Identification and molecular evolution of the *GLX* genes in 21 plant species: a focus on the *Gossypium hirsutum*

Menglin Xu[1,2], Dongyun Zuo[2], Qiaolian Wang[2], Limin Lv[2], Youping Zhang[2], Huixin Jiao[1,2], Xiang Zhang[1,2], Yi Yang[1,2], Guoli Song[1,2*] and Hailiang Cheng[1,2*]

Abstract

Background The glyoxalase system includes glyoxalase I (GLXI), glyoxalase II (GLXII) and glyoxalase III (GLXIII), which are responsible for methylglyoxal (MG) detoxification and involved in abiotic stress responses such as drought, salinity and heavy metal.

Results In this study, a total of 620 *GLX* family genes were identified from 21 different plant species. The results of evolutionary analysis showed that *GLX* genes exist in all species from lower plants to higher plants, inferring that *GLX* genes might be important for plants, and *GLXI* and *GLXII* account for the majority. In addition, motif showed an expanding trend in the process of evolution. The analysis of cis-acting elements in 21 different plant species showed that the promoter region of the *GLX* genes were rich in phytohormones and biotic and abiotic stress-related elements, indicating that *GLX* genes can participate in a variety of life processes. In cotton, *GLXs* could be divided into two groups and most *GLXIs* distributed in group I, *GLXIIs* and *GLXIIIs* mainly belonged to group II, indicating that there are more similarities between *GLXII* and *GLXIII* in cotton evolution. The transcriptome data analysis and quantitative real-time PCR analysis (qRT-PCR) show that some members of *GLX* family would respond to high temperature treatment in *G.hirsutum*. The protein interaction network of GLXs in *G.hirsutum* implied that most members can participate in various life processes through protein interactions.

Conclusions The results elucidated the evolutionary history of *GLX* family genes in plants and lay the foundation for their functions analysis in cotton.

Keywords Glyoxalase family, Cotton, Phylogeny, Heat treatment

*Correspondence:
Guoli Song
sglzms@163.com
Hailiang Cheng
pser2010@163.com
[1]Zhengzhou Research Base, State Key Laboratory of Cotton Biology, School of Agricultural Sciences, Zhengzhou University, Zhengzhou 450001, Henan, China
[2]State Key Laboratory of Cotton Biology, Cotton Research Institute of Chinese Academy of Agricultural Sciences, Anyang 455000, Henan, China

Background

Abiotic stresses such as drought, salinity and extreme temperature seriously threaten the growth of plants and limit the development of modern agriculture [1–3]. Under abiotic stress, photosynthesis and respiration of plants can produce excessive toxic aldehydes, such as methylglyoxal (MG), glyoxal (GO) and 3-deoxyglucosone (DOG), through enzymatic and non enzymatic pathways [4]. Biomacromolecules will be damaged by higher concentrations of MG, resulting in carbonyl stress. MG can react with arginine, lysine and cysteine to produce irreversible advanced glycation end products (AGEs) [5–7]; MG can also react with membrane lipids to produce irreversible advanced lipid peroxidation end products (ALEs) [8, 9]; In addition, MG can induce the production of ROS, and then generate oxidative stress, which damages proteins, DNA, RNA, lipids and biofilms [9–13]. At low concentration MG can also act as a signal molecule. It forms a signal network through interaction with other signal molecules, such as Ca^{2+}, H_2O_2, nitric oxide (NO), hydrogen sulfide (H_2S), to further regulate a variety of physiological processes and stress tolerance such as seed germination, plant growth, development, reproduction [4]. Therefore, it is particularly important to maintain the dynamic balance of MG in cells.

Plants have evolved many effective detoxification mechanisms, including glyoxalase system and non glyoxalase system [14]. The glyoxalase system is the main defense line, accounting for 99% of the total MG clearance [4]. There are three kinds of glyoxalase: glyoxalase I (GLXI), glyoxalase II (GLXII) and glyoxalase III (GLXIII) [15–17]. GLXI removes excess MG with glutathione (GSH) as a cofactor to produce the intermediate S-d lactoylglutathione (SLG), which is then converted into lactic acid by GlXII, and regenerates GSH [4, 18–21]. It can be seen that glyoxalase system can not only remove excess MG to maintain its homeostasis, but also maintain redox homeostasis in cells by regulating GSH regeneration. GLXIII is a glutathione independent glyoxalase, which can directly catalyze the irreversible conversion of MG to lactic acid without the participation of glutathione or other cofactors [22–25].

GLXs are important enzymes for plants to cope with the stress of aldehydes and ketones and abiotic stress [26–29]. So far, there have been some studies on the molecular function of the *GLX* genes, for example, *GLXI* can participate in cell proliferation in soybean [30] and Amaranthus paniculatus [31], be involved in abiotic stress in tomato [32], pumpkin [33], onion [33] and wheat [34], improve stress resistance in tobacco [35], black gram [36], Arabidopsis [37], mustard [38] and rice [39]. *GLXIs* and *GXIIs* can participate in salt stress in rice [40, 41], but not on its molecular evolution. Evolutionary studies of gene families can not only systematically elucidate gene members, sequence structure and molecular functions, but also help to explore biological characteristics of species, such as growth and development, transcriptional regulation and environmental adaptation, laying a solid foundation for biological breeding research [42–45]. In recent years, global water shortage, soil salinization and frequent extreme weather have seriously affected the growth environment of crops [46–50]. Cotton is an important Cash crop in the world, but systematic identification and analysis of *GLX* family in cotton have not been reported yet. In this study, *GLX* genes were identified in different green plants, and an evolutionary analysis was performed. We focused on the identification, evolutionary relationship and function of *GLX* genes in *Gossypium hirsutum*. The results of this study provide a theoretical basis for the evolution and biological functions of the *GLX* genes in plants and the study of *GLX* gene family will help us understand its potential application in cotton.

Results

Identification of *GLX* genes in green plants

In order to study the evolutionary relationship of *GLX* genes in green plants, 620 *GLX* genes were identified in 21 plant species, which represent the evolutionary relationship of 10 kinds of green plants, including green algae, charophytes, bryophytes, lycophytes, pteridophytes, gymnosperms, basal angiosperms, monocots, basal eudicots, and core eudicots (Fig. 1 and Table S2). The *GLX* gene family has already appeared in green algae, with 14 and 15 *GLX* genes in *O. lucimarinus* and *V. carteri*, respectively. *GLX* genes in other plants were more than green algae (Fig. 1). Notably, from Algae to Bryophytes, there was an explosion of *GLX* number, leading to 47 members in *P. patens*. There were the most *GLX* genes (80) in *G.hirsutum*, which could be due to it is allotetraploid. This indicates that *GLX* family undergo expansion in the process of evolution. Existence of *GLXs* in both lower and higher plants suggested that *GLX* genes could be important for plants. In most species, *GLXI* and *GLXII* are more abundant than *GLXIII*, and gene numbers varied more in different plants. In order to study the homologous domain and preservation degree of GLX protein in evolution, MEME (http://meme-suite.org/tools/meme, accessed on 18 November 2022) was used for motif analysis, and 20 conservative motifs were identified in *GLXI*, *GLXII*, and *GLXIII* respectively. The results showed that motif generally showed an expansion trend in the evolutionary process. For example, motif 3, motif 9 and motif 15 in *GLXI* were not found in green algae, and gradually appeared from charophytes (Figure S1). In *GLXII*, motif 19 and motif 20 were not found in green algae, and began to appear in charophytes (Figure S2). In *GLXIII*, motif 2, motif 3, motif 8, motif 9, motif 14, motif 16 and motif 20
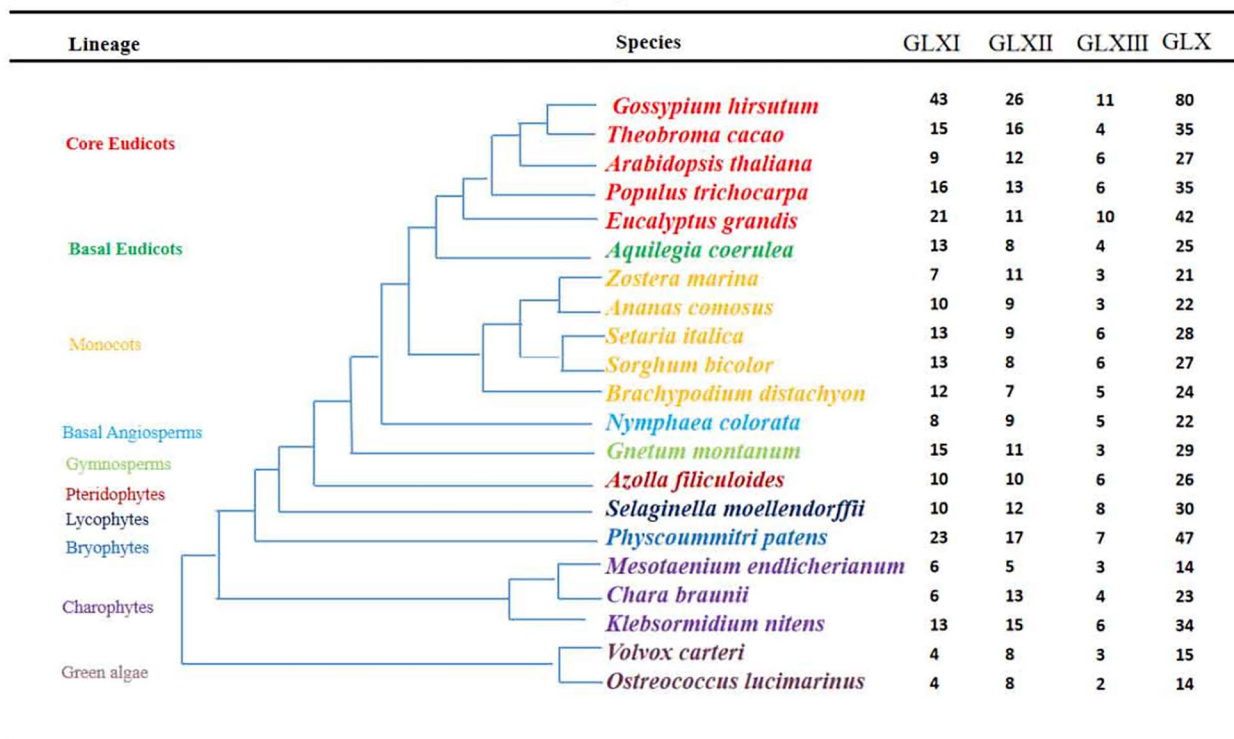
Fig. 1 Phylogenetic relationships among 21 plant species. Different background colors in the phylogenetic tree show different lineages. The right side shows the numbers of *GLXI, GLXII,* and *GLXIII* genes in 21 different plant species, and the total number of *GLX* genes in each species
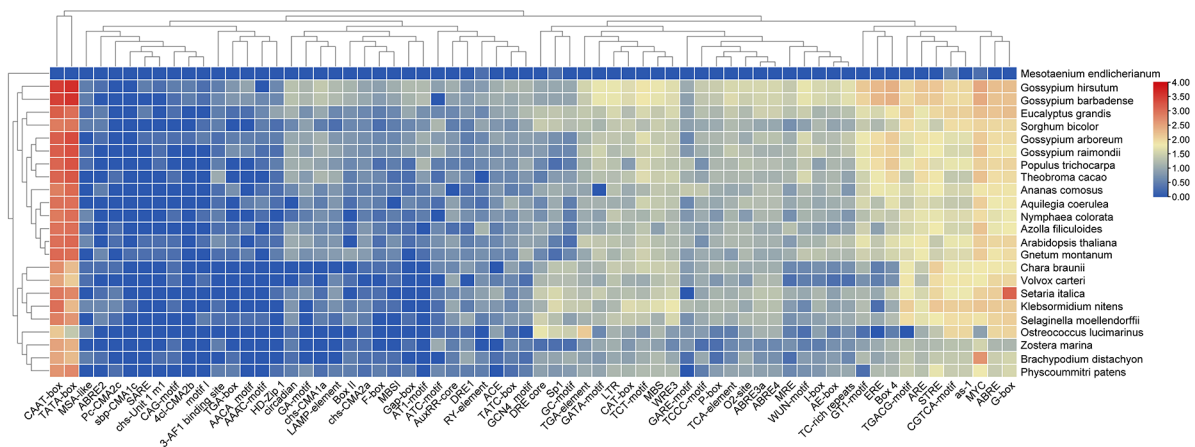

Fig. 2 Predicted cis-elements in the promoter regions of *GLX* genes in 21 species. The species names were shown on the right, and the cis-acting elements were displayed at the bottom. Scale bars at the right represented log2 (FPKM + 1). Different colours represent the different numbers of cis-acting elements

did not appear in green algae, and began to be found in charophytes with evolution (Figure S3).

Analysis of putative cisacting elements in *GLX* promoters
The cis-acting element is essentially a DNA sequence, which is a binding site to transcription factors, and regulates the precise initiation and conversion efficiency of gene transcription by binding to transcription factors.

The 2-kb promoter sequence of *GLX* genes upstream in 21 plant species was analyzed using PlantCARE database [51]. The results showed that different types of Cis-acting element were found in the *GLX* genes promoter (Fig. 2 and Table S3). As shown in Fig. 2, among them, the core element CAAT box which related to the regulation of nopaline synthase, and the core element TATA box which related to transcription, are the most abundant

and exist in most species [52]. Various phytohormone responsive elements have also been found, such as ABRE, an ABA response element [53], ERE, an ethylene-responsive element [54], CGTCA-motif, MYC, chs-CMA1a and TGACG-motif which belong to MeJA-responsive element [55], GARE-box, P-box, TATC-box and Pc-CMA2c which belong to GA-responsive element [56], as-1, SARE, sbp-CMA1c and TCA-element which belong to salicylic acid-responsive element [57], and TGA-element, AuxRR-core and GA-box which belong to IAA-responsive element, this suggests that different phytohormones can regulate the expression of *GLXs*. Various light-related elements such as G-box, Box 4, GT1-motif, GATA-motif, Sp1, I-box, 3-AF1 binding site, TCT-motif, and Gap-box were also observed [58]. The stress-responsive elements were detected in the promoter of some *GLXs*, such as F-box [59], MBS (MYB binding site involved in drought-inducibility), ARE (anaerobic induction element) [60], STRE (stressresponsive element) [61], LTR (low temperature-responsive elements) [62], dehydration-responsive element (DRE1 and DRE-core) [63], TC-rich repeats (defence and stress-responsive element), wound-responsive element (WRE3 and WUN-motif) [64], and GC-motif (enhancer-like element involved in anoxic specific inducibility). Some growth and development elements were also predicted, such as CAT-box (cis-acting regulatory element related to meristem expression), GCN4_motif (cis-regulatory element involved in endosperm expression) [65], circadian (cis-acting regulatory element involved in circadian control) [66] and RY-element (cis-acting regulatory element involved in seed-specific regulation) [67]. The results indicated that *GLXs* are involved in plant growth and development as well as environmental stress responses.

Phylogenetic analysis of *GLX* genes family in cotton

In order to further study the evolutionary relationship of GLX proteins in cotton. Total 235 *GLX* genes were identified in *G.hirsutum*, *G. arboreum*, *G.raimondii* and *G.barbadense.* Phylogenetic tree was constructed using the amino acid sequences of all the GLX proteins from four cotton species (Fig. 3). The phylogenetic tree showed that all GLX proteins in cotton were divided into two groups. The first group consists of 127 members and is divided into two subgroups. The second group was divided into two subgroups and contained 108 members. All GLXI are distributed in group I, GLXII and GLXIII are mostly distributed in Group II, indicating that the members of GLXII and GLXIII are more similar in evolution.

Collinearity analysis of the *GLX* gene family in cotton

The *G. hirsutum* and *G.barbadense* are evolved due to hybridization between an A-genome species (ancestor of *G.herbaceum* or *G.arboreum*) and a D-genome species (ancestor of *G.raimondii*) [68]. To understand
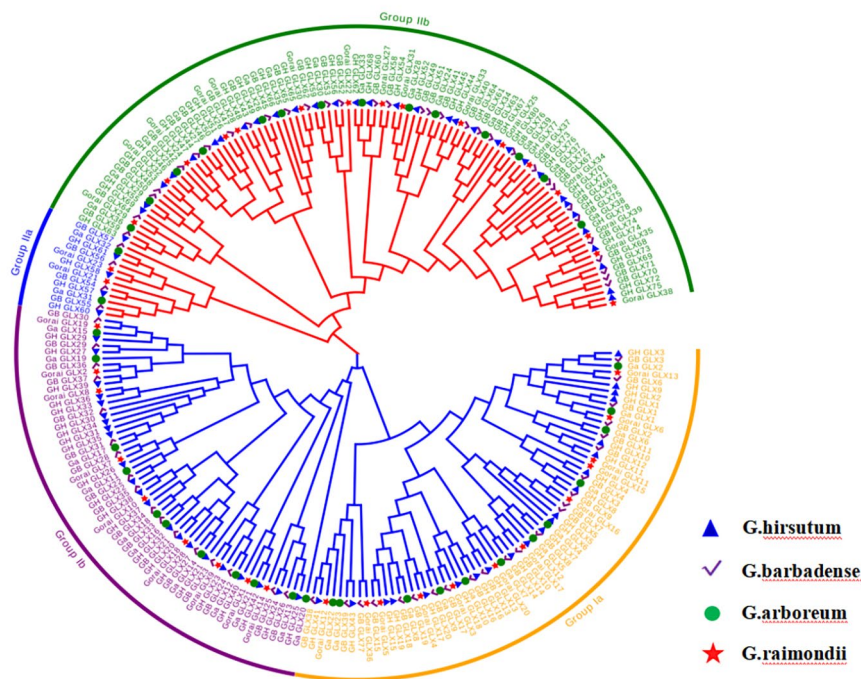

Fig. 3 Phylogenetic analysis and subfamily classification of the *GLX* genes in *G. hirsutum* (*GH_GLX*), *G. barbadense* (*GB_GLX*), *G. arboreum* (*Ga_GLX*) and *G. raimondii* (*Gorai_GLX*). The phylogenetic tree was constructed with MEGA 6.0 using the neighbor-joining model with 1000 bootstrap replicates. All 235 *GLXs* in cotton were divided in to four subgroups, which were highlighted by different colors

the evolutionary relationships of *GLX* genes in cotton, a relative syntenic map of *GLX* genes from the four cotton species was fabricated (Fig. 4). According to MCScan analysis, 103 duplicate gene pairs were found between diploid *G. arboreum* and tetraploid *G.hirsutum*, and also 103 between diploid *G.arboreum* and tetraploid *G.barbadense*. Meanwhile, 113 and 112 duplication gene pairs were found between diploid *G.raimondii* and tetraploid *G.hirsutum*, diploid *G.raimondii* and tetraploid *G.barbadense*, respectively. The <u>GLX</u> genes of *G.Aimondii* has a good collinear relationship with the *GLX* genes of *G.hirsutum* and *G.barbadense* on chr03, chr09, chr10 and chr11. The *GLX* genes of *G.arboreum* has a good collinear relationship with the *GLX* genes of *G.hirsutum* and *G.barbadense* on chr10. The result shows that during the evolution of *GLXS*, chromosome may have undergone small deletion, duplication and reshuffling [69].

Expression analysis of *GLX* genes under stress in *G.hirsutum*
In previous studies, it was reported that members of the *GLX* family were associated with stress response in rice. In order to explore whether the *GLXs* of *G. hirsutum* were also related to stress responses, gene expression profiles after stresses treatment were obtained from public RNA-seq data. It can be seen that the expression levels

of *GLX15*, *GLX19*, *GLX20*, *GLX24*, *GLX25*, *GLX28* and *GLX38* changed under 37℃ treatment (Fig. 5, Figure S4 and Table S4), suggesting that these seven genes may respond to high temperature.

GLX20, *GLX28* and *GLX38* have the highest expression level in anthers, *GLX19* is mainly expressed in roots, *GLX15* is mainly expressed in leaves and roots, and *GLX24* and *GLX25* are mainly highly expressed in leaves(Figure S5 and Table S5). To validate the expression change of 7 *GLX* genes under heat treatment, qRT-PCR was performed to verify their expression in leaves at 0 h, 1 h, 3 h, 6 h, 12 h, and 24 h after heat treatment. The qRT-PCR results were almost consistent with transcriptome data. The trend of *GLX15*, *GLX19*, *GLX20*, *GLX28* and *GLX38* was roughly the same, the expression levels of all genes were the highest at 1 h after heat treatment, and began to decrease significantly at 3 h, and then almost leveled off, *GLX38* was almost not expressed after 3 h treatment. The expression trend of *GLX24* and *GLX25* was almost the same, the expression decreased after 1 h heat treatment, reached the highest level at 3 h, and then tended to be stable. In short, the expression of these seven genes roughly decreased with the passage of heat treatment time, and responded to heat treatment (Fig. 6).
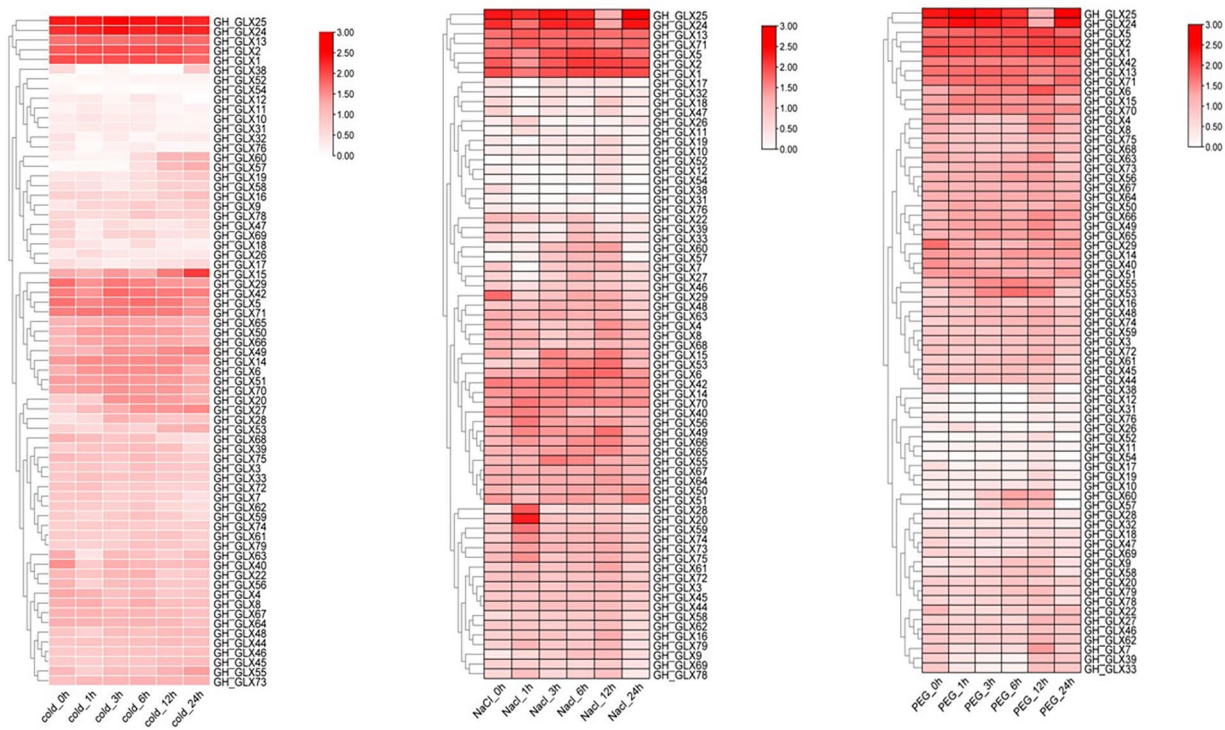

Fig. 4 The sub-genome distribution and synteny analysis of cotton *GLX* genes. The blue lines represents duplicated *GLX* pairs, the gray lines represents collinear blocks
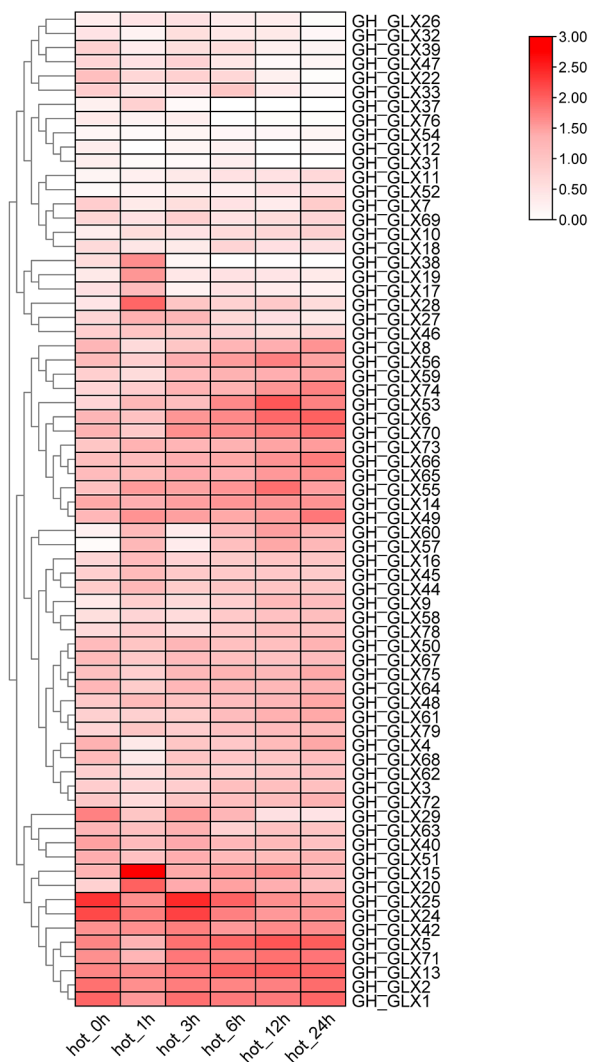
Xu *et al. BMC Genomics*   (2023) 24:474

Page 6 of 11

Fig. 5 Expression of *GLX g*enes in *G.hirsutum* under heat treatment. The phylogenetic relationships were displayed on the left, the gene names were shown on the right, and the time after heat treatment is displayed at the bottom. Scale bars at the right represented log2 (FPKM+1). Different colours represent the different expression levels of *GH_GLXs*

Prediction of GLX protein interactions in *G.hirsutum*

To further detect the potential roles of GLX family members, we investigated the interaction of the GLX proteins in *G.hirsutum* using the STRING database [70]. The results showed that most members of GLX family can interact with each other, and a few could interact with other proteins (Fig. 7). For instance, GLX61 and GLX58 can interact with A0A1U8KBF4, A0A1U8HUQ9, A0A1U8PER4 and A0A1U8NZA9 which related to cleavage. GLX52 and GLX45 can interact with A0A1U8N6I4, A0A1U8KCE2, A0A1U8HY12 and A0A1U8J6K2, which are related to segmentation, flowering time, nucleotide polymerization and phosphorylation, respectively. This implies that GLX proteins may function as homologous

protein complex and participate in various biological processes through protein interactions.

Discussion

Glyoxalase was discovered in animal in 1913 and not reported in plants until the end of the 20th century [28]. In plants, the *GLX* gene family has an important impact on growth, development and coping with abiotic stresses [71]. In this study, a total of 620 *GLX* genes were identified in 21 green plants, and *GLX* genes were distributed in all the 21 species, but the number of *GLXI* and *GLXII* in each species was far greater than that of *GLXIII* (Fig. 1 and Table S2), this indicates that *GLX* genes are widely distributed in plants, but the number of *GLXIII* is relatively small.

The promoter is DNA sequence located in the upstream of the structural gene and recruit RNA polymerase for transcription initiation [72]. The specificity of gene expression depends on cis-regulatory elements [73]. Therefore, 2000 bp promoter sequences of all the *GLX* genes were extracted, and the cis-acting elements were predicted using the PlantCRAE. Various types of elements were found in the promoter sites of *GLX* genes in green plants. The distribution of cis-acting elements among different species was different, as shown in a large number of cis-acting elements were predicted to be related to transcription, various hormones, and stresses response, cell cycle, and development. The results indicated that *GLXs* are might involved in plant growth and development as well as environmental stress responses (Fig. 2).

Cotton is one of the most important cash crops in the world. At the same time, cotton is also a crop with strong drought resistance and salt tolerance, which is suitable for planting in moderate drought conditions or mild saline-alkali land. Based on this characteristic of cotton, we mainly studied *GLX* gene in cotton. A total of 235 *GLXs* family members were identified in four cultivated cotton species, and the phylogenetic tree divided all *GLXs* into two groups, *GLXI* distributed in the first subgroup, *GLXII* and *GLXIII* mostly distributed in the second subgroup (Fig. 3). This suggests that *GXLII* and *GLXIII a*re relatively close in evolution. According to previous studies, although *GLXII* and *GLXIII* are close in evolution, there are still great differences in function. In terms of function, *GLXII* and *GLXI* have similar modes of action, both catalyzing cascade reactions [18–21]. Unlike *GLXI* and *GLXII*, *GLXIII* can directly convert MG irreversibly into non-toxic lactic acid without glutathione glyoxylase [22–25]. The relative syntenic map of *GLX* genes from the four cotton species indicate that chromosome may have undergone small deletion, duplication and reshuffling during the evolution of *GLXs* (Fig. 4).
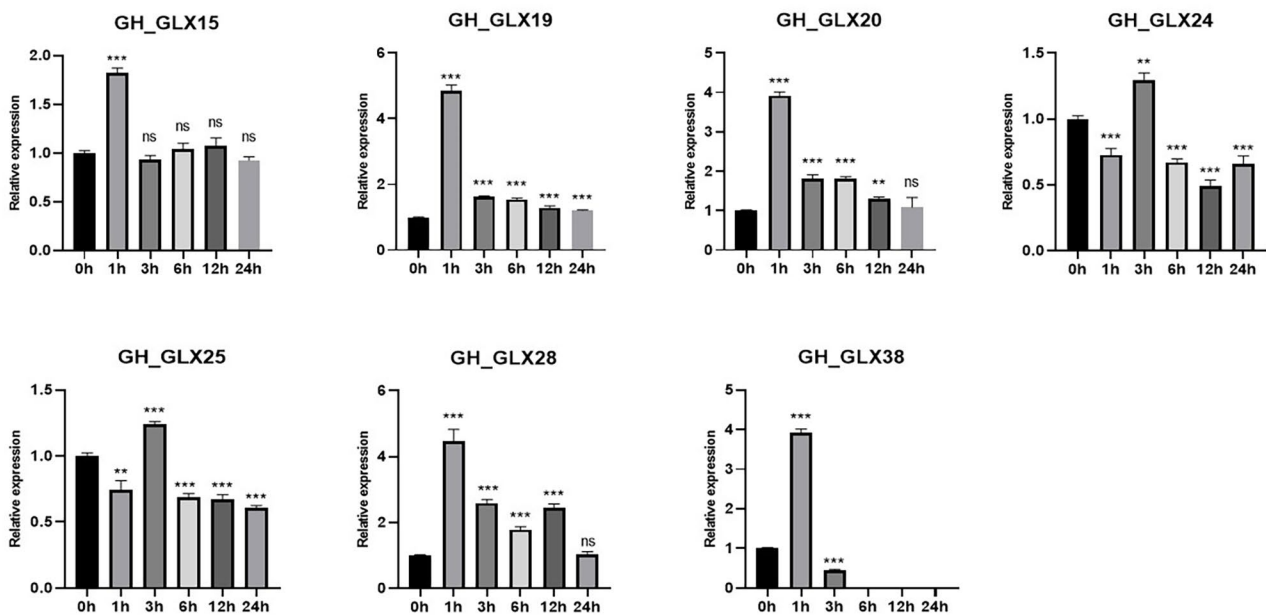
Fig. 6 The expression patterns of 7 *GH_GLXs* at different time after heat treatment by qRT-PCR. The values are standardized. *GhUBQ7* (DQ116441) was used as the internal control. Each experimnet was performed in three biological replicates, and the error bars represent mean±SD; n=3. (**$p < 0.01$, ***$p < 0.001$, Student's *t*-test)
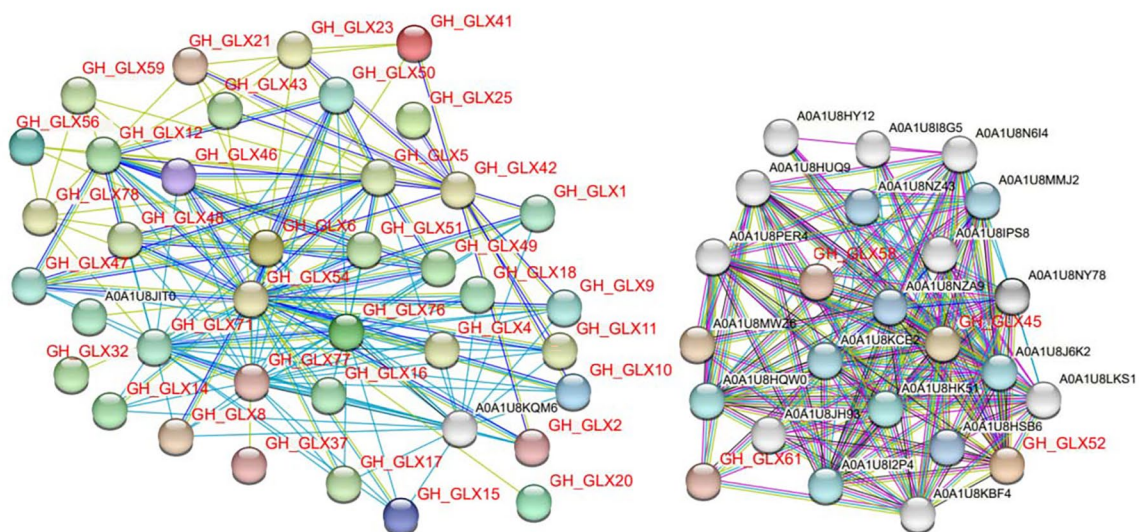

Fig. 7 Functional network assembly of the GLX proteins in *G.hirsutum*. Red letters represent the GLX proteins. Light blue and purple lines stand for represent known interactions determined by the database and experiments, respectively. Green, red, and blue lines stand for predicted interactions from gene proximity, fusion, and symbiosis, respectively. Light green, black, and gray lines represent other interactions from text mining, co-expression, and protein homology, respectively. Empty nodes: proteins of unknown three-dimensional structure

The potential role of the *GLX* family members was further explored by predicting *G.hirsutum* transcriptome data, quantitative real-time PCR analysis (qRT-PCR) and protein interaction networks. The results showed that the expression levels of *GLX15*, *GLX19*, *GLX20*, *GLX24*, *GLX25*, *GLX28*, and *GLX38* changed under high temperature treatment, indicating that they were responsive to high temperature treatment (Figs. 5 and 6). It was also found that the expression levels of these seven genes were different in different tissues. *GLX20*, *GLX28*, and *GLX38* were mainly expressed in anthers, *GLX19* was highly expressed in roots, *GLX15* was expressed in leaves

Xu *et al. BMC Genomics*   (2023) 24:474

Page 8 of 11

and roots, and *GLX24* and *GLX25* have high expression level mainly in leaves. (Figure S4 and Table S5). This indicates that although *GLX* exists widely in nature, its expression levels in different tissues and organs are different, which is consistent with the results of previous studies. Meanwhile, the protein interaction prediction indicated that most *GLX* genes interact with each other. *GLX61*, *GLX58*, *GLX52* and *GLX45* may be involved in cleavage, flowering time, nucleotide polymerization and phosphorylation by different protein interactions (Fig. 7).

Conclusion

In this study, 620 *GLX* genes were identified from 21 plants. It was found that *GLX* genes are distributed in lower and higher plants and widely exist in nature, but their expression levels are different in different tissues and organs. We focused our analysis on *GLX* genes in cotton and found that some of them might be involved in abiotic stress and some other life processes through interacting proteins.

Materials and methods

Genome-wide identification of *GLX* family genes in plants

In this study, the genome-wide data of 21 plant species were analyzed. The genomic, CDS and protein sequences of *G. hirsutum* (ZJU), *G. arboreum* (CRI), *G. barbadense* (ZJU) and *G.raimondii* (JGI) were downloaded from CottonGen (https://www.cottongen.org, accessed on 18 November 2022) [74–76]. The genomic, CDS and protein sequences of Arabidopsis was downloaded from TAIR (https://www.arabidopsis.org/, accesses on 18 November 2022) [77]. The genomic, CDS and protein sequences of *Klebsormidium nitens* and *Theobroma cacao* were downloaded from NCBI. The sources of genomic data for others species are listed in Table S1. *GLXI* genes domain information (PF00903, PF12681), *GLXII* genes domain information (PF00753), and *GLXIII* genes domain information (PF01965) were downloaded from Pfam database(https://pfam.xfam.org, accesses on 18 November 2022) [78], Then, we used HMMER3.0 software (http://www.hmmer.org/, accessed on 18 November 2022) with an e-value of 1×10^{-5} as the threshold to acquire GLX protein sequence. Use online tool SMART (http://smart.embl-heidelberg.de/, accessed on 18 November 2022) to further analyze the domain, remove the sequence without *GLX* domain, and finally obtain all genes of *GLX* family [79].

Analysis of Cis-acting elements in the promoter regions of *GLX* genes

A 2.0 kb of promoter sequence upstream from the transcription start site in each *GLX* genes were extracted from the genome database and analyzed using PlantCare (http://bioinformatics.psb.ugent.be/webtools/plantcare/ html, accesses on 18 November 2022) online software to predict the putative cis-acting regulatory elements [51]. After classifying and counting the results [80], TBtools software was used for visualization [81].

Construction of phylogenetic tree of *GLX* family in cotton

Multi-sequence alignment of all GLX protein sequences was carried out using Clustal X [82], and phylogenetic trees were constructed using the MEGA software proximity method (version 6.0) (Neighbor-Joining, NJ) [83] with the calibration parameter Bootstrap being set to 1000. The online software Evolview (https://www.omicsclass.com/article/671, accessed on 10 July 2021) was used to modify the evolutionary tree [84].

Collinearity analysis of the *GLX* genes family in cotton

The gene duplication analysis of four cotton species were conducted using MCScanX software [85]. The synonymous substitution (Ks) and nonsynonymous substitution (Ka) of each duplicated gene pairs were calculated in TBtools with default parameters [81]. Gene duplication and synteny relationship were visualized by TBtools [81].

Expression of *GLX* genes under heat treatment in *G.hirsutum*, Sample collection, RNA extraction, cDNA synthesis, and qRT-PCR analysis

Transcriptome data were downloaded from the Cotton Omics Datebase (http://cotton.zju.edu.cn/10.rnasearch. html, accessed on 18 November 2022). The heat map generated by TBtools software is used to display the relative expression level [81]. Upland cotton cultivar Z12 was planted under controlled conditions. Heat treatment was applied to plants at three leaf stage. Leaves were collected from treated plants after 0 h, 1 h, 3 h, 6 h, 12 and 24 h of treatment. Three biological repeats were collected in each stage. All samples were immediately frozen in liquid nitrogen and stored at -80 ℃. Plant total RNA was extracted by the RNAprep Pure Plant Plus Kit (Polysaccharides&Polyphenolicsrich) (DP441) (Tiangen, Beijing, China). The reverse transcription kit was the Prime Script™ II 1st Strand cDNA Synthesis Kit (Genstar, Beijing, China). It was used to reverse the extracted RNA to obtain the first-strand cDNA for transcriptomic and qRT-PCR analysis. Specific primers were designed according to the CDS sequence of the *GH_GLX* genes on the primer-blast of NCBI website (Table S6). *GhUBQ7* (DQ116441) was used as the internal control. qRT-PCR was performed with LightCycler 480 Real-Time PCR system (Roche, Switzerland). Obtaining gene expression data using three biological replicates. Moreover, the relative quantitative analysis of gene expression was carried out by the $2^{-\Delta\Delta \text{ Ct}}$ method with three independent replicates [86].

Xu *et al. BMC Genomics*   (2023) 24:474

Page 9 of 11

Prediction of GLX protein interactions in *G.hirsutum*

The prediction of GLX protein interactions was established using the STRING database (https://cn.string-db.org/, accessed on 18 November 2022) [70], Parameter uses the default values.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-023-09524-w.

Additional file: Figure S1. Motif analysis of GLXI gene in 21 plant species

Additional file: Figure S2. Motif analysis of GLXII gene in 21 plant species

Additional file: Figure S3. Motif analysis of GLXIII gene in 21 plant species

Additional file: Figure S4. Expression of GLX genes under abiotic stress in *G.hirsutum*

Additional file: Figure S5. Tissue-specific expression of seven differentially expressed genes

Additional file: Table S1. The genome data sources for the species used in this study

Additional file: Table S2. GLX gene IDs in 21 plant species

Additional file: Table S3. List of identified cis-elements in the putative promoter region of *GLX* genes

Additional file: Table S4. Expression of GLX genes under abiotic stress in *G.hirsutum*

Additional file: Table S5. Tissue-specific expression of seven differentially expressed genes

Additional file: Table S6. List of primers used in qRT-PCR experiments

---

Data Availability
The following information was supplied regarding data availability: Data is available at NCBI SRA, accession numbers: PRJNA490626(https://www.ncbi.nlm.nih.gov/sra?linkname=bioproject_sra_all&from_uid=490626) and PRJNA248163 (https://www.ncbi.nlm.nih.gov/sra?linkname=bioproject_sra_all&from_uid=248163). Extra data has been appended as supplementary Tables.

Declarations

References

1. Rhaman MS, Imran S, Karim MM, Chakrobortty J, Mahamud MA, Sarker P, et al. 5-aminolevulinic acid-mediated plant adaptive responses to abiotic stress. Plant Cell Rep. 2021;40(8):1451–69.
2. Waadt R, Seller CA, Hsu PK, Takahashi Y, Munemasa S, Schroeder JI. Plant hormone regulation of abiotic stress responses. Nat Rev Mol Cell Biol. 2022;23(10):680–94.
3. Fang S, Hou X, Liang X. Response mechanisms of plants under saline-alkali stress. Front Plant Sci. 2021;12:667458.
4. Mostofa MG, Ghosh A, Li ZG, Siddiqui MN, Fujita M, Tran LP. Methylglyoxal - a signaling molecule in plant abiotic stress responses. Free Radic Biol Med. 2018;122:96–109.
5. Borysiuk K, Ostaszewska-Bugajska M, Vaultier MN, Hasenfratz-Sauder MP, Szal B. Enhanced formation of Methylglyoxal-Derived Advanced Glycation End Products in Arabidopsis under ammonium Nutrition. Front Plant Sci. 2018;9:667.
6. Kold-Christensen R, Johannsen M. Methylglyoxal metabolism and aging-related disease: moving from correlation toward causation. Trends Endocrinol Metab. 2020;31(2):81–92.
7. vander Bruggen MM, Spronck B, Delhaas T, Reesink KD, Schalkwijk CG. The putative role of Methylglyoxal in arterial stiffening: a review. Heart Lung Circ. 2021;30(11):1681–93.
8. Moldogazieva NT, Mokhosoev IM, Mel'nikova TI, Porozov YB, Terentiev AA. Oxidative stress and Advanced Lipoxidation and Glycation End Products (ALEs and AGEs) in aging and age-related Diseases. Oxid Med Cell Longev. 2019;2019:3085756.
9. Stopper H, Schupp N, Bahner U, Sebekova K, Klassen A, Heidland A. Genomic damage in end-stage renal failure: potential involvement of advanced glycation end products and carbonyl stress. Semin Nephrol. 2004;24(5):474–8.
10. Kalapos MP. Methylglyoxal toxicity in mammals. Toxicol Lett. 1994;73(1):3–24.
11. Bellahcène A, Nokin MJ, Castronovo V, Schalkwijk C. Methylglyoxal-derived stress: an emerging biological factor involved in the onset and progression of cancer. Semin Cancer Biol. 2018;49:64–74.
12. Hasanuzzaman M, Alam MM, Nahar K, Mohsin SM, Bhuyan MHMB, Parvin K, et al. Silicon-induced antioxidant defense and methylglyoxal detoxification works coordinately in alleviating nickel toxicity in Oryza sativa L. Ecotoxicology. 2019;28(3):261–76.
13. Majláth I, Éva C, Hamow K, Kun J, Pál M, Rahman A, et al. Methylglyoxal induces stress signaling and promotes the germination of maize at low temperature. Physiol Plant. 2022;174(1):e13609.
14. Morgenstern J, Campos Campos M, Nawroth P, Fleming T. The glyoxalase system-new Insights into an ancient metabolism. Antioxid (Basel). 2020;9(10):939.
15. Yumnam S, Subedi L, Kim SY. Glyoxalase system in the progression of skin aging and skin malignancies. Int J Mol Sci. 2020;22(1):310.
16. Wang J, Yang X, Wang Z, Wang J. Role of the glyoxalase system in breast Cancer and Gynecological Cancer-Implications for therapeutic intervention: a review. Front Oncol. 2022;12:857746.
17. Inoue Y, Maeta K, Nomura W. Glyoxalase system in yeasts: structure, function, and physiology. Semin Cell Dev Biol. 2011;22(3):278–84.
18. Lee DY, Lin YC, Chang GD. Biochemical regulation of the glyoxalase system in response to insulin signaling. Antioxid (Basel). 2021;10(2):326.

19. Stratmann B, Goldstein B, Thornalley PJ, Rabbani N, Tschoepe D. Intracellular Accumulation of Methylglyoxal by glyoxalase 1 knock down alters collagen homoeostasis in L6 myoblasts. Int J Mol Sci. 2017;18(3):480.

20. Scott GF, Nguyen AQ, Cherry BH, Hollrah RA, Salinas I, Williams AG Jr, et al. Featured article: pyruvate preserves antiglycation defenses in porcine brain after cardiac arrest. Exp Biol Med (Maywood). 2017;242(10):1095–103.

21. Rose IA, Nowick JS. Methylglyoxal synthetase, enol-pyruvaldehyde, glutathione and the glyoxalase system. J Am Chem Soc. 2002;124(44):13047–52.

22. Mohanan MV, Pushpanathan A, Padmanabhan S, Sasikumar T, Jayanarayanan AN, Selvarajan D, et al. Overexpression of *glyoxalase III* gene in transgenic sugarcane confers enhanced performance under salinity stress. J Plant Res. 2021;134(5):1083–94.

23. Kumar B, Kaur C, Pareek A, Sopory SK, Singla-Pareek SL. Tracing the evolution of Plant Glyoxalase III enzymes for structural and functional divergence. Antioxid (Basel). 2021;10(5):648.

24. Jana GA, Krishnamurthy P, Kumar PP, Yaish MW. Functional characterization and expression profiling of *glyoxalase III* genes in date palm grown under abiotic stresses. Physiol Plant. 2021;172(2):780–94.

25. Zhao Q, Su Y, Wang Z, Chen C, Wu T, Huang Y. Identification of glutathione (GSH)-independent glyoxalase III from Schizosaccharomyces pombe. BMC Evol Biol. 2014;14:86.

26. Hoque TS, Hossain MA, Mostofa MG, Burritt DJ, Fujita M, Tran LS. Methylglyoxal: an Emerging Signaling Molecule in Plant Abiotic stress responses and tolerance. Front Plant Sci. 2016;7:1341.

27. Parvin K, Hasanuzzaman M, Mohsin SM, Nahar K, Fujita M. Coumarin improves tomato plant tolerance to salinity by enhancing antioxidant defence, glyoxalase system and ion homeostasis. Plant Biol (Stuttg). 2021;23(Suppl 1):181–92.

28. Kaur C, Ghosh A, Pareek A, Sopory SK, Singla-Pareek SL. Glyoxalases and stress tolerance in plants. Biochem Soc Trans. 2014;42(2):485–90.

29. Kaya C. Nitrate reductase is required for salicylic acid-induced water stress tolerance of pepper by upraising the AsA-GSH pathway and glyoxalase system. Physiol Plant. 2021;172(2):351–70.

30. Paulus C, Köllner B, Jacobsen HJ. Physiological and biochemical characterization of glyoxalase I, a general marker for cell proliferation, from a soybean cell suspension. Planta. 1993;189(4):561–6.

31. Chakravarty TN, Sopory SK. Blue light stimulation of cell proliferation and glyoxalase I activity in callus cultures of Amaranthus paniculatus[J]. Plant Sci. 1998;132(1):63–9.

32. Espartero J, Sánchez-Aguayo I, Pardo JM. Molecular characterization of glyoxalase-I from a higher plant; upregulation by stress. Plant Mol Biol. 1995;29(6):1223–33.

33. Hossain MA, Fujita M. Purification of glyoxalase I from onion bulbs and molecular cloning of its cDNA. Biosci Biotechnol Biochem. 2009;73(9):2007–13.

34. Lin F, Xu J, Shi J, Li H, Li B. Molecular cloning and characterization of a novel glyoxalase I gene TaGly I in wheat (Triticum aestivum L). Mol Biol Rep. 2010;37(2):729–35.

35. Veena, Reddy VS, Sopory SK. Glyoxalase I from Brassica juncea: molecular cloning, regulation and its over-expression confer tolerance in transgenic tobacco under stress. Plant J. 1999;17(4):385–95.

36. Prasanna B, Chandrama P, Upadhyay, Mukesh S, et al. Salt stress alleviation in transgenic Vigna mungo L. Hepper (blackgram) by overexpression of the *glyoxalase I* gene using a novel Cestrum yellow leaf curling virus (CmYLCV) promoter[J]. Mol Breeding. 2008;22(2):169–81.

37. Deb RS, Mukesh S, Bhomkar PS, Mikhail P, Thomas H, Neera BS. Generation of marker free salt tolerant transgenic plants of Arabidopsis thaliana using the gly I gene and cre gene under inducible promoters[J]. Plant Cell. 2008(1):95.

38. Rajwanshi R, Kumar D, Yusuf MA, DebRoy S, Sarin NB. Stress-inducible overexpression of glyoxalase I is preferable to its constitutive overexpression for abiotic stress tolerance in transgenic Brassica juncea[J]. Mol Breeding. 2016;36(6):1–15.

39. Verma M, Verma D, Jain RK, Sopory SK. R WU. Overexpression of *glyoxalase I* gene confers salinity tolerance in transgenic japonica and indica rice plants[J]. 2005.

40. Zimaro T, Gottig N, Garavaglia BS, Gehring C, Ottado J. Unraveling plant responses to bacterial pathogens through proteomics. J Biomed Biotechnol. 2011;2011:354801.

41. Liu S, Liu W, Lai J, Liu Q, Zhang W, Chen Z. *OsGLYI3*, a glyoxalase gene expressed in rice seed, contributes to seed longevity and salt stress tolerance. Plant Physiol Biochem. 2022;183:85–95.

42. Jackson AP. Gene family phylogeny and the evolution of parasite cell surfaces. Mol Biochem Parasitol. 2016;209(1–2):64–75.

43. Grice LF, Gauthier MEA, Roper KE, Fernàndez-Busquets X, Degnan SM, Degnan BM. Origin and evolution of the sponge aggregation factor Gene Family. Mol Biol Evol. 2017;34(5):1083–99.

44. Rijnkels M. Multispecies comparison of the casein gene loci and evolution of casein gene family. J Mammary Gland Biol Neoplasia. 2002;7(3):327–45.

45. Faddeeva-Vakhrusheva A, Derks MF, Anvar SY, Agamennone V, Suring W, Smit S, et al. Gene Family Evolution reflects adaptation to Soil Environmental Stressors in the genome of the Collembolan Orchesella cincta. Genome Biol Evol. 2016;8(7):2106–17.

46. Zhou M, Sun G, Sun Z, Tang Y, Wu Y. Cotton proteomics for deciphering the mechanism of environment stress response and fiber development. J Proteom. 2014;105:74–84.

47. Naeem M, Iqbal M, Ul-Allah S, Chaudhary HJ, Nazeer W, Ashraf J, et al. Expression studies of stress responsive genes in cotton *Gossypium hirsutum L*. Mol Biol Rep. 2021;48(11):7077–85.

48. Khan A, Pan X, Najeeb U, Tan DKY, Fahad S, Zahoor R, et al. Coping with drought: stress and adaptive mechanisms, and management through cultural and molecular alternatives in cotton as vital constituents for plant stress resilience and fitness. Biol Res. 2018;51(1):47.

49. Ahmad I, Zhu G, Zhou G, Song X, Hussein Ibrahim ME, Ibrahim Salih EG, et al. Pivotal role of Phytohormones and their responsive genes in Plant Growth and their signaling and transduction pathway under salt stress in cotton. Int J Mol Sci. 2022;23(13):7339.

50. An M, Wang X, Chang D, Wang S, Hong D, Fan H, et al. Application of compound material alleviates saline and alkaline stress in cotton leaves through regulation of the transcriptome. BMC Plant Biol. 2020;20(1):462.

51. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.

52. Savinkova LK, Ponomarenko MP, Ponomarenko PM, Drachkova IA, Lysova MV, Arshinova TV, et al. TATA box polymorphisms in human gene promoters and associated hereditary pathologies. Biochem (Mosc). 2009;74(2):117–29.

53. Nakashima K, Yamaguchi-Shinozaki K. ABA signaling in stress-response and seed development. Plant Cell Rep. 2013;32(7):959–70.

54. Wang C, Lin T, Wang M, Qi X. An AC-Rich Bean element serves as an Ethylene-Responsive element in Arabidopsis. Plants (Basel). 2020;9(8):1033.

55. Ruíz-Rivero OJ, Prat S. A-308 deletion of the tomato LAP promoters is able to direct flower-specific and MeJA-induced expression in transgenic plants. Plant Mol Biol. 1998;36(5):639–48.

56. Bastian R, Dawe A, Meier S, Ludidi N, Bajic VB, Gehring C. Gibberellic acid and cGMP-dependent transcriptional regulation in Arabidopsis thaliana. Plant Signal Behav. 2010;5(3):224–32.

57. Stange C, Ramírez I, Gómez I, Jordana X, Holuigue L. Phosphorylation of nuclear proteins directs binding to salicylic acid-responsive elements. Plant J. 1997;11(6):1315–24.

58. Gangappa SN, Maurya JP, Yadav V, Chattopadhyay S. The regulation of the Z- and G-box containing promoters by light signaling components, SPA1 and MYC2, in Arabidopsis. PLoS ONE. 2013;8(4):e62194.

59. Zhou SM, Kong XZ, Kang HH, Sun XD, Wang W. The involvement of wheat F-box protein gene TaFBA1 in the oxidative stress tolerance of plants. PLoS ONE. 2015;10(4):e0122117.

60. Olive MR, Walker JC, Singh K, Dennis ES, Peacock WJ. Functional properties of the anaerobic responsive element of the maize Adh1 gene. Plant Mol Biol. 1990;15(4):593–604.

61. Chatterjee MT, Khalawan SA, Curran BPG. Cellular lipid composition influences stress activation of the yeast general stress response element (STRE). Microbiol (Reading). 2000;146(Pt 4):877–84.

62. Wang CT, Ru JN, Liu YW, Li M, Zhao D, Yang JF, et al. Maize WRKY transcription factor ZmWRKY106 confers Drought and Heat Tolerance in transgenic plants. Int J Mol Sci. 2018;19(10):3046.

63. Li W, Chen Y, Ye M, Lu H, Wang D, Chen Q. Evolutionary history of the C-repeat binding factor/dehydration-responsive element-binding 1 (CBF/DREB1) protein family in 43 plant species and characterization of CBF/DREB1 proteins in Solanum tuberosum. BMC Evol Biol. 2020;20(1):142.

64. Tanin MJ, Saini DK, Sandhu KS, Pal N, Gudi S, Chaudhary J, et al. Consensus genomic regions associated with multiple abiotic stress tolerance in wheat and implications for wheat breeding. Sci Rep. 2022;12(1):13680.

65. Wu CY, Suzuki A, Washida H, Takaiwa F. The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants. Plant J. 1998;14(6):673–83.

66. Voigt RM, Forsyth CB, Green SJ, Engen PA, Keshavarzian A. Circadian rhythm and the gut Microbiome. Int Rev Neurobiol. 2016;131:193–205.

67. Guerriero G, Martin N, Golovko A, Sundström JF, Rask L, Ezcurra I. The RY/Sph element mediates transcriptional repression of maturation genes from late maturation to early seedling growth. New Phytol. 2009;184(3):552–65.

68. Fang DD. Cotton fiber: physics, chemistry and biology. 2018.

69. Zafar MM, Rehman A, Razzaq A, Parvaiz A, Mustafa G, Sharif F, et al. Genome-wide characterization and expression analysis of *erf* gene family in cotton. BMC Plant Biol. 2022;22(1):134.

70. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45(D1):D362–8.

71. Hasanuzzaman M, Nahar K, Hossain MS, Mahmud JA, Rahman A, Inafuku M, et al. Coordinated actions of glyoxalase and antioxidant defestems in conferring abiotic stress tolerance in plants. Int J Mol Sci. 2017;18(1):200.

72. Hernandez-Garcia CM, Finer JJ. Identification and validation of promoters and cis-acting regulatory elements. Plant Sci. 2014;217–8:109 – 19.

73. Zhou DX. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. Trends Plant Sci. 1999;4(6):210–4.

74. Hu Y, Chen J, Fang L, Zhang Z, Ma W, Niu Y, et al. *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. Nat Genet. 2019;51(4):739–48.

75. Du X, Huang G, He S, Yang Z, Sun G, Ma X, et al. Resequencing of 243 diploid cotton accessions based on an updated a genome identifies the genetic basis of key agronomic traits. Nat Genet. 2018;50(6):796–802.

76. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, et al. Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature. 2012;492(7429):423–7.

77. Zhao W, Shafiq S, Berr A, Shen WH. Genome-wide gene expression profiling to investigate molecular phenotypes of Arabidopsis mutants deprived in distinct histone methyltransferases and demethylases. Genom Data. 2015;4:143–5.

78. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The pfam protein families database in 2019. Nucleic Acids Res. 2019;47(D1):D427–32.

79. Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, et al. SMART 4.0: towards genomic data integration. Nucleic Acids Res. 2004;32(Database issue):D142–4.

80. Li W, Sun K, Ren Z, Song C, Pei X, Liu Y, et al. Molecular evolution and stress and phytohormone responsiveness of *SUT* genes in *Gossypium hirsutum*. Front Genet. 2018;9:494.

81. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative Toolkit developed for interactive analyses of big Biological Data. Mol Plant. 2020;13(8):1194–202.

82. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. Curr Protoc Bioinformatics. 2002;Chap. 2:Unit 2.3.

83. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–9.

84. Subramanian B, Gao S, Lercher MJ, Hu S, Chen WH. Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. Nucleic Acids Res. 2019;47(W1):W270–5.

85. Cui Y, Su Y, Wang J, Jia B, Wu M, Pei W, et al. Genome-wide characterization and analysis of *CIPK* Gene Family in two cultivated Allopolyploid Cotton Species: sequence variation, association with seed oil content, and the role of *GhCIPK6*. Int J Mol Sci. 2020;21(3):863.

86. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods. 2001;25(4):402–8.

Publisher's Note