scientific reports

Check for updates

OPEN

Genome-wide identification and expression profiling of growth-regulating factor (*GRF*) and GRF-interacting factor (*GIF*) gene families in chickpea and pigeonpea

Mitesh Khisti, Tejaswi Avuthu, Kalenahalli Yogendra, Vinod Kumar Valluri, Himabindu Kudapa, Palakolanu Sudhakar Reddy & Wricha Tyagi ✉

The growth-regulating factor (*GRF*) and GRF-interacting factor (*GIF*) families encode plant-specific transcription factors and play vital roles in plant development and stress response processes. Although *GRF* and *GIF* genes have been identified in various plant species, there have been no reports of the analysis and identification of the GRF and GIF transcription factor families in chickpea (*Cicer arietinum*) and pigeonpea (*Cajanus cajan*). The present study identified seven *CaGRFs*, eleven *CcGRFs*, four *CaGIFs*, and four *CcGIFs*. The identified proteins were grouped into eight and three clades for GRFs and GIFs, respectively based on their phylogenetic relationships. A comprehensive *in-silico* analysis was performed to determine chromosomal location, sub-cellular localization, and types of regulatory elements present in the putative promoter region. Synteny analysis revealed that *GRF* and *GIF* genes showed diploid-polyploid topology in pigeonpea, but not in chickpea. Tissue-specific expression data at the vegetative and reproductive stages of the plant showed that *GRFs* and *GIFs* were strongly expressed in tissues like embryos, pods, and seeds, indicating that *GRFs* and *GIFs* play vital roles in plant growth and development. This research characterized *GRF* and *GIF* families and hints at their primary roles in the chickpea and pigeonpea growth and developmental process. Our findings provide potential gene resources and vital information on *GRF* and *GIF* gene families in chickpea and pigeonpea, which will help further understand the regulatory role of these gene families in plant growth and development.

Keywords Growth-regulating factor, GRF-interacting factor, Pigeonpea, Chickpea, Climate-resilience

Transcription factors (TFs) are proteins that bind to specific DNA sequences and regulate the ability of cells to express different genes and thereby control development[1]. Interactions between TFs and their co-activators are essential to regulate the expression of associated downstream genes. *GRF* and *GIF* genes are a family of plant-specific transcription factors that play essential roles in plant growth and development. GRF proteins bind to DNA and regulate the expression of other genes, while GIF proteins interact with GRFs to modulate their activity[2,3]. GRFs are a complex and diverse family of transcription factors, and their roles in plant growth and development are still being actively investigated. Earlier studies showed GRF's role in plant stem and leaf development[4–6]. Further studies reported that GRF is also involved in regulating the growth and development of other plant tissues, including root development[7], flower organ development[8,9], seed oil content[10], and stress responses[11,12]. Moreover, expression of the *GRFs* was reported to be at higher levels in young tissues than in mature tissues[13]. Therefore, understanding the role of *GRFs* is essential for plant growth research and genetic improvement.

Most GRFs contain unique QLQ (Gln, Leu, Gln) and WRC (Trp, Arg, Cys) domains in the N-terminal region[13,14]. The *WRC* domain can interact with the *cis*-acting regions of associated genes to regulate their

Research Program-Accelerated Crop Improvement, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, Patancheru, Telangana 502324, India. ✉email: wricha.tyagi@icrisat.org

expression. The *QLQ* domain can interact with GIF protein to form a transcription activator complex[15]. Genes encoding GIF proteins containing unique *SSXT* domains are functionally homologous to human synovial sarcoma translocation protein (*SYT* or *SSXT*)[16]. GIF protein has strong transcriptional activity and cell division ability[4]. The *Arabidopsis* GIF family contains three proteins, *GIF1*, *GIF2*, and *GIF3,* and they play essential roles in vegetative and reproductive organ development[17,18]. In other studies, knockout mutants of *AtGIF1* result in a decreased cell number with narrow-leaf phenotypes, whereas enhanced *AtGIF1* expression leads to increased leaf area with increased cell numbers in leaf primordial[4]. In addition, *AtGIF1* is involved in the establishment of cotyledon by suppressing ectopic root formation[19] and functions in adaxial/abaxial patterning and leaf growth[20]. Although GRF or GIF alone can regulate plant growth and development, they are more effective when they form a complex.

The *GRF-GIF* complex is a plant-specific transcriptional complex that regulates various aspects of plant growth and development, including leaf, stem, root, seed, and flower development[21]. Expression of a *GRF4–GIF1* fusion protein significantly increased the regeneration efficiency, speed of regeneration, and somatic embryogenesis of wheat and rice[2]. Furthermore, a dicot *GRF-GIF* chimera was also found to increase the regeneration efficiency of citrus[2] and watermelon[22] and transformation efficiency in lettuce[23], suggesting that this strategy can also be applied to dicot crops.

The chickpea (*Cicer arietinum*) and pigeonpea (*Cajanus cajan*) are protein-rich grain legumes used for human consumption in many countries. These rich protein sources serve as an alternative to meat for the vegetarian populations of many developing countries[24]. Both crops are environmentally friendly and sustain soil productivity, as they fix nitrogen in the soil. The productivity of these pulses has remained static over the past several decades because of the adverse effects of biotic and abiotic stresses in the changing climate scenario. Because of the drastically changing climate and increasing population, it is of utmost importance to generate crops that are resilient to the changing conditions with high nutrition quality and greater yield. Despite the development in plant regeneration methods, chickpea and pigeonpea are lagging behind other crops because of the complex nature of their genome and the general tendency of recalcitrance to regeneration[25]. However, there are multiple literature reports on the functions of GRF and GIF family members to increase plant regeneration and transformation efficiency in dicots[2,22,23] and detailed genome-wide phylogenetic and functional studies, of *GRF* and *GIF* genes are yet to be available for chickpea and pigeonpea.

To better understand the evolutionary dynamics of *GRF* and *GIF* genes in chickpea and pigeonpea and explore their potential regulatory roles in vegetative and reproductive stages of plant growth and development, we identified eleven *CcGRF*, seven *CaGRF,* and four *GIF* genes each in chickpea and pigeonpea. We performed *in-silico* analyses to document their chromosome locations, *cis*-acting elements, evolutionary relationships, genomic collinearity, and selection pressure. We also investigated the potential roles of these GRFs and GIFs using protein interaction network prediction and gene expression analysis. These results provide insights into the functions of *GRF* and *GIF* family members in chickpea and pigeonpea growth and development.

Materials and method

Sequence retrieval and identification of *GRF-GIF* genes

The genomic information as well as gene annotations of chickpea and pigeonpea were obtained from the Phytozome database (https://phytozome-next.jgi.doe.gov/) and CajanusMine (part of Legume Information System) (https://mines.legumeinfo.org/cajanusmine/begin.do), respectively. GRF and GIF protein sequences of soybean (*Glycine max*)[26,27], *Arabidopsis thaliana*[13,28] and *Medicago truncatula*[29] were obtained from previous studies. To identify the GRF sequences in chickpea and pigeonpea, a local BLAST was performed using 22 soybean, nine *Arabidopsis,* and eight *Medicago* GRF sequences using the local protein library of chickpea and pigeonpea downloaded separately (https://mines.legumeinfo.org) in TBtools[30]. Similarly, to identify GIF sequences in both the species, eight soybean, and three *Arabidopsis* GIF sequences were considered as a query to perform a local BLAST in TBtools. The e-value for the BLAST search was set to 1^{e-5} with 500 total hits and 250 total aligns. To confirm the identified protein sequences, a comprehensive HMM-based search was also performed in TBtools using Pfam IDs of *QLQ* and *WRC* domains for GRFs (PF08879 and PF08880) and *SSXT* domain (PF05030) for GIFs using an HMM-library downloaded from Pfam database (https://www.ebi.ac.uk/interpro/download/Pfam/). After removing the redundant sequences, the obtained sequences were searched for the presence of *QLQ* along with *WRC* domains (PF08879 and PF08880) in the case of GRFs and *SSXT* domain (PF05030) in the case of GIFs using the Interpro scan tool (https://www.ebi.ac.uk/interpro/search/sequence/).

ExPasy ProtParam tool (https://web.expasy.org/protparam/)[31] was used to predict the physiochemical properties of the obtained sequences, such as predicted theoretical isoelectric point (pI), molecular weight, and aliphatic index. Subcellular localization of the obtained proteins was predicted using the WoLF-PSORT tool (https://www.genscript.com/tools/wolf-psort)[32].

Phylogenetic tree construction and motif analysis

To construct a phylogenetic tree for establishing evolutionary relations between the identified genes, multiple sequence alignment of peptide sequences of *G. max*, *A. thaliana,* and *M. truncatula* for GRFs and *G. max*, *A. thaliana* for GIFs was done along with identified peptide sequences from chickpea and pigeonpea using ClustalW tool from MEGA v10. Phylogenetic trees for GRF and GIFs were constructed using the Maximum-likelihood (ML) method with 1000 bootstraps using the MEGA v10 software package. The obtained trees were visualized using iTOL v6 online tool (https://itol.embl.de/)[33]. Motifs of the obtained protein sequences were analysed using a Multiple Em for Motif Elicitation (MEME) online tool (https://meme-suite.org/meme/tools/meme) with a Zero or One Occurrence Per Sequence (ZOOPS) distribution, a number of motifs to be identified set to 8 and a motif range of 6–50 amino acids wide. The obtained motif analysis was visualised using TBtools.

Collinearity and evolutionary analysis of identified *GRF* and *GIF* genes

Collinearity analysis of chickpea and pigeonpea was performed with soybean, *M. truncatula,* and *Arabidopsis* separately using the Multiple Collinearity Scan toolkit (MCScanX) program of TBtools with parameters as follows: CPU for BLASTp-2; Number of BLAST hits- 5; E-value- 1^{e-10}. To identify orthologous genes in the considered organisms, the advanced circos tool of TBtools was employed, and orthologous *GRF*, as well as *GIF* genes of chickpea and pigeonpea in soybean, *M. truncatula,* and *Arabidopsis,* were connected in the form of curved arches. The non-synonymous substitution rates and synonymous substitution rates of the obtained orthologous pairs of genes were determined using a simple Ka/Ks calculator tool of TBtools.

Chromosomal location and gene structure analysis of *GRF* and *GIF* genes

The information regarding the chromosomal location and intron–exon structures for *GRF* and *GIF* genes of chickpea and pigeonpea was obtained from the respective annotation files. Based on the chromosomal locations, the identified genes were mapped onto chromosomes using the MapGene2Chromosome v2.1 online tool[34] (http://mg2c.iask.in/mg2c_v2.1/). Gene structures of the identified *GRF* and *GIF* genes were visualized using Gene Structure Display Server 2.0[35] (http://gsds.gao-lab.org/index.php).

Cis-regulatory element analysis of the identified *GRF-GIF* genes

To identify the *cis-regulatory* elements in the promoter region of identified *GRF* and *GIF* genes, 2000 bp upstream region of *GRF* and *GIF* genes were obtained from Phytozome v13 (https://phytozome-next.jgi.doe.gov/) for chickpea and from Cajanusmine (https://mines.legumeinfo.org/cajanusmine/begin.do) for pigeonpea. The *cis-regulatory* elements were identified by submitting the identified promoter sequences to the PlantCARE online tool[36] (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The representation for *cis-regulatory* element analysis was done using TBtools.

miR396 target site prediction in the identified *GRF* sequences of chickpea and pigeonpea

miR396 RNA sequences for chickpea and pigeonpea were retrieved from the PmiREN: Plant microRNA Encyclopedia (https://pmiren.com/database). The target sites of the obtained *miRNA396* in chickpea and pigeonpea GRF sequences were determined using psRNATarget[37] (https://www.zhaolab.org/psRNATarget/). The predicted target sites were visualized using Geneious Prime software package v2022.1.1.

Protein interaction network prediction

To predict the potential protein-interaction network, the identified CcGRF/GIF and CaGRF/GIF proteins were submitted to the STRING online tool[38] (http://string-db.org). The orthologs of the submitted proteins were identified in *A. thaliana* and rice using BLAST and the ortholog with the highest bit score was considered further. The protein–protein interaction network was generated with high confidence level of 0.07 and non-interacting proteins were removed.

Expression analysis of *GRF* and *GIF* genes

The *GRF* and *GIF* gene sequences of chickpea and pigeonpea were blasted against the CDS sequences of chickpea[39] and pigeonpea[40] using BLASTn program from NCBI Blast version 2.7.1 with default parameters. Considering alignment length, percent identity, and e-value from the BLASTn output results, *GRF* and *GIF* gene sequences of chickpea and pigeonpea were selected. Later, corresponding gene expression data was obtained from gene expression atlas datasets of chickpea breeding cultivar ICC4958[41] and pigeonpea genotype ICPL87119 (Asha)[42]. The expression data for different tissues in vegetative and reproductive developmental stages of the considered crops was represented in the form of heatmaps drawn using TBtools.

Functional validation of expression data using qRT-PCR

Pigeonpea genotype TS3R and chickpea genotype ICC4958 were used to analyse the expression pattern of identified *GRF* and *GIF* genes. For the vegetative stage of the plants, we collected embryos (24 h after germination), radicle (3 days after germination), root (10 days after germination), leaf (10 days after germination) and shoot (10 days after germination). Similarly, to analyse the expression of *GRF* and *GIF* genes in the reproductive stage of the plant, 90 days old glasshouse-grown plants were utilized to collect root, shoot, leaf, and mature seeds. All the samples were immediately frozen in liquid nitrogen after collection and were stored at −80 °C.

Total RNA was isolated using NucleoSpin RNA Plant Mini Kit (Macherey–Nagel, Düren, Germany) following the manufacturer's protocol. The integrity of extracted RNA was assessed via 2% agarose gel electrophoresis, while the quantity of the RNA samples was determined using an Invitrogen Qubit fluorometer (Fisher Scientific, Loughborough, UK). A total 2 µg of RNA was utilized for first-strand cDNA synthesis using protoScript II reverse transcriptase (NEW ENGLAND *Biolabs*, United Kingdom), which is a recombinant Moloney Murine Leukemia Virus (M-MuLV) reverse transcriptase following the manufacturer's protocol. The primers specific to each targeted gene from both pigeonpea and chickpea, were designed using the online primer-3 tool (Table S12). A concentration of 10 µM for each primer set was used per reaction. To ensure the specificity of primers, they were checked in NCBI-BLAST against the respective legume genome. The qRT-PCR was performed using SensiFAST™ No-ROX SYBR green master mix on a CFX96TM Real-Time System (Bio-Rad, Gurugram, India). One gene each, *CaUCP* in chickpea[43] and *CcGAPDH* in pigeonpea[44] were used as housekeeping genes. Two biological and three technical replicates were performed for each qRT-PCR. To assess the amplification specificity of each gene, a melting curve step was integrated post-reaction. Relative expression levels of identified *GRF* and *GIF* genes were calculated by the comparative Ct method[45].

Results

Identification and physiochemical properties of *CaGRFs* and *CcGRFs*

Based on the BLAST results and domain analysis, we identified seven *GRF* genes in chickpea and eleven *GRF* genes in pigeonpea (Fig. S1a and b). For the seven identified GRFs in chickpea, the molecular weight ranged from 37.99 kDa to 69.55 kDa, while the amino acid length ranged from 344 to 633. The predicted theoretical pI for five of the *CaGRF*s was in the basic region (*CaGRF1, CaGRF2, CaGRF4*, *CaGRF5*, and *CaGRF6)*, while two of the *CaGRF*s showed predicted theoretical pI in the acidic range (6.66 for *CaGRF3* and 6.51 for *CaGRF7*). The aliphatic index of the identified *CaGRF*s was 42.42 to 58.26 (Table S1). The molecular weight of *CcGRF*s ranged from 31.45 to 64.9 kDa, while the amino acid length ranged from 282 to 598. The predicted theoretical pI for most of the *CcGRF*s was in the basic range except for *CcGRF5* and *CcGRF7* (6.67 and 6.62), which were in the acidic range. The aliphatic index for *CcGRF*s ranged from 37.44 to 64.55 (Table S2). Other important properties, including coding region (CDS), number of exons/introns, chromosomal location, and cellular localization of the identified genes, are summarized in Tables S1 and S2.

Identification and physiochemical properties of *CcGIFs* and *CaGIFs*

We identified four *GIF* genes, each in chickpea and pigeonpea (Fig. S1c). The molecular weight of *CaGIF*s ranged from 22.14 to 24.15 kDa, while *CcGIF*s ranged from 16.73 kDa to 23.44 kDa. Most *CcGIF*s had acidic predicted theoretical pI except for *CcGIF4*, which had a basic predicted pI of 7.96. Similarly, predicted theoretical pI for all of the *CaGIF*s was in the acidic range. The aliphatic index for *CaGIF*s was in the range of 44.43 to 61.16, while for *CcGIF*s it was in the range of 51.82 to 72.24 (Tables S1 and S2). The genes' physiochemical properties and genomic information are summarised in Tables S1 and S2.

Phylogenetic and motif analysis of the *GRF* and *GIF* genes

To establish a phylogenetic relationship between *GRF*s, a Maximum-likelihood phylogenetic tree was constructed among *GRF*s from soybean (22), *Arabidopsis* (9), *Medicago* (8), pigeonpea (11) and chickpea (7) (Fig. 1a). A total of 57 genes clustered into eight clades (I-VIII). Among all clades, clade II, IV, and VIII had the highest number of members (10 members), followed by clades III and VI (7 members), clade VII (6), clade V (5) and clade I (2 members). Except for clade V and VI all the other clades contained at least one AtGRF protein, this can be attributed to the distinction of identified GRF proteins in designated clades based on their similarity with AtGRF proteins. All the CcGRFs and CaGRFs were closely associated with GmGRFs and AtGRFs, indicating the evolutionary conservation of *GRF* genes among dicots.

Similarly, a Maximum-Likelihood phylogenetic tree was constructed for *GIF*s from soybean (8), *Arabidopsis* (3), pigeonpea (4), and chickpea (4) (Fig. 1b). A total of 19 *GIF* genes clustered into three clades (I-III). Clade I contained the fewest members (3), while clades II and III had an equal number of members (8 each). No clades specific to particular species were observed in the trees. Except for Clade I the other two clades contained at

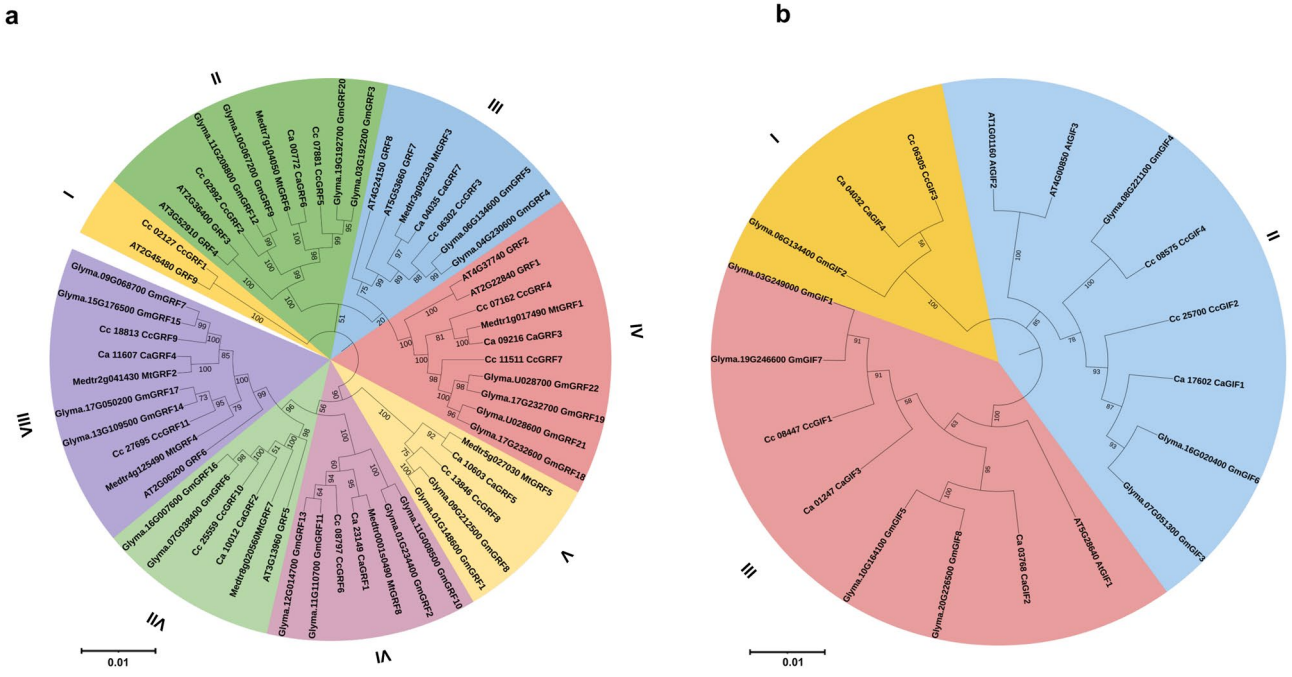

Figure 1. Phylogenetic analysis of *GRF* and *GIF* genes. (**a**) Phylogenetic analysis of *GRF* genes of *Cicer arietinum*, *Cajanus cajan*, *Glycine max*, *Arabidopsis thaliana* and *Medicago truncatula*. Eight clades are represented in eight colours. (**b**) Phylogenetic analysis of *GIF* genes of *Cicer arietinum*, *Cajanus cajan*, *Glycine max* and *Arabidopsis thaliana*. Three clades are represented in three different colours. Trees were constructed using the Maximum-Likelihood (ML) method with 1000 bootstraps.

least one member of *Arabidopsis* GIF protein family supporting the distinction of GIF proteins in clades based on their homology with *AtGIF* proteins. Motif analysis indicated conservation of motif 1 belonging to the *SSXT* domain across all GIF proteins (Fig. S2). Clade-specific conservation of motifs was also noted in the GIF proteins' phylogenetic tree, with all motifs conserved in Clade III, whereas the number of conserved motifs ranged from 3 to 7 in Clade II. Clade I exhibited the least detected motifs, with *CaGIF4* showing presence of only one conserved motif (motif 1) (Fig. 1b and Fig. S2).

The specific conservation patterns of motifs across GRF sequences contribute significantly to the delineation of distinct clades in the phylogenetic tree (Fig. 1a and Fig. S2). GRF sequences examined shared two common motifs, belonging to the *QLQ* and *WRC* domains, which are evolutionarily conserved (Fig. S1). Notably, there was consistency in the presence of conserved motifs among GRF protein sequences within specific clades. Members of clades VII and I exhibited conservation of all observed motifs (Fig. S2). Furthermore, with the exception of *AtGRF6*, all other members in clade VIII displayed identical motifs. Similarly, most motifs were conserved in clade II, except for motif 5, which was absent in *CcGRF2* and *GmGRF20*. In clade IV, the number of conserved motifs ranged from three to five, with motifs 1, 2, and 3 conserved across all members. Clade III showed variability in the number of conserved motifs, ranging from four to six.

Collinearity and evolutionary analysis of *GRF* and *GIF* genes

To better understand the evolutionary relationships between *GRF* and *GIF* genes of chickpea and pigeonpea with other dicots, we performed a collinearity analysis with two members of the Fabaceae family, viz. Soybean, *M. truncatula* and model dicot plant *Arabidopsis* (Fig. 2). Out of seven *CaGRF* genes five showed collinearity with 7 orthologous *GRF* genes in *Medicago* and 14 orthologous *GRF* genes in soybean, while only two *CaGRF* genes showed collinearity with 4 orthologous *GRF* genes in Arabidopsis (Fig. 2a, Table S3). This shows the convergence of the *GRF* gene family in chickpea in the course of evolution.

Similarly, the collinearity analysis identified 19 unique collinear genes in soybean for 11 *CcGRF* genes with a total of 33 collinear orthologous pairs of genes (Fig. 2b, Table S5), while the number of unique collinear genes for *CcGRF* was 8 with 16 collinear orthologous pairs of genes in *M. truncatula* and 7 with 11 collinear orthologous pairs of genes in *Arabidopsis* (Fig. 2b, Table S5). Interestingly, the number of collinear genes in soybean for each *CcGRF* gene varied from 2 to 5, supporting the diploid-polyploid topology (Table S5).

We identified seven collinear genes in soybean, four collinear genes in *M. truncatula,* and only one collinear gene in Arabidopsis for four orthologous *CaGIFs* (Fig. 2c, Table S4). It is interesting to note that the identified collinear gene in Arabidopsis showed the presence of two orthologous gene pairs in chickpea, namely *CaGIF2* and *CaGIF3,* while, each identified *CaGIF* gene had one or more than one collinear pair in soybean and *Medicago truncatula* (Table S4). Similar to *CaGIFs*, three orthologous *CcGIFs* also had 7, 4, and 3 collinear genes in soybean, *Medicago truncatula,* and *Arabidopsis,* respectively (Fig. 2d, Table S6).

To understand the selection pressure acting on the considered genes throughout the evolutionary time frame, we calculated the ratios of synonymous to non-synonymous mutations (Ka/Ks) in the identified gene pairs in chickpea and pigeonpea (Tables S3 and S4). All the calculated Ka/Ks ratio values were below one, indicating that the GRF and GIF protein family experienced purifying solid selection during evolution in chickpea and pigeonpea (Fig. S3, Tables S3, S4, S5 and S6).

Chromosomal location and gene structure analysis of *GRF* and *GIF* genes

A total of 7 *CaGRF* genes were distributed over six chickpea chromosomes (Chr01, Chr02, Chr04, Chr06, Chr07 and Chr08), with Chr02 containing two *CaGRF* genes (*CaGRF3* and *CaGRF4*). Four *CaGIF* genes were distributed on four chickpea chromosomes (Fig. 3a). In the case of pigeonpea, 11 *CcGRFs* and 4 *CcGIF* genes were distributed over eight chromosomes (Chr 01–06; Chr 08, and Chr 11), with chromosome 03 and chromosome 11 containing multiple copies of *GRF* and *GIF* genes (Fig. 3b).

The gene structure analysis for *CaGRFs* revealed that the exon numbers ranged from 3 to 5, and all *CaGIFs* had four exons each (Fig. S4a and b). For *CcGRF* and *CcGIF* genes, the exon numbers ranged from 3 to 4 (Fig. S4c and d).

Cis-regulatory elements analysis of the upstream region

Cis-regulatory elements analysis is essential to understand the interactions of the 5′ UTR region of a gene with regulatory proteins, which in turn controls gene expression. To analyse the *cis-elements* associated with the putative promoter regions of the identified *GRF* and *GIF* genes in both species, we extracted the 2000 bp upstream sequence of the *GRF* and *GIF* genes and analysed the *cis-regulatory* elements in the region (Fig. 4, Tables S7 and S8).

The identified *cis-regulatory* elements associated with *CcGRFs* and *CaGRFs* were divided into three types (Fig. 4a and c). Hormone-responsive elements included elements reported in response to gibberellic acid, auxin, methyl jasmonic acid, abscisic acid, and salicylic acid. The presence and number of these elements in the promoter regions of the identified genes varied from gene to gene. For example, gibberellin responsive element was present in upstream region of all *CaGRFs* except *CaGRF4* (Fig. 4a). Similarly, in *CcGRF* genes, *CcGRF8* and *CcGRF11* had auxin-responsive elements in their promoter region, while all the other genes lacked this element (Fig. 4c). The number of methyl-jasmonate responsive elements was higher in *CaGRF* gene promoters than in *CcGRF* promoters. Another type of *cis-regulatory* element is growth and development-related elements, including light responsiveness, circadian control, meristem, endosperm expression, and cell cycle regulation. Out of all the elements, the putative promoter regions of the *GRF* genes had the highest number of light-responsive elements, accounting for more than half of the identified elements. The presence of elements responsible for differentiating meristem and mesophyll palisade cells in the upstream region of the *GRF* gene highlights the role of these genes in

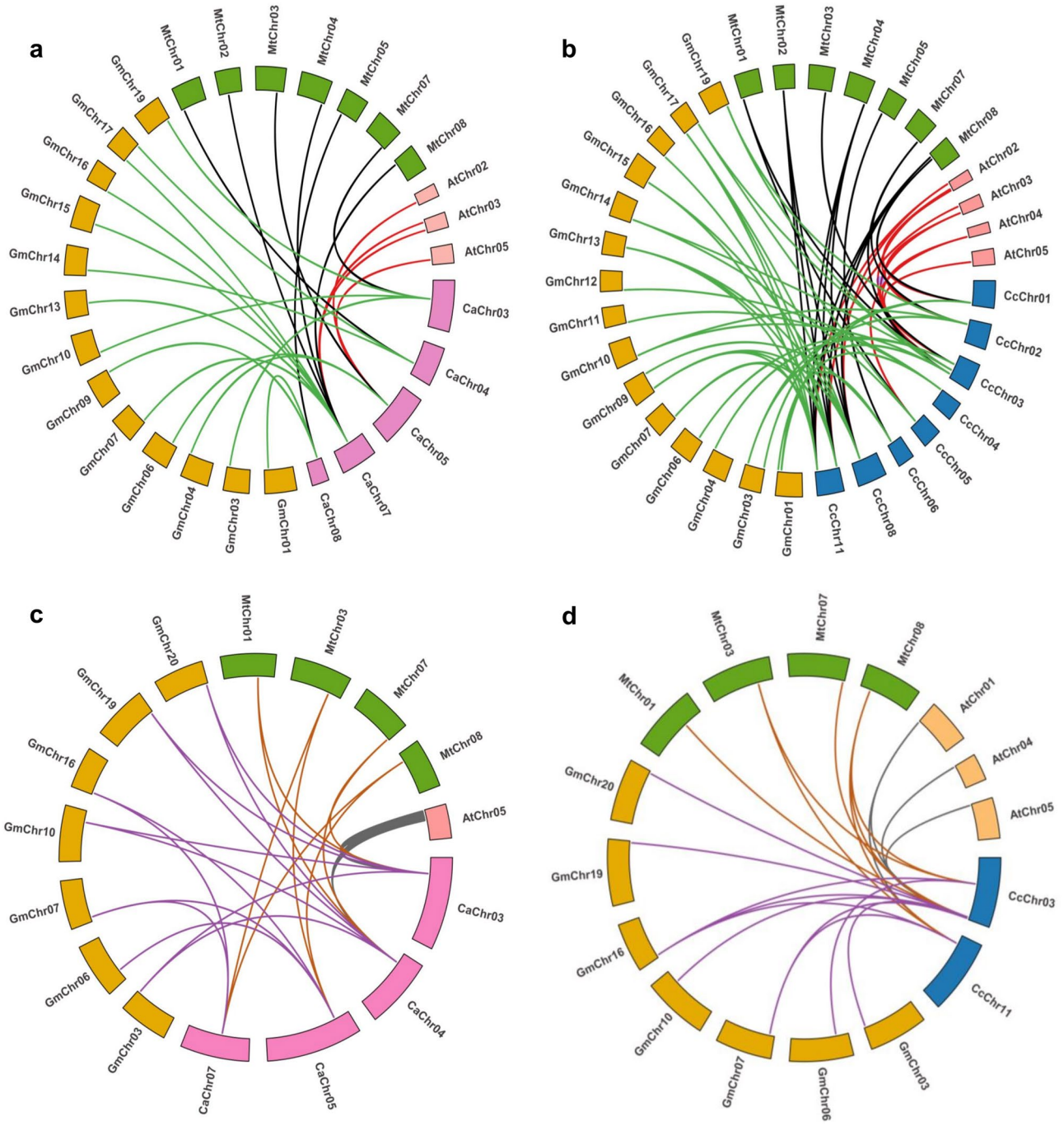

Figure 2. Collinearity analysis of *GRF* and *GIF* genes. Syntenic *GRF* gene pairs of (**a**) *Cicer arietinum* and (**b**) *Cajanus cajan*. Orthologous syntenic gene pairs for *Glycine max*, *Medicago truncatula* and *Arabidopsis thaliana* are represented in green, black and red colour, respectively. Syntenic *GIF* gene pairs of (**c**) *Cicer arietinum* and (**d**) *Cajanus cajan*. Orthologous syntenic gene pairs for *Glycine max*, *Medicago truncatula* and *Arabidopsis thaliana* are represented in violet, orange and grey colours, respectively.

plants' growth and morphology. Cell growth-promoting morphogenic genes have recently been actively employed to increase recalcitrant crops' regeneration potential. Interestingly, the 5′ upstream region of *CcGRF6* showed the presence of all the major growth development-related elements, including meristem differentiation, cell cycle control, and circadian control, making it a potential candidate for a morphogenic gene (Fig. 4c). The third type of elements identified included elements related to stress responses, such as defence and stress-responsive elements, elements associated with anaerobic induction, drought-inducible elements, wound-responsive elements, low-temperature responsive elements, and elicitor-mediated activation elements for biotic stresses (Fig. 4a and c).

Similarly, for *CaGIF* and *CcGIF* putative promoter sequences, the *cis-regulatory* elements identified were divided into three types (Fig. 4b and d)—hormone-responsive elements comprised elements responsive to gibberellins, abscisic acid, methyl-jasmonate, and salicylic acid. Salicylic acid-responsive elements were observed

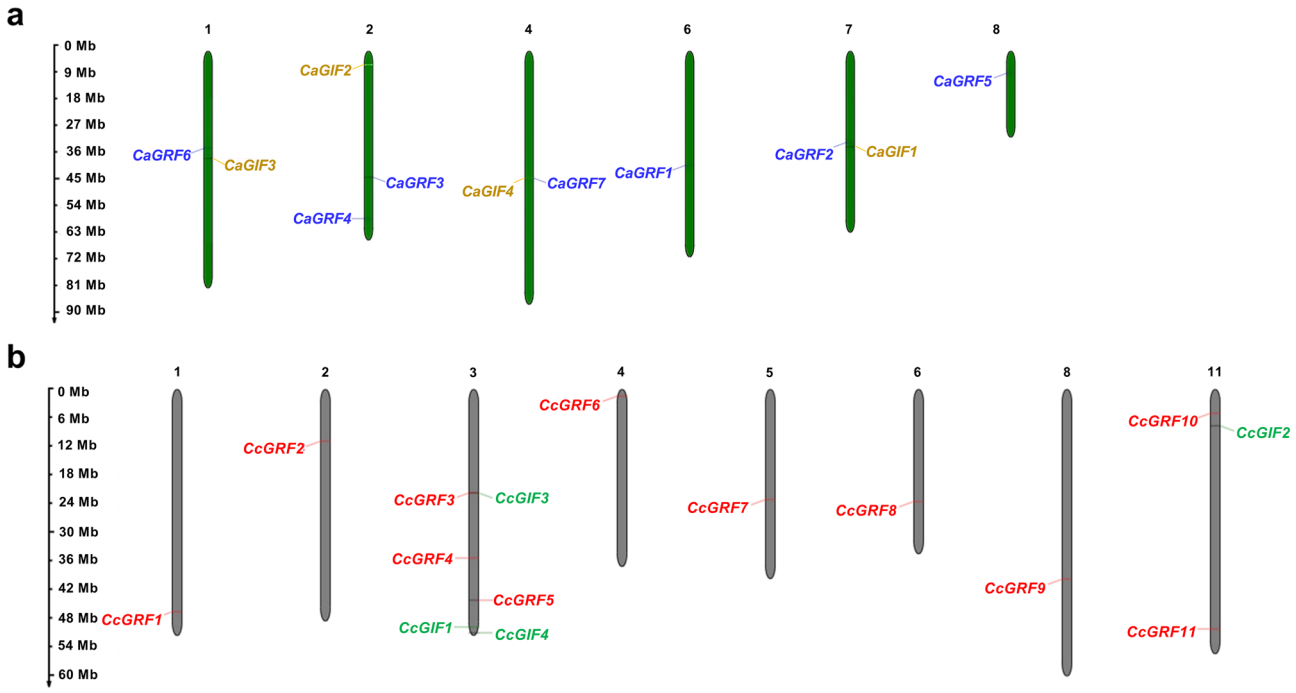

Figure 3. Chromosomal locations of *GRF* and *GIF* genes. (**a**) Chromosomal locations of *CaGRF* and *CaGIF* genes; *CaGRF* genes are represented in blue while *CaGIF* genes are represented in yellow. (**b**) Chromosomal locations of *CcGRF* and *CcGIF* genes; *CcGRF* genes are represented in red while *CcGIF* genes are represented in green. The scale on the left side represents the length of the chromosome in Mb.

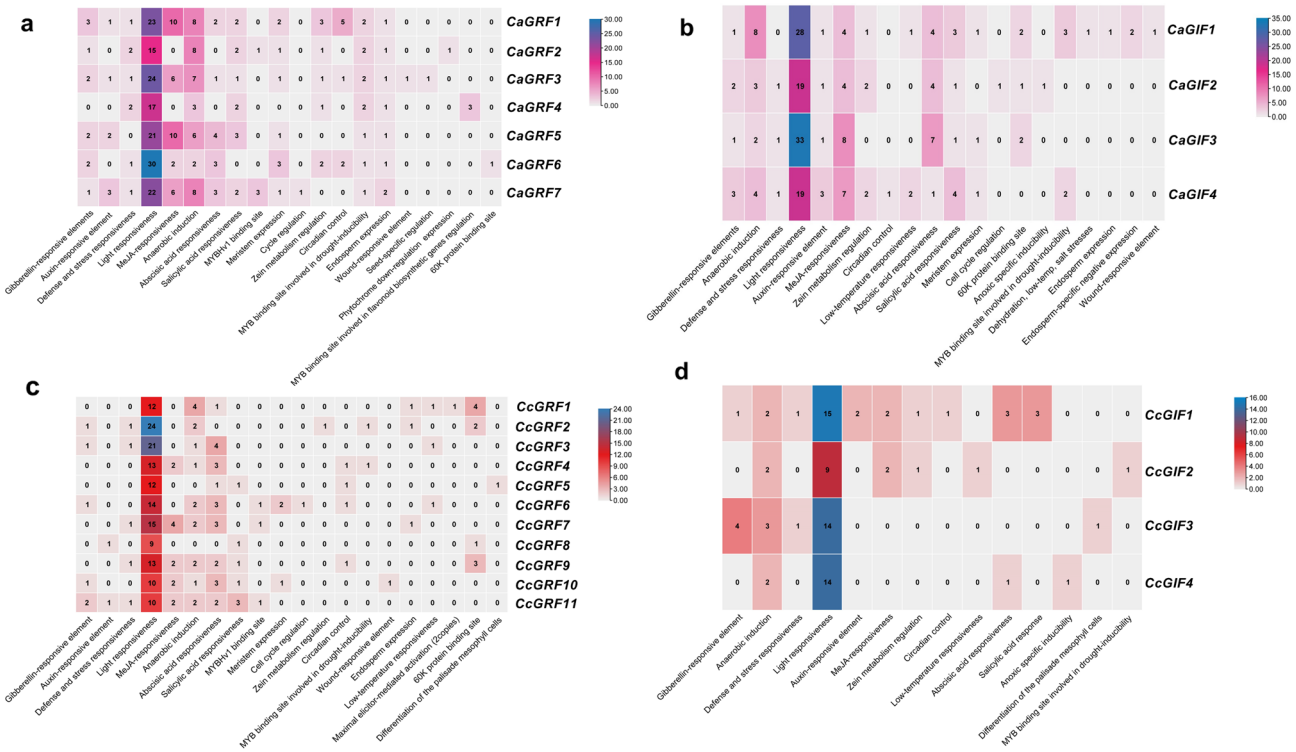

Figure 4. Heatmaps representing *cis*-regulatory elements analysis in 5′ upstream region of *GRF* and *GIF* genes. Category of *cis*-regulatory elements present in *GRF* genes in (**a**) *Cajanus cajan* and (**c**) *Cicer arietinum*. *Cis*-regulatory elements analysis of *GIF* genes in (**b**) *Cajanus cajan* and (**d**) *Cicer arietinum*. The scale on the right represents the color scale based on the number of *cis*-regulatory elements.

to be more in the upstream region of *CaGIF* genes, with *CcGIF1* being the only gene whose 5′ UTR showed the occurrence of salicylic acid-responsive *cis*-element. In the case of growth and development-related elements, *CcGIF* and *CaGIF* genes had certain elements like light-responsiveness and circadian control common to both. In contrast, some elements like meristem expression, cell-cycle regulation, and endosperm expression were unique to *CaGIFs*. *CaGIF1* displayed regulatory elements responsible for negative and positive endosperm expression (Fig. 4b). *CcGIF3* sequence carries unique growth and developmental elements responsible for the differentiation of the mesophyll cells (Fig. 4d). *CcGIFs* and *CaGIFs* showed the presence of stress-related elements similar to *GRF* genes, including elements like anaerobic induction, low-temperature response, anoxic specific inducibility, dehydration, and salt-inducible elements. Out of all the identified *GIF* genes, only *CaGIF1* showed the presence of a wound-responsive element specific to biotic stresses (Fig. 4b).

The presence of all three types of elements (hormone, growth, and development, as well as stress-responsive) in the 5′ UTR region of both *GRF* and *GIF* genes indicates the comprehensive nature of their biological function as transcription factor families in regulating the expression of essential genes.

miRNA target site prediction of *CaGRF* and *CcGRF* genes

miRNA396 and GRF transcription factors are involved in a growth-regulatory module in which *miRNA396* targets the GRF mRNA to reduce its expression and modulate growth[46]. We predicted the presence of these *miRNA396* target sites in 7 *CaGRFs* and 11 *CcGRFs* identified in this study (Fig. S5, Table S9). In chickpea, *miRNA396* is predicted to be encoded by two loci (*CaMiR396a* and *CaMiR396b*), while in pigeonpea, by four loci (*CcMiR396a, CcMiR396b, CcMiR396c, CcMiR396d*). We obtained the positions of *miRNA396* target sites using the psRNATarget online tool (Table S9). The prediction revealed that the *CaGRFs* and *CcGRFs* had the target sites for *miRNA396*. The target sites were annotated to be present in the *WRC* domains of the gene. Prediction revealed a bulge at the eighth position in the case of *CaGRFs*, and a bulge at the seventh position (from the 5′ end of miRNA) in *CcGRFs* (Fig. S5), inferring the role of *miRNA396* in the transcriptional regulation of *GRF* genes.

Protein–Protein interaction

The protein–protein interaction network of high confidence level (0.700) was generated based on the homology of GRF-GIF proteins of pigeonpea and chickpea with *Arabidopsis* predicted their interaction with multiple proteins. The PPI enrichment p-value for the predicted network is 3.01^{e-4} and showed significantly more interactions than expected with an average local clustering coefficient of 0.353. In the case of chickpea, interactions were observed between five CaGRF proteins viz. CaGRF1, CaGRF3, CaGRF5, CaGRF6, and CaGRF7, and two CaGIF proteins with CaGIF3 and CaGIF4 show strong interaction with all other GRF proteins, revealing their role as transcription factors for GRF gene expression. CaGRF proteins were also predicted to interact with other CaGRFs, while no such interactions were predicted within the CaGIF proteins (Fig. 5a). Apart from these interactions, other proteins like F4HT79_ARATH, F2P16.100, which are transmembrane proteins showed interaction with CaGRF6 and CaGIF3 respectively. Both MCK7.15 and MJB21.7 (putative B3 domain-containing proteins) showed their interactions with CaGRF1 protein while MCK7.15 interacted with CaGRF3 as well. ROT4 (DVL family protein) showed interaction with CaGIF3 and CaGIF4 proteins, while MOB 1B (MOB kinase activator-like 1B; belongs to the MOB1/phocein family) interacted with CaGRF5 protein.

Furthermore, the protein–protein interaction analysis of pigeonpea was also predicted with a high confidence level (0.700). The PPI enrichment p-value for the predicted network is 1.22^{e-4} and showed significantly more interactions than expected with an average local clustering coefficient of 0.407. The predicted network showed interaction between two GIF proteins viz. CcGIF1 and CcGIF4 with six GRF proteins viz. CcGRF1, CcGRF3, CcGRF5, CcGRF7, CcGRF9 and CcGRF10 (Fig. 5b). Similar to chickpea, all the CcGIF proteins were predicted to interact with all CcGRF proteins. Interactions were also predicted within the CcGRF proteins, while no interactions were observed within the CcGIF proteins (Fig. 5b). Other proteins like F4HT79_ARATH, F2P16.100 which are transmembrane proteins, showed interaction with CcGRF5 and CcGIF1 respectively. ROT4 (DVL family protein) showed interaction with both CcGIF1 and CcGIF4 proteins, while MOB1B (MOB kinase activator-like 1B; belongs to the MOB1/phocein family) interacted with CcGRF10. F28I8.6 (uncharacterized protein) and F28B23.2 (DNA binding protein) showed their interaction with the CcGRF9 protein.

The protein–protein interaction network was also generated for chickpea and pigeonpea GRF-GIF proteins based on their homology with *Oryza sativa* GRF-GIF proteins as these two genes were first reported from rice (Fig. S6). The protein–protein interaction network of CaGRF and CaGIF proteins predicted their interaction with multiple proteins. CaGIF3 and CaGIF4 interacted with other GRF proteins, uncharacterized proteins, J domain-containing protein, BAF250_C domain-containing protein, and SWIB domain-containing protein, while CaGRF5 interacted with RING-CH-type domain-containing protein and protein kinase domain-containing protein; that belongs to the protein kinase superfamily. The predicted network has a PPI enrichment p-value of 3.71^{e-05} and an average local clustering coefficient of 0.656 (Fig. S6).

Similarly, in the case of pigeonpea CcGIF1, CcGIF3 and CcGIF4 interacted with other GRF proteins, uncharacterized proteins, J domain-containing protein, BAF250_C domain-containing protein and SWIB domain-containing protein, while CcGRF3 interacted with RING-CH-type domain-containing protein and protein kinase domain-containing protein; that belongs to the protein kinase superfamily. The predicted network has a PPI enrichment p-value of 3.6^{e-07} and an average local clustering coefficient of 0.595 (Fig. S6).

Expression analysis of *GRF-GIF* genes in chickpea and pigeonpea growth and development

The expression of *GRF* and *GIF* genes was checked across various tissues in two developmental stages, viz., vegetative stage and reproductive stage in chickpea and pigeonpea. The tissues considered for vegetative stages included radicle, embryo, leaf, and root, while the tissues nodule, petiole, leaf, bud, flower, mature pod, and

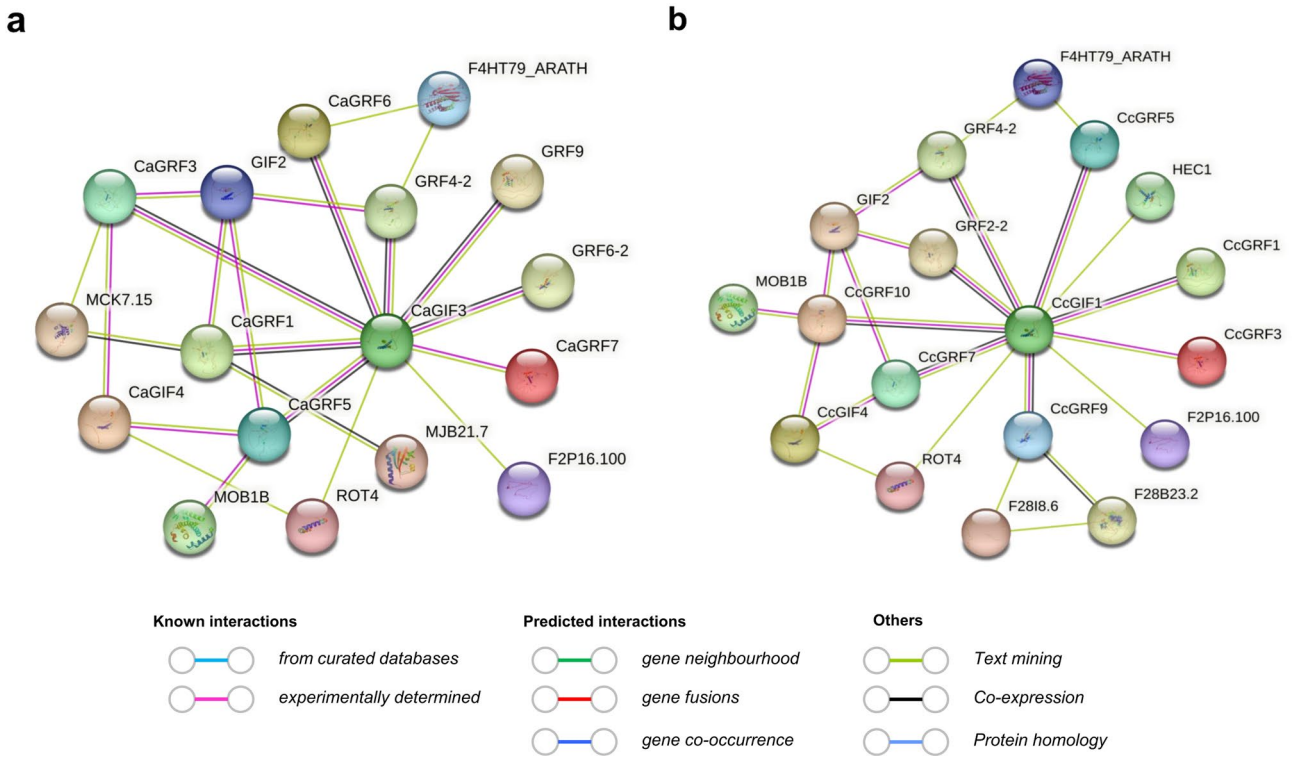

Figure 5. Predicted protein interaction networks of (**a**) CaGRF and GIF proteins, and (**b**) CcGRF and GIF proteins based on their homology with *Arabidopsis thaliana* GRF-GIF proteins obtained from STRING database. The type of interactions between the proteins is represented by different coloured lines as described in the figure legend. Red line indicates the presence of fusion evidence; Green line is neighborhood evidence; Blue line is co-occurrence evidence; Purple line is experimental evidence; Yellow line—text mining evidence; Light blue line—database evidence; Black line—co-expression evidence.

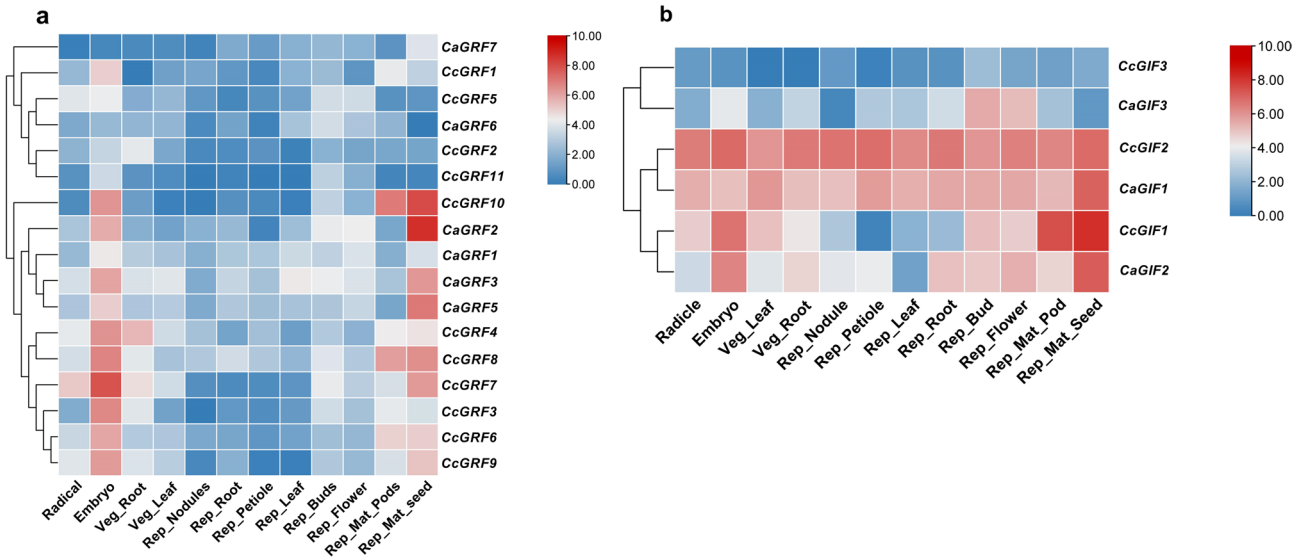

Figure 6. Expression atlas for chickpea and pigeonpea *GRF* and *GIF* genes. **a**, Expression atlas of *GRF* genes. (**b**) Expression atlas of *GIF* genes. The heat map was created according to the Fragments per kilobase of transcript per million fragments mapped (FPKM) value of *GRF* and *GIF* genes. The changes in expression are shown as colour scale.

mature seed were considered for the reproductive stages (Tables S10 and S11). The comparison of gene expression is represented in heatmaps for *GRF* and *GIF* genes of chickpea and pigeonpea (Fig. 6).

For chickpea, all the *CaGRFs* had lower expression in the radical tissue, while in embryonic tissue, except for *CaGRF6* and *7*, all the other *CaGRFs* had enhanced expression. For root and leaf tissue in the vegetative stage of the crop, all the *CaGRFs* followed a lower expression pattern. In the reproductive stage, matured seeds had the highest expression of *CaGRF* genes except for *CaGRF6*, which showed no expression in mature seeds. Among all the *CaGRFs*, *CaGRF2* had the highest expression in mature seeds, followed by *CaGRF5*. In general, *CaGRF6* and *CaGRF7* consistently expressed low across all tissues of different developmental stages. Out of all the identified *CaGRF* genes, the expression patterns of *CaGRF4* were not detectable in our transcriptome data (Fig. 6a).

For pigeonpea, certain *GRFs* (*CcGRFs 1, 2, 3, 10,* and *11*) had low expression in the radicle tissue, while *CcGRF 4*, *5*, *6*, *7*, *8*, and *9* displayed relatively higher expression. All the *CcGRF* genes expressed highly in the embryonic tissue, with *CcGRF7* expressing the highest. The expression of *GRF* genes in leaf and root was observed to be moderate, with *CcGRFs 1*, *2*, *5*, and *11* showing moderately low expression as compared to other *CcGRFs* (*CcGRF3*, *4*, *6*, *7*, *8*, and *9*) having moderately higher expression (Fig. 6a). *CcGRFs* showed lower expression in tissues like nodules and petioles. The expression of all *CcGRFs* was higher in flower and flower buds. Except for certain *CcGRFs* like *CcGRF1*, *2*, *5,* and *11*, all other *CcGRFs* expressed very highly in pod and seed tissue of the reproductive stage, with *CcGRF10* expressing the highest followed by *CcGRF8* (Fig. 6).

In the case of chickpea *GIF*, expression of *CaGIF4* was not detected in the transcriptome data. Out of the three *CaGIFs*, expressions of *CaGIF1* and *CaGIF2* were relatively higher than *CaGIF3*. *CaGIF1* was detected to have the highest expression in all chickpea tissues followed by *CaGIF2* (Fig. 6).

In the case of pigeonpea *GIF* genes, out of the four identified *CcGIFs*, expressions of three *CcGIF* genes viz. *CcGIF1*, *2*, and *3* were observed across all the considered tissues while *CcGIF4* showed no detectable expression. Out of the three *GIFs*, expressions of *CcGIF1* and *CcGIF2* were observed to be significantly higher than those of the *CcGIF3* gene, which was similar to the expression pattern of *CaGIF* genes. *CcGIF2* showed the highest expression across all tissues. Expression of *CcGIF1* decreased during the initialization of the reproductive stage of the plant, which was evident in tissues such as the petiole and nodule. The expression increased gradually with the onset of floral organ development (or sexual organ development) across tissues like bud, flower, mature pod, and mature seed, with the highest expression observed in mature seeds (Fig. 6b).

Further, to validate the expression of the identified *GRF-GIF* genes, a comprehensive qRT-PCR analysis was conducted across various developmental stages of chickpea and pigeonpea plants. The tissues considered for qRT-PCR analysis included radicle, embryo, leaf, shoot, and root at the vegetative stage while root, shoot, leaf, and mature seeds at the reproductive stage. The *CcGAPDH* and *CaUCP* genes were employed as the reference housekeeping genes for pigeonpea and chickpea, respectively to calculate the relative expression levels, presented as fold-change (FC), across the tissues analyzed (Table S13).

The qRT-PCR data for chickpea *GRFs* revealed that *CaGRF2* (7.7 FC) had the highest relative expression in embryos, while *CaGRF1, 4*, *6*, and *7* expressed high in vegetative roots (Fig. 7a). *CaGRF5* had high relative expression in both vegetative root and embryo.

In pigeonpea, eight *CcGRFs* (*CcGRF1* (17.6 FC), *CcGRF3* (51.5 FC)*, CcGRF6* (49.2 FC), and *CcGRF10* (88.5 FC) exhibited notably high relative expression levels in mature seed tissue during the reproductive stage (Fig. 7b). Conversely, the remaining *CcGRFs* (*CcGRF4* (16.1 FC)*, CcGRF7* (61.3 FC), *CcGRF9* (11.1 FC)) showed significant relative expression in the embryo during the vegetative stage (Fig. 7b). Overall, the identified *CcGRF* genes displayed higher expression levels in key tissues such as mature seeds, vegetative embryos, and reproductive roots compared to other tissues collected from various plant development stages (Fig. 7b).

The relative expression of *CaGIF* and *CcGIF* genes with reference to *CaUCP* and *CcGAPDH* housekeeping genes, respectively across the nine different tissues by qRT-PCR was also studied (Fig. 8). The chickpea GIFs (*CaGIFs)* expressed in all nine tissues considered in the present study with the relative expression being the highest in embryo tissue for *CaGIF2* (1.4 FC) and *CaGIF3* (0.5 FC) while *CaGIF4* (1.3 FC) expressed at higher levels in vegetative roots (Fig. 8a). Aligning with the transcriptome data, we could not detect the expression of *CcGIF4* gene by qRT-PCR analysis, may be due to the presence of splicing variant in the primer binding region of the *CcGIF4*. The expression pattern of *CcGIF* genes revealed high expression in mature seeds similar to *CcGRF* genes (Fig. 8b).

Discussion

Growth-regulating factors (*GRF*s) are a family of plant-specific transcription factors that play pivotal roles in plant growth and development, including cell proliferation, organogenesis, and stress responses[3,47,48]. The *GRF* gene family has been identified and described in detail in model plants including *Arabidopsis*[13], rice[14], soybean[27], and *Medicago*[29]. Furthermore, *GIF*s are a family of proteins that interact with *GRF*s and act as co-factors that modulate the activity of *GRF*s. Existing studies have confirmed that the *GRF* and *GIF* families can regulate the development of roots, stems, leaves, flowers, and fruits and maintain shoot apical meristems[3,23,49,50]. Although *GRF* and *GIF* genes have been studied in other plant species, they have yet to be characterized in the chickpea and pigeonpea. We, therefore, identified the *GRF* and *GIF* genes in chickpea and pigeonpea, analysed their evolutionary relationships, compared their sequence features, and analysed their expression patterns in different plant tissues during different stages of plant growth.

Identification, evolution, and characteristics of the *GRF*s and *GIF*s in chickpea and pigeonpea

Based on sequence analysis in the genome database and domain analysis, we identified eleven *GRF* genes in the pigeonpea and seven in the chickpea genome. Chickpea and pigeonpea have fewer *GRF* genes than the dicot legume plants, soybean (22), and *Medicago* (27). This suggests that *GRF* family expansion in dicots may

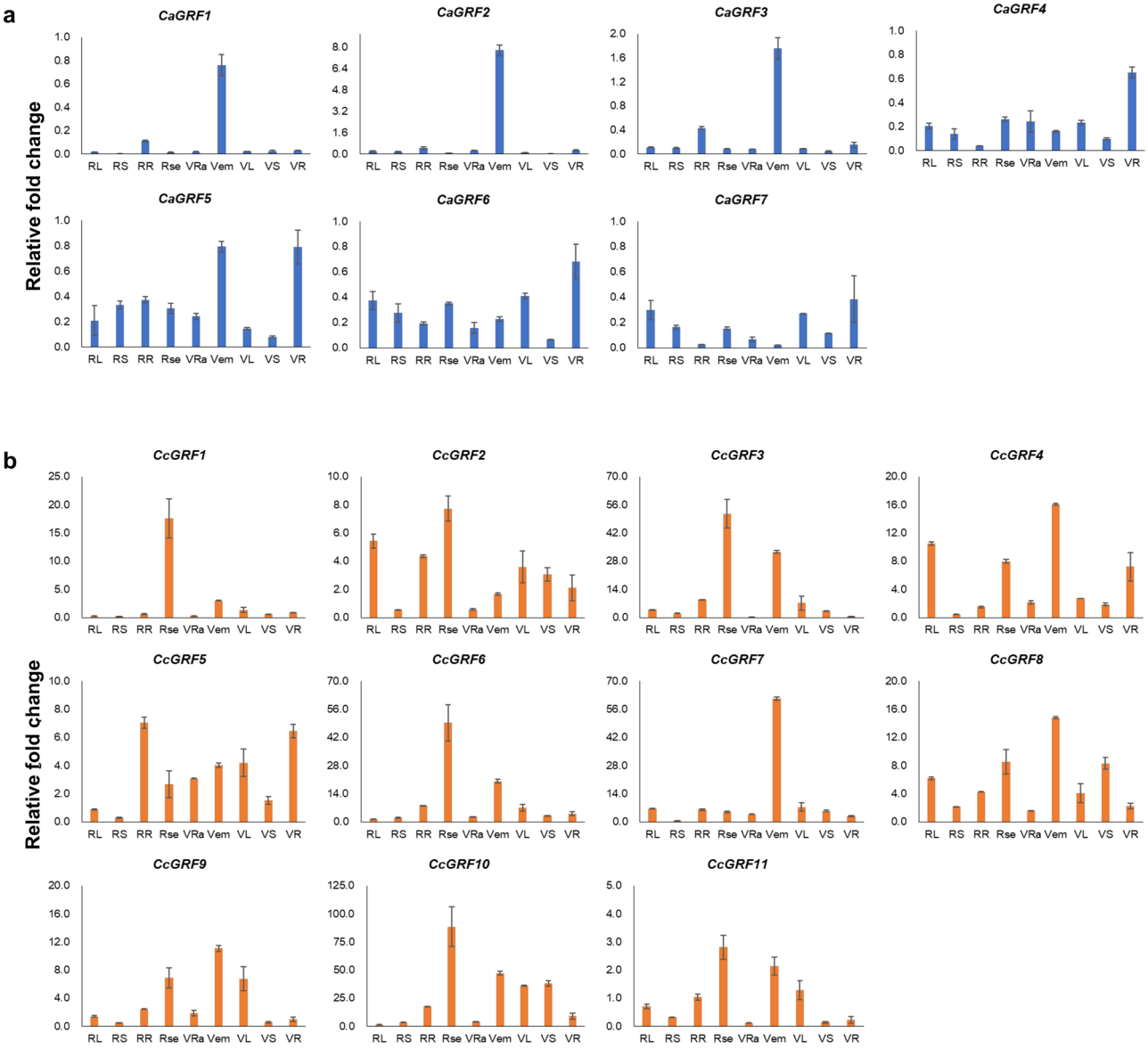

Figure 7. Functional validation of *GRF* genes in (**a**) chickpea and (**b**) pigeonpea by qRT-PCR. X-axis represents various considered tissues RL: Reproductive leaf; RS: Reproductive shoot; RR: Reproductive root; Rse: Reproductive mature seed VRa: Vegetative radicle; Vem: Vegetative embryo; VL: Vegetative leaf; VS: Vegetative shoot; VR: Vegetative root. Y-axis represents normalized relative fold change of gene expression in the respective tissues. Chickpea *UCP* and pigeonpea *GAPDH* gene transcript level were used for normalization and transcript abundance was expressed as a ratio relative to the house keeping gene (Mean ± SD, n = 3).

be associated mainly with lineage-specific whole genome duplication (WGD) events[51,52]. The *GRF* family is characterized by a conserved *QLQ* domain, which is responsible for DNA binding, and a *WRC* domain, which is involved in protein–protein interactions. The *QLQ* domain can interact with GIF to form a transcriptional activator and play a regulatory role[14]. We further identified four *GIF*s, each in chickpea and pigeonpea. The *GIF*s are characterized by a conserved SNH (SYT N-terminal homology or *SSXT*) domain involved in binding with the *QLQ* domain of *GRF* proteins (Kim et al. 2004).

Phylogenetic analysis showed that chickpea members were distributed in seven clades. In contrast, pigeonpea members were distributed in eight clades. Previously, for soybean[27], *Arabidopsis*[13], *M. sativa*[53] and cotton[54] distribution of members in different number of clades has been reported. These contrasting results suggest that the preservation of specific clades may vary among species with different evolutionary histories. Similarly, we observed three clades in chickpea and pigeonpea for GIFs, similar to *Arabidopsis*[13], tomato[55], and soybean[27]. Several genomic evolutionary models proposed in model species indicated that exon–intron structure is vital in understanding the relationships between evolutionary and functional differentiation[56–58]. Furthermore, exon or intron gain/loss events create gene structure divergence and functional differentiation[59]. However, the number and length of introns are not conserved in *GRF* genes from chickpea and pigeonpea, similar to *Arabidopsis* and soybean, even within the same clade. *GRF* genes are randomly distributed in the genome. The exon numbers

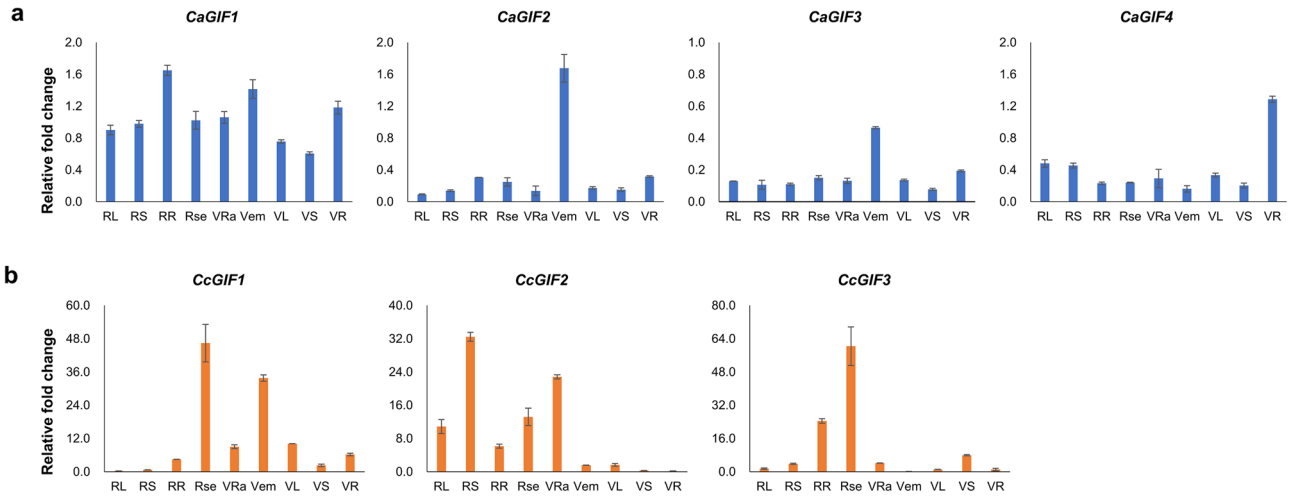

Figure 8. Functional validation of *GIF* genes in (**a**) chickpea and (**b**) pigeonpea by qRT-PCR. X-axis represents various considered tissues *RL* Reproductive leaf, *RS* Reproductive shoot, *RR* Reproductive root, *Rse* Reproductive mature seed, *VRa* Vegetative radicle, *Vem* Vegetative embryo, *VL* Vegetative leaf, *VS* Vegetative shoot, *VR* Vegetative root. Y-axis represents normalized relative fold change of gene expression in the respective tissues. Chickpea *UCP* and pigeonpea *GAPDH* gene transcript level was used for normalization and transcript abundance was expressed as a ratio relative to the house keeping gene (Mean ± SD, n = 3).

ranged from 3 to 5 for *CaGRFs* while the exon numbers ranged from 3 to 4 for *CcGRFs* and *CcGIFs*. However, the same exon number of 4 is present in *CaGIF* genes. Differences in the number and length of exons/introns can result from chromosome rearrangement and fusion. It may lead to different biological functions of these genes.

Previous studies have reported that the expansion of the *GRF* and *GIF* families mainly occurs through gene duplication, especially large-scale duplication, including whole genome or tandem duplication, to enhance plant adaptation to environmental changes[21]. *GRF* and *GIF* genes were distributed on eight chromosomes in pigeonpea and six chromosomes in chickpea, and no *GRF* family members were identified on chromosomes 07 and 09 in pigeonpea and chromosomes 03 and 05 in chickpea. Considering the WGD event in soybean[27], *GRF* and *GIF* genes followed a general trend in pigeonpea, where the gene number showed diploid-polyploid topology. The number of identified *GRF* genes (11), as well as *GIF* genes (4), are precisely half of soybean *GRF* (22) and *GIF* (8) genes. The trend is also evident for chickpea *GIF* genes (4 *GIF*s in chickpea compared to 8 *GIF*s in soybean). We also found that the calculated Ka/Ks ratio values were less than 1, indicating that the GRF and GIF protein family experienced purifying solid selection during evolution in chickpea and pigeonpea.

Cis-acting elements are specific DNA sequences located upstream of a gene, and their primary role is to regulate gene expression. Unlike coding sequences, *cis-acting* elements do not encode proteins; instead, they serve as binding sites for various regulatory factors and proteins, influencing the gene's transcription and expression[60]. The 5′ UTR analysis of the identified *GRF* and *GIF* genes in chickpea and pigeonpea suggested an abundance of various light response elements involved in the role of *GRF-GIF* transcriptional complexes in light-mediated response. Along with light-responsive elements, the upstream region of *GRF* and *GIF* genes displayed the presence of many *cis-regulatory* elements involved in hormonal responses such as responses to methyl-jasmonic acid (MeJA), abscisic acid (ABA) and jasmonic acid (JA) highlighting their involvement in hormonal signalling pathways. The 5′ UTR regions of these genes also contained elements responsive to biotic and abiotic stresses, suggesting that the expression of the *GRF-GIF* transcriptional complex can be subject to environmental cues[61].

Interaction of *miRNA396*, *GRF* and *GIF*, a plant development regulatory system

It has been established that various regulatory components regulate plant development, evolution and domestication. *miR396*, *GRF*, and *GIF* form one such regulatory system in plants regulating various aspects of plant development[62]. The last few amino acid sequences of the *WRC* domain of the GRF protein form a binding site for *miR396*, and binding of *miR396* to the mRNA of the GRF proteins results in its cleavage[47]. The *in-silico* analysis for the binding site of *miR396* in identified *GRF* in chickpea and pigeonpea showed that all identified *GRF* genes had *miR396* binding sites, unlike *Arabidopsis* and rice, which showed the presence of naturally occurring *GRF* genes with disrupted *miR396* binding sites[63]. The presence of *miR396* sites in all the identified *GRF* genes suggests that *miR396* might act as a negative regulator of all the *GRFs* in chickpea and pigeonpea and might have a role in modulating developmental processes.

Regulatory relationship between GRFs and GIFs in chickpea and pigeonpea

The regulatory relationship between GRFs and GIFs is complex and still being studied. However, it is known that GIFs interact with the *QLQ* domain of GRFs to form a transcriptional complex. This complex binds to target genes and regulates their expression, affecting plant growth and development[2,15,64]. Previous studies have shown that AtGIF1 interacts with six GRF proteins in *Arabidopsis*[47], while OsGIF1 interacts with three GRF proteins[65]. We observed that all the *CaGRF*s in the predicted interaction network interacted with CaGIF4 and 3 while all

CcGRF proteins in the predicted network interacted with *CcGIF1* and 4. In addition, it has been reported that several rice[66] and maize[67] GRF proteins are located downstream of the *GIF* gene, and increasing the expression of the *GIF* gene can increase the transcription level of the *GRF* gene. The specific functions and modes of action of the different GRFs and GIFs in chickpea and pigeonpea may require further research. By understanding how GRFs and GIFs interact to regulate gene expression, new strategies can be developed to improve plant growth and productivity.

*GRF*s and *GIF*s are involved in the growth and development of chickpea and pigeonpea

Previous studies have confirmed that *GRF*s and *GIF*s are expressed in different tissues throughout the plant growth and developmental cycle, such as vegetative and reproductive tissues[13,14,27,53]. It has been established that *GRF-GIF* transcription factor complexes regulate various aspects of plant growth and development[21,27,53]. To assess this, we studied the expression of identified *GRF* and *GIF* genes using transcriptome data and validated using qRT-PCR across various developmental stages in two leguminous plant species, chickpea and pigeonpea. Transcriptome data suggested the preferential expression of *GRF-GIF* genes in tissues like mature seeds, pods, and embryos, which was consistent with the qRT-PCR expression analysis suggesting the higher expression of few *CaGRF* genes in embryos and *CcGRF* genes in mature seeds. A similar trend was also followed for *GIF* genes, with *CaGIF* genes expressing the highest in embryos and *CcGIF* genes expressing in mature seeds. The co-expression of *GRF-GIF* genes in similar tissues supports the growth regulatory module in plants involving the interaction between GRF and GIF TFs.

In soybean, *GRF* genes such as *GmGRF6*, *GmGRF18*, and *GmGRF16* expressed the highest in the seed tissue[27]. In the case of chickpea, *CaGRF2* was highly expressed in reproductive seed tissue, which also belonged to the same clade as *GmGRF16*. Similarly, in the case of pigeonpea, transcriptome data suggested that *CcGRF10* expressed the highest in reproductive seed tissue, which belonged to the same clade as *GmGRF16* in the phylogenetic tree. Expression analysis using qRT-PCR suggested higher expression of eight *CcGRFs* genes (*CcGRF 1, 3,* and *6–11*) in seed tissue. The results were consistent with the expression patterns of *M. truncatula GRF* genes, which also exhibited higher expression in seeds and buds[68]. The consistency in the expression of these *GRF* genes in seed tissue hints towards their role in seed development.

Previous studies have also reported the strong expression of *GRF* genes in young tissue and relatively weaker expression in matured tissue across many plant species such as *Triticum aestivum*[69], *Jatropha curcas*[70] and *M. truncatula*[53] which was consistent with the observed results of lower expression of *CcGRF* and *CaGRF* genes across reproductive tissues as compared to vegetative tissues.

Along with *GRFs*, *GIFs* also play a crucial role in regulating seed development[26,65]. Consistent with these findings, qRT-PCR expression data of specific *CcGIFs* genes showed the preferential expression in reproductive seed tissue while specific *CaGIFs* expressed the highest in embryo tissue. Similarly, two members each from the identified *CaGIFs* viz. *CaGIF1* and *CaGIF2* as well as from *CcGIFs* viz. *CcGIF1* and *CcGIF2* showed preferential expression in reproductive seed tissue in transcriptome data, suggesting their involvement in seed development.

Conclusion

Our study identified seven *CaGRF*, eleven *CcGRF*, four *CaGIF*, and four *CcGIF* family members in the chickpea and pigeonpea genomes, and their essential characteristics and functions were subjected to preliminary analysis. *QLQ* and *WRC* are conserved domains unique to the GRF, and the *SSXT* domain is unique to the *GIF* gene family that helped us identify GRFs and GIFs in chickpea and pigeonpea. Phylogenetic analysis classified the GRFs into eight clades and GIFs into three clades. The synteny analysis revealed the collinearity between the orthologous genes of chickpea, pigeonpea, soyabean, Medicago and Arabidopsis revealing the evolutionary selection pressure in the *GRF-GIF* gene families. The upstream regions of the promoters of *GRF*s and *GIF*s contain one or more hormone or stress-related *cis-acting* elements. A significant variation was recorded in the expression profiles of *GRF-GIF* genes at the vegetative and reproductive stages of chickpea and pigeonpea growth and development. Identifying and characterizing *GRF* and *GIF* gene family members provides potential targets for further functional studies for the genetic improvement of chickpea and pigeonpea.

Data availability

The dataset supporting the findings of this article are included within the article.

Received: 19 January 2024; Accepted: 18 July 2024
Published online: 26 July 2024

References

1. Zhang, L., Fang, W., Chen, F. & Song, A. The role of transcription factors in the regulation of plant shoot branching. *Plants* **11**(15), 1997. https://doi.org/10.3390/plants11151997 (2022).
2. Debernardi, J. M. *et al.* A GRF–GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nat. Biotechnol.* **38**(11), 1274–1279. https://doi.org/10.1038/s41587-020-0703-0 (2020).
3. Liu, Y., Guo, P., Wang, J. & Xu, Z. Y. Growth-regulating factors: conserved and divergent roles in plant growth and development and potential value for crop improvement. *Plant J.* **113**(6), 1122–1145. https://doi.org/10.1111/tpj.16090 (2023).
4. Horiguchi, G., Kim, G. T. & Tsukaya, H. The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of *Arabidopsis thaliana*. *Plant J.* **43**(1), 68–78. https://doi.org/10.1111/j.1365-313X.2005.02429.x (2005).
5. Kim, J. H. & Lee, B. H. GROWTH-REGULATING FACTOR4 of Arabidopsis thaliana is required for development of leaves, cotyledons, and shoot apical meristem. *J. Plant Biol.* **49**, 463–468. https://doi.org/10.1007/BF03031127 (2006).

6. Zhang, B. *et al.* Identification of GROWTH-REGULATING FACTOR transcription factors in lettuce (Lactuca sativa) genome and functional analysis of LsaGRF5 in leaf size regulation. *BMC Plant Biol.* **21**, 1–13. https://doi.org/10.1186/s12870-021-03261-6 (2021).
7. Bao, M. *et al.* miR396a-mediated basic helix–loop–helix transcription factor bHLH74 repression acts as a regulator for root growth in Arabidopsis seedlings. *Plant Cell Physiol.* **55**(7), 1343–1353. https://doi.org/10.1093/pcp/pcu058 (2014).
8. Lee, S.-J. *et al.* GROWTH-REGULATING FACTOR and GRF-INTERACTING FACTOR specify meristematic cells of gynoecia and anthers. *Plant Physiol.* **176**(1), 717–729. https://doi.org/10.1104/pp.17.00960 (2018).
9. Liu, H. *et al.* OsmiR396d-regulated OsGRFs function in floral organogenesis in rice through binding to their targets OsJMJ706 and OsCR4. *Plant Physiol.* **165**(1), 160–174. https://doi.org/10.1104/pp.114.235564 (2014).
10. Liu, J. *et al.* The BnGRF2 gene (GRF2-like gene from Brassica napus) enhances seed oil production through regulating cell number and plant photosynthesis. *J. Exp. Bot.* **63**(10), 3727–3740. https://doi.org/10.1093/jxb/ers066 (2012).
11. Kim, J.-S. *et al.* Arabidopsis GROWTH-REGULATING FACTOR7 functions as a transcriptional repressor of abscisic acid–and osmotic stress–responsive genes, including DREB2A. *Plant Cell* **24**(8), 3393–3405. https://doi.org/10.1105/tpc.112.100933 (2012).
12. Liu, J., Rice, J. H., Chen, N., Baum, T. J. & Hewezi, T. Synchronization of developmental processes and defense signaling by growth regulating transcription factors. *PloS One* **9**(5), e98477. https://doi.org/10.1371/journal.pone.0098477 (2014).
13. Kim, J. H., Choi, D. & Kende, H. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. *Plant J.* **36**(1), 94–104. https://doi.org/10.1046/j.1365-313X.2003.01862.x (2003).
14. Choi, D., Kim, J. H. & Kende, H. Whole genome analysis of the OsGRF gene family encoding plant-specific putative transcription activators in rice (*Oryza sativa* L.). *Plant Cell Physiol.* **45**(7), 897–904. https://doi.org/10.1093/pcp/pch098 (2004).
15. Wang, P. *et al.* Whole-genome identification and expression profiling of growth-regulating factor (GRF) and GRF-interacting factor (GIF) gene families in Panax ginseng. *BMC Genom.* **24**(1), 1–15. https://doi.org/10.1186/s12864-023-09435-w (2023).
16. Kim, J. H. & Kende, H. A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in Arabidopsis. *Proc. Natl. Acad. Sci.* **101**(36), 13374–13379. https://doi.org/10.1073/pnas.040545010 (2004).
17. Lee, B. H. *et al.* The Arabidopsis GRF-INTERACTING FACTOR gene family performs an overlapping function in determining organ size as well as multiple developmental properties. *Plant Physiol.* **151**(2), 655–668. https://doi.org/10.1104/pp.109.141838 (2009).
18. Lee, B. H. *et al.* The Arabidopsis thaliana GRF-INTERACTING FACTOR gene family plays an essential role in control of male and female reproductive development. *Dev. Biol.* **386**(1), 12–24. https://doi.org/10.1016/j.ydbio.2013.12.009 (2014).
19. Kanei, M., Horiguchi, G. & Tsukaya, H. Stable establishment of cotyledon identity during embryogenesis in Arabidopsis by ANGUSTIFOLIA3 and HANABA TARANU. *Development* **139**(13), 2436–2446. https://doi.org/10.1242/dev.081547 (2012).
20. Horiguchi, G. *et al.* ANGUSTIFOLIA3 plays roles in adaxial/abaxial patterning and growth in leaf morphogenesis. *Plant Cell Physiol.* **52**(1), 112–124. https://doi.org/10.1093/pcp/pcq178 (2011).
21. Omidbakhshfard, M. A., Proost, S., Fujikura, U. & Mueller-Roeber, B. Growth-regulating factors (GRFs): A small transcription factor family with important functions in plant biology. *Mol. Plant* **8**(7), 998–1010. https://doi.org/10.1016/j.molp.2015.01.013 (2015).
22. Feng, Q. *et al.* Highly efficient, genotype-independent transformation and gene editing in watermelon (*Citrullus lanatus*) using a chimeric ClGRF4-GIF1 gene. *J. Integr. Plant Biol.* **63**(12), 2038–2042. https://doi.org/10.1111/jipb.13199 (2021).
23. Bull, T. *et al.* GRF–GIF chimeric proteins enhance in vitro regeneration and Agrobacterium-mediated transformation efficiencies of lettuce (Lactuca spp.). *Plant Cell Rep.* **42**(3), 629–643. https://doi.org/10.1007/s00299-023-02980-4 (2023).
24. Henchion, M., Hayes, M., Mullen, A. M., Fenelon, M. & Tiwari, B. Future protein supply and demand: Strategies and factors influencing a sustainable equilibrium. *Foods* **6**(7), 53. https://doi.org/10.3390/foods6070053 (2017).
25. Pratap, A. *et al.* Potential, constraints and applications of in vitro methods in improving grain legumes. *Plant Breed.* **137**(3), 235–249. https://doi.org/10.1111/pbr.12590 (2018).
26. Alam, I., Wu, X. & Ge, L. Comprehensive genomic survey, evolution, and expression analysis of GIF gene family during the development and metal ion stress responses in soybean. *Plants* **11**(4), 570. https://doi.org/10.3390/plants11040570 (2022).
27. Chen, F. *et al.* Genome-wide identification of GRF transcription factors in soybean and expression analysis of GmGRF family under shade stress. *BMC Plant Biol.* **19**(1), 1–13. https://doi.org/10.1186/s12870-019-1861-4 (2019).
28. Cao, Y., Han, Y., Jin, Q., Lin, Y. & Cai, Y. Comparative genomic analysis of the GRF genes in Chinese pear (Pyrus bretschneideri Rehd), poplar (Populous), grape (*Vitis vinifera*), Arabidopsis and rice (*Oryza sativa*). *Front. Plant Sci.* **7**, 1750. https://doi.org/10.3389/fpls.2016.01750 (2016).
29. Li, H. *et al.* Genome-wide analysis of the growth-regulating factor family in *Medicago truncatula*. *J. Plant Growth Regul.* **42**(4), 2305–2316. https://doi.org/10.1007/s00344-022-10704-3 (2023).
30. Chen, C. *et al.* TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **13**(8), 1194–1202. https://doi.org/10.1016/j.molp.2020.06.009 (2020).
31. Gasteiger, E. *et al.* Protein identification and analysis tools on the Expasy server. In *The Proteomics Protocols Handbook* (ed. Walker, J.) 571–607 (Springer, 2005).
32. Horton, P. *et al.* WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* **35**(2), W585–W587. https://doi.org/10.1093/nar/gkm259 (2007).
33. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**(W1), W293–W296. https://doi.org/10.1093/nar/gkab301 (2021).
34. Chao, J. *et al.* MG2C: A user-friendly online tool for drawing genetic maps. *Mol. Hortic.* **1**(1), 1–4. https://doi.org/10.1186/s43897-021-00020-x (2021).
35. Hu, B. *et al.* GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **31**(8), 1296–1297. https://doi.org/10.1093/bioinformatics/btu817 (2015).
36. 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.* **30**(1), 325–327. https://doi.org/10.1093/nar/30.1.325 (2002).
37. Dai, X., Zhuang, Z. & Zhao, P. X. psRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* **46**(W1), W49–W54. https://doi.org/10.1093/nar/gky316 (2018).
38. Szklarczyk, D. *et al.* STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**(D1), D447–D452. https://doi.org/10.1093/nar/gku1003 (2015).
39. Varshney, R. K. *et al.* Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. *Nat. Biotechnol.* **31**(3), 240–246. https://doi.org/10.1038/nbt.2491 (2013).
40. Varshney, R. K. *et al.* Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nature biotechnology* **30**(1), 83. https://doi.org/10.1038/nbt.2022 (2012).
41. Kudapa, H., Garg, V., Chitikineni, A. & Varshney, R. K. The RNA-Seq-based high resolution gene expression atlas of chickpea (*Cicer arietinum* L.) reveals dynamic spatio-temporal changes associated with growth and development. *Plant Cell Environ.* **41**(9), 2209–2225. https://doi.org/10.1111/pce.13210 (2018).
42. Pazhamala, L. T. *et al.* Gene expression atlas of pigeonpea and its application to gain insights into genes associated with pollen fertility implicated in seed formation. *J. Exp. Bot.* **68**(8), 2037–2054. https://doi.org/10.1093/jxb/erx010 (2017).
43. Sinha, P. *et al.* Evaluation and validation of housekeeping genes as reference for gene expression studies in pigeonpea (*Cajanus cajan*) under drought stress conditions. *PloS One* **10**(4), e0122847. https://doi.org/10.1371/journal.pone.0122847 (2015).

44. Reddy, D. S. *et al.* Identification and validation of reference genes and their impact on normalized gene expression studies across cultivated and wild cicer species. *PloS One* **11**(2), e0148451. https://doi.org/10.1371/journal.pone.0148451 (2016).
45. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **6**, 1101–8. https://doi.org/10.1038/nprot.2008.73 (2008).
46. Liu, D., Song, Y., Chen, Z. & Yu, D. Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in Arabidopsis. *Physiologia Plantarum* **136**(2), 223–236. https://doi.org/10.1111/j.1399-3054.2009.01229.x (2009).
47. Debernardi, J. M. *et al.* Post-transcriptional control of GRF transcription factors by micro RNA miR396 and GIF co-activator affects leaf size and longevity. *Plant J.* **79**(3), 413–426. https://doi.org/10.1111/tpj.12567 (2014).
48. Wang, W. *et al.* Comprehensive evolutionary analysis of growth-regulating factor gene family revealing the potential molecular basis under multiple hormonal stress in Gramineae crops. *Front. Plant Sci.* **14**, 1174955. https://doi.org/10.3389/fpls.2023.1222765 (2023).
49. Kim, Y., Takahashi, S. & Miyao, M. Relationship between reduction in rice (Nipponbare) leaf blade size under elevated CO2 and miR396–GRF module. *Plant Signal. Behav.* **17**(1), 2041280. https://doi.org/10.1080/15592324.2022.2041280 (2022).
50. Lu, Y. *et al.* Coordination between GROWTH-REGULATING FACTOR1 and GRF-INTERACTING FACTOR1 plays a key role in regulating leaf growth in rice. *BMC Plant Biol.* **20**(1), 1–12. https://doi.org/10.1186/s12870-020-02417-0 (2020).
51. Schmutz, J. *et al.* Genome sequence of the palaeopolyploid soybean. *Nature* **463**(7278), 178–183. https://doi.org/10.1038/nature08670 (2010).
52. Young, N. D. *et al.* The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature* **480**(7378), 520–524 (2011).
53. Sun, Y. *et al.* Genome-wide identification of growth-regulating factor transcription factor family related to leaf and stem development in alfalfa. *Front. Plant Sci.* **13**, 964604. https://doi.org/10.3389/fpls.2022.964604 (2022).
54. Cao, J.-F. *et al.* Genome-wide characterization of the GRF family and their roles in response to salt stress in Gossypium. *BMC Genom.* **21**, 1–16. https://doi.org/10.1186/s12864-020-06986-0 (2020).
55. Ai, G. *et al.* Genome-wide identification and molecular characterization of the growth-regulating factors-interacting factor gene family in tomato. *Genes* **11**(12), 1435. https://doi.org/10.3390/genes11121435 (2020).
56. Cao, D. *et al.* Regulations on growth and development in tomato cotyledon, flower and fruit via destruction of miR396 with short tandem target mimic. *Plant Sci.* **247**, 1–12. https://doi.org/10.1016/j.plantsci.2016.02.012 (2016).
57. Hurley, I., Hale, M. & Prince, V. Duplication events and the evolution of segmental identity. *Evol. Dev.* **7**(6), 556–567. https://doi.org/10.1111/j.1525-142X.2005.05059.x (2005).
58. Moore, R. C. & Purugganan, M. D. The early stages of duplicate gene evolution. *Proc. Natl. Acad. Sci.* **100**(26), 15682–15687. https://doi.org/10.1073/pnas.2535513100 (2003).
59. Xu, G., Guo, C., Shan, H. & Kong, H. Divergence of duplicate genes in exon–intron structure. *Proc. Natl. Acad. Sci.* **109**(4), 1187–1192. https://doi.org/10.1073/pnas.1109047109 (2012).
60. Hernandez-Garcia, C. M. & Finer, J. J. Identification and validation of promoters and cis-acting regulatory elements. *Plant Sci.* **217**, 109–119. https://doi.org/10.1016/j.plantsci.2013.12.007 (2014).
61. Kim, J. H. Biological roles and an evolutionary sketch of the GRF-GIF transcriptional complex in plants. *BMB Rep.* **52**(4), 227. https://doi.org/10.5483/BMBRep.2019.52.4.051 (2019).
62. Liebsch, D. & Palatnik, J. F. MicroRNA miR396, GRF transcription factors and GIF co-regulators: A conserved plant growth regulatory module with potential for breeding and biotechnology. *Curr. Opin. Plant Biol.* **53**, 31–42. https://doi.org/10.1016/j.pbi.2019.09.008 (2020).
63. Beltramino, M. *et al.* Robust increase of leaf size by Arabidopsis thaliana GRF3-like transcription factors under different growth conditions. *Sci. Rep.* **8**(1), 13447. https://doi.org/10.1038/s41598-018-29859-9 (2018).
64. Ma, C. *et al.* PeGRF6-PeGIF1 complex regulates cell proliferation in the leaf of Phalaenopsis equestris. *Plant Physiol. Biochem.* **196**, 683–694. https://doi.org/10.1016/j.plaphy.2023.02.026 (2023).
65. Li, S. *et al.* The OsmiR396c-OsGRF4-OsGIF1 regulatory module determines grain size and yield in rice. *Plant Biotechnol. J.* **14**(11), 2134–2146. https://doi.org/10.1111/pbi.12569 (2016).
66. Li, S. *et al.* Modulating plant growth–metabolism coordination for sustainable agriculture. *Nature* **560**(7720), 595–600. https://doi.org/10.1038/s41586-018-0415-5 (2018).
67. Nelissen, H. *et al.* Dynamic changes in ANGUSTIFOLIA3 complex composition reveal a growth regulatory mechanism in the maize leaf. *Plant Cell* **27**(6), 1605–1619. https://doi.org/10.1105/tpc.15.00269 (2015).
68. Du, W. *et al.* Genome-wide identification and characterization of growth regulatory factor family genes in Medicago. *Int. J. Mol. Sci.* **23**(13), 6905. https://doi.org/10.3390/ijms23136905 (2022).
69. Zan, T., Zhang, L., Xie, T. & Li, L. Genome-wide identification and analysis of the growth-regulating factor (GRF) gene family and GRF-interacting factor family in *Triticum aestivum* L.. *Biochem. Genet.* **58**, 705–724. https://doi.org/10.1038/nature10625 (2020).
70. Tang, Y. *et al.* Genome-wide identification and expression analysis of the growth regulating factor (GRF) family in Jatropha curcas. *PLoS One* **16**(7), e0254711. https://doi.org/10.1371/journal.pone.0254711 (2021).

Acknowledgements

MK acknowledges Junior Research Fellowship (DBT/2022-23/ICRISAT/2065) support from the Department of Biotechnology, Government of India for his doctoral research work. This work is supported by Novo Nordisk Foundation (Grant number- NNF23SA0084140). The funding agency had no role in experimental design, data collection and analyses, the decision to publish, or the preparation of the manuscript. All the authors have read and approved the manuscript.

Author contributions

Conceptualization: WT; methodology: KY, PSR, HK; software, MK, TA, VKV; data generation and analysis: TA, MK, VKV; writing—original draft preparation, MK, KY; writing—review and editing: MK, KY, HK, WT; supervision and funding acquisition: WT; All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-68033-2.

Correspondence and requests for materials should be addressed to W.T.