

Genome-wide identification and expression pattern analysis of the GRF transcription factor family in *Astragalus mongholicus*

panpan Wang

Heilongjiang University of Chinese Medicine

Zhen Wang

Heilongjiang University of Chinese Medicine

Huiyan Cao

Heilongjiang University of Chinese Medicine

Jiajun He

Heilongjiang University of Chinese Medicine

Chen Qin

Heilongjiang University of Chinese Medicine

Lianqing He

Heilongjiang University of Chinese Medicine

Bo Liu

Harbin Cambridge University

Jiamei Wang

Heilongjiang University of Chinese Medicine

Lingyang Kong

Heilongjiang University of Chinese Medicine

Weichao Ren

Heilongjiang University of Chinese Medicine

Xiubo Liu

Heilongjiang University of Chinese Medicine

Wei Ma

mawei@hljucm.net

Heilongjiang University of Chinese Medicine

Research Article

Keywords: Astragalus mongholicus, Growth-regulating factor, Expression pattern, Cis-acting elements, Evolutionary analysis

Posted Date: February 12th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-3940307/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Molecular Biology Reports on May 6th, 2024. See the published version at <https://doi.org/10.1007/s11033-024-09581-8>.

Abstract

Background *Astragalus membranaceus* is a plant of the *Astragalus* genus, which is used as a traditional Chinese herbal medicine with extremely high medicinal and edible value. *Astragalus mongholicus*, as one of the representative medicinal materials with the same origin of medicine and food, has a rising market demand for its raw materials, but the quality is different in different production areas. Growth-regulating factors are transcription factors unique to plants that play important roles in plant growth and development. Up to now, there is no report about GRF in *A. mongholicus*.

Methods and results This study conducted a genome-wide analysis of the *AmGRF* gene family, identifying a total of nine *AmGRF* genes that were classified into subfamily V based on phylogenetic relationships. In the promoter region of the *AmGRF* gene, we successfully predicted *cis*-elements that respond to abiotic stress, growth, development, and hormone production in plants. Based on transcriptomic data and real-time quantitative polymerase chain reaction (qPCR) validation, the results showed that *AmGRFs* were expressed in the roots, stems, and leaves, with overall higher expression in leaves, higher expression of *AmGRF1* and *AmGRF8* in roots, and high expression levels of *AmGRF1* and *AmGRF9* in stems.

Conclusions The results of this study provide a theoretical basis for the further exploration of the functions of *AmGRFs* in plant growth and development.

Introduction

Growth regulating factor (*GRF*) is a transcription factor that exists in plants and plays a crucial role in the growth and development process of plants (Lee et al., 2018). The first *GRF* were discovered in rice, and numerous *GRFs* have since been successfully isolated and identified. These genes have played important roles in the evolution of plants. Nine genes have been identified in *Arabidopsis thaliana* (Kim et al., 2003), 11 in *Oryza sativa* (Choi et al., 2004), 25 in *Nicotiana benthamiana* (Zhang et al., 2018), and eight in *Sorghum bicolor* (Shi et al., 2022). *GRF* has been confirmed to be involved in plant growth and development. The expression of the *GRF* gene is higher in actively growing tissues, such as plant root tips, young shoots, and leaves; however, its expression is relatively low in mature tissues or organs. *GRF* determines cell volume, positively regulates leaf size, and participates in chloroplast proliferation, photosynthesis, and osmotic stress (Lu et al., 2023; Wang et al., 2023b). Therefore, it is important to explore the function of *GRFs* in plant growth and genetic improvement (Choi et al., 2004).

GRF contains specific QLQ (Gln, Leu, and Gln) and WRC (Trp, Arg, and Cys) conserved structural domains. Both conserved domains are present in the N-terminal region of the protein sequence. QLQ promotes the interaction between *GRF* and *GRF*-interacting factor (GIF), thereby forming *GRF*-GIF complexes. This complex is referred to in the literature as a transcriptional co-activator (Kim, 2019). The QLQ domain is highly similar to the N-terminus of the SWI2/SNF2 protein in yeast, and QLQ binds to SNF11 to form chromatin-remodeling complexes (Fonini et al., 2020). WRC can participate in the DNA-binding process of DNA, consists of a specific zinc finger structure and a nuclear localization signal region (NLS) (Meng et al., 2022). In *AtGRF9* of *A. thaliana* and *BrGRF12* of *Brassica campestris*, a second WRC domain exists at the C-terminus (van der Knaap et al., 2000). The evolutionary differences in *GRF* between various family members and species in plants, as well as the diversity of the C-terminal domains of different *GRF* proteins, indicate that their biological functions may be diverse or highly specific.

Astragalus mongholicus is a plant of the genus *Astragalus* in the family Leguminosae, which has high medicinal and edible properties. It has been recorded as early as Shennong Ben Cao Jing for more than 2,000 years and exhibits anti-inflammatory, antiviral, antitumor, and other pharmacological effects (Zhang et al., 2020). The primary chemical components of *A. mongholicus* are polysaccharides, saponins, and flavonoids, and more than 200 compounds have been isolated from *Astragalus membranaceus* (Dong et al., 2023). Astragalus injection is clinically used to treat viral myocarditis (Guo et al., 2022), and astragalus polysaccharide can be used to treat severe acute respiratory syndrome coronavirus-2 (SARS-CoV2) infection (Yeh et al., 2021). There have been several reports on the analysis of the *GRF* gene family; however, *A. mongholicus* has not yet been studied (Lee et al., 2018)..

Currently, regarding *A. mongholicus* genome has been published, and this study focuses on *A. mongholicus GRFs* gene family was subjected to whole-genome identification and analysis, and nine *AmGRFs* genes were identified. The physicochemical properties, chromosomal position, motif composition, evolutionary relationships, and genomic collinearity of *AmGRFs* were studied, and the expression of *AmGRFs* in different varieties, tissues, and leaves was determined using real-time quantitative reverse transcription-polymerase chain reaction (qRT-PCR). These results provide insights into the roles of *GRF* family members in *A. mongholicus* in growth and development provides new insights.

Materials And Methods

2.1 Plant material:

The seeds of *A. mongholicus* used in this study were identified by Ma Wei of the Heilongjiang University of Chinese Medicine. The *A. mongholicus* samples were grown in a light incubator at the laboratory of Heilongjiang University of Chinese Medicine with a photoperiod of 14 h at a temperature cycle of 24 °C (light) and 10 h at a temperature of 22 °C (dark). After 40 d of *A. mongholicus*, the roots, stems, and leaves were collected as samples. Samples for each treatment were obtained from three individual plants with average growth. All samples were frozen in liquid nitrogen immediately after harvesting and stored at 80 °C for qRT-PCR analysis.

2.2 Identification of AmGRF gene family

The genome and gene annotation files of *A. mongholicus* used in this study were obtained from the Global Pharmacopoeia Genome Database (GPGD) (<http://www.gpgenome.com/species/109>) (Liao et al., 2022). We extracted the CDS sequences of these data using TBtools (Chen et al., 2023) and converted them into protein sequences, which were submitted to the Plant Transcription Factor Database (PlantTFDB) (<http://planttfdb.gao-lab.Org/index.php>) for transcription factor prediction. Candidate sequences or conserved structural domains were further screened using the CD-Search tool (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) and the Pfam databases (<https://ngdc.cncb.ac.cn/>). Redundant and incomplete sequences were removed and members of the *AmGRFs* gene family were identified. The physicochemical properties of the proteins were predicted using the ExPASy website (<http://web.expasy.org/protparam/>), including the isoelectric point (PI), molecular weight (MW), grand average of hydropathicity (GRAVY), and chromosome position (CP) of the *AmGRFs* proteins. We also predicted the subcellular localization of *AmGRF* proteins from the website (<http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/>).

2.3 Phylogenetic analysis of the AmGRF gene family

To obtain an understanding of the phylogenetic relationships of the *AmGRF* gene family, we merged the nine GRF proteins of *A. thaliana* and 12 GRF proteins of *O. sativa* with the nine protein sequences of AmGRF (Choi et al., 2004). Next, we used the Muscle program for multiple sequence alignment and constructed an evolutionary tree of the *AmGRF* gene family system using the MEGA-X11 software (Kumar et al., 2016). Using the adjacency method, we set the bootstrap value to 1000 and identified it through the Evolview website ([www://evolgenius.info/evolview-v3/](http://www.evolgenius.info/evolview-v3/)).

2.4 Analysis of *AmGRF* gene structure and motifs

The AmGRF sequence was submitted to the MEME (<https://www.example.com/tools/meme>), the conserved motifs present in *AmGRF* were predicted, and a position distribution with up to one occurrence per sequence was selected. The conservative number of motifs was set to ten, with other default values. We utilized the Gene Structure View (Advanced) feature in TBtools software and combined it with the phylogenetic tree to visualize the annotation file of *A. mongholicus*. These files include coding and non-coding regions in the gene structure as well as internal exons and conserved structural domains.

2.5 Chromosomal localization and collinearity analysis of the *AmGRF* gene family

The information required for chromosome localization of the *AmGRF* gene family, including the start and end positions, length, and number of chromosomes, was obtained using TBtools software in *A. mongholicus* genome annotation file. The chromosomal distribution map of the *AmGRF* gene family was visualized using the TBtools software Advance Circle. It was named based on the order of the chromosomes. From the National Biotechnology Information Center (<https://www.example.com>) The genome sequences of *Solanum lycopersicum* (GCF_000188115.5) *Codonopsis lanceolata* (GCA_013146195.2) *Cannabis sativa* (GCF_029168945.1) *Vitis vinifera* (GCF_030704535.1), and *O. sativa* (GCF001433935.1) were downloaded. *A. thaliana* and *O. sativa* were downloaded from the National Center for Biotechnology Information (NCBI) in the United States (<https://www.ncbi.nlm.nih.gov>). Visualization (default parameters) was achieved using the TBtools software.

2.6 Analysis of cis-elements and GO analysis of the *AmGRF* promoter

We extracted 2000 bp upstream of the start codon (ATG) of the *AmGRF* gene sequence as the promoter region and submitted it to PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Cis-acting elements were predicted on the website, default parameters were set, and the Plantcare analysis results was simplified by including the stress, hormone, and growth-related categories. TBtools software was used for visualization and the evolview website was used for beautification (<http://evolgenius.info/evolview-v3/>). The AmGRF protein sequence was submitted to the David database (<http://david.ncifcrf.gov/>), Gene Ontology (GO) analysis was performed, and the GO analysis results were visualized through the website (<http://www.bioinformatics.com.cn/>).

2.7 Analysis of *AmGRF* gene expression patterns and qRT-PCR

Transcriptome sequencing and qRT-PCR validation of *AmGRF* gene expression profiles were performed using an RNA extraction kit (Simgen Biotechnology Co. Ltd. Hangzhou, China). RNA was extracted from the roots, stems, and leaves of 50-d-old *A. mongholicus* hydroponic seedlings and measured using A Nanodrop 1000 ultraviolet (UV) spectrophotometer. qRT-PCR was performed to verify the accuracy of transcriptome data. ReverTra Ace™ qPCR RT Master Mix TOYOBO with gDNA Remover (Orient Biosciences Co., LTD, Shanghai, China) was used for reverse

transcription of the first-strand cDNA strand, and this cDNA was used as a reference template. qRT-PCR experiments were performed using the ChamQ Universal SYBR qPCR Master Mix (Yeastar Biotechnology, Shanghai, China). RT-PCR was performed using an AriaMx Real-Time Quantitative Polymerase Chain Reaction System (Agilent Technologies). The procedure was described as follows: 95 °C was applied for 30 s, followed by 95 °C for 10 s and 60 °C for 30 s. This process was divided into 40 stages. We performed three biological replicates for each sample and used the $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen, 2001) method to estimate relative gene expression. Primer3web software (<https://primer3.ut.ee/>) was used design specific primers for qRT-PCR, and these primer sequences can be found in Supplementary Table 1.

Results

3.1 Identification and physicochemical property analysis of *AmGRF* gene

All protein sequences in *A. mongholicus* genome were extracted and submitted to the PlantTFDB website to obtain candidate transcription factor sequences. After further retrieval and elimination of redundant sequences in *A. mongholicus* genome, nine *AmGRF* genes were identified and named *AmGRF1-AmGRF9* (Supplementary Table 2). Physicochemical analysis showed that the sequence length of all *AmGRF* proteins ranged 320 (*AmGRF4*) to 642 (*AmGRF2*) amino acids, and the molecular mass ranged 36.1–69.9 kDa. The isoelectric points ranged 6.74 (*AmGRF8*) to 9.16 (*AmGRF4*). The mean value of *AmGRF* hydrophilicity was < 0, indicating that *AmGRF* is a hydrophobic protein (Table1). Subcellular localization revealed that *AmGRF* proteins were located in the nucleus.

Table 1. Physicochemical properties of the *AmGRF* gene

Gene ID	Rename	pI	MW	AA	GRAVY	Subcellular Localization	Chromosome position
Am01G001010.1	AmGRF1	8.66	37790.22	330	-1.092	Nucleus	3920824–3922818
Am02G003770.1	AmGRF2	7.77	69908.11	642	-0.690	Nucleus	14766959,14771445
Am03G012480.1	AmGRF3	7.00	55486.16	512	-0.751	Nucleus	33675371,33677913
Am04G009960.1	AmGRF4	9.16	36089.21	320	-0.839	Nucleus	28771317,28774332
Am05G010550.1	AmGRF5	7.72	36018.75	325	-0.759	Nucleus	83747571,83751451
Am05G010590.1	AmGRF6	7.72	36824.78	332	-0.720	Nucleus	84151650,84155421
Am06G015050.1	AmGRF7	8.67	40208.41	353	-0.906	Nucleus	54343425,54346356
Am06G023050.1	AmGRF8	6.74	41999.23	368	-0.941	Nucleus	137513488,137515282
Am09G017960.1	AmGRF9	8.34	39125.40	357	-0.727	Nucleus	64370726,64372998

3.2 Phylogenetic analysis of the *AmGRF* gene family

To further understand the evolutionary relationships among members of the *AmGRF* gene family, MEGA-X11 software was used to analyze *A. mongholicus*, *A. thaliana*, and *O. sativa* GRF were compared and analyzed for a total of 29 protein sequences, and a neighbor-joining (NJ) phylogenetic tree was constructed (Fig.1). The results showed that the nine identified *AmGRF* genes were divided into five of six subfamilies (I–VI). Group I contained *AmGRF1*, *AmGRF7*, and *AmGRF8*; Group II contained *AmGRF4*, *AmGRF5*, and *AmGRF6*; Group III contained

AmGRF2; Group IV contained *AmGRF3*; and Group V contained *AmGRF6*. However, no *AmGRF* gene was classified in Group VI. Among these, *AtGRF* and *AmGRF*, such as *AtGRF7*, *AmGRF3*, *AtGRF5* and *AmGRF1*.

Fig.1 NJ phylogenetic tree constructed with *A. thaliana*, *O. sativa*, and *AmGRF* proteins. The bootstrap value was set to 1000 replicates. The red star represents the *AmGRF* protein, the blue triangle represents the *OsGRF* protein, and the green circle represents the *AtGRF* protein. Different colors represent the various subfamilies.

3.3 *AmGRF* gene structure, domain, and motifs

To further understand the diversity and similarity of *AmGRF*, the CD Search tool and Pfam data were used to screen the conserved domains of the candidate sequences. Based on the QLQ (PF08880) and WRC (PF08879) conserved domains of the *GRF* gene, we analyzed the *AmGRF* protein using the MEME tool and identified ten motifs (Fig.3B). Among these, nine *AmGRF* proteins containing motif1, motif2, motif1, and motif2 were QLQ and WRC domains, respectively (Fig.3C), with lengths of 45 and 41 amino acids present in the *GRF* protein, and were highly conserved in the *AmGRF* protein sequence (Fig.2A&B). The *GRF* protein contains 2–9 different conserved motifs, and members grouped in the same branch have similar conserved motifs, such as *AmGRF7/AmGRF8* and *AmGRF4/AmGRF5*. In subfamilies I, II, and III containing motif4 and motif6, only *AmGRF6* in subfamily II exhibited the phenomenon of motif6 preceding motif4. Simultaneously, motif3 was missing during the evolutionary process in *AmGRF3* and *AmGRF6*. *AmGRF4* and *AmGRF5* have the highest number of conserved motifs. *AmGRF3* contains the least number of conserved motifs, with only two conserved motifs, motif1 and motif2. Among the three subfamilies, motif1, motif2, motif3, motif4, and motif6 are common. We further analyzed the distribution of introns and exons in the *AmGRF* gene family and found that the number of introns and exons in all *AmGRF* is approximately 2–3 (Fig.3D). The *AmGRF* protein contains UTR and CDS.

Fig.2 Conservative sequence alignment of *AmGRF* gene family proteins. A: QLQ conservative domain. B: WRC conserved domain.

Fig.3 f the *AmGRF* protein. (A) Representation of the *AmGRF* protein phylogenetic tree. Green, pink, yellow, light blue, and dark blue backgrounds represent groups I, II, III, IV, and V, respectively. (B) The MEME website was used to identify the composition of *AmGRF* protein motifs, with different colors representing the various motifs. (C) The *AmGRF* gene has a conserved domain, with green representing the WRC domain and yellow representing the QLQ domain. (D) *AmGRF* gene structure. CDS: encoding sequence, UTR: In the untranslated region; lines represent introns. (E) The conservative structural domains motif1 and motif2 of *AmGRFs*.

3.4 Chromosome location, duplication, and syntenic analysis of *AmGRF* gene family

The distribution of *AmGRF* genes on the chromosomes was visualized using gene annotation (Fig.4). Nine *AmGRF* genes were randomly distributed across seven chromosomes. *AmGRF* was primarily distributed at both ends of the chromosome, with less distribution in the middle. The *AmGRF* gene has up to two genes distributed on the fifth chromosome. There was no chromosomal distribution on the seventh and eighth chromosomes, whereas there was gene distribution on all the other chromosomes (Fig.5A). Three pairs of gene replicates were identified on chromosomes one and six, namely *AmGRF1/AmGRF8*, *AmGRF1/AmGRF7*, and *AmGRF7/AmGRF8*. The *AmGRF* gene replication events may have played a crucial role in the evolution of *AmGRF*.

Fig.4 Chromosomal distribution and localization of *AmGRFs*. All nine *AmGRF* genes are distributed on the chromosomes, and the scale on the left represents the length of the chromosome.

To further understand the homologous relationships between different species of *AmGRFs*, we compared five dicotyledonous plants and one monocotyledonous plant *O. Sativa*, as well as five dicotyledonous plants, including two medicinal plants, *C. lanceolata* and *C. sativa*, and three types of fruit, *V. vinifera*, *S. lycopersicum*, and *M. domestica* (Fig.5B). The results showed the closest collinearity between the *AmGRF* and *MdGRF* genes, with 17 pairs of collinearity genes, 12 pairs of *AmGRF* and *SlbGRF*, ten pairs of *AmGRF* and *VvGRF*, six pairs of *AmGRF* and *CIGRF*, five pairs of *AmGRF* and *OsGRF*, and four pairs of collinearity between *AmGRF* and *CsGRF*. The number of collinear genes was lower in monocotyledonous plants is less than in dicotyledonous plants.

Fig.5 A Chromosomal distribution and gene replication events of the *AmGRF* gene. The red line represents the tandem replication relationship between the *AmGRF* genomes. The two outer circles show gene density information, with red indicating high gene density and blue indicating low gene density.**B** Collinear relationship between *A. mongholicus* and other species. The light blue lines represent collinear gene pairs. The green section represents the chromosome information of *A. mongholicus*, and the blue sections represent the chromosome information of *M. domestica*, *S. lycopersicum*, *V. vinifera*, *C. lanceolata*, *O. sativa*, and *C.sativa*.

3.5 Cis-acting element analysis of the *AmGRF* gene

To understand the functions of members of the *AmGRF* gene family, we used the upstream 2000 bp region of nine *AmGRF* motifs as the promoter region to perform a predictive analysis of cis-acting elements. We obtained the cis-elements of the *AmGRF* gene using the PlantCRAE website, summarized their possible functions, and plotted them (Fig.6). The 211 cis-acting elements identified could be divided into three categories: stress-, hormone-, and growth-related. Among these, the number of light-responsive elements is the highest, with a total of 121 active elements. Endosperm expression is a cis-acting element of *AmGRF7*, whereas flavonoid biosynthetic gene regulation is a cis-acting element of *AmGRF3*, which may be involved in flavonoid synthesis. These results indicate that the *AmGRF* gene may be involved in various biological processes in plants and plays an indispensable role.

Fig.6 Prediction of the cis-acting elements in the upstream 2000-bp promoter of the nine *AmGRF* genes.

3.5 GO analysis of the *AmGRF* gene family

GO classification was used to identify the possible functions of the nine *AmGRFs*. Analysis of the annotation results showed that *AmGRF* gene functions could be divided into three categories: 25 functional groups of cellular components (CC), molecular functions (MF), and biological processes (BP) (Fig.7). The primary biological processes involving these genes are affected by hormones, light, and low temperatures. *AmGRF* genes were primarily found in the Golgi cytoplasm, plasma membrane, nuclear envelope, chloroplast matrix, and nucleus. In the molecular function category, these genes promoted the transcription and binding of protein sites at the molecular level and promoted their expression.

Fig.7 GO annotation of members of the *AmGRF* gene family.

3.6 qRT-PCR verification of *AmGRF* gene expression patterns in different tissues of *A. mongolicus*

To explore the expression patterns of *AmGRF* genes in different tissues, the expression levels of nine *AmGRF* genes were analyzed based on the transcriptome data of the leaves, roots, and stems of *A. mongholicus*. Five genes were expressed in the stems and leaves (RPKM>0.5), and all genes showed low expression in the

roots. *AmGRF4* and *AmGRF8* were highly expressed in the stems. *AmGRF2*, and *AmGRF4* are expressed in various tissue sites, and these genes may be involved in the growth and development cycle of *A. mongholicus* (Figure 8). The expression of the *AmGRF* gene was higher in the leaves than in the roots. Three genes (*AmGRF5*, *AmGRF6* and *AmGRF8*) were not listed as expressed in all tissue sites (RPKM<0.5), which may be pseudogenes or require specific conditions to activate their expression (Fig.8A).

Based on the transcriptome data, these nine *AmGRF* genes were selected as candidate genes for qRT-PCR analysis. The results showed that the expression patterns of the five candidate genes were consistent with the expression trends obtained from RNA sequencing (RNA-Seq) data. Notably, compared to the expression of *AmGRF5* and *AmGRF6* in the transcriptome data, the RT-qPCR results indicated that the *AmGRF5* and *AmGRF6* gene was expressed in all root leaves. *AmGRF1* and *AmGRF8* exhibited higher expression in roots, and *AmGRF1* and *AmGRF9* exhibited higher expressed in stems. These results indicate that these nine *AmGRF* genes may be closely related to the growth and development of *A. mongolicus* (Fig.8B).

Fig.8 A. Expression patterns of *AmGRF* genes in roots, stems, and leaves. Expression patterns were generated based on Roes scale 1 and analyzed using heatmap hierarchical clustering. Different color codes represent the relative expression levels of genes, and a color gradient from purple to yellow indicates a decrease in gene expression levels. **B.** The relative expression levels of differentially expressed genes (DEGs). *AmGRFs* in the roots, stems, and leaves of 50-d hydroponic seedlings of *A. mongolicus* under normal growth conditions. The X-axis represents the location of plant tissue, whereas the Y-axis represents the relative expression level of genes. The error bars represent standard deviations (SD; n = 3). The significance analysis was performed using a t-test. The *, **, and *** show significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

Discussion

A. mongolicus is a medicinal plant belonging to the *Astragalus* genus, with dry roots as its medicinal part. It is commonly used as a tonifying medicine for the treatment of various diseases. *A. mongolicus* is used in both medicines and food products. It is often used as a tonic diet for making soup, tea, wine, decoctions, and congees (Jiayu et al., 2023). With the continuous expansion of research on *A. mongolicus*, the demand for *A. mongolicus* is increased, leading to a gradual decrease in wild *A. mongolicus* resources and an uneven quality of *A. mongolicus* on the market. Therefore, it is important to explore the molecular mechanisms underlying the growth and development of *A. mongolicus* and improve its quality and yield (Jialin, 2021).

GRF is a transcription factor unique to plants. *GRF* transcription factors exhibit a broad impact on several biological processes such as plant organ development, stress response, flowering time, and yield (Omidbakhshfard et al., 2015; Qin et al., 2022). The *GRF* gene family has been extensively studied in *A. thaliana* (Kim et al., 2003), *O. sativa* (Choi et al., 2004), *Lycopersicon esculentum* (Chen et al., 2019), *N. nucifera* (Wang et al., 2020), *Setaria italica* (Wang et al., 2023a), and other crops. Interestingly, eight (*Hordeum vulgare*)-19 (*Musa nana*) *GRF* family members have been found in monocot plants, and two (*Lotus corniculatus*)-26 (*Glycine max*) *GRF* family members have been found in dicot plants. Only 2 members have been found in lower plant mosses (Chen and Ge, 2021). The number of *GRF* family members in dicots is generally higher than that in monocots. However, there are no reports on the *GRF* gene family of *A. mongholicus*. We conducted a comprehensive analysis of the *AmGRF* gene family using the whole genome of *A. mongholicus*, and identified nine *AmGRF* gene family members. Because the genomic data of *A. mongolica* are quite different from

those of *A. thaliana* (1.4G:125MB) (Kim et al., 2003), the number of *GRF* genes in *A. mongolica* is the same as that of *AtGRF* genes, indicating that the *A. mongolica* genome was deleted during duplication.

We constructed a phylogenetic tree of the *AmGRF*, *AtGRF*, and *OsGRF* proteins. Nine *AmGRF* proteins, nine *AtGRF* proteins, and 12 *OsGRF* proteins were divided into six groups (I–VI), whereas *AmGRF* was only divided into five groups, which is consistent with the results of the *O. sativa* evolutionary tree (Choi et al., 2004). Based on the intra-species evolutionary relationships, we termed *AmGRF1–AmGRF9* according to their chromosomal positions, and the *AmGRF* gene family members were evenly distributed. Duplications between *AmGRF1*, *AmGRF7*, and *AmGRF8* were observed. Based on the evolutionary relationships among various species, dicots are more closely related to *A. mongolicus* than the monocots.

From the perspective of phylogenetic relationships, by comparing the subfamily positions of different *AmGRF* gene family members, we found that the gene family members clustered in the same branch were highly homologous and structurally similar. Among these, *AtGRF* and *AmGRF* genes have the highest homology. In subfamily I, *AtGRF7*, and *AmGRF3* are homologous, and in subfamily V, *AtGRF5* and *AmGRF1*, indicating that *AtGRF5* and *AmGRF1* have positive regulatory effects on plant leaf development and seed growth (Choi et al., 2004). This further indicated that the classification results of *A. mongolicus* and *Arabidopsis* were consistent because they are dicots, whereas rice is a monocot. In the *Arabidopsis* subfamily VI, no *AmGRF* protein exists, and such a situation also exists in soybeans (Kim and Kende, 2004) such as *Liriodendron chinense* and *Setaria italica* (Chen et al., 2019; Wang et al., 2023a), which is consistent with our analysis results, indicating that good gene structure stability of *AmGRF*.

The cis-acting elements in the promoter region play a crucial role in regulating gene expression, as they can participate in regulating downstream-related gene expression at the transcriptional level, thereby improving plant resistance to external environmental pressures. In the upstream 2000 bp of the promoter, we identified multiple cis-responsive elements of *AmGRF*, including various responsive elements, such as light, abscisic acid, MeJA, salicylic acid, and gibberellin. These elements are involved in plant hormone production, stress responses, and regulatory mechanisms related to plant growth and development. Non-biological stress can constrain plant growth and development and affect crop yield and quality (Wang et al., 2023c). Previous studies have confirmed that *GRF* are expressed in different tissues, such as germinated seeds, callus tissues, and buds, generally in the growth zone where cell proliferation occurs (Kim et al., 2003; Kim and Lee, 2006; Pajoro et al., 2014). Studies have reported that the *AmGRF* gene contains elements expressed in the endosperm (*AmGRF7*) and meristem (*AmGRF7*, *AmGRF4*, and *AmGRF3*), indicating that these genes may be involved in plant growth and development processes. We found that the promoters contained abscisic acid, MeJA, salicylic acid, and gibberellin elements. Appropriate application of these hormones to plants may affect the expression of the *AmGRF* gene (Khatun et al., 2017). The response of tobacco *NtabGRF* to ABA, GA3, IAA, BR, and 6-BA was significant. However, moderate spraying of GA3 had a particularly significant effect on the *VvGRF* gene, with most genes showing decreased expression in grape plants (Chen and Ge, 2021). GA3 treatment enhanced the expression of *OsGRF* in rice (Choi et al., 2004). In addition, studies have shown that ABA, MeJA, and GA3 treatments cause significant differences in the expression levels of *CsGRF* genes in most *Melastoma dodecandrum* (Huang et al., 2023). We found a regulatory element for flavonoid biosynthesis genes in *AmGRF3*, and overexpression of the *AmGRF3* gene leads to flavonoid accumulation. This suggests that the *AmGRF* gene family participates in stress response processes through various hormone signaling pathways. The *AmGRF* gene primarily regulates the growth and development of *A. mongolicus* by responding to environmental factors such as light and hormones,

confirming that the *GRF* gene plays an important role in plant growth and development in plant tissues and organs.

The GO annotation data revealed that the *AmGRF* gene family is involved in 14 biological activities, eight molecular functions, and the regulation of three cellular components and is widely involved in plant growth and development. The primary biological activities of these genes are influenced by multiple factors such as hormone regulation, light conditions, and low-temperature environments. In the classification of molecular functions, these genes contribute to the binding of transcription and protein sites at the molecular level, as well as to further promote gene expression.

The *OsGRF* gene is highly expressed in buds, immature leaves, and flower buds and participates in plant growth and development by regulating cell growth in actively growing tissues (Wang et al., 2023b). In the present study, we analyzed the expression profile of the *AmGRF* gene family from the transcriptome data of *A. mongholicus* tissue, indicating that the expression level of this gene family is low in the roots of *A. mongholicus*, but high in the stems and leaves of *AmGRF2*, *AmGRF4*, and *AmGRF9*. We validated the expression of the nine *AmGRF* genes using qRT-PCR (Figure 9). *AmGRF1* and *AmGRF8* showed higher expression in the roots, and the expression patterns of the five candidate genes with higher expression levels of *AmGRF1* and *AmGRF9* in the stems were consistent with the expression trends obtained from the RNA-Seq data. Interestingly, *AmGRF5* and *AmGRF6* were the opposite of the transcriptome data and was expressed in the roots, stems, and leaves.

Conclusion

This study identified, for the first time, nine members of the *AmGRF* gene family distributed on seven chromosomes across the entire genome of *A. mongholicus*, each containing QLQ and WRC domains. Based on the evolutionary relationship of the species, it is divided into six subgroups, and members of the same family have conserved exon-intron structures, conserved motifs, and categories. The reliability of the transcriptome data was validated simultaneously. This study enriches our understanding of the *AmGRF* gene and lays a theoretical foundation for further functional research on the *AmGRF* gene.

Declarations

Funding

This study was funded by the talent training project supported by the central government for the reform and development of local colleges and Universities(ZYRCB2021008); National Key Research and development Project, research and demonstration of collection, screening, and breeding technology of ginseng and other genuine medicinal materials, Project No.2021YFD1600901; National Natural Science Foundation of China Mining and functional study of key enzyme genes of Astragalus isoflavone biosynthesis821713927; Heilongjiang Touyan Innovation Team Program (Grant Number: [2019] No. 5)

Ethics declarations

Competing interests

The authors have no competing interests to declare.

Ethical approval

Not applicable for this study.

Informed consent

Not applicable for this study.

Consent to participate

Not applicable for this study.

Consent to Publish

Not applicable for this study.

Credit authorship contribution statement

Panpan Wang: Conceptualization, Methodology, Visualization, Writing – review & editing. Zhen Wang: Software, Formal analysis. Huiyan Cao: Data curation. Jiajun He: Data curation. Chen Qin: Data curation. LianqingHe: Data curation. Bo Liu: Data curation. Lingyang Kong: Validation. Weichao Ren: Validation. Xiubo Liu: Project administration, Writing – review & editing. Ma: Project administration, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this publication.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files. The raw sequencing data used during the study have been deposited in NCBI SRA with the accession number PRJNA1064679.

References

1. Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., Liu, Y., Feng, J., Chen, H., He, Y. and Xia, R., 2023. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol Plant* 16, 1733-1742.
2. Chen, F., Yang, Y., Luo, X., Zhou, W., Dai, Y., Zheng, C., Liu, W., Yang, W. and Shu, K., 2019. Genome-wide identification of GRF transcription factors in soybean and expression analysis of GmGRF family under shade stress. *BMC Plant Biol* 19, 269.
3. Chen, H. and Ge, W., 2021. Identification, Molecular Characteristics, and Evolution of GRF Gene Family in Foxtail Millet (*Setaria italica* L.). *Front Genet* 12, 727674.
4. Choi, D., Kim, J.H. and Kende, H., 2004. Whole genome analysis of the OsGRF gene family encoding plant-specific putative transcription activators in rice (*Oryza sativa* L.). *Plant Cell Physiol* 45, 897-904.
5. Dong, M., Li, J., Yang, D., Li, M. and Wei, J., 2023. Biosynthesis and Pharmacological Activities of Flavonoids, Triterpene Saponins and Polysaccharides Derived from *Astragalus membranaceus*. *Molecules* 28.

6. Fonini, L.S., Lazzarotto, F., Barros, P.M., Cabreira-Cagliari, C., Martins, M.A.B., Saibo, N.J.M., Turchetto-Zolet, A.C. and Margis-Pinheiro, M., 2020. Molecular evolution and diversification of the GRF transcription factor family. *Genet Mol Biol* 43, 20200080.
7. Guo, J., Zhao, N., Jin, P. and Yin, Y., 2022. Effect of Astragalus injection on inflammatory mediators in patients with viral myocarditis: A systematic review and meta-analysis. *Phytomedicine* 107, 154436.
8. Huang, J., Chen, G.Z., Ahmad, S., Hao, Y., Chen, J.L., Zhou, Y.Z., Lan, S.R., Liu, Z.J. and Peng, D.H., 2023. Genome-Wide Identification and Characterization of the GRF Gene Family in *Melastoma dodecandrum*. *Int J Mol Sci* 24.
9. jialin, Z., 2021. The identification method of easy to confuse Chinese medicine decoction pieces in Chinese medicine mixing. *Shenzhen J. Inte Tradi Chin and Wes Med* 31, 101-102.
10. Jiayu, X., Yuanjin, S., Kunpeng, Z., Jinliang, F. and Fude, Y., 2023. Research progress in medicinal and food homologous of *Astragali Radix*. *J Pharm Sci* 38, 718-724.
11. Khatun, K., Robin, A.H.K., Park, J.I., Nath, U.K., Kim, C.K., Lim, K.B., Nou, I.S. and Chung, M.Y., 2017. Molecular Characterization and Expression Profiling of Tomato GRF Transcription Factor Family Genes in Response to Abiotic Stresses and Phytohormones. *Int J Mol Sci* 18.
12. Kim, J.H., 2019. Biological roles and an evolutionary sketch of the GRF-GIF transcriptional complex in plants. *BMB Rep* 52, 227-238.
13. Kim, J.H., Choi, D. and Kende, H., 2003. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in *Arabidopsis*. *Plant J* 36, 94-104.
14. Kim, J.H. and Kende, H., 2004. A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in *Arabidopsis*. *Proc Natl Acad Sci U S A* 101, 13374-9.
15. Kim, J.H. and Lee, B.H., 2006. GROWTH-REGULATING FACTOR4 of *Arabidopsis thaliana* is required for development of leaves, cotyledons, and shoot apical meristem. *Journal of Plant Biology* 49, 463-468.
16. Kumar, S., Stecher, G. and Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33, 1870-4.
17. Lee, S.J., Lee, B.H., Jung, J.H., Park, S.K., Song, J.T. and Kim, J.H., 2018. GROWTH-REGULATING FACTOR and GRF-INTERACTING FACTOR Specify Meristematic Cells of Gynoecia and Anthers. *Plant Physiol* 176, 717-729.
18. Liao, B., Hu, H., Xiao, S., Zhou, G., Sun, W., Chu, Y., Meng, X., Wei, J., Zhang, H., Xu, J. and Chen, S., 2022. Global Pharmacopoeia Genome Database is an integrated and mineable genomic database for traditional medicines derived from eight international pharmacopoeias. *Sci China Life Sci* 65, 809-817.
19. Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402-8.
20. Lu, J., Wang, Z., Li, J., Zhao, Q., Qi, F., Wang, F., Xiaoyang, C., Tan, G., Wu, H., Deyholos, M.K., Wang, N., Liu, Y. and Zhang, J., 2023. Genome-Wide Analysis of Flax (*Linum usitatissimum* L.) Growth-Regulating Factor (GRF) Transcription Factors. *Int J Mol Sci* 24.
21. Meng, L., Li, X., Hou, Y., Li, Y. and Hu, Y., 2022. Functional conservation and divergence in plant-specific GRF gene family revealed by sequences and expression analysis. *Open Life Sci* 17, 155-171.
22. Omidbakhshfard, M.A., Proost, S., Fujikura, U. and Mueller-Roeber, B., 2015. Growth-Regulating Factors (GRFs): A Small Transcription Factor Family with Important Functions in Plant Biology. *Mol Plant* 8, 998-1010.
23. Pajoro, A., Madrigal, P., Muiño, J.M., Matus, J.T., Jin, J., Mecchia, M.A., Debernardi, J.M., Palatnik, J.F., Balazadeh, S., Arif, M., Ó'Maoiléidigh, D.S., Wellmer, F., Krajewski, P., Riechmann, J.L., Angenent, G.C. and

- Kaufmann, K., 2014. Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome Biol* 15, R41.
24. Qin, L., Chen, H., Wu, Q. and Wang, X., 2022. Identification and exploration of the GRF and GIF families in maize and foxtail millet. *Physiol Mol Biol Plants* 28, 1717-1735.
25. Shi, Y., Wang, X., Wang, J., Niu, J., Du, R., Ji, G., Zhu, L., Zhang, J., Lv, P. and Cao, J., 2022. Systematical characterization of GRF gene family in sorghum, and their potential functions in aphid resistance. *Gene* 836, 146669.
26. van der Knaap, E., Kim, J.H. and Kende, H., 2000. A novel gibberellin-induced gene from rice and its potential regulatory role in stem growth. *Plant Physiol* 122, 695-704.
27. Wang, J., Tu, Z., Wang, M., Zhang, Y., Hu, Q. and Li, H., 2023a. Genome-wide identification of GROWTH-REGULATING FACTORS in *Liriodendron chinense* and functional characterization of LcGRF2 in leaf size regulation. *Plant Physiol Biochem* 206, 108204.
28. Wang, J., Zhou, H., Zhao, Y., Sun, P., Tang, F., Song, X. and Lu, M.Z., 2020. Characterization of poplar growth-regulating factors and analysis of their function in leaf size control. *BMC Plant Biol* 20, 509.
29. Wang, P., Xiao, Y., Yan, M., Yan, Y., Lei, X., Di, P. and Wang, Y., 2023b. Whole-genome identification and expression profiling of growth-regulating factor (GRF) and GRF-interacting factor (GIF) gene families in *Panax ginseng*. *BMC Genomics* 24, 334.
30. Wang, Z., Zhang, Z., Wang, P., Qin, C., He, L., Kong, L., Ren, W., Liu, X. and Ma, W., 2023c. Genome-wide identification of the NAC transcription factors family and regulation of metabolites under salt stress in *Isatis indigotica*. *Int J Biol Macromol* 240, 124436.
31. Yeh, Y.C., Doan, L.H., Huang, Z.Y., Chu, L.W., Shi, T.H., Lee, Y.R., Wu, C.T., Lin, C.H., Chiang, S.T., Liu, H.K., Chuang, T.H., Ping, Y.H., Liu, H.S. and Huang, C.F., 2021. Honeysuckle (*Lonicera japonica*) and Huangqi (*Astragalus membranaceus*) Suppress SARS-CoV-2 Entry and COVID-19 Related Cytokine Storm in Vitro. *Front Pharmacol* 12, 765553.
32. Zhang, J., Li, Z., Jin, J., Xie, X., Zhang, H., Chen, Q., Luo, Z. and Yang, J., 2018. Genome-wide identification and analysis of the growth-regulating factor family in tobacco (*Nicotiana tabacum*). *Gene* 639, 117-127.
33. Zhang, J., Wu, C., Gao, L., Du, G. and Qin, X., 2020. Astragaloside IV derived from *Astragalus membranaceus*: A research review on the pharmacological effects. *Adv Pharmacol* 87, 89-112.

Figures

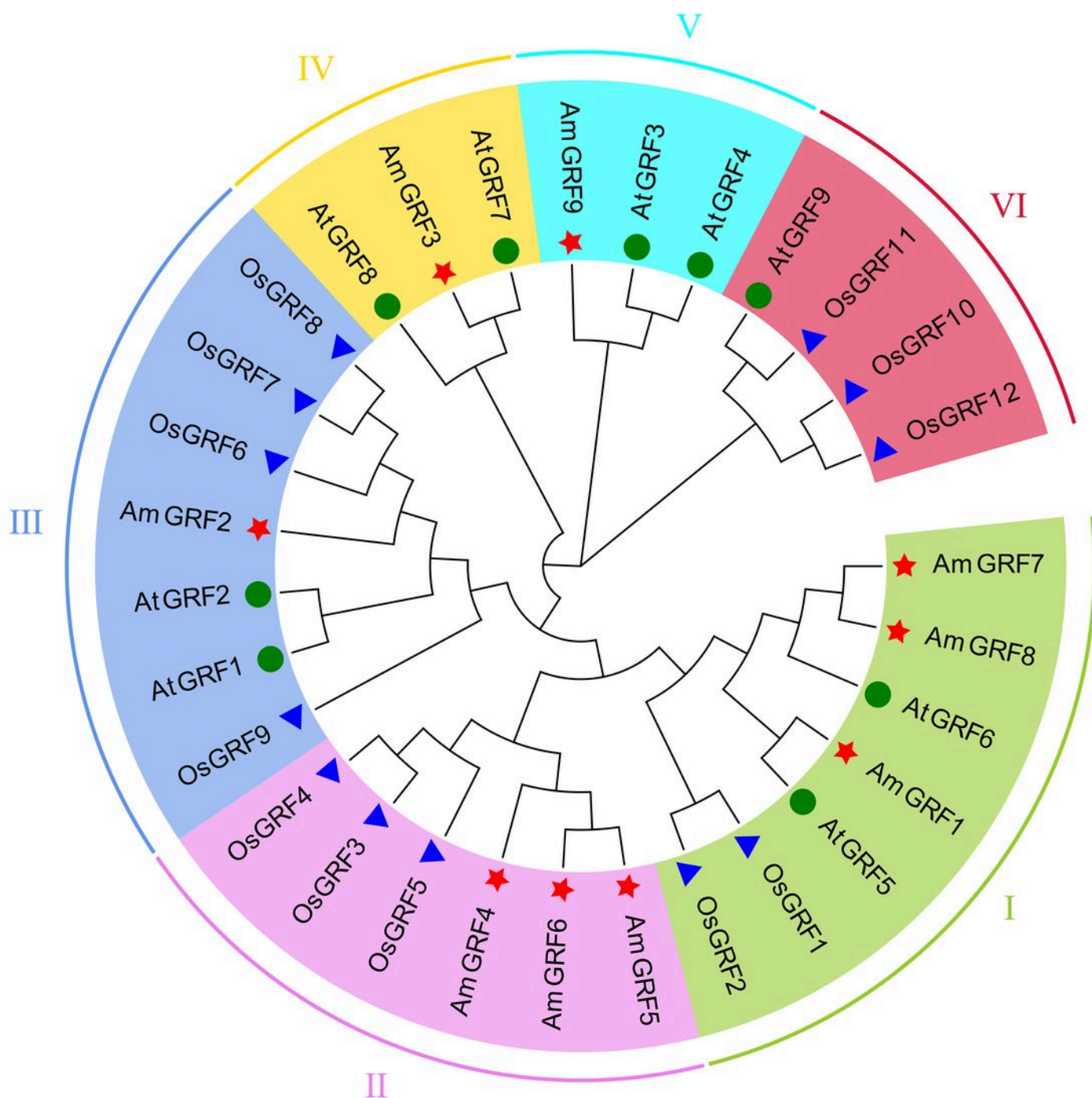


Figure 1

NJ phylogenetic tree constructed with *A. thaliana*, *O. sativa*, and AmGRF proteins. The bootstrap value was set to 1000 replicates. The red star represents the AmGRF protein, the blue triangle represents the OsGRF protein, and the green circle represents the AtGRF protein. Different colors represent the various subfamilies.

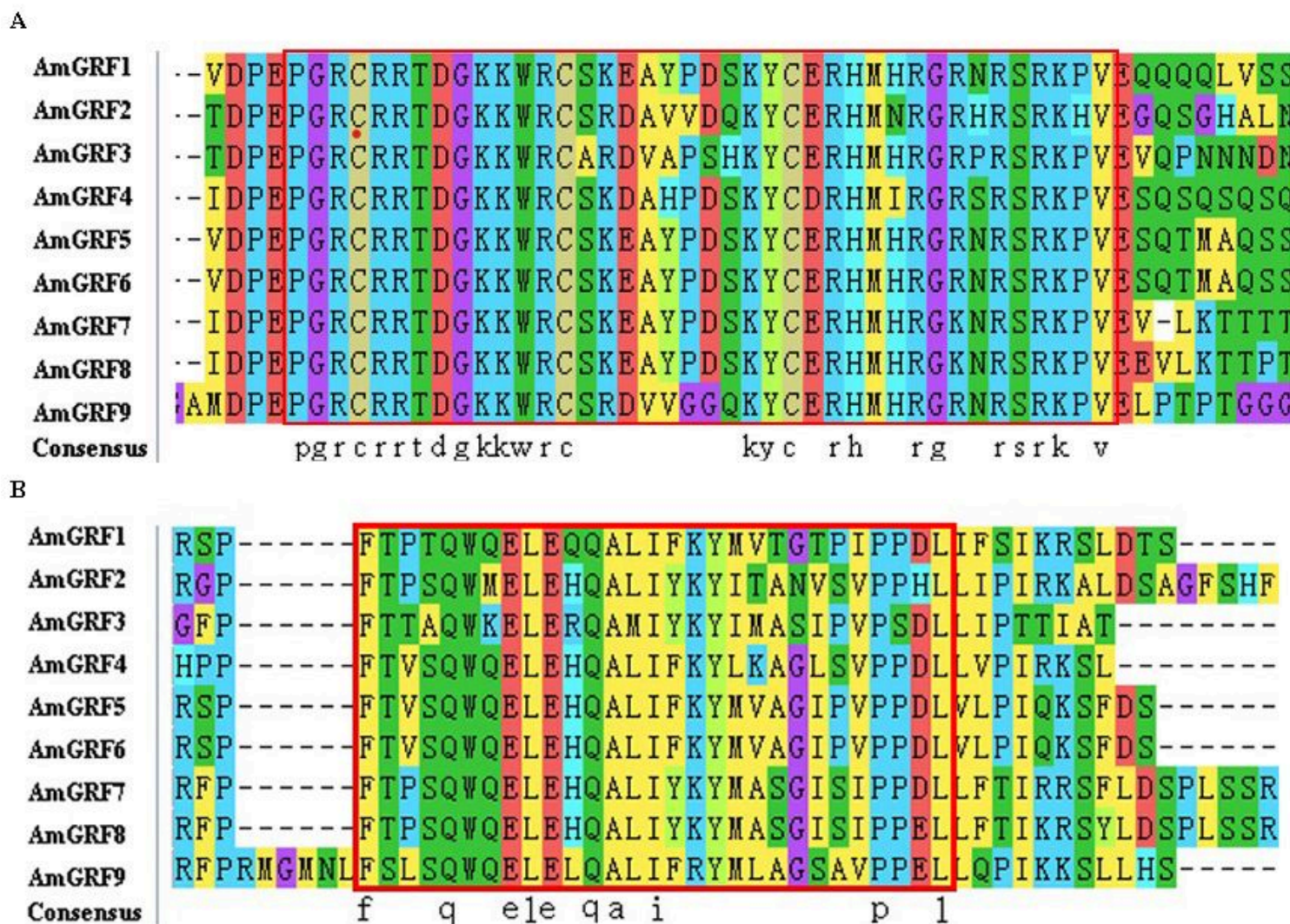


Figure 2

Conservative sequence alignment of AmGRF gene family proteins. A: QLQ conservative domain. B: WRC conserved domain.

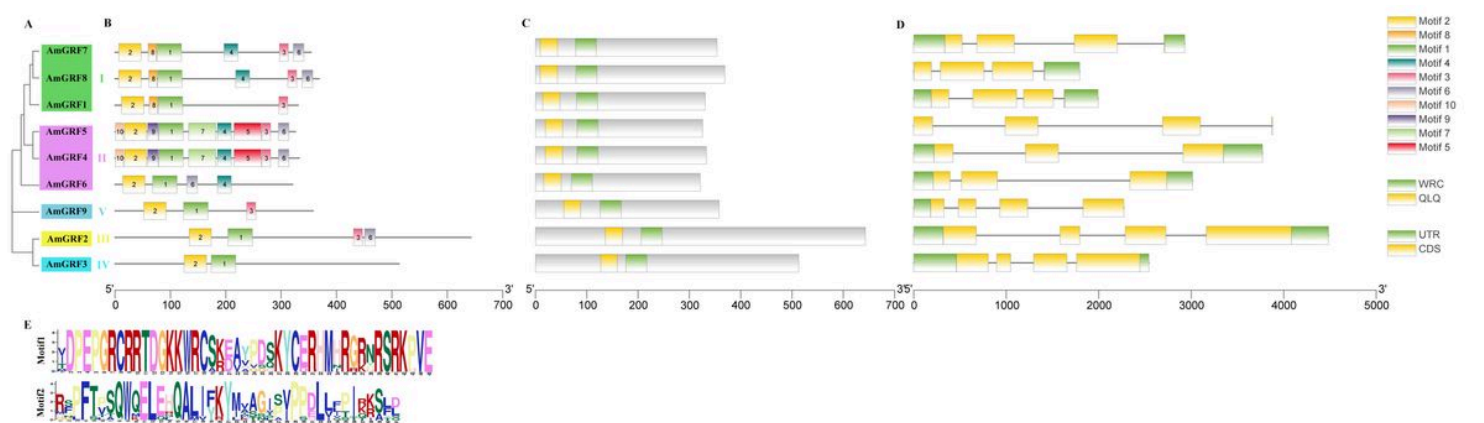


Figure 3

f the AmGRF protein. (A) Representation of the AmGRF protein phylogenetic tree. Green, pink, yellow, light blue, and dark blue backgrounds represent groups I, II, III, IV, and V, respectively. (B) The MEME website was used to identify the composition of *AmGRF* protein motifs, with different colors representing the various motifs. (C) The *AmGRF* gene has a conserved domain, with green representing the WRC domain and yellow representing the QLQ domain. (D) *AmGRF* gene structure. CDS: encoding sequence, UTR: In the untranslated region; lines represent introns. (E) The conservative structural domains motif1 and motif2 of *AmGRFs*.

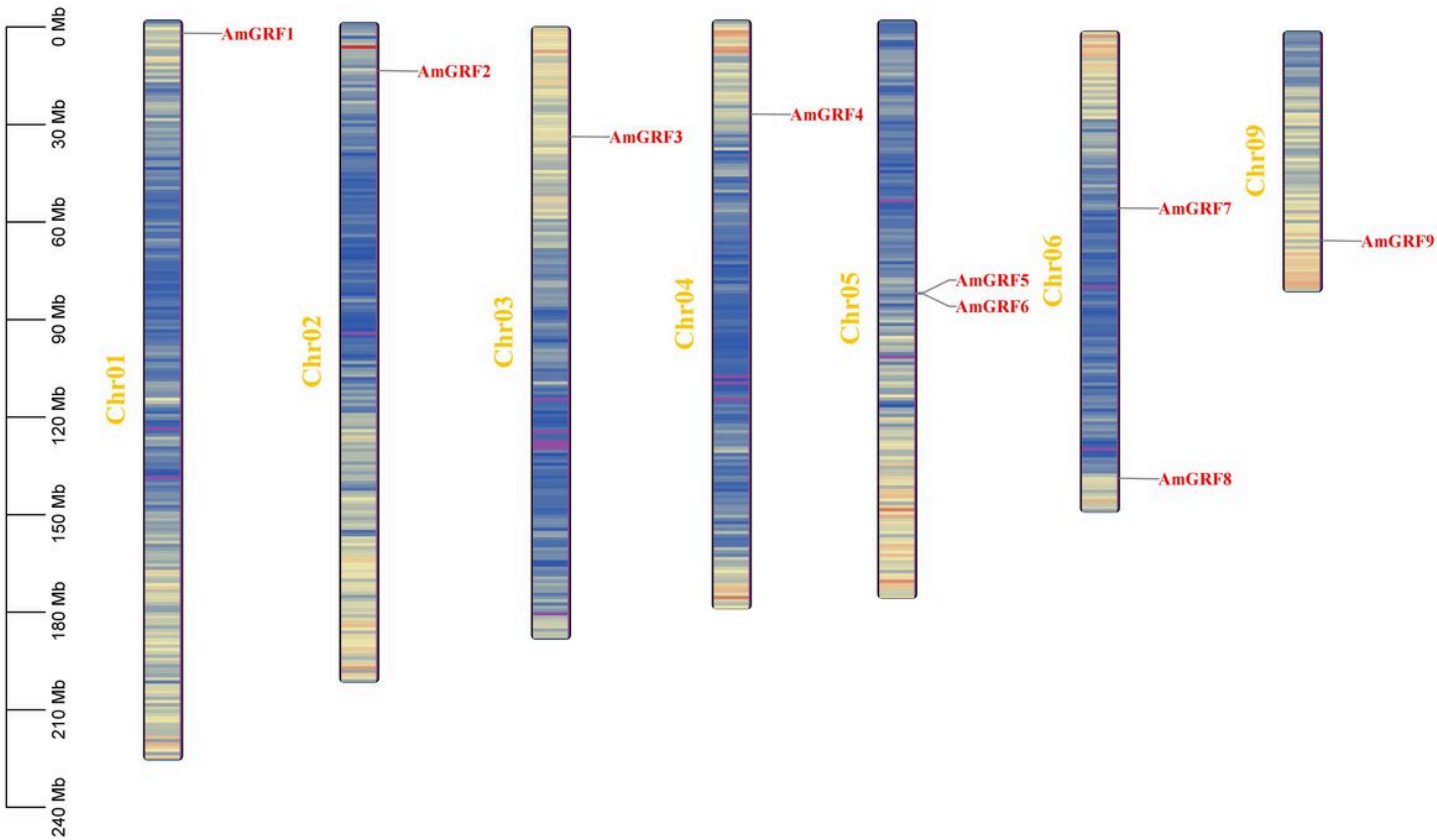


Figure 4

Chromosomal distribution and localization of *AmGRFs*. All nine *AmGRF* genes are distributed on the chromosomes, and the scale on the left represents the length of the chromosome.

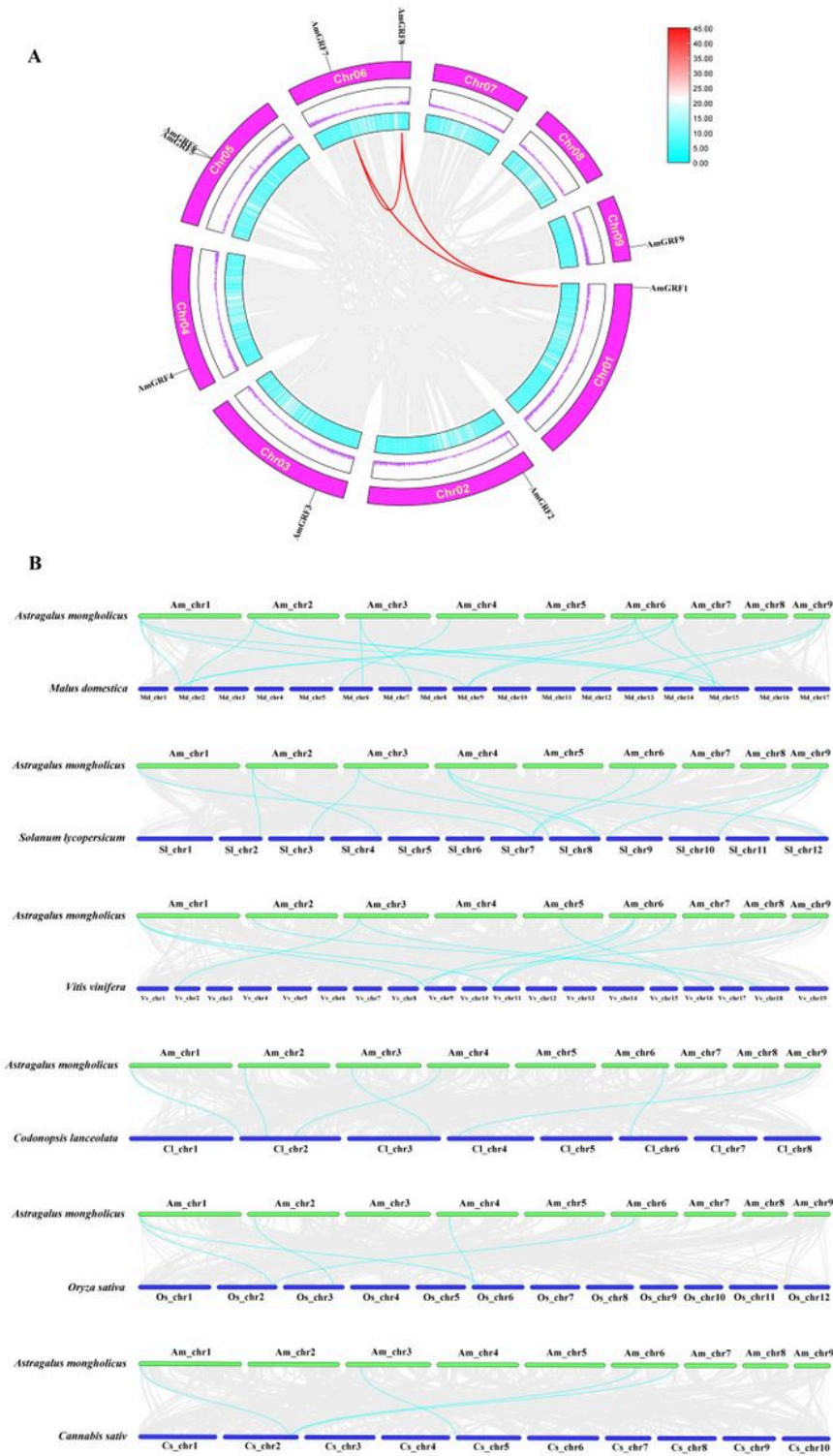


Figure 5

AChromosomal distribution and gene replication events of the *AmGRF* gene. The red line represents the tandem replication relationship between the *AmGRF* genomes. The two outer circles show gene density information, with red indicating high gene density and blue indicating low gene density. **B**Collinear relationship between *A. mongholicus* and other species. The light blue lines represent collinear gene pairs. The green section represents

the chromosome information of *A. mongholicus*, and the blue sections represent the chromosome information of *M. domestica*, *S. lycopersicum*, *V. vinifera*, *C. lanceolata*, *O. sativa*, and *C.sativa*.

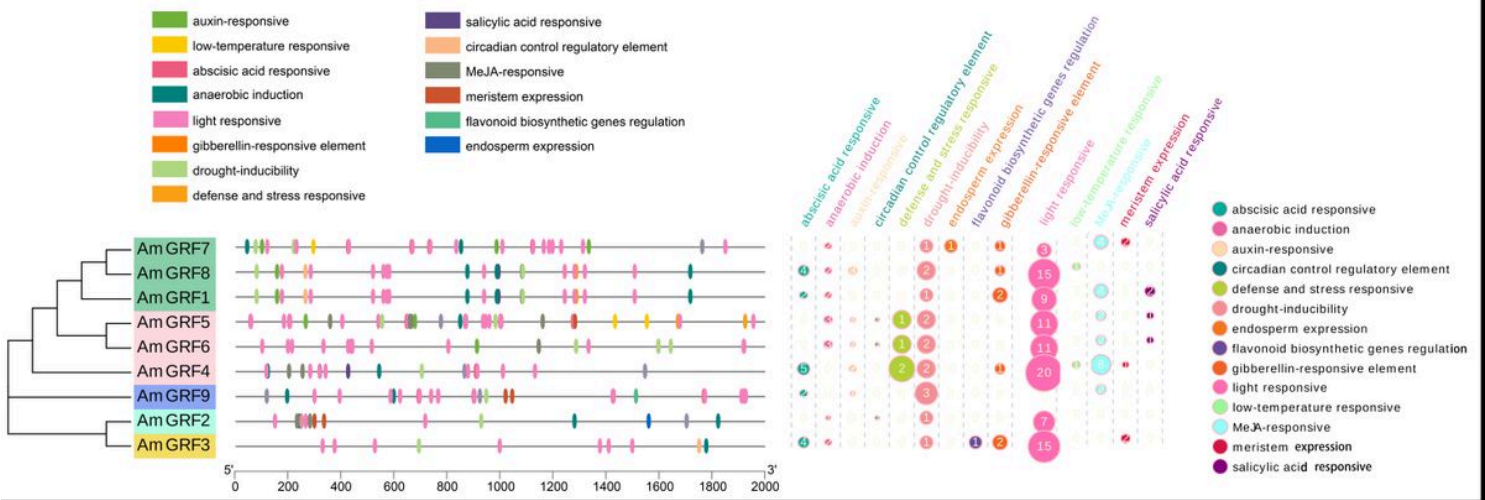


Figure 6
Prediction of the cis-acting elements in the upstream 2000-bp promoter of the nine *AmGRF* genes.

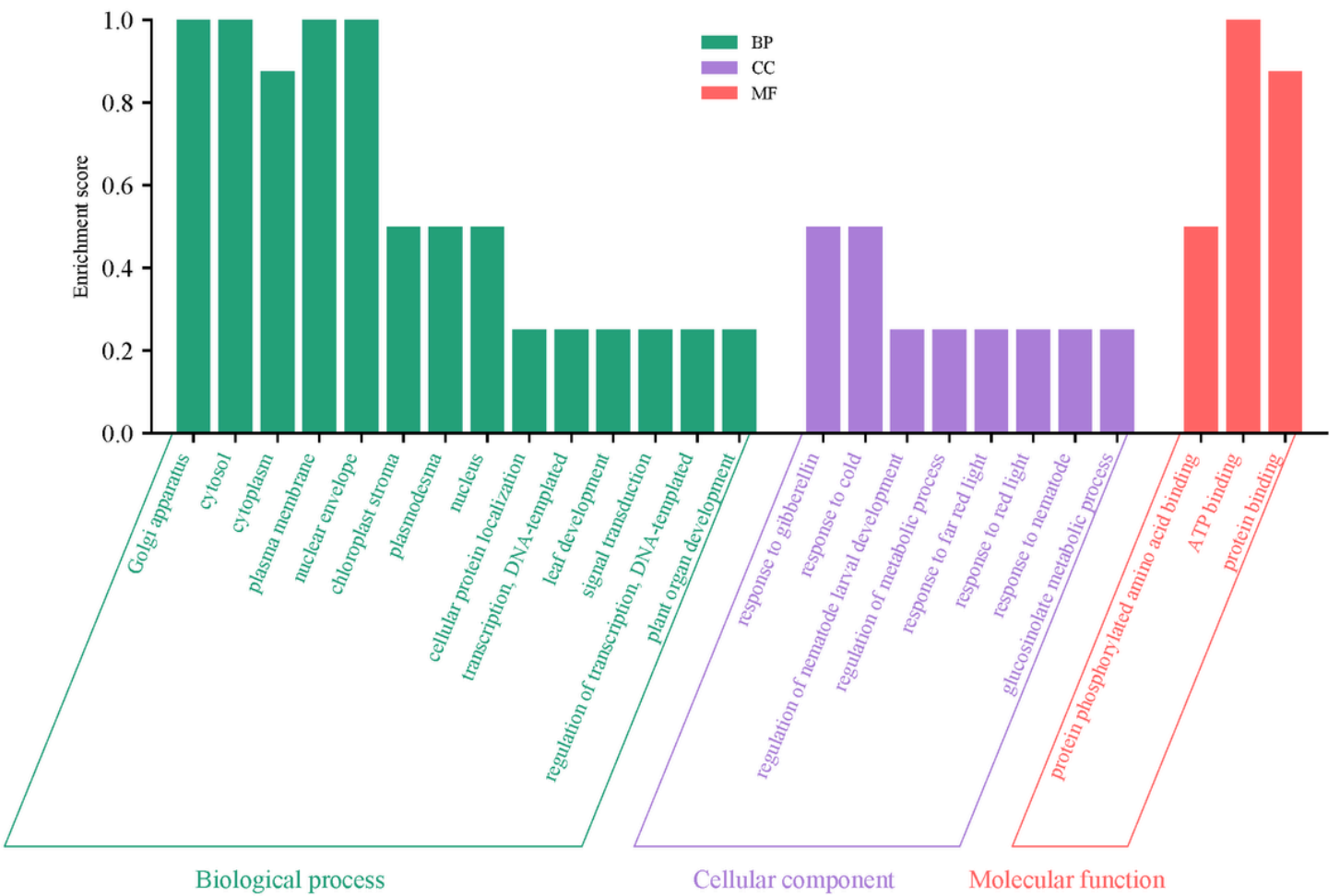


Figure 7

GO annotation of members of the *AmGRF* gene family.

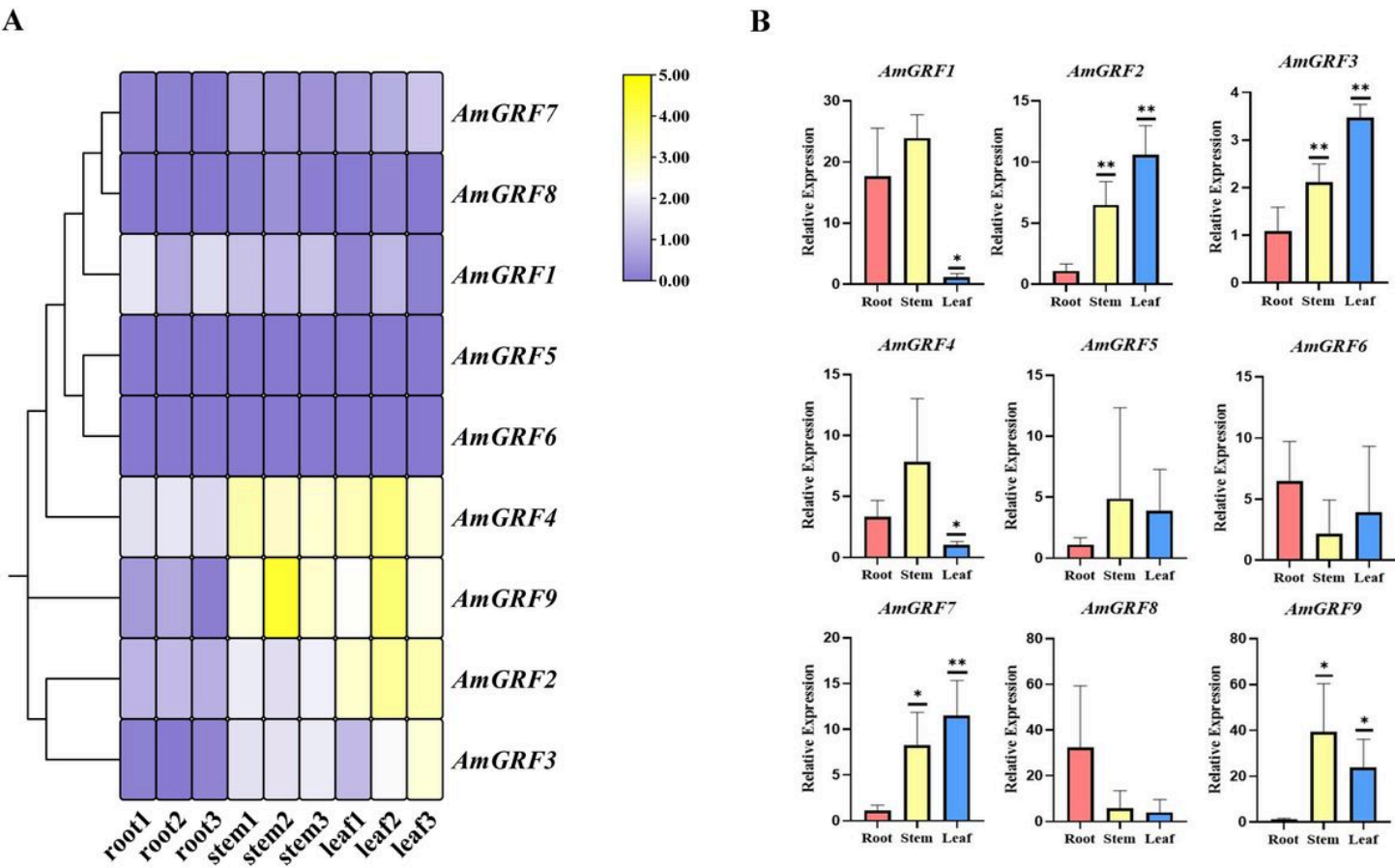


Figure 8

A. Expression patterns of *AmGRF* genes in roots, stems, and leaves. Expression patterns were generated based on Roes scale 1 and analyzed using heatmap hierarchical clustering. Different color codes represent the relative expression levels of genes, and a color gradient from purple to yellow indicates a decrease in gene expression levels. **B.** The relative expression levels of differentially expressed genes (DEGs). *AmGRF*s in the roots, stems, and leaves of 50-d hydroponic seedlings of *A. mongolicus* under normal growth conditions. The X-axis represents the location of plant tissue, whereas the Y-axis represents the relative expression level of genes. The error bars represent standard deviations (SD; n = 3). The significance analysis was performed using a t-test. The *, **, and *** show significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementfile1.docx](#)
- [Supplementaryfile2.docx](#)