.01	6.09	−0.944	Nucleus
AcGATA23	g18777.t1	7	27,761,858	27,762,343	-ve	3	105	11.54	9.22	−0.98	Nucleus
AcGATA24	g351433.t1	7	59,596,968	59,598,289	+ ve	2	174	19.88	9.19	−0.443	Nucleus

CDS Coding DNA sequence, *MW* Molecular weight, *pI* Isoelectric point, *GRAVY* Grand average of hydropathicity

the de novo motif elucidation of the *AcGATA*s resulted in predicting ten structural motifs, and their distribution among the *AcGATA*s was not found to be uniform (Fig. 2A). However, all the *AcGATA*s contained the conserved signature GATA motif (C-X_2-C-X_{17-20}-C-X_2 − C) (Fig. 2B). Lastly, the peptide properties of the *AcGATA*s, such as molecular weight, isoelectric point (pI), hydrophobicity, etc., were predicted. For instance, the molecular weight of the *AcGATA*s varied from 11.54 KDa (*AcGATA23*) to 61.12 KDa (*AcGATA11*). Similarly, the theoretical pI value ranged from 5.23 (*AcGATA7*) to 9.86 (*AcGATA4*). All *AcGATA*s were found to have a negative predicted GRAVY score, indicating their hydrophilic nature. The subcellular localization prediction of the *AcGATA*s revealed that all of them may localize in the nucleus (Table 1).

Phylogenetic analysis and chromosomal mapping of *AcGATA*s

A phylogenetic analysis was performed to establish the evolutionary relationship of the identified *AcGATA*s among other plant GATAs. The resultant phylogenetic tree was subdivided into five groups containing the *AcGATA*s and GATAs from *A. thaliana* (*AtGATA*s) (Fig. 3A). The onion GATAs were distributed into all five groups. However, the numbers per group were not the same. The alignment of *AcGATA*s revealed their conservancy levels within each subgroup (Fig. 3B). For instance, *AcGATA*s distributed in group I-IV contained the canonical GATA motif C-X_2-C-X_{18}-C-X_2 − C, whereas the group V *AcGATA*s possessed the C-X_2-C-X_{20}-C-X_2 − C motif (Fig. 3C).

In our attempt to map the identified *AcGATA*s onto onion chromosomes, only six were mapped onto different chromosomes (Fig. 4, Table 1). However, the majority of the genes were mapped onto different onion scaffolds. In this regard, we have combined all the scaffolds, hypothesized it as one single chromosome (chromosome 0), and mapped all the respective genes onto it. As the onion genome draft is not fully characterized, the availability of newer versions of the onion genome sequence may improve the chromosomal mapping.

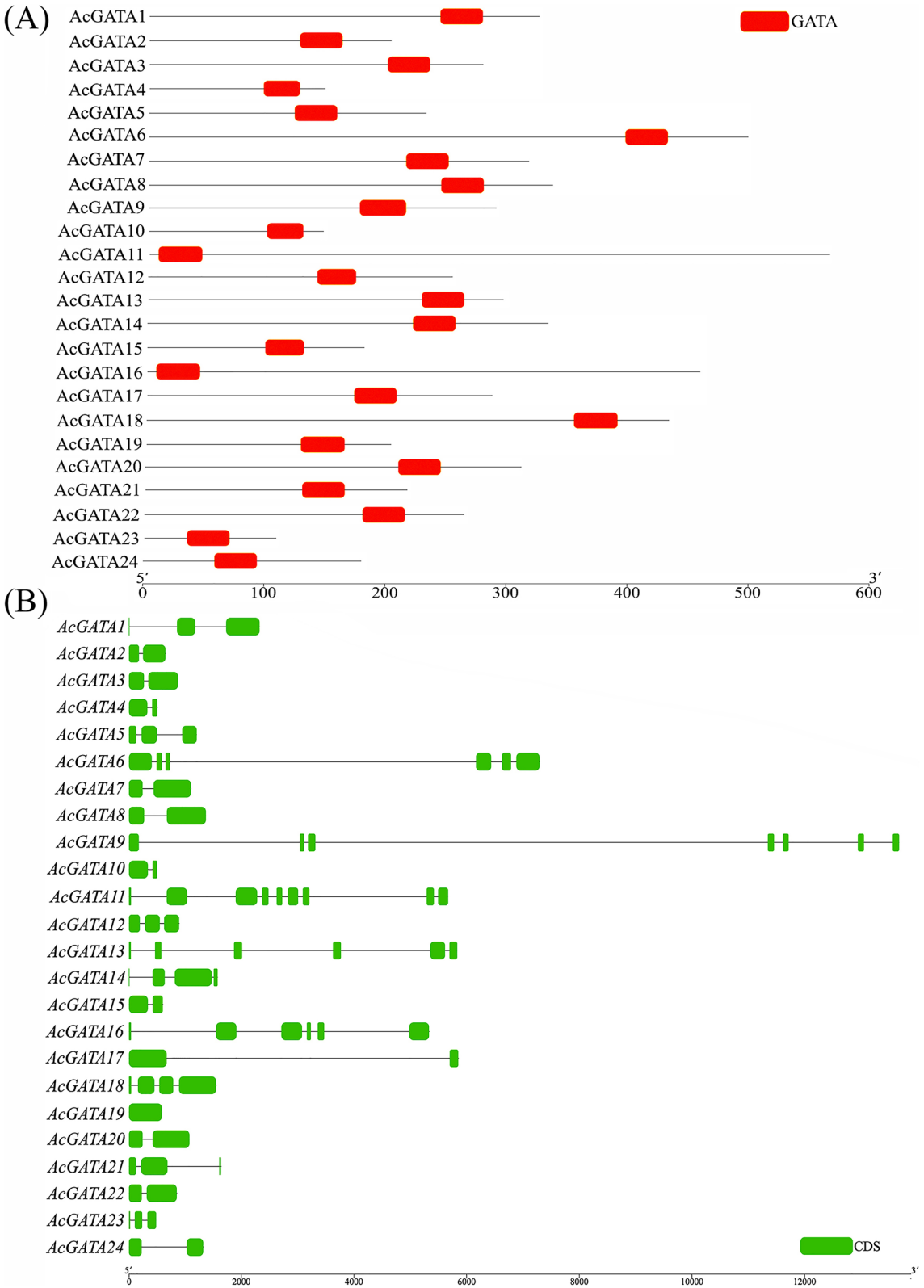

Fig. 1 **A** The presence of the conserved GATA domain in the *AcGATA*s. The red box indicates the GATA domain and the solid black lines indicate the individual protein lengths. **B** The exon–intron organization in the *AcGATA*s. The green boxes indicate the exons while the introns are indicated by the solid lines. The ruler at the bottom indicates the gene lengths

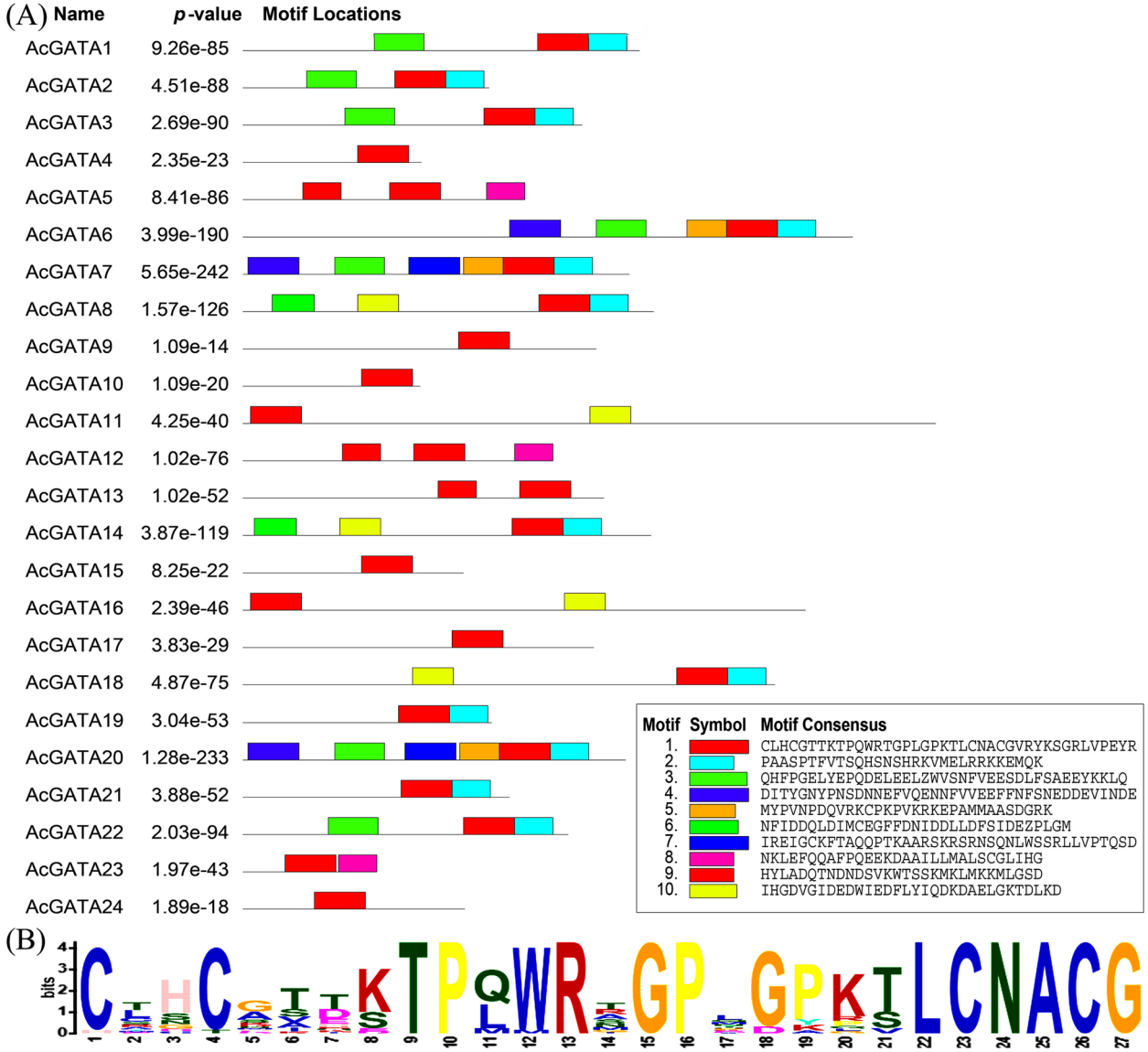

Fig. 2 **A** Distribution of the identified motifs in the *AcGATA*s. The solid lines indicate the protein lengths, and the colored boxes indicate the different motifs. The motif sequences are given at the bottom-right of the box. **B** The conserved C-X_2-C-X_{17-20}-C-X_2–C motif of the *AcGATA*s obtained from the MEME analysis

Gene duplication, synteny, collinearity, CREs, and GO analysis

The Ka/Ks ratio and synteny analysis revealed the gene duplication events, mechanism of amplification, expansion, and evolution of the *AcGATA*s in onion. The Ka/Ks comparison among the 24 *AcGATA*s ranged between 0.105 to 0.578 with five possible duplicated gene pairs, including *AcGATA2-AcGATA3, AcGATA7-AcGATA20, AcGATA8-AcGATA14, AcGATA4-AcGATA10,* and *AcGATA11-AcGATA16*, indicating a purifying selection in the onion GATA gene family (Table S1). Additionally, the synteny and collinearity analysis among the *AcGATA*s in onion and between the onion and *A. thaliana* GATAs, respectively, were performed (Fig. 5A, B). Subsequently, the CRE analysis of the *AcGATA* promoter sequences revealed the presence of several classes of CREs, including stress-responsive, phytohormone-responsive, and growth-related CREs (Fig. 6). The distribution of the CREs was not uniform across the *AcGATA*s. The putative function of the *AcGATA*s was predicted by the GO analysis. The results revealed the major probable biological functions of the *AcGATA*s to be response to stimulus, cell differentiation, germination and development, and circadian rhythm. Similarly, the major molecular functions

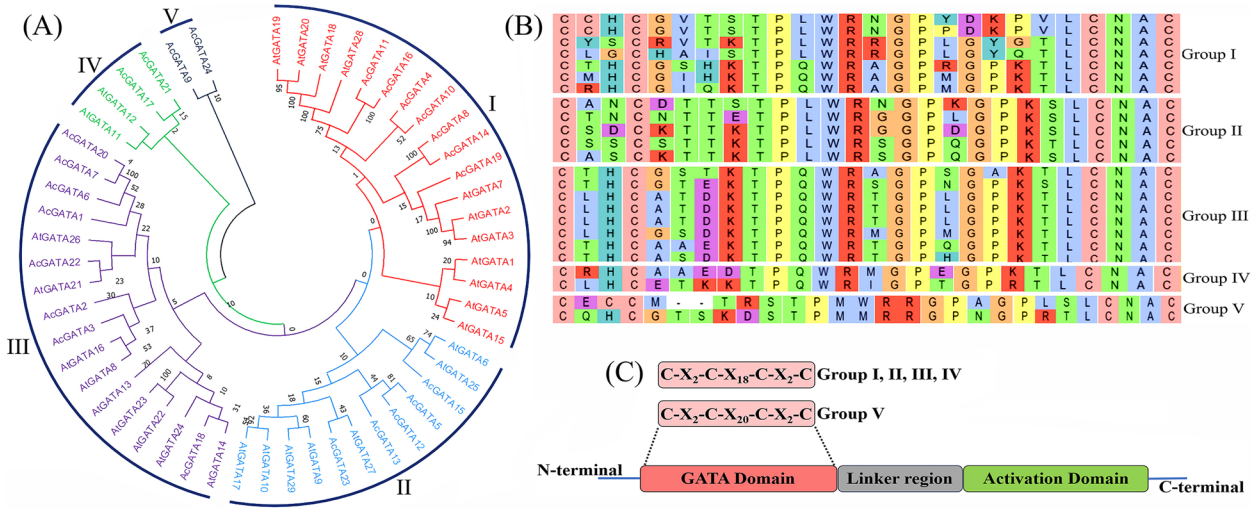

Fig. 3 **A** Phylogenetic analysis of the *AcGATA*s through the neighbour-joining method with 1000 bootstraps using MEGA. Different groups have been marked with the Roman numerals. **B** Multiple sequence alignment of the five phylogenetic groups of GATAs from onion and *A. thaliana*. **C** The domain structures of the *AcGATA*s showing the variation in the GATA domain and the corresponding phylogenetic groups

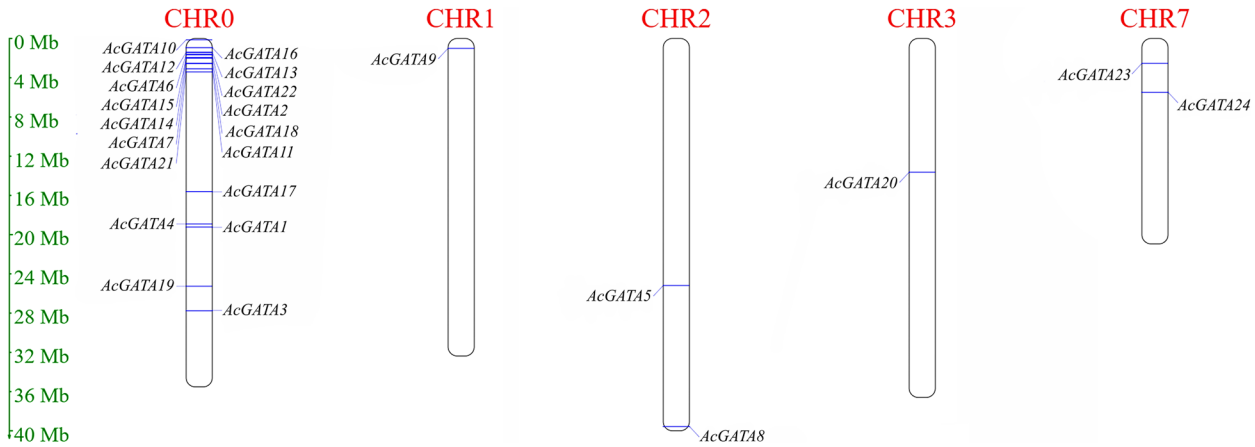

Fig. 4 Mapping of the *AcGATA*s on different onion chromosomes. The identified *AcGATA*s are mapped onto individual chromosomes as per their specific positions. The side ruler indicates the chromosome size

were predicted to be zinc ion binding, DNA binding, and protein binding (Fig. 7).

Expression analysis of *AcGATA*s

To get an insight into the expression of *AcGATA*s in different parts of the onion, the tissue-specific expressions were estimated by performing the real-time quantitative PCR (RT-qPCR) analysis. The results indicated that *AcGATA*s were expressed in the leaf, bulb, and roots. Most of the *AcGATA*s showed an upregulated expression in the bulb tissues, while the least were found in the roots (Fig. 8). Similarly, to deduce the expression dynamics of the *AcGATA*s under different abiotic stresses, their transcript abundances were analyzed and out of 24 *AcGATA*s, 19 (*AcGATA2, AcGATA3, AcGATA4, AcGATA5, AcGATA7, AcGATA8, AcGATA9, AcGATA11, AcGATA13, AcGATA14, AcGATA15, AcGATA16, AcGATA17, AcGATA18, AcGATA19, AcGATA20, AcGATA22, AcGATA23*, and *AcGATA24*) of them exhibited induced expression at different time points under chromium stress (Fig. 9). Only *AcGATA10* showed downregulated expression, while *AcGATA1, AcGATA6, AcGATA12*, and *AcGATA21* showed no significant change in the expression at all time points (6, 12, and 24 h) under chromium stress.

Under salinity stress, 22 (*AcGATA1, AcGATA2, AcGATA3, AcGATA4, AcGATA5, AcGATA7, AcGATA8, AcGATA9, AcGATA10, AcGATA11, AcGATA12,*

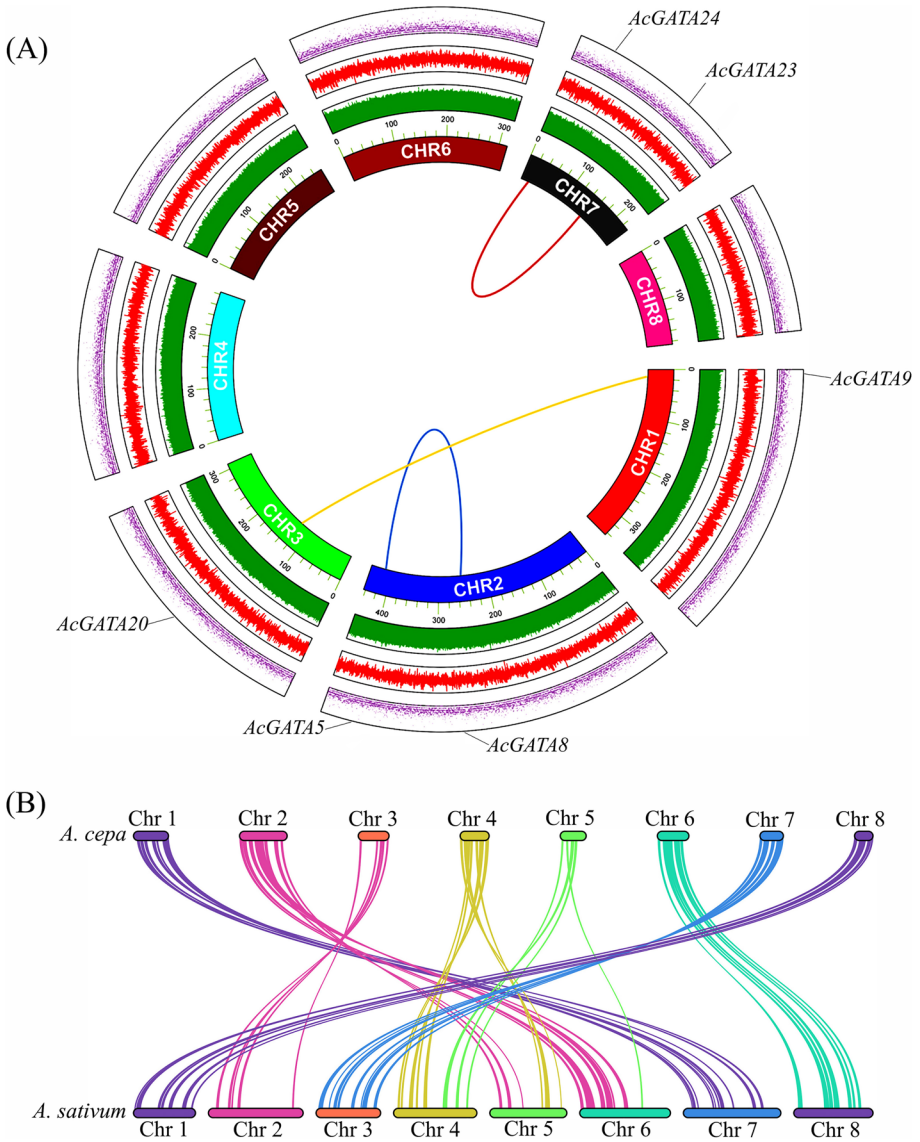

Fig. 5 **A** Synteny and chromosomal level analysis of the *AcGATA*s. The colored lines indicate duplication of *AcGATA*s among the chromosomes. The innermost ring represents the chromosome length; the green-colored rings represent the nucleotide density per chromosome; the red-colored ring represents the GC content per chromosome; the purple-colored ring represents the gene density per chromosome. **B** Collinearity analysis between onion and *A. thaliana* chromosomes. Different colored lines denote the collinear genes. The chromosome numbers are indicated above each chromosome

AcGATA13, AcGATA14, AcGATA15, AcGATA16, AcGATA19, AcGATA20, AcGATA21, AcGATA22, AcGATA23, and *AcGATA24*) out of the 24 *AcGATA*s exhibited upregulated expressions at different time points (Fig. 10). On the contrary, *AcGATA17* and *AcGATA18* were found to get downregulated at 12-, 24-, and 6-h post-treatment, respectively. Moreover, the number of *AcGATA*s exhibited differential expressions under salinity stress was found to be more than under chromium stress.

Similarly, 17 of the *AcGATA*s were found to be induced by both salinity and chromium stress, showing their possible significance in regulating abiotic stress response in onions.

Protein–protein interaction network prediction

The protein–protein interaction (PPI) analysis revealed the possible interactions of *AcGATA*s with other onion proteins in regulating different physiological processes. The result indicated that the *AcGATA*s might interact

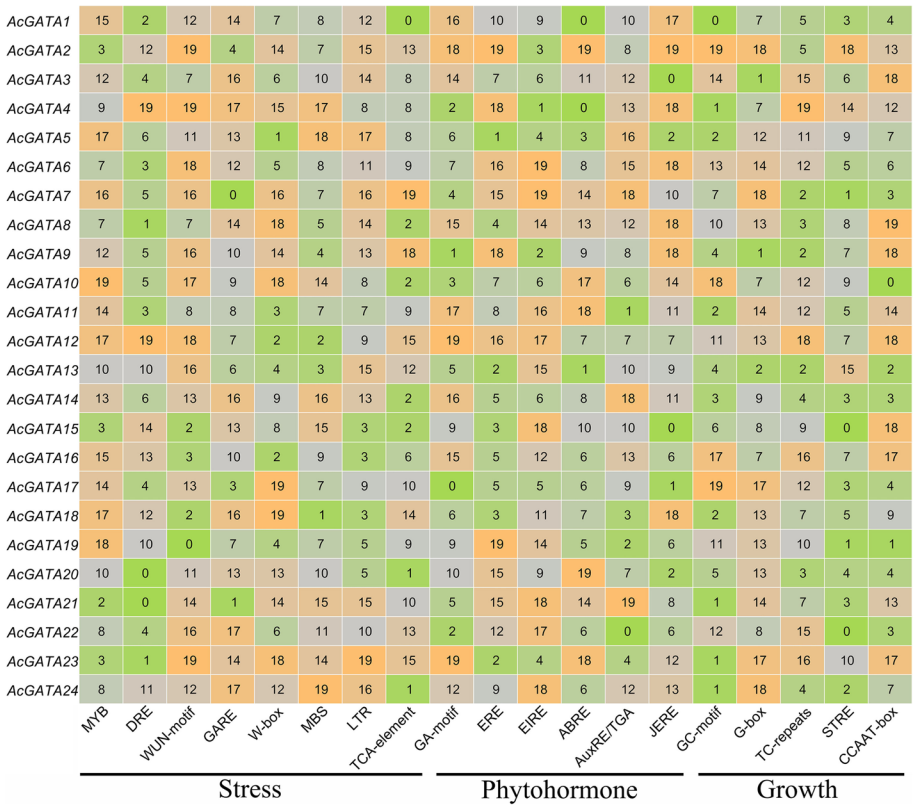

Fig. 6 Presence of the CREs in the promoter sequences of the *AcGATAs*. In the colour gradient, the green colour represents a lower, orange represents a higher number of cis-elements, and the respective number of each category of CREs are represented as numbers

with crucial proteins and transcription factors, including membrane-associated kinase regulators (MAKR1 and MAKR4), callose synthases (CALS9 and CALS10), cryptochrome (CRY1 and CRY2), de-etiolated (DET1), DNA-binding with one finger (DOF3.2), nitrilases (NIT2), bZIP transcription factor (HY5), GOLDEN2-LIKE transcription factor (GLK2), and bHLH transcription factor (bHLH118) (Fig. 11). In addition, the *AcGATA*s were found to interact with each other to regulate onion physiological processes probably.

Discussion

The GATA transcription factors are involved in crucial physiological processes in plants. The GATA gene family has been explored in several plant species, including *A. thaliana*, wheat, maize, rice, rapeseed, poplar, and buckwheat [4, 6, 10, 13, 28, 29]. However, the GATA family has not been well-explored in onions. In this study, 24 *AcGATA*s were identified and characterized in the onion genome. The exon–intron organization of the *AcGATA* s (1–9 exons) was found to be similar to other plant GATAs, including rapeseed (1–9) and rice (2–9) [4, 10]. The *AcGATA*s contained two types of GATA domains with 18 (group I, II, III, IV) and 20 (group V) amino acid residues. The difference in the protein structures can contribute to the diverge functionality of the GATAs. For instance, *A. thaliana* GATAs having different lengths of the consensus sequence regulate various physiological responses, such as hypocotyl growth, root development, flowering, lateral root formation, branching of roots, and cell differentiation, respectively [30, 31]. Thus, the *AcGATA*s could be involved in different physiological processes in onions.

A GO analysis of the *AcGATA*s was performed to predict the putative functions. The results suggested that the *AcGATA*s are mostly involved in processes like response to stimulus, response to phytohormones, hormonal signaling, cell differentiation, and plant development. Earlier reports have confirmed the role of GATAs in regulating the aforementioned processes in plants. For example, GATAs in *A. thaliana* get induced by light and regulate the development of hypocotyls and stomata [32]. Similarly, two GATAs, GNC and CGA1/GNL, are involved in the cytokinin-induced expression of photosynthesis-related genes [33]. *GmGATA58* in soybeans regulates chlorophyll biosynthesis and chlorophyll and nitrogen metabolism [34]. Recently, the role of *AtGATA25* has been established in regulating the circadian rhythm

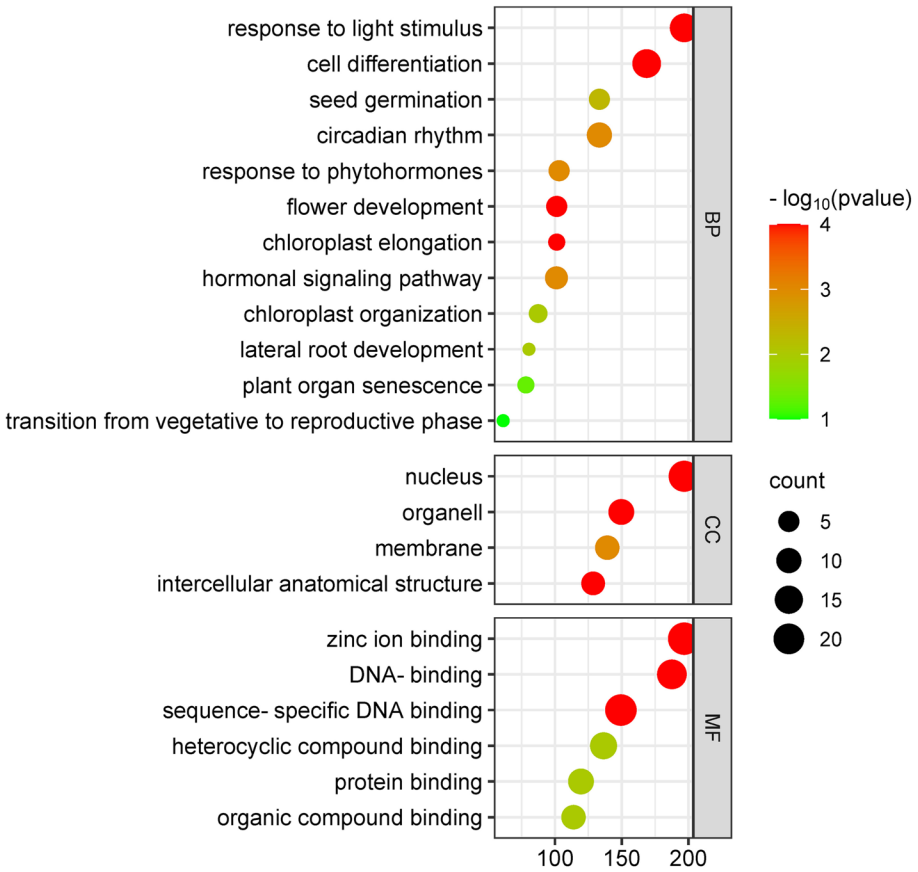

Fig. 7 GO analysis of the *AcGATA*s. The most enriched predicted biological processes (BP), cellular components (CC), and molecular functions (MF) have been listed. The colour gradient bar represents the relative upregulation (red colour gradient) of the functions, whereas the size of the circle indicates the number of hits obtained for a specific function

and hypocotyl elongation in *A. thaliana* [35]. In addition, analysis of the promoter sequences of the *AcGATA* s revealed that several CREs associated with stress response, phytohormone response, and growth-related are present. These findings suggest the probable multifaceted roles of *AcGATA*s in regulating the onion physiological processes.

In the present study, the expression analysis of the *AcGATA*s provided meaningful insights into their involvement in different tissues and under various stresses. The tissue-specific expression analysis confirmed that more *AcGATA*s are expressive in the bulb and leaf tissues compared to the roots. Under chromium stress, multiple *AcGATA*s exhibited induced expressions at different time points. Thirteen of the 24 *AcGATA*s (*AcGATA2, AcGATA3, AcGATA5, AcGATA8, AcGATA9, AcGATA11, AcGATA13, AcGATA14, AcGATA15, AcGATA18, AcGATA19, AcGATA20*, and *AcGATA22*) showed an early response to chromium stress by getting induced at 6 h post-treatment. On the other hand, 5 *AcGATA*s (*AcGATA4, AcGATA7, AcGATA17, AcGATA23*, and *AcGATA24*) exhibited a late induced response by getting induced either at 12- or 24-h post-treatment. Early and late gene expressions can be seen in several genes that respond to specific stresses. For instance, several differential expressive genes were reported in *Olea europaea* in response to cold stress [36]. Similarly, the zinc-regulated transporter and iron-regulated transporter genes exhibited both early and late expressions under cadmium stress [37]. Only a few reports are available on the expression of GATAs in response to heavy metal stress [38, 39]. To the best of our knowledge, this is the first report of the involvement of GATAs in response to chromium stress. The results of this study will help get newer insights into the involvement of GATAs in heavy metal stress in onion, particularly against chromium stress. In addition, several *AcGATA*s were found to be induced under salinity stress. The participation of GATAs in regulating salt stress response in plants is a well-studied phenomenon. Multiple GATAs have been reported to be induced under salt stress in several plants, including

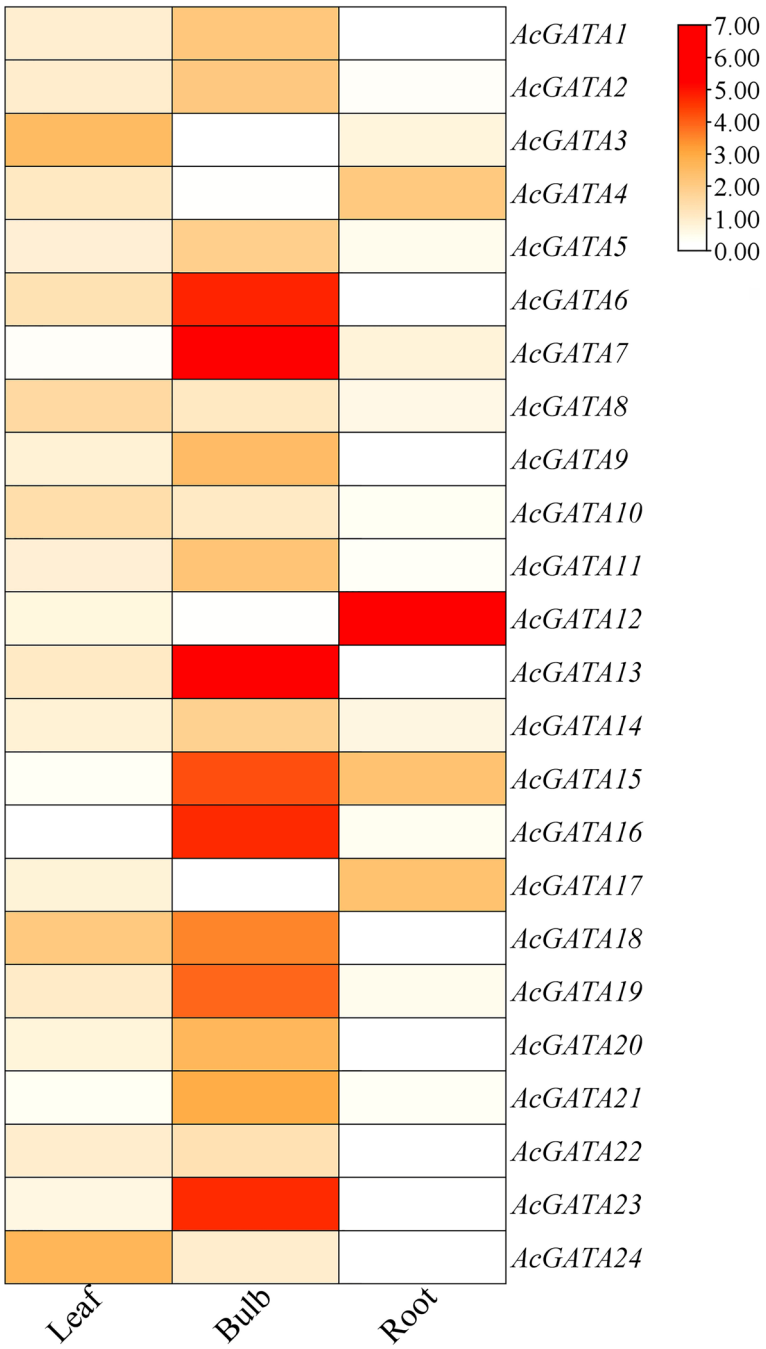

Fig. 8 Heat map exhibiting the expression of the *AcGATA*s in different onion tissues. The red and white colours indicate the up and down-regulation of the gene expressions

Phoebe bournei, *Vitis vinifera*, and *Triticum aestivum* [12, 15, 40]. Recently, a tomato GATA, *SlGATA17*, has been reported to modulate tolerance against salt stress in tomatoes by interacting with SlHY5 [24]. Moreover, the induced expression of the *AcGATA*s under chromium and salt stress suggests that multiple GATAs might be involved in the stress response in onions.

Lastly, the PPI network prediction results suggested the probable key interactions of GATAs and other proteins in onion. The results revealed that the *AcGATA*s could be interacting with nine different types of proteins apart from interacting with each other. For instance, one of such predicted interacting partners is the membrane-associated kinase regulators (MAKR1

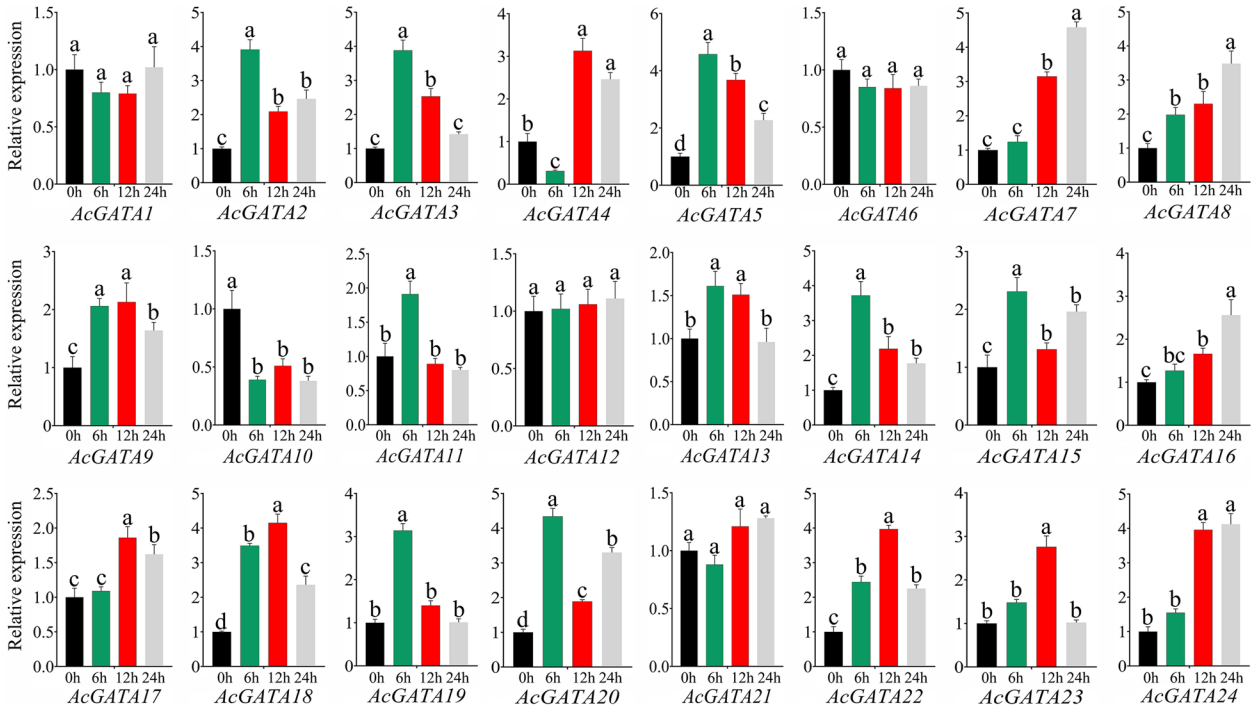

Fig. 9 Differential expression analysis of the *AcGATA*s under chromium stress. The relative expressions (fold changes) are shown on the Y-axis, whereas the different time points are shown on the X-axis. The expression data are represented as mean ± SE. The lowercase alphabets represent the statistical significance of the data

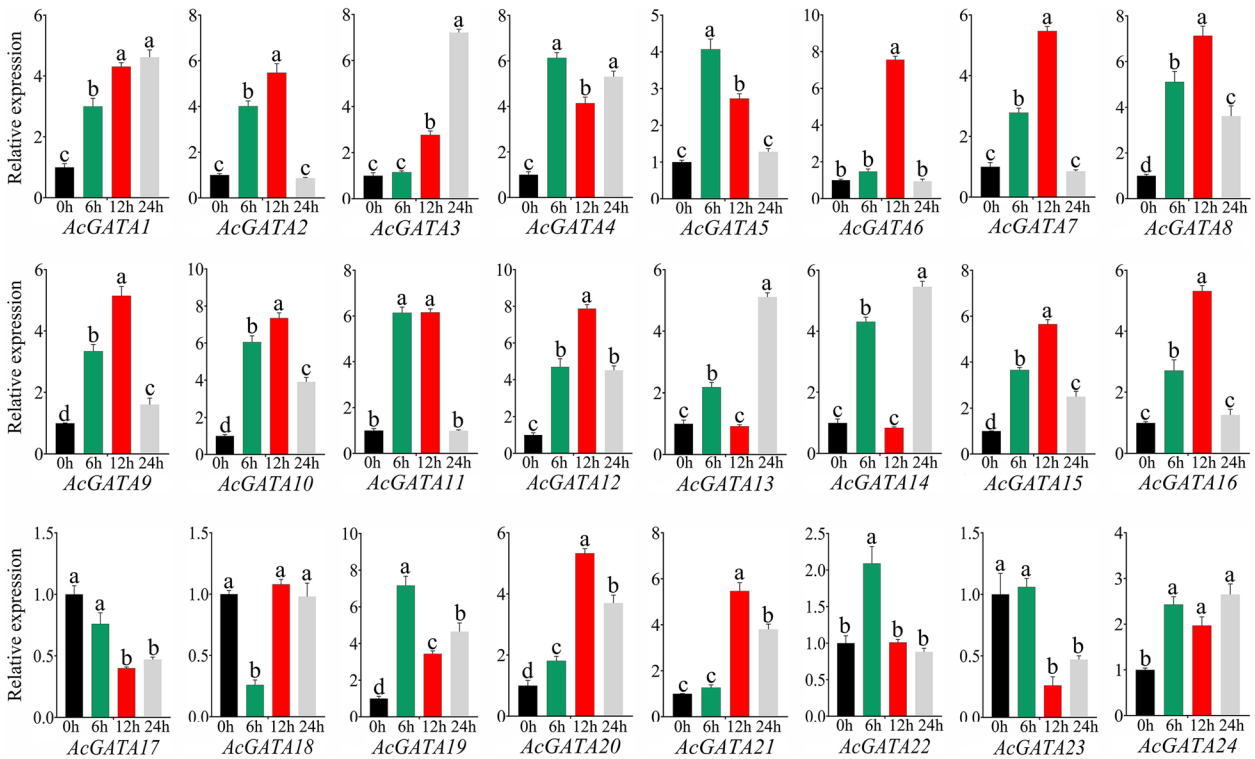

Fig. 10 Differential expression analysis of the *AcGATA*s under salinity stress. The relative expressions (fold changes) are shown on the Y-axis, whereas the different time points are shown on the X-axis. The expression data are represented as mean ± SE. The lowercase alphabets represent the statistical significance of the data

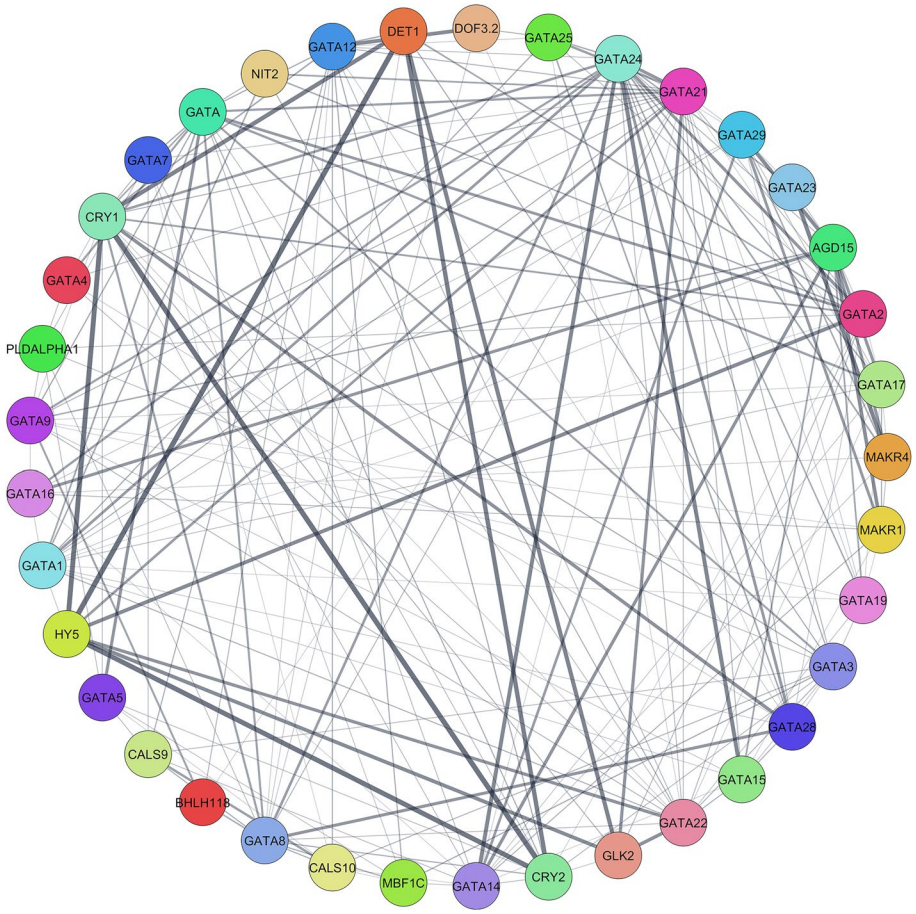

Fig. 11 The protein–protein interaction network analysis of the *AcGATAs*. The thick lines indicate a stronger interaction between the proteins, whereas the thin lines indicate a weaker interaction. The dark lines represent high-confidence interactions, while light lines indicate low-confidence interactions

and MAKR4). The MAKRs are vital players in regulating important plant functions, such as phytohormonal signaling and root development [41]. Similarly, *AcGATA*s were predicted to interact with callose synthases (CALS9 and CALS10), which play pivotal roles in plant development and environment interactions, including stress response [42]. Additionally, other possible interacting proteins to be predicted were cryptochrome (CRY1 and CRY2) and de-etiolated (DET1), which control the plant circadian rhythm and photomorphogenesis, respectively [43, 44]. Moreover, the *AcGATA*s were predicted to interact with other transcription factors, such as bZIP transcription factor (HY5), GOLDEN2-LIKE transcription factor (GLK2), and bHLH transcription factor (bHLH118), which have instrumental roles in regulating various crucial processes in plants, including growth, development, cell differentiation, and stress responses. Overall, these results suggested that the *AcGATA*s could modulate multiple physiological processes in onion by interacting with several upstream and downstream targets. However, functional validation and in-depth research are required to corroborate the same.

Conclusion

The present study identified 24 GATA genes in the onion genome. Their structural characterizations revealed that all of them possessed the canonical conserved GATA signature motif and exhibited similarity to other plant GATAs. Functional predictions through GO analysis, CRE analysis, and PPI network analysis suggested the putative multifaceted roles of the *AcGATA*s in regulating various physiological activities in onions. The RT-qPCR-based expression analysis of *AcGATA*s revealed that multiple GATAs could be involved in abiotic stress response, particularly against salt and chromium stress. Overall, the findings of this study provide basic understating and important insights into the onion GATAs and their possible roles in stress response.

Materials and methods

Plant material and treatments

The onion variety "Arka Kalyan" (originally obtained from the Indian Institute of Horticulture Research, Bangalore, India) has been used as the plant material for this study. Briefly, the onion seeds were initially sown in a nursery bed in the shed-net at Centurion University, Bhubaneswar, Odisha, India. Later, the onion plants with ~ 10 cm shoot were transplanted into individual pots and transferred to the climate control greenhouse at 24 ± 1 °C temperature and 60 ± 5% relative humidity. For salinity stress, the onion plants were treated with 100 mM NaCl solution, whereas plants treated with normal water served as the control for the experiment. Similarly, the onion plants were treated with a 10 mM chromium (Cr^{VI}) solution to induce chromium stress, while plants treated with normal water served as the control for this experiment. All the experiments were performed in triplicates.

Total RNA isolation and cDNA synthesis

After the treatments, the leave samples were collected from the plants (test and control) at 6, 12, and 24 h and snap-frozen using liquid nitrogen. The total RNA was extracted from the samples using the TRIzol™ reagent (Thermo Fisher, Waltham, Massachusetts, USA) following the manufacturer's instructions. Subsequently, the first strand cDNA was synthesized from the isolated RNA using the Verso cDNA synthesis kit following the manufacturer's protocol (Thermo Fisher, Waltham, Massachusetts, USA). The synthesized cDNA was diluted 10 times and stored at −80 °C until further use.

Identification of *AcGATA*s in onion

The onion genome files, including genome sequence, coding DNA sequences, proteins, and annotation files, were downloaded from the onion genome sequence project website (https://www.oniongenome.wur.nl/). The homologous onion GATA sequences were searched in the downloaded file using the *A. thaliana* CRK sequences as bait using the Blast tool in TBtools program [45]. The retrieved sequences were further screened for the presence of signature GATA domain by using the Simple Modular Architecture Research Tool and Conserved Domain Database [46, 47]. Finally, the peptide properties of the identified *AcGATA*s were predicted using the ProtParam tool [48].

Gene structure, motifs, multiple sequence alignment, and phylogenetic analysis

The gene structure organization of the *AcGATA*s was analyzed using the TBtools [45]. Different motifs in *AcGATA*s were predicted using the Multiple Expectation Maximization for Motif Elicitation (MEME) tool [49]. Similarly, the *AcGATA* sequences were aligned using Clustal Omega. The subsequent phylogenetic analysis was estimated by constructing a neighbor-joining tree with 1000 bootstraps using Molecular Evolutionary Genetics Analysis (MEGA) *version* 11 [50].

Chromosomal mapping, gene duplication, and synteny analysis

The identified *AcGATA*s were mapped onto respective chromosomes using the MapGene2Chrom (MG2C) tool [51]. Further, the gene duplication analysis among the *AcGATA*s was calculated using the Ka/Ks Calculator in TBtools [45]. Subsequently, the synteny and collinearity analysis were performed on TBtools using the Circos and one-step MCScanX programs, respectively, and visualized using the dual synteny plot.

CREs, subcellular localization, GO, and protein interaction network analysis

For the identification of the CREs, about 2 Kb upstream sequences of the *AcGATA*s were taken and analyzed by using the PlantCARE tool [52]. To predict the subcellular localization of the *AcGATA*s, the mGOASVM plant dedicated server was used [53]. Subsequently, the GO analysis of the *AcGATA*s was done using the Blast2GO tool and visualized on SRplot [54]. Lastly, the protein–protein interaction network was predicted using the STRING database and visualized using Cytoscape [55].

Expression analysis of *AcGATA*s

The expression profiles of the *AcGATA*s were evaluated by performing RT-qPCR on a Roche Light Cycler (Basel, Switzerland) by using gene-specific primers (Table S2), as described in Nanda et al. [56]. *AcAct1* was used as the reference gene to deduce the differential expressions in different stress conditions and tissues. The relative expression was estimated using the $2^{-\Delta\Delta Ct}$ method [57]. All experiments were conducted with three independent biological replicates. Finally, the statistical significance of the expression data was checked by forming a One-way ANOVA at $P \leq 0.05$ using the Data Processing System package [58]. The expression analysis results were visualized as bar diagrams under chromium and salinity stress, whereas as a heatmap for the tissue-specific expressions. For constructing the heatmap TBtool software has been used.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11251-3.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Acknowledgements
SN thanks the Science and Engineering Research Board (SERB), Govt. of India, for supporting his research works via the Start-up Research Grant (SRG/2021/000077). CB acknowledges the Institutional Fellowship for research from Centurion University of Technology and Management.

Authors' contributions
SN and SG conceived and designed the project. CB, PR, and PR did the in silico characterization works. PKD did the stress assay experiments. CB and KK did the RT-qPCR validation and data analysis works. SN and SG supervised the work, wrote the manuscript, and acquired funding.

Funding
The research funding is supported by ICAR-DOGR under the project Biotechnological Approaches for Biotic Stress Management (052/CI/IPP/DOGR/21–26/CPT/002).

Data availability
All data related to this study are present in this article (Table 1) and the corresponding supplementary materials (accession numbers and sequences). The accession numbers of the AcGATA1-14 are g17314.t1, g46105.t1, g79291.t1, g79608.t1, g113541.t1, g175187.t1, g175189.t1, g177201.t1, g251888.t1, g291540.t1, g361876.t1, g374580.t1, g379039.t1, g380862.t1, g393953.t1, g403354.t1, g424699.t1, g440419.t1, g483965.t1, g523096.t1, g499934.t1, g207213.t1, g18777.t1, and g351433.t1, respectively. The datasets generated and analysed during the current study are available in the Onion Genome Sequencing Project repository (https://www.oniongenome.wur.nl/, accessed on 04 February 2024).

Declarations

Ethics approval and consent to participate
No experiments were performed on animals or humans. The experiments conducted on plants, including plant stress subjection and sample collection were done in compliance with the institutional, national, and international guidelines.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Biotechnology, Centurion University of Technology and Management, Bhubaneswar, Odisha 752050, India. [2]Department of Phytopharmaceuticals, School of Agriculture and Bioengineering, Centurion University of Technology and Management, Paralakhemundi, Odisha 761211, India. [3]Department of Botany, S.N. Arts, D.J.M. Commerce and B.N.S. Science College, Sangamner, Maharashtra 422605, India. [4]ICAR-Directorate of Onion and Garlic Research, Pune, Maharashtra 410505, India. [5]Department of Zoology, S.N. Arts, D.J.M. Commerce and B.N.S. Science College, Sangamner, Maharashtra 422605, India.

Received: 10 May 2024 Accepted: 16 January 2025
Published online: 27 February 2025

References

1. Strader L, Dolf W, Doris W. Plant transcription factors—being in the right place with the right company. Curr op in plant bio. 2022;65: 102136.
2. Franco-Zorrilla JM, López-Vidriero I, Carrasco JL, Godoy M, Vera P, Solano R. DNA-binding specificities of plant transcription factors and their potential to define target genes. Proc of the Nat Aca of Sci. 2014;111(6):2367–72.
3. Hong JC. General aspects of plant transcription factor families. In Plant transcription fact. Amsterdam: Academic Press. 2016;35–56.
4. Reyes JC, Muro-Pastor MI, Florencio FJ. The GATA family of transcription factors in Arabidopsis and rice. Plant physio. 2004;134(4):1718–32.
5. He P, Wang X, Zhang X, Jiang Y, Tian W, Zhang X, Li Y, Sun Y, Xie J, Ni J, He G. Short and narrow flag leaf1, a GATA zinc finger domain-containing protein, regulates flag leaf size in rice (Oryza sativa). BMC plant bio. 2018;18:1–1.
6. Feng X, Yu Q, Zeng J, et al. Genome-wide identification and characterization of GATA family genes in wheat. BMC Plant Biol. 2022;22:372.
7. Zhang C, Hou Y, Hao Q, Chen H, Chen L, Yuan S, Shan Z, Zhang X, Yang Z, Qiu D, Zhou X. Genome-wide survey of the soybean GATA transcription factor gene family and expression analysis under low nitrogen stress. PLoS ONE. 2015;10(4): e0125174.
8. Yu C, Li N, Yin Y, Wang F, Gao S, Jiao C, et al. Genome-wide identification and function characterization of GATA transcription factors during development and in response to abiotic stresses and hormone treatments in pepper. J Appl Genet. 2021;62:265–80.
9. Zhang K, Jia L, Yang D, Hu Y, Njogu MK, Wang P, et al. Genome wide identification, phylogenetic and expression pattern analysis of GATA family genes in cucumber (Cucumis sativus L.). Plants-Basel. 2021;10:1626.
10. Zhao K, Nan S, Li Y, Yu C, Zhou L, Hu J, Jin X, Han Y, Wang S. Comprehensive Analysis and Characterization of the GATA Gene Family, with Emphasis on the GATA6 Transcription Factor in Poplar. Int J Mol Sci. 2023;24:14118.
11. Zhang Z, Zou X, Huang Z, Fan S, Qun G, Liu A, et al. Genome-wide identifification and analysis of the evolution and expression patterns of the GATA transcription factors in three species of Gossypium genus. Gene. 2019;680:72–83.
12. Yin Z, Liao W, Li J, Pan J, Yang S, Chen S, Cao S. Genome-Wide Identification of GATA Family Genes in Phoebe bournei and Their Transcriptional Analysis under Abiotic Stresses. Int J Mol Sci. 2023;24(12):0342.
13. Yao X, Zhou M, Ruan J, He A, Ma C, Wu W, Lai D, Fan Y, Gao A, Weng W, Cheng J. Genome-wide identification, evolution, and expression pattern analysis of the GATA gene family in tartary buckwheat (Fagopyrum tataricum). Int J Mol Sci. 2022;23(20):12434.
14. Zheng K, Lu J, He X, Lan S, Zhai T, Cao S, Lin Y. Genome-Wide Identification and Expression Analysis of GATA Family Genes in Dimocarpus longan Lour. Int J Mol Sci. 2024;25:731.
15. Zhang X, Ma J, Yang S, et al. Analysis of GATA transcription factors and their expression patterns under abiotic stress in grapevine (Vitis vinifera L.). BMC Plant Biol. 2023;23:611.
16. An Y, Zhou Y, Han X, Shen C, Wang S, Liu C, et al. The GATA transcription factor GNC plays an important role in photosynthesis and growth in poplar. J Exp Bot. 2020;71:1969–84.
17. Ravindran P, Verma V, Stamm P, Kumar PP. A novel RGL2-DOF6 complex contributes to primary seed dormancy in Arabidopsis thaliana by regulating a GATA transcription factor. Mol Plant. 2017;10:1307–20.
18. Liu X, Zhu X, Wei X, Lu C, Shen F, Zhang X, et al. The wheat LLM-domain containing transcription factor TaGATA1 positively modulates host immune response to Rhizoctonia cerealis. J Exp Bot. 2020;71:344–55.
19. Zhang Y, Zhang Y, Zhang L, Huang H, Yang B, Luan S, et al. OsGATA7 modulates brassinosteroids-mediated growth regulation and influences architecture and grain shape. Plant Biotechnol J. 2018;16:1261–4.
20. Lu G, Casaretto JA, Ying S, Mahmood K, Liu F, Yong B, et al. Overexpression of OsGATA12 regulates chlorophyll content, delays plant senescence and improves rice yield under high density planting. Plant Mol Biol. 2017;94:215–27.
21. Nutan KK, Singla-Pareek SL, Pareek A. The Saltol QTL-localized transcription factor OsGATA8 plays an important role in stress tolerance and seed development in Arabidopsis and rice. J Exp Bot. 2019;71:684–98.
22. Zhang H, Wu T, Li Z, et al. OsGATA16, a GATA Transcription Factor, Confers Cold Tolerance by Repressing OsWRKY45–1 at the Seedling Stage in Rice. Rice. 2021;14:42.
23. Zhu H, Zhai H, He S, Zhang H, Gao S, Liu Q. A novel sweet potato GATA transcription factor, IbGATA24, interacting with IbCOP9-5a positively regulates drought and salt tolerance. Environ Exp Bot. 2021;194: 104735.

24. Wang Y, Cao X, Zhang D, Li Y, Wang Q, Ma F, Xu X, Zhan X, Hu T. SlGATA17, A tomato GATA protein, interacts with SlHY5 to modulate salinity tolerance and germination. Environ Exp Bot. 2023;206: 105191.
25. Chaudhry UK, Junaid MD, Gökçe ZN, Gökçe AF. Impact of Biotic and Abiotic Stresses on Onion Production: Potential Mitigation Approaches in Modern Era. In Smart Plant Breeding for Vegetable Crops in Post-genomics Era. Singapore. 2023;143–162.
26. Pattnaik S, Mohapatra S, Pati S, Dash D, Devadarshini D, Tanaya K, Mishra BB, Samantaray D. Microbial bioremediation of Cr (VI)-contaminated soil for sustainable agriculture. In Microbial Biodeg and Biorem. Amsterdam: Elsevier. 2022;395–407.
27. Finkers R, van Kaauwen M, Ament K, et al. Insights from the first genome assembly of Onion (Allium cepa). G3 (Bethesda). 2021;11(9):jkab243.
28. Long J, Yu X, Chen D, Hu F, Li J. Identification, phylogenetic evolution and expression analysis of GATA transcription factor family in maize (Zea mays). Int J Agric Biol. 2020;23:637–43.
29. Zhu W, Guo Y, Chen Y, Wu D, Jiang L. Genome-wide identification, phylogenetic and expression pattern analysis of GATA family genes in Brassica napus. BMC Plant Biol. 2020;20:543.
30. Richter R, Bastakis E, Schwechheimer C. Cross-repressive interactions between SOC1 and the GATAs GNC and GNL/CGA1 in the control of greening, cold tolerance, and flowering time in Arabidopsis. Plant Physiol. 2013;162:1992–2004.
31. Rybel BD, Vassileva V, Parizot B, et al. A novel aux/IAA28 signaling Cascade activates GATA23-dependent specification of lateral root founder cell identity. Curr Biol. 2010;20:1697–706.
32. Klermund C, Ranftl QL, Diener J, Bastakis E, Richter R, Schwechheimer C. LLM-domain B-GATA transcription factors promote stomatal development downstream of light signaling pathways in Arabidopsis thaliana hypocotyls. Plant Cell. 2016;28:646–60.
33. Chiang YH, Zubo YO, Tapken W, Kim HJ, Lavanway AM, Howard L, et al. Functional characterization of the GATA transcription factors GNC and CGA1 reveals their key role in chloroplast development, growth, and division in Arabidopsis. Plant Physiol. 2012;160:332–48.
34. Zhang C, Huang Y, Xiao Z, Yang H, Hao Q, Yuan S, Chen H, Chen L, Chen S, Zhou X, et al. A GATA Transcription Factor from Soybean (Glycine max) Regulates Chlorophyll Biosynthesis and Suppresses Growth in the Transgenic Arabidopsis thaliana. Plants. 2020;9:1036.
35. Kim K, Shin J, Kang TA, et al. CRISPR/Cas9-mediated AtGATA25 mutant represents a novel model for regulating hypocotyl elongation in Arabidopsis thaliana. Mol Biol Rep. 2023;50:31–41.
36. María OLP, Antonio VC, Raquel V, Jaime JR, Antonio MM, Oswaldo T, Juan BB, Jesús MB, Francisco L. Early and delayed long-term transcriptional changes and short-term transient responses during cold acclimation in olive leaves. DNA Res. 2015;22(1):1–11.
37. Singh S, Parihar P, Singh R, Singh VP, Prasad SM. Heavy Metal Tolerance in Plants: Role of Transcriptomics, Proteomics, Metabolomics, and Ionomics. Front Plant Sci. 2016;6:1143.
38. Manzoor Z, Hassan Z, Ul-Allah S, Khan AA, Sattar A, Shahzad U, Amin H, Hussain M. Transcription factors involved in plant responses to heavy metal stress adaptation. In Plant perspectives to global climate changes. 2022;221–231.
39. Rutherford JC, Bird AJ. Metal-Responsive Transcription Factors That Regulate Iron, Zinc, and Copper Homeostasis in Eukaryotic Cells. Eukaryot Cell. 2004;3:1.
40. Du X, Lu Y, Sun H, Duan W, Hu Y, Yan Y. Genome-Wide Analysis of Wheat GATA Transcription Factor Genes Reveals Their Molecular Evolutionary Characteristics and Involvement in Salt and Drought Tolerance. Int J Mol Sci. 2023;24:27.
41. Novikova DD, Korosteleva AL, Mironova V, Jaillais Y. Meet your MAKR: the membrane-associated kinase regulator protein family in the regulation of plant development. FEBS J. 2022;289(20):6172–86.
42. Li N, Lin Z, Yu P, Zeng Y, Du S, Huang LJ. The multifarious role of callose and callose synthase in plant development and environment interactions. Front Plant Sci. 2023;14:1183402.
43. Liu H, Liu B, Zhao C, Pepper M, Lin C. The action mechanisms of plant cryptochromes. Trends Plant Sci. 2011;16(12):684–91.
44. Li K, Gao Z, He H, Terzaghi W, Fan LM, Deng XW, Chen H. Arabidopsis DET1 represses photomorphogenesis in part by negatively regulating DELLA protein abundance in darkness. Mol Plant. 2015;8(4):622–30.
45. Chen C, Wu Y, Li J, Wang X, Zeng Z, Xu J, Liu Y, Feng J, Chen H, He Y, Xia R. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. Mol Plant. 2023;16:1733–42.
46. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. Nucleic Acids Res. 2021;49(D1):D458-60.
47. Marchler-Bauer A, et al. CDD: NCBI's conserved domain database. Nuclei Acids Res. 2015;43(D1):D222-6.
48. Gasteiger E, et al. Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook. 2005;571–607.
49. Bailey TL, et al. The MEME suite. Nucleic Acids Res. 2015;43(W1):W39–49.
50. Koichiro T, Glen S, Sudhir K. MEGA11: Molecular Evolutionary Genetics Analysis version 11. Mol Bio and Evo. 2021;38:3022–7.
51. Chao JT, Kong YZ, Wang Q, Sun YH, Gong DP, Lv J, Liu GS. MapGene-2Chrom, a tool to draw gene physical map based on Perl and SVG languages. Yi chuan. 2015;37(1):91–7.
52. Lescot M, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.
53. Wan S, Mak M-W, Kung S-Y. mGOASVM: multi-label protein subcellular localization based on gene ontology and support vector machines. BMC Bioinformatics. 2012;13(1):1–16.
54. Tang D, Chen M, Huang X, Zhang G, Zeng L, Zhang G, Wu S, Wang Y. SRplot: A free online platform for data visualization and graphing. PLoS ONE. 2023;18(11): e0294236.
55. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.
56. Nanda S, Rout P, Ullah I, Nag SR, Reddy VV, Kumar G, Kumar R, He S, Wu H. Genome-wide identification and molecular characterization of CRK gene family in cucumber (Cucumis sativus L.) under cold stress and Sclerotium rolfsii infection. BMC Genomics. 2023;24(1):219.
57. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods. 2001;25(4):402–8.
58. Tang QY, Zhang CX. Data Processing System (DPS) software with experimental design, statistical analysis and data mining developed for use in entomological research. Insect science. 2013;20(2):254–60.

Publisher's Note