Environmental Sciences Europe

RESEARCH

Open Access

A flavonol synthase *(FLS)* gene, *GhFLS1*, was screened out increasing salt resistance in cotton

Mingge Han[1,2†], Ruifeng Cui[2†], Yupeng Cui[2†], Junjuan Wang[2†], Shuai Wang[2†], Tiantian Jiang[2], Hui Huang[2], Yuqian Lei[2], Xiaoyu Liu[1,2], Cun Rui[2], Yapeng Fan[2], Yuexin Zhang[2], Kesong Ni[1,2], Liangqing Sun[2], Xiugui Chen[2], Xuke Lu[2], Delong Wang[2], Zujun Yin[2], Chao Chen[2], Lixue Guo[2], Lanjie Zhao[2], Quanjia Chen[1*] and Wuwei Ye[1,2*]

Abstract

Background Flavonols play important roles in antioxidation and anticancer activities, longevity, and cardiovascular protection. *Flavonol synthase* (*FLS*) is a key enzyme for flavonol synthesis.

Result Phenotypic, transcriptional and metabolic data were analyzed, which showed that there was a close relationship between salt stress and flavonoids, and flavonols were significantly upregulated under salt stress. Nine, seven, four, and four *FLS* genes were identified in *Gossypium hirsutum*, *Gossypium barbadense*, *Gossypium arboreum*, and *Gossypium raimondii*, respectively. The results of subcellular localization showed that FLS existed in the nucleus and cytoplasmic. Through phylogenetic analysis, 24 *FLS* genes were divided into three subfamilies. The results of the RNA sequencing showed that the expression of *GhFLS* genes was mainly induced by salt, drought, low temperature, and heat stress. *GhFLS* promoter mainly comprised plant hormone response elements and abiotic stress elements, indicating that the *GhFLS* gene may play a key role in abiotic stress response. The proline contents of pYL156:GhFLS1 was reduced significantly compared to pYL156 under salt stress, thereby reducing the resistance of cotton to salt stress.

Conclusion This study will lay a foundation for further study on the antioxidant regulation mechanism of the *FLS* gene under abiotic stress.

Keywords *GhFLS1*, NaCl stress, Flavonol, Expression level, VIGS

†Mingge Han, Ruifeng Cui, Yupeng Cui, Junjuan Wang and Shuai Wang contributed equally to this work.

*Correspondence:
Quanjia Chen
chqjia@126.com
Wuwei Ye
yew158@163.com
[1] Engineering Research Centre of Cotton, Ministry of Education, College of Agriculture, Xinjiang Agricultural University, Urumqi 830052, China
[2] Institute of Cotton Research of Chinese Academy of Agricultural Sciences/Research Base, Anyang Institute of Technology, National Key Laboratory of Cotton Bio-Breeding and Integrated Utilization, Anyang 455000, China

Introduction

Soil salinization is an important problem facing agriculture worldwide. Salt stress is a major environmental factor limiting plant distribution, growth, and crop production [1]. Salt accumulation in cultivated land soil mainly comes from irrigation water and seawater containing trace NaCl [2]. Plants absorb high concentrations of salt through their roots, which then causes damage to plants. High salinity stress induces osmotic stress, ionic stress, and excessive production of reactive oxygen species (ROS) [3, 4]. Therefore, plants resist salt stress mainly through antioxidation, ion antagonism, and efflux. Using exogenous substances to resist salt stress is an important

Springer Open

© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

way for plants to resist salt stress. In plants, antioxidant enzymes and specialized metabolites with antioxidant activity play a key role in allowing productive ROS signaling by preventing ROS from reaching damaging levels [5].

Flavonoids have antioxidant, anti-inflammatory, and anti-proliferative properties, which help protect human beings from cancer and cardiovascular disease [6, 7]. Flavonoids are also polyphenolic secondary metabolites with C6–C3–C6 carbon skeletons that are biosynthetic through the phenylpropanoid pathway [8]. Flavonoids play an important role in regulating development and stress responses [5, 9]. Flavonoids can improved salt tolerance by removing excess ROS in soybeans [10]. Flavonoids with radical scavenging activity mitigate against oxidative and drought stress in *Arabidopsis thaliana* [11]. Flavonols are the main flavonoid compound in plants. It is reported that flavonols have significant health-related biological activities, including antioxidation [12], anticancer [13], increased longevity [14], and cardiovascular protection properties [15]. Flavonols can be divided into three subclasses, namely, kaempferol, quercetin, and myricetin, according to the hydroxylation mode of its flavonol bring. The content of kaempferol significantly increased under salt-alkali stress [16]. Most flavonols undergo various modifications, such as glycosylation and methylation, resulting in a large number of different molecules [17]. Quercetin is an effective anti-osmotic agent that can reduce the adverse effects of mannitol-induced osmotic stress on seed germination and seed vitality [18]. Each subclass of flavonol shows different spatial and temporal distribution and accumulation patterns, and is affected by environmental factors [19]. Flavonols have a variety of physiological functions in plants, such as ultraviolet protection, regulating auxin transport, the promotion of male fertility [20, 21], and the deposition of pigment and production of anthocyanin [22]. Flavonol could reduce the level of ROS and affect the development of guard cell, roots and leaves [23–25]. Flavonol regulates lateral root germination through scavenging reactive oxygen species in *Arabidopsis thaliana* [26]. Proline has been proven to be responsible for scavenging ROS and other free radicals [27]. There was a positive correlation between ABA content and proline synthesis [28]. Abscisic acid-induced reactive oxygen species were modulated by flavonols to control stomata aperture [29].

Flavonol synthase (*FLS*) is a key enzyme specific to the flavonol pathway. It competes with dihydroflavonol 4-reductase (*DFR*) for dihydroflavonol as a substrate [30]. Therefore, the competition between *FLS* and *DFR* enzyme activities regulates different branches of the flavonoid biosynthesis pathway [31, 32]. The *FLS* genes convert dihydroflavonol into the corresponding flavonol by introducing a double bond between C-2 and C-3 of the C-ring [33]. Most genes of the central enzyme in flavonoid biosynthesis are encoded by single-copy genes. Only *FLS1* encodes a functional *FLS* and thus is the major contributor to flavonol production in *A. thaliana* [34]. *FLS* is classified as a 2-hydroxyglutarate-dependent dioxygenase (2OGD), similar to flavonoid 3-hydroxylase (*F3H*) and anthocyanin reductase (*ANS*). These three enzymes showed partial amino acid sequence similarity and overlapping functions.

The first *FLS* cDNA was cloned from petunia (*Petunia hybrida*), and was functionally expressed in yeast and plants [35]. The other *FLS* genes have been identified and characterized in various plant species, including *A. thaliana* [36, 37], *Citrus unshiu* [38], *Glycine max* [39], *Zea mays* [40], *Camellia nitidissima* [41], and *Litchi chinensis* [42]. For example, *OsFLS* was expressed in both non-pigmented and pigmented rice seeds and was subject to developmental regulation during seed maturation [43]. Flavonol synthesized by nucleus *FLS1* was found to play a role in *Arabidopsis* resistance to Pb stress [44]. The heterotopic expression of *DoFLS1* in *Dendrobium candidum* enhanced the accumulation of flavonols in *A. thaliana* and the tolerance to abiotic stress [45]. *FLS* was involved in auxin transport and protection against environmental stresses [46].

Cotton is a crop with strong salt tolerance and an ideal pioneer crop for improving saline-alkali land [47]. However, the flavonol synthesis-related *FLS* gene in cotton has not been systematically analyzed and identified. The measurement of flavonoid content, analysis of existing transcriptome and metabolome data, bioinformatics analysis, subcellular localization and virus induced gene silencing (VIGS) were used to reveal the function of *FLS* under salt stress. The physicochemical properties, gene structure, phylogenetic evolution, and *cis*-acting elements were analyzed. These will provide molecular basis and reference for further exploring the relationship between the synthesis gene *FLS* of flavonols and salt stress in cotton.

Materials and methods

Plant materials and treatment methods

The plant material was the salt-tolerant cotton variety Zhong9807 [48]. The cultivation of cotton adopted the sand culture method. Cotton seeds were sown in 3/2 of the seedling bowls filled with sand. The seedling bowl was placed in a growth box with an alternating cycle of 28 ℃/14 h light and 25 ℃/10 h darkness. When the cotton seedlings grew to the three-leaf stage, 400 mM NaCl was applied to the bowl until it was saturated.

Determination of total flavonoid content under salt stress

The leaves were cut and sampled 0 h and 48 h after the beginning of salt stress, and then placed in an oven at 115 ℃ for 10 min. The leaves were then dried at 80℃ until the weight was constant. The flavonoid content in cotton leaves was measured according to the instructions of the Solarbio kit (BC1330) from Beijing. First, the dried leaves were ground into powder and then weighed to 0.1 g through a 50-mesh sieve. One milliliter extraction solution was added and extraction was performed with the ultrasonic extraction method. Extraction was performed for 30 min, 120,000 rpm, and 25 ℃, centrifugation was conducted for 10 min, and then the supernatant was collected and the extraction solution was added to the supernatant until it reached a volume of 1 ml for determination. The flavonoid content was calculated as follows: flavonoid content (mg/g) $= x/w$, where x represents the sample concentration and w represents the sample quality.

Data analysis

For the analysis of differential genes, the transcriptome data obtained in the laboratory under salt stress were used [48]. A powerful analytical method called gene set enrichment analysis (GSEA) was used to interpret the gene expression data [49]. This method derives its power from focusing on gene sets, or groups of genes that share common biological functions. GSEA was performed using the GSEA software (https://www.broadinstitute.org/gsea/). The metabolic expression data analysis of 0 h and 48 h leaves was also based on the metabolic data obtained in our laboratory. A widely targeted metabolite analysis was performed for comprehensive of flavonoids in cotton leaves based on LC–MS/MS. Flavonoids are the primary classification of substances, and the secondary classification of flavonoids includes several major categories: dihydroflavonols, flavanols, flavanols, chalcones, anthocyanins, flavonoid, and flavonoid carbonoside. We compared and analyzed the metabolomic data of salt treatment for 48 h (L48) and 0 h (L0). All histograms were designed with GraphPad Prism v8.0.2.263 software. The SPSS 26.0 statistical software was used to analyze the significant differences between the control and treatments based on one-way analysis of variance at $p < 0.05$ (*) or $p < 0.01$ (**).

Database download source

The genome files and protein sequences of *Gossypium arboreum* (*Ga*) (CRI), *Gossypium barbadense* (*Gb*) (ZJU), *Gossypium hirsutum* (*Gh*) (ZJU), and *Gossypium raimondii* (*Gr*) (JGI) were downloaded from the Cotton Functional Genomics Database (CottonFGD) (https://cottonfgd.net/). Genome data of other 10 species including *Arabidopsis thaliana* (*At*), *Arabidopsis halleri* (*Ah*), *Vitis vinifera* (*Vv*), *Arabidopsis lyrata* (*Al*), *Capsella grandiflora* (*Cg*), *Theobroma cacao* (*Tc*), *Glycine max* (*Gm*), *Boechera stricta* (*Bs*), *Oryza sativa* (*Os*), and *Zea mays* (*Zm*) were obtained from the Phytozome plant database (https://phytozome-next.jgi.doe.gov/blast-search). The Local Blast Alignment Tool was downloaded from the National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov/). The hmmer program, which was used to search the domain, was downloaded from Hmmer (http://www.hmmer.org/).

Identification of *FLS* gene family members

Six *Arabidopsis FLS* family members have been published [50]. This study downloaded the *FLS* family members from the *Arabidopsis* website (https://www.arabidopsis.org/), namely, *AtFLS1*, *AtFLS2*, *AtFLS3*, *AtFLS4*, *AtFLS5*, and *AtFLS6*. According to the protein sequence of *A. thaliana*, it was known that the *FLS* family members contained the PF03171 and PF14226 domains. The Hidden Markov Model (HMM) of FLS protein (PF03171 and PF14226 in PFAM) was downloaded from the PFAM database (http://pfam.xfam.org/), and Hmmer was used to search all possible members of the *FLS* gene family. A local BLAST search was used to identify family members (E-value $< e^{-7}$). Then, the common genes obtained using the two methods were selected as candidate genes. To confirm these genes, these sequences were further verified using the NCBI CD Search Tool (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and the SMART database (http://smart.embl-heidelberg.de/). Finally, sequences that did not belong to the conservative binding domain and that had incomplete C and N terminals were manually deleted.

Sequence alignment and phylogenetic analysis

The full-length amino acid sequences of 14 plants, namely, *G. arboreum*, *G. barbadense*, *G. hirsutum*, *G. raimondii*, *A. thaliana*, *A. halleri*, *V. vinifera*, *A. lyrata*, *C. grandiflora*, *T. cacao*, *G. max*, *B. stricta*, *O. sativa*, and *Z. mays* were aligned using the Cluster W program with the default settings. Then, a neighborhood-joining (NJ) tree was built with 1000 bootstrap replicates using the p-distance model with default parameters in MEGA 7.0 [51]. The EvolView website (https://www.evolgenius.info/evolview/#/treeview) was used to beautify the phylogenetic tree [52].

FLS chromosome positions in four species of cotton

The chromosome positions of *G. arboreum*, *G. barbadense*, *G. hirsutum*, and *G. raimondii* were mapped

with TBtools software [53]. The reference genome GFF3 file was downloaded from CottonFGD.

Collinearity analysis of the *FLS* family in four species of cotton

To study the collinearity of four cotton *FLS* families and analyze their collinearity relationships, the whole genome sequence and genome annotation files of these cotton varieties were obtained using the MCScanX tool. The collinear and homologous chromosome regions of four species of cotton were visualized using the advanced circos software package. Gene duplication was evaluated using MCScanX. To visualize the duplicated regions of four species of cotton, TBtools was used to draw spectral lines between the repetitive genes in Circos [53].

Calculation of selective pressure

To study the selection pressure experienced by *FLS* repeat gene pairs from the four cotton species, the *Ka/Ks* calculator in TBtools was used to calculate synonymous (*Ks*) and non-synonymous (*Ka*) replacement rates and their ratios.

Analysis of the conserved protein motifs and gene structures

The MEME website (https://meme-suite.org/meme/tools/meme) was used to predict gene motifs. The parameters were as follows: the maximum number of motifs was 15, and the other parameters were set by default. The files for this domain were obtained from the hmmersearch online website (https://www.ebi.ac.uk/Tools/hmmer/). The evolutionary relationship, gene structure, domain, and motif composition of genes were mapped using TBtools software.

Analysis of the *GhFLS* promoter region and different expression patterns

The 2000 bp DNA sequence of the upstream region of GhFLSs was derived from the CottonFGD database (https://cottonfgd.net/). The predicted *cis*-acting elements related to abiotic stresses and plant hormones in the promotor regions of the GhFLSs were obtained from the PlantCARE website (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) for further analysis. RNA Sequencing data (PRJNA490626) from the NCBI (https://www.ncbi.nlm.nih.gov/) were used to analyze the expression level (Fragments Per Kilobase of exon model per Million mapped fragments, FPKM) of GhFLSs in 4-week-old seedlings under cold (4 °C), heat (37 °C), salt (0.4 M NaCl), and polyethylene glycol (PEG, 200 g/L) stress for 1 h, 3 h, 6 h, and 12 h, respectively [54]. RNA Sequencing data from the NCBI (PRJNA559592) [48] were analyzed to determine the FPKM of GhFLSs in different time periods under 400 mM NaCl treatment. Finally, TBtools software was used to draw an image of an evolutionary tree, *cis*-acting elements, and an expression level heat map for visual observation.

Verification of relative gene expression using by qRT-PCR

RNA was extracted and reversed-transcribed into cDNA as a template for qRT-PCR. The primers for the qRT-PCR of GhFLSs were designed on the GenScript website (https://www.genscript.com/tools/real-time-pcr-taqman-primer-design-tool) (Additional file 1: Table S1). According to the instructions provided by the manufacturer of TransStart Top Green qPCR Supermix reagent (TransGene Biotechnology Co., Ltd, Beijing, China), qRT-PCR was performed on an Applied Biosystem ABI7500 Fast Real-Time PCR platform, and the experiment was conducted in three independent replicates. The *GhActin* (AY305733) gene was used as the internal reference gene, and $2^{-\Delta\Delta Ct}$ was used to calculate the relative expression level of *GhFLS*.

Interaction network analysis and subcellular localization of GhFLS1

The STRING database (https://cn.string-db.org/) was used to analyze the GhFLS protein interaction network. The interactions between GhFLS1 and other genes in cotton were predicted based on *A. thaliana* homology. The Wolf-PSORT and ProtComp 9.0 websites were used to predict the subcellular location of GhFLS. A GhFLS1:121-GFP vector was constructed, and transformed into tobacco leaves.

VIGS of *GhFLS1*

To verify the function of *FLS* genes, this study selected a highly expressed gene, *GhFLS1* (GH_A05G2328). VIGS fragments of 300 bp were designed using SGN-VIGS (https://vigs.solgenomics.net/). The fragment was connected to the pYL156 vector and the recombinant vector was transformed into *Agrobacterium tumefaciens* GV3101. GV3101 bacterial solution containing the control pYL156 (empty vector), pYL156: GhFLS1, pYL156: PDS (positive control), and pYL192 (auxiliary carrier) was injected into the cotyledons of cotton variety Zhong9807. After dark treatment for 24 h, cotton was grown in an incubator containing 25 °C/16 h light and 23 °C/8 h dark circulation culture. NaCl stress treatments were performed at the three-leaf stage and samples were quickly frozen with liquid nitrogen.

After 400 mM NaCl treatment, 0.1 g samples were obtained and ground with 1:9 medium homogenate. The proline content was determined using a proline content detection kit (Nanjing Jiancheng Bioengineering Research Institute, A107-1-1).

Results

Phenotype, total flavonoid content, and expression analysis of cotton under salt stress

Cotton cotyledons showed wilting and water loss under salt stress (Fig. 1A). Analysis of the total flavonoids content in leaves revealed, significant differences between the treatment and the control (Fig. 1B). GSEA analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) enriched with the transcriptome data showed that flavonoid metabolism-related genes were significantly

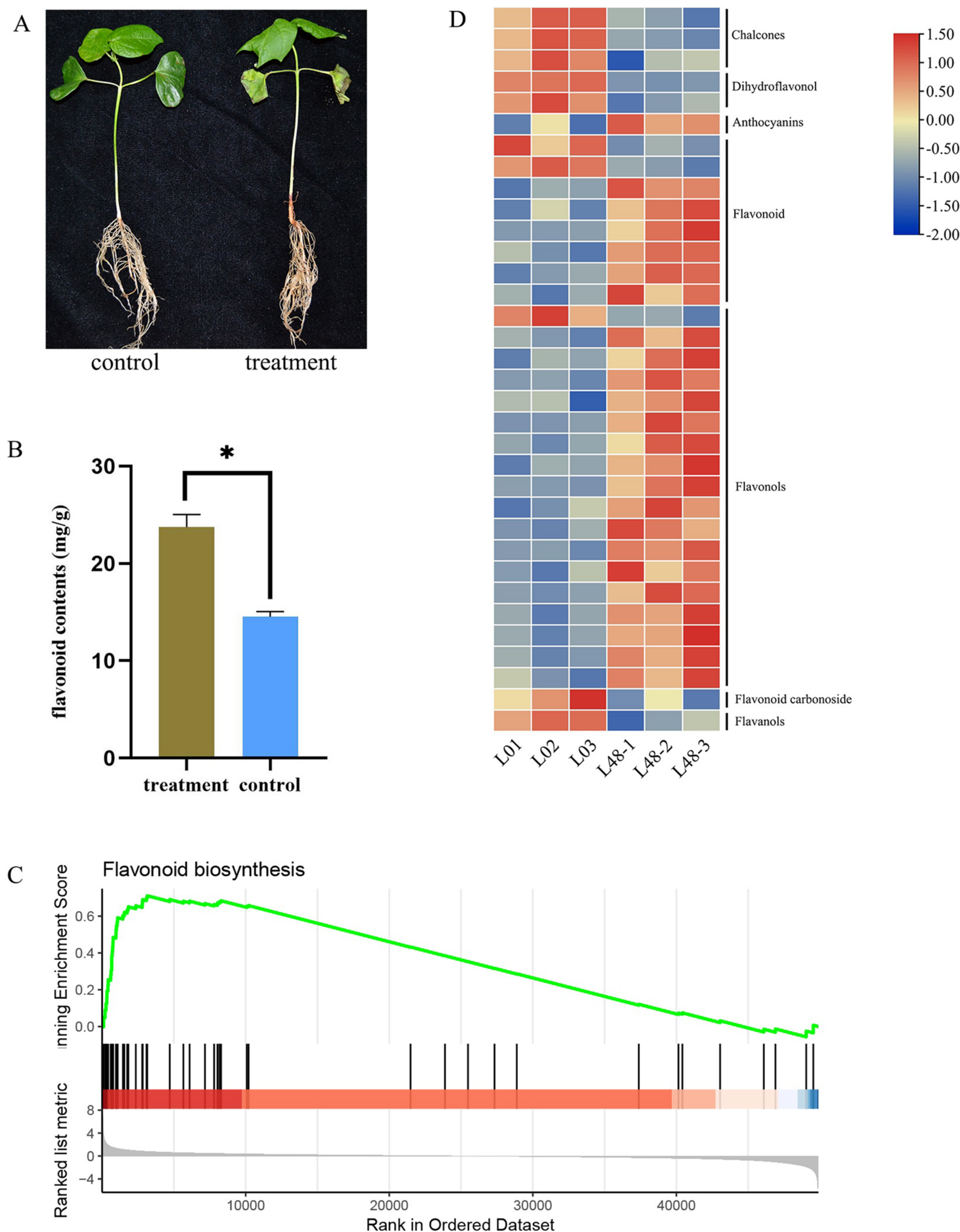

Fig. 1 Phenotype, total flavonoid content, and expression analysis of cotton under salt stress. **A** Phenotype of cotton under salt stress. **B** Total flavonoid content of cotton under salt stress. **C** Gene set enrichment analysis (GSEA) based on RNA sequencing. **D** Content of flavonoid-related substances under salt stress. Flavonoids include dihydroflavonols, flavanols, flavanols, chalcones, anthocyanins, flavonoid, and flavonoid carbonoside. L01, L02, and L03 represented three replicates of the control, respectively. L48-1, L48-2, and L48-3 represented the replicates of the treatment, respectively

expressed (Fig. 1C). Analysis of the metabolomics data showed that flavonols were significantly upregulated in flavonoid-related substances (Fig. 1D). This indicated that there was a close relationship between flavonols and salt stress. *FLS* is a key enzyme involved in the formation of flavonols, which are a subclass of flavonoids [55]. This study speculated that the expression of the *FLS* gene could improve salt stress resistance, and the *FLS* family members were analyzed.

Identification of *FLS* family members

The *FLS* enzyme belongs to the 2OGD superfamily. A total of 24 family members have been identified from the four *Gossypium* species, including four in *G. arboreum*, seven in *G. barbadense*, nine in *G. hirsutum*, and four in *G. raimondii*, via tree building. Genes were renamed according to chromosome position information (Additional file 1: Table S2). The open reading frame (ORF) of 24 *FLS* family members of four cotton species ranged from 981 (*GbFLS7*) to 2270 (*GrFLS1*) bp. The number of amino acid-encoding protein ranged from 326 (*GbFLS7*) to 498 (*GbFLS5*). The isoelectric point (pI) ranged from 5.13 to 7.28, and the molecular weights (MWs) ranged from 37.08 to 56.62 kDa. The number of exons varied from 2 to 5.

A. thaliana, *A. halleri*, *V. vinifera*, *A. lyrata*, *C. grandiflora*, *T. cacao*, *G. max*, *B. stricta*, *O. sativa*, and *Z. mays* were selected to identify *FLS* family members that were closely related to cotton and had been studied more than cotton *FLS* family genes. *FLS* family genes also have been identified in 10 additional plants, including six in *A. thaliana*, six in *G. max*, three in *O. sativa*, three in *Z. mays*, five in *V. vinifera*, three in *B. stricta*, twelve in *C. grandiflora*, four in *T. cacao*, five in *A. lyrata* and nine in *A. halleri*. Then, *FLS* family members were renamed according to their position on the chromosome. The *FLS* family genes of two tetraploid cotton species, *G. barbadense* and *G. hirsutum*, were about twice as many as those of the two diploid cotton species, *G. arboreum* and *G. raimondii*. The *FLS* family members in dicotyledons were more abundant than those in monocotyledons.

Evolutionary tree analysis of *FLS*

To understand the evolutionary relationship of the *FLS* family, 80 protein sequences were used to construct the phylogenetic trees of *G. arboreum*, *G. barbadense*, *G. hirsutum*, *G. raimondii*, *A. thaliana*, *A. halleri*, *V. vinifera*, *A. lyrata*, *C. grandiflora*, *T. cacao*, *G. max*, *B. stricta*, *O. sativa*, and *Z. mays* (Fig. 2). According to sequence similarity, tree topology, gene structure characteristics, and each motif, the *FLS* family was divided into three branches. The results showed that the *FLS* branch had the largest number of group C (31), followed by group A (25), and finally group B (24). Five *GhFLS* genes were distributed in branch A, and group B had three *GhFLS* genes. *AtFLS* only existed in branch group C, and most of *A. halleri* and *A. lyrata* were also distributed in group C, indicating that the FLSs in branch group C were closely related. The branches of *TcFLS* and cotton *FLS* were similar, indicating that cacao was closely related to cotton and may have originated from the same ancestor. The *FLS* gene family members of the four cotton species always gathered together, which indicated that the four cotton species had a close evolutionary relationship.

Chromosome mapping of *FLS* gene family

To understand the distribution of genes on chromosomes more intuitively, this work constructed a physical map of the chromosome distribution of *FLS* gene family members in four cotton species (Fig. 3). Chromosome mapping analysis showed that the distribution of chromosome positions was uneven. FLSs in *G. hirsutum* were distributed in chromosome 5, 8, and 12 of subgenome A and in chromosome D groups 4, 5, and 8. *FLS* gene members of *G. arboreum* were distributed on chromosomes 4, 5, 8, and 12. *FLS* family members of *G. raimondii* were distributed on chromosomes 4, 9, and 12. Compared with *G. hirsutum*, the *FLS* genes were missing at the end of the A05 and D04 chromosomes in *G. barbadense*. It is speculated that this may be caused by gene loss or incomplete genome assembly during the evolution of *G. barbadense*.

Analysis of motifs, domain structures and exon–intron structures of *FLS* genes

This study analyzed the evolutionary relationships, motifs, domains, exons, and introns to study the conservative structure of *FLS* family genes (Fig. 4). All *FLS* family genes had the *DIOX_N* domain and the *20G-FeII_Oxy* domain. This domain could combine with plant hormones to achieve its catalytic function (Fig. 4C). GbFLS5 had two *DIOX_N* institutional domains.

FLSs were classified according to the topology of the evolutionary tree (Fig. 4A). The distribution patterns of exons and introns were related to their biological functions. Their arrangement could be used to analyze the evolutionary relationship between members of different gene families. Interestingly, an intron in the *FLS* in branch I (*GbFLS5*) and an intron in the *FLS* in branch III (*GrFLS1*) were significantly longer than other introns (Fig. 4D). Except for *GbFLS5*, the motifs of *FLS* family gene members were relatively consistent, which means that they have similar functions at the protein level (Fig. 4B). All FLSs contained Motif 1, Motif 2, Motif 3, Motif 4, and Motif 6, which were common conservative domains of the *FLS* family. Motif 9 in branch I and

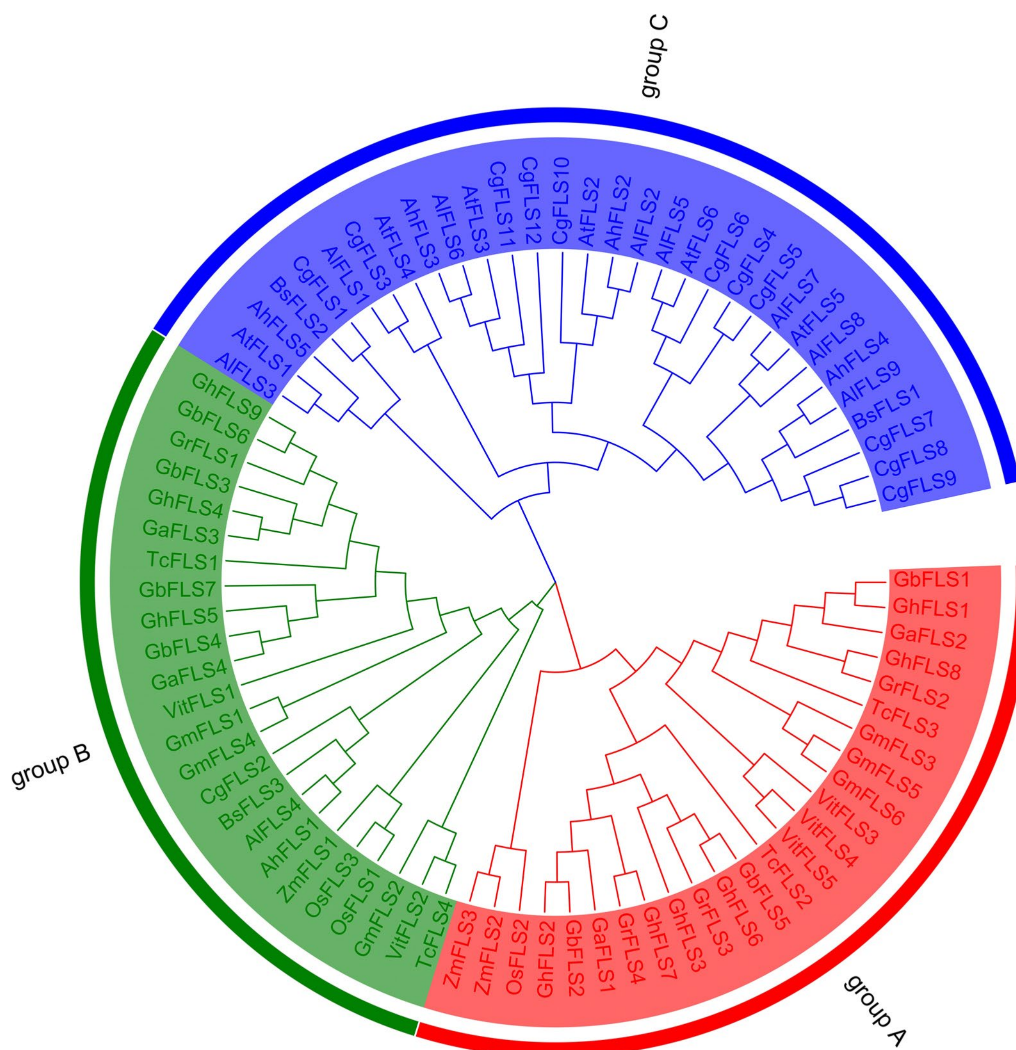

Fig. 2 Phylogeny tree constructed using MEGA 7 by the Neighbor-Joining (NJ) method. Phylogenetic relationship of the 80 identified FLSs from *G. arboreum*, *G. barbadense*, *G. hirsutum*, *G. raimondii*, *A. thaliana*, *A. halleri*, *V. vinifera*, *A. lyrata*, *C. grandiflora*, *T. cacao*, *G. max*, *B. stricta*, *O. sativa*, and *Z. mays*

branch II was at the N-terminus, and Motif 9 of *GaFLS4–GbFLS7* in branch III was at the C-terminus. It was speculated that this may have been due to functional changes during evolution.

Gene replication and collinearity analysis

Gene replication events are considered to play an important role in the amplification of gene families. Gene replication events include whole-genome duplication (WGD), fragment replication, and tandem replication. Most plants have experienced an ancient whole-genome replication event or polyploidy. The replication region caused by WGD is usually the large-scale replication of all genes, rather than the replication of a single gene or multiple genes. Large-scale whole-genome replication and small-scale tandem replication and fragment replication can be identified from collinear fragments, which can be used as inferential data on species evolution. To explore the amplification mechanism of the *FLS* gene family, 95 pairs of homologous gene pairs were identified by comparing the genomes of Ga–Ga, Ga–Gb, Ga–Gh, Gb–Gb, Gb–Gr, Gr–Gr, Gr–Ga, and Gh–Gh. Genes connected by lines of the same color represent the same gene. In Fig. 5, the GhA/GhD and GbA/GbD subgenomes, as well as many chromosomes in the A and D genomes, are connected by lines of the same color, indicating that the GhA/GhD and GbA/GbD subgenomes have FLS homologs in the A and D genomes. These results indicate that these genomes/subgenomes are evolutionarily related and that most *FLS* genes have been

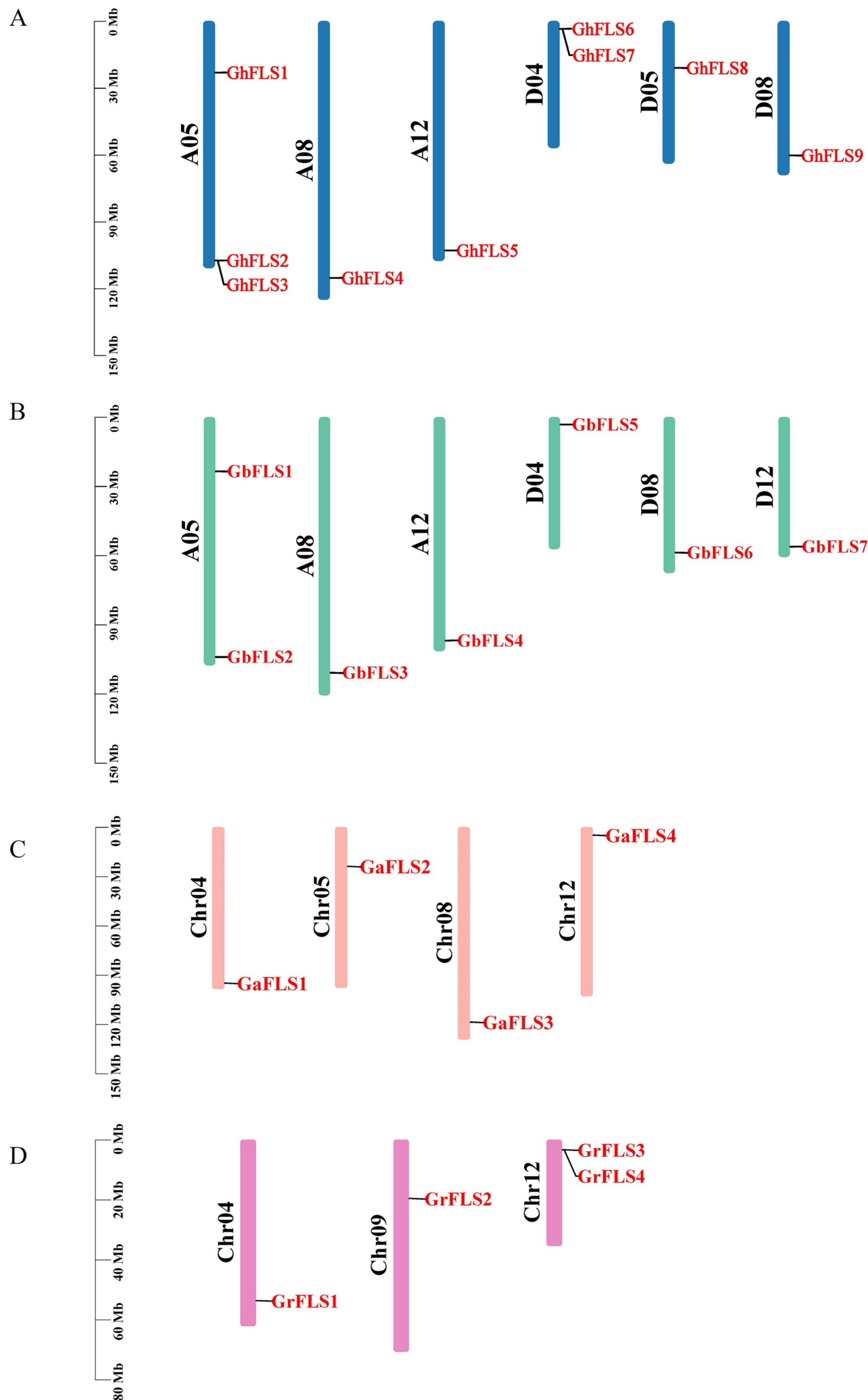

Fig. 3 The chromosome distribution of *FLS* gene family in four cotton *Gossypium*. **A** Chromosomal location of FLSs on chromosomes in *G. hirsutum*. **B** Chromosomal location of FLSs on chromosomes in *G. barbadense*. **C** Chromosomal location of FLSs on chromosomes in *G. arboreum*. **D** Chromosomal location of FLSs on chromosomes in *G. raimondii*. The scale of the genome size was given on the left

Fig. 4 Conservative motifs, domain structures, and exon–intron structures of *FLS* genes from *G. hirsutum*, *G. barbadense*, *G. raimondii*, and *G. arboreum*

retained during polyploid evolution. Eight, eight, one, and one pairs of duplicate *FLS* gene pairs were found in Gh–Gh, Gb–Gb, Ga–Ga, and Gr–Gr, respectively. The 18 pairs of homologous gene pairs were predicted to be fragment duplicates according to chromosome location. Seventy-seven gene pairs had undergone WGD, with 12, 12, 11, and 11 Ga–Gb, Ga–Gh, Gr–Gb, and Gh–Gr gene pairs, respectively. Based on these results, it was speculated that fragment duplication and WGD were the main reasons for the evolution of the *FLS* genes from diploid to tetraploid.

Calculation of selective pressure (*Ka/Ks*) during evolution

During evolution, duplicate gene pairs may deviate from their original functions, resulting in new functionalization (the loss of original functions), sub-functionalization (a division of the original functions), and new functionalization (the acquisition of new functions). To study the driving force of FLS family genes in the evolutionary process, this study calculated the values of *Ka/Ks* synonymous substitution of 70 duplicate gene pairs from four cotton species. The selection pressure of duplicate gene pairs can be inferred according to the *Ka/Ks* ratio. It is generally believed that $Ka/Ks = 1$ represents neutral selection (pseudogene), $Ka/Ks < 1$ represents purifying or negative selection (purifying selection), and $Ka/Ks > 1$ represents positive selection. There were 164 duplicate gene pairs in Ga–Ga, Ga–Gb, Gb–Gb, Gb–Gr, Gh–Ga, Gh–Gb, Gh–Gh, Gh–Gr, and Gr–Gr (Fig. 6). The *Ka/Ks* values of four pairs of genes were greater than 1, indicating that these genes experienced positive selection during evolution (Table 1). Fifty-eight (83%) duplicate genes had a *Ka/Ks* ratio < 0.5, and five (7%) duplicate genes had a *Ka/Ks* ratio between 0.5 and 0.99. These genes were *GbFLS1–GaFLS2*, *GhFLS1–GaFLS2*, *GrFLS1–GbFLS6*, *GrFLS2–GhFLS8*, and *GbFLS3–GhFLS4*. This showed that these FLSs evolved slowly, underwent strong purifying selection pressure, and had limited functional differences after fragment replication and WGD. In Ga–Gb, Gh–Ga, and Gh–Gr, the logarithms of genes with *Ka/Ks* values greater than 1 were 1, 2, and 1, respectively, indicating that these genes were actively selected during evolution and recently underwent rapid evolution. Whether it will result in harmful or beneficial characteristics remains to be studied.

Promoter and expression analysis of *GhFLS* under salt stress

The online promoter website revealed the response of GhFLSs to hormonal and abiotic stresses, which was helpful to further analyze their regulatory networks. GhFLSs were related to plant hormones (abscisic acid, methyl jasmonate, gibberellic acid, auxin, salicylic acid) and various stresses (low temperature, drought, hypoxia, defense, and stress response). GhFLSs also had many MYB binding sites and zein metabolism regulatory elements (Fig. 7). In addition, there were light-regulated promoters.

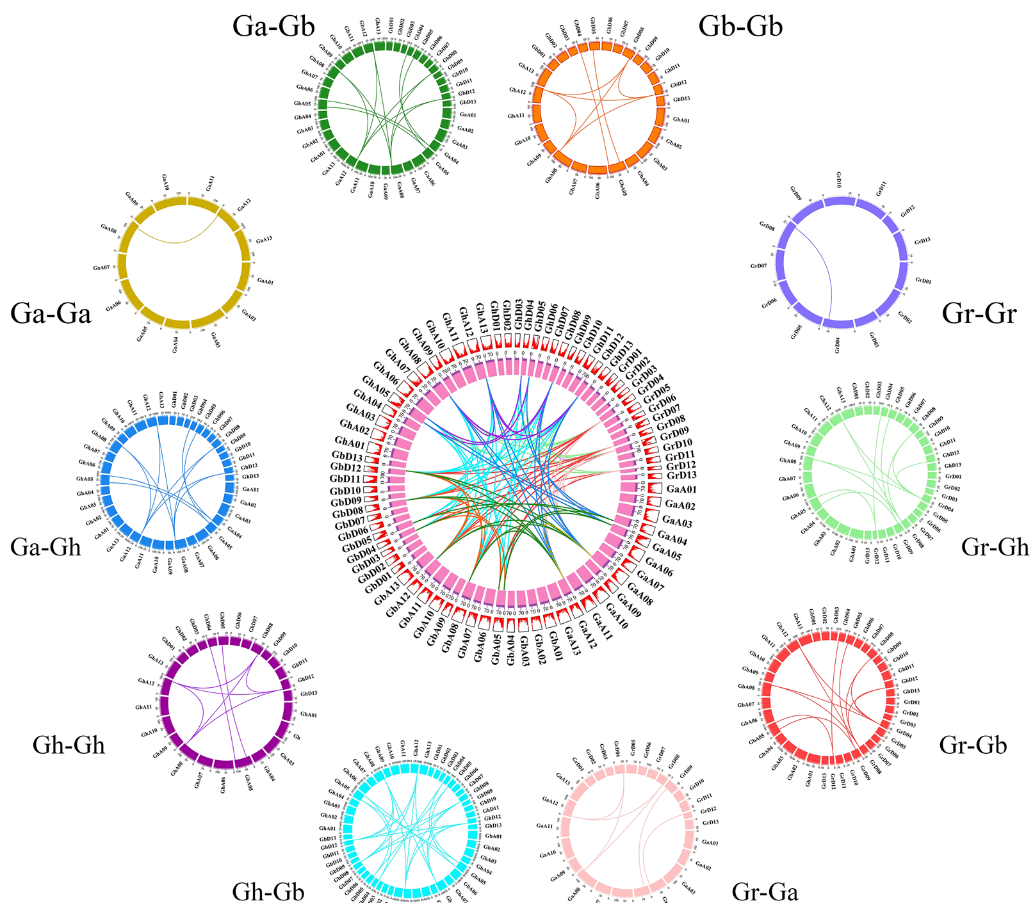

Fig. 5 Collinearity relationship of FLSs duplicated genes pairs from four *Gossypium* (*G. hirsutum*, *G. barbadense*, *G. arboreum*, and *G. raimondii*). Chromosomal lines represented by various colors indicate the syntenic regions around the FLSs. The heatmap and line map of the outer ring represented the density of genes on chromosomes

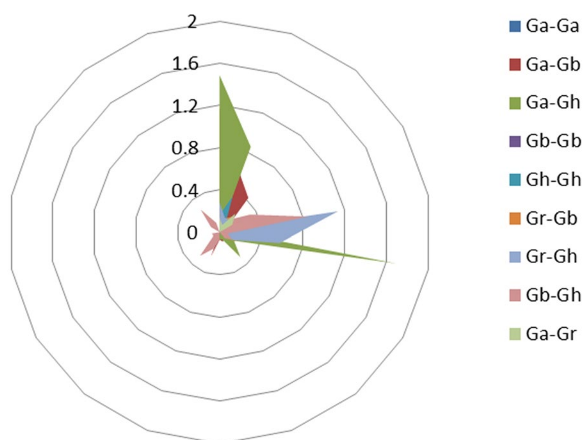

Fig. 6 Prediction of a number of duplicated gene pairs involved indifferent combinations from four *Gossypium* species. Gh represented *G. hirsutum*, Gb represented *G. barbadense*, Ga represented *G. arboreum*, Gr represented *G. raimondii*. Different colors represented *Ka/Ks* gene pairs between Ga–Ga, Ga–Gb, Ga–Gh, Gb–Gb, Gh–Gh, Gr–Gb, Gr–Gh, Gb–Gh, Ga–Gr

Transcriptome data were used to analyze the FPKM values of eight cotton tissues ((root, stem, leaf, torus, petal, stamen, pistil, and calycle tissues). The results showed that FLSs had different expression patterns in different tissues. *GhFLS4* and *GhFLS9* were highly expressed in all tissues, and their expression levels were significantly higher than those of other *FLS* family genes, which may be necessary for maintaining the normal living activities of cotton. *GhFLS1* was specifically expressed in the stamen tissue, and it speculated that it may play an important role in stamen development. These results indicate that FLSs have tissue-specific expression under normal growth conditions.

The expression level of *GhFLS1*, *GhFLS8*, and *GhFLS5* changed significantly at 12 h after exposure to salt stress. The expression of *GhFLS1* and *GhFLS8* decreased at 12 h, while the expression of *GhFLS5* increased at 12 h (Additional file 1: Fig. S1). The expression of *GhFLS7* was significantly decreased after exposure to cold stress, the expression *GhFLS4* and *GhFLS9* was significantly

Table 1 Prediction of the number of duplicated gene pairs involved in different genomes of four cotton species

Pairs	Positive selection No.	Neutral selection No.	Pure selection No.		Total	Purify %
			0.5–0.99	0–0.49		
Ga–Ga	0	0	0	1	1	100
Ga–Gb	1	0	1	9	11	90.91
Ga–Gh	2	0	1	7	10	80
Gr–Gr	0	0	0	0	0	0
Gh–Gh	0	0	0	5	5	100
Gr–Gb	0	0	1	6	7	100
Gr–Gh	1	0	1	5	7	85.71
Gb–Gh	0	0	1	14	15	100
Ga–Gr	0	0	0	4	4	100
Gb–Gb	0	0	0	7	7	100
Total	4	0	5	58	67	94.03

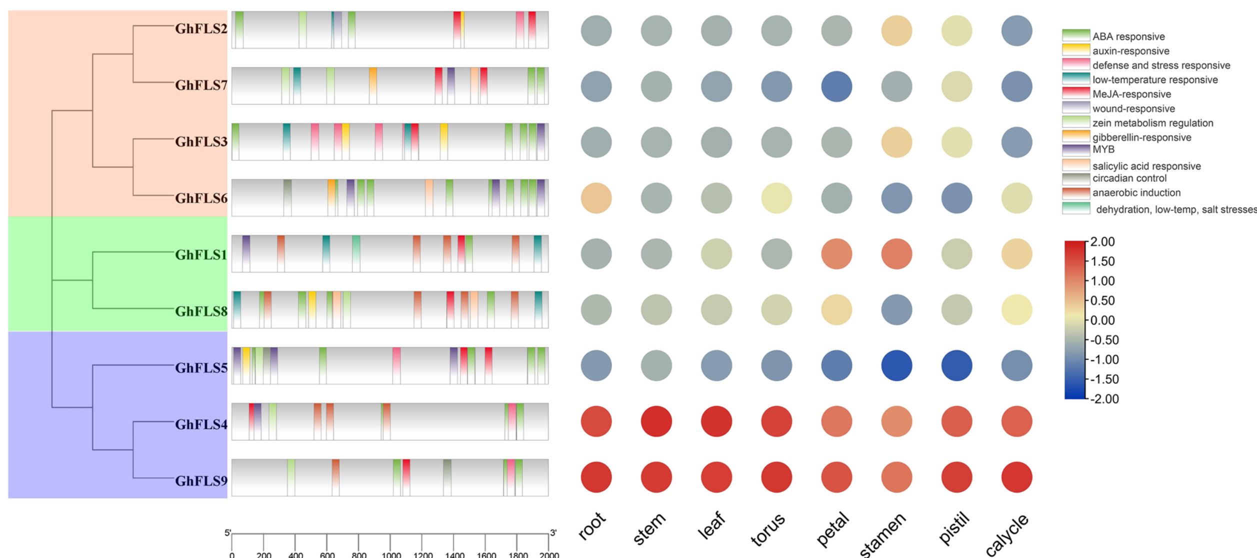

Fig. 7 Analysis of promoters and differentially expressed of *GhFLS* gene family

increased after exposure to PEG stress, and the expression of *GhFLS4* and *GhFLS9* was significantly decreased after exposure to heat stress. This demonstrated that different *FLS* family members showed different expression patterns under different abiotic stresses.

Expression patterns of GhFLSs under 400 mM NaCl stress

To study the response of GhFLSs to abiotic stress, this study examined the expression changes of the *FLS* genes in *G. hirsutum* leaves under 400 mM NaCl stress (Fig. 8). Under NaCl stress, the expression levels of *GhFLS1* and *GhFLS8* were significantly increased at 3 h, 12 h, 24 h, and 48 h after exposure to salt stress, and these genes continued to be overexpressed. The expression of the *GhFLS2* and *GhFLS4* genes increased significantly only 24 h and 48 h after exposure to stress. The expression of the *GhFLS3*, *GhFLS5*, and *GhFLS7* genes did not change significantly under salt stress. The expression of *GhFLS6* was significantly reduced at 12 h, 24 h, and 48 h under salt stress. The results showed that different *FLS* family members had different expression patterns under NaCl stress, and the expression patterns were different under different salt concentrations.

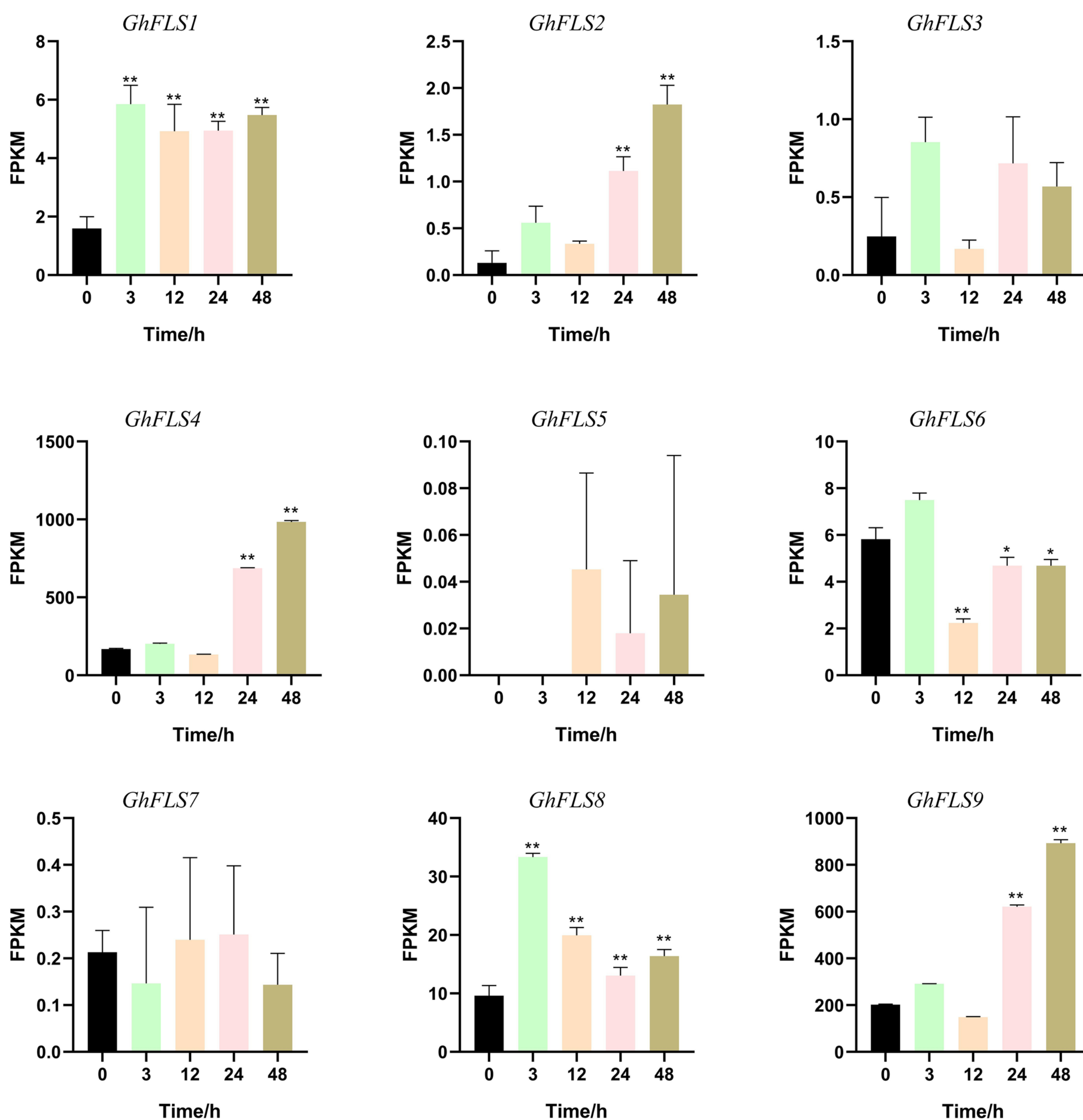

Fig. 8 Analysis the expression level of *GhFLS* genes under 400 mM NaCl in different time treatment. *Represented $0.01 < p < 0.05$, **represented $p < 0.01$

Interaction network and subcellular localization of GhFLS1 protein

To further understand the functions of GhFLS proteins, this study compared the GhFLS1 protein with *Arabidopsis*, obtained the *Arabidopsis* homologous protein AtFLS1 (AT5G08640), and used the online STRING tool to predict the interaction protein of the FLS protein (Additional file 1: Fig. S2). AtFLS1 interacts with dihydroflavonol 4-reductase (DFR), 4-coumaric acid: coenzyme A ligase (4CL), flavanone-3-hydroxylase (F3H), glycosyltransferase (UGT78D2), and cytochrome P450 (TT7) proteins. Analysis of the KEGG pathway enrichment of *GhFLS1* based on transcriptome data revealed, that *GhFLS1* was mainly involved in flavonoid metabolism (ko00250), and *FLS* was involved in the biosynthesis of flavonoids and flavonols through catalyzing the conversion of

dihydroflavonols into flavonols. The 4CL3, F3H, and DFR were involved in the synthesis of flavonoids. It was speculated that FLS interacts with these proteins and responds to NaCl stress through regulating flavonol content.

According to the prediction of subcellular localization in four cotton species, nine GhFLS proteins were most located in the cytoplasmic (cyto), chloroplast (chlo) and nucleus (nucl), three were located in the chloroplast (chlo), eight were located in the cytoplasmic (cyto), and three were located in the nucleus (nucl). Due to GhFLS1 protein sequence was highly similar to AtFLS1 sequence, and only *AtFLS1* was a functional gene encoding flavonol [34]. GhFLS1 protein was selected for subcellular localization verification. The results showed that GhFLS1 protein was located in the nucleus and cytoplasmic (Fig. 9).

VIGS of *GhFLS1* in cotton

The VIGS experiment was performed to verify the role of *GhFLS1* under NaCl stress (Fig. 10). PYL156: PDS showed obvious albino phenotypes (Fig. 10A). The relative expression level of *GhFLS1* was detected using qRT-PCR. The results showed that the expression level of pYL156: GhFLS1 was 67% lower than pYL156 (Fig. 10B), indicating that it had a good silencing effect. Under NaCl stress, the cotton leaves wilted and the stems bent, and pYL156: GhFLS1 leaves wilted more severely than those of pYL156. The expression of pYL156: GhFLS1 was significantly lower than that of pYL156 under NaCl stress. In addition, the proline content of pYL156: GhFLS1 decreased significantly under NaCl stress (Fig. 10C).

Discussion

Flavonols respond to salt stress in cotton

As a cash crop that is widely planted around the world, upland cotton is faced with various biotic and abiotic stresses. In plants, flavonols are significantly involved in plant growth and development, and they have been identified as the most active flavonoid, as well as a regulator of polar auxin transport [56, 57]. *FLS* is a key enzyme for the synthesis of flavonol substances, but systematic reports on FLSs are still lacking in cotton. In this study, the transcriptome data showed significant enrichment of genes related to flavonoid synthesis under salt stress. The metabolomics data showed that flavonols were significantly upregulated under salt stress. We measured and analyzed the total flavonoid content of cotton leaves, and the results showed a significant increase in total flavonoids under salt stress. This result further verified the correctness of the metabolomics data. Due to the significant increase in the content of flavonols under salt stress, we further conducted bioinformatics analysis on the key gene for synthesizing flavonols, *FLS*. This study provides important reference information for further understanding the functions of *FLS*.

FLS genes showed evolutionary conservation in cotton

In this study, the *FLS* genes of *G. arboreum*, *G. barbadense*, *G. hirsutum*, and *G. raimondii* were comprehensively identified. A total of 24 FLS genes were identified in the four cotton species, and 56 *FLS* family members were identified in *A. thaliana*, *A. halleri*, *V. vinifera*, *A. lyrata*, *C. grandiflora*, *T. cacao*, *G. max*, *B. stricta*, *O. sativa*, and *Z. mays*.

All *FLS* family members had the *2OG-FeII_Oxy* domain. The *2OG-FeII_Oxy* oxidase domain was composed of 2-ketoglutarate and the Fe (II)-dependent oxidase superfamily. The domain of 2-ketoglutarate (2OG)/Fe(II) dependent dioxygenase (2OGDD) plays an important role in plant primary and secondary metabolism. The 2OGD gene family is extremely large. According to the sequence similarity, the members of the 2OGDD gene family could be divided into three categories: DOXA,

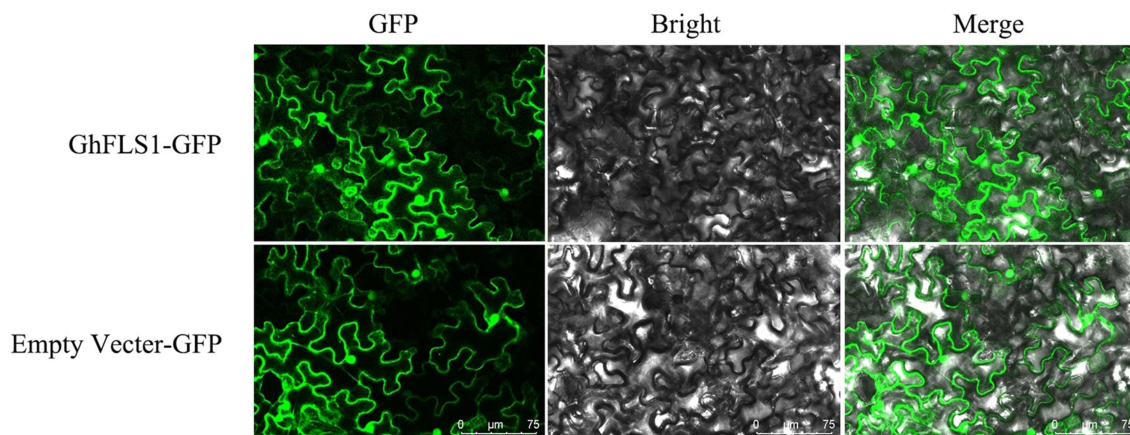

Fig. 9 Subcellular localization verification of GhFLS1 protein (bar = 75 μm)

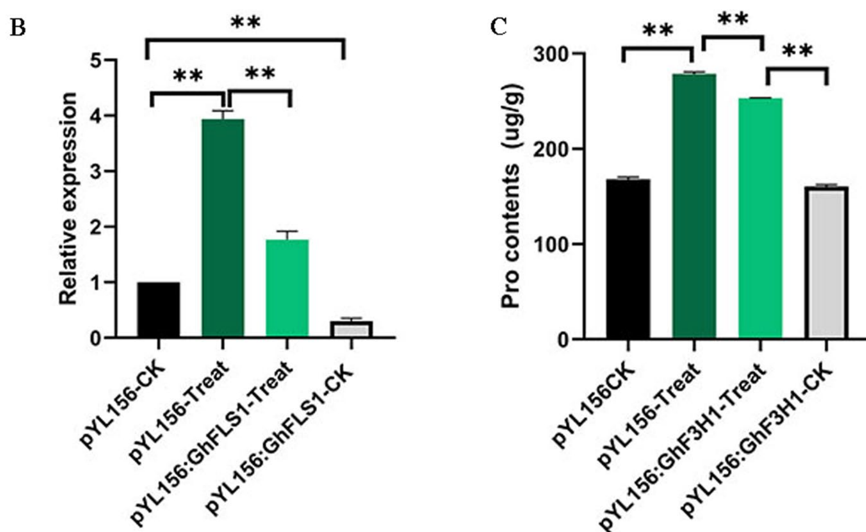

Fig. 10 Silencing *GhFLS1* via VIGS increased sensitivity to NaCl stress. **A** The phenotype of cotton after *GhFLS1* gene silencing under NaCl stress. pYL156: PDS as a positive control, pYL156 was an empty vector as control, and pYL156:GhFLS1 was *GhFLS1* silenced lines. **B** The relative expression level of *GhFLS1* under water and NaCl stress. **C** Proline content under water and NaCl stress. *Represented $0.01 < p < 0.05$, **Represented $p < 0.01$

DOXB, and DOXC. Members of the DOXA subfamily evolved from the DNA repair protein *Alkb* in *Escherichia coli* [58]. The DOXA subfamily mainly participates in primary metabolic processes. The members of the DOXB subfamily were relatively conservative, encoding prolyl 4-hydroxylases (*P4Hs*), which are mainly involved in the post-translational modification of polypeptide chains and are of great significance for plant signal peptide hormones and cell wall formation [59]. The members of the DOXC subfamily are complex, and their functions and numbers in different plants vary greatly, showing obvious species specificity. Members of this subfamily are mainly involved in the secondary metabolism of plants, including the biosynthesis of terpenoids, alkaloids, plant hormones, flavonoids, and phenolic acids. Most of the 2OGD genes involved in plant secondary metabolism can be classified into this subfamily [58]. This indicates that *FLS* family members belong to a branch of the DOXC subfamily. In addition, *FLS* family members were also found to contain the *DIOX_N* domain, a highly conserved N-terminal region of a protein with 2-ketoglutarate/Fe (II) dependent dioxygenase activity [60]. These two domains together constituted the unique gene function of *FLS* family members.

Replication events are one of the main driving factors for the evolution and diversification of genomes and genetic systems [61]. Environmental conditions and artificial selection affect the number of gene family members.

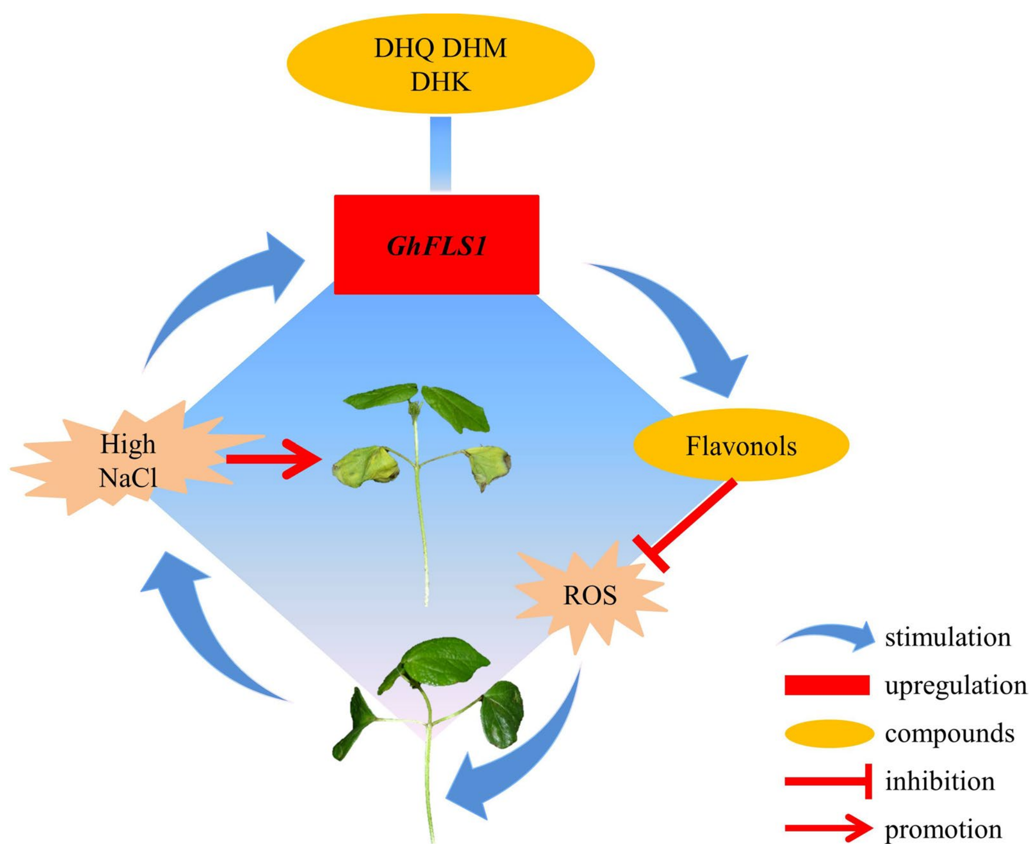

Fig. 11 A model illustrating the role of *GhFLS1* in cotton providing salinity stress tolerance. *DHQ* dihydroquercetin, *DHK* dihydrokaempferol, *DHM* dihydromyricetin, *ROS* reactive oxygen species

With the occurrence of repeated WGD events, the gene sequence of cotton was doubled, and some redundant genes were selectively lost or recombined [62]. The research results showed that four, seven, nine and four *FLS* family members were identified in *G. arboreum*, *G. barbadense*, *G. hirsutum*, and *G. raimondii*, respectively, and the number of genes in each cotton species explained the evolutionary origin of diploid and tetraploid cotton to some extent.

The uneven distribution of genes on each chromosome indicated that there was genetic variation during evolution [63]. The chromosome location of FLSs clearly showed the physical location distribution of each *FLS* gene in the genome and the evolutionary relationships of some genes. There was one tandem-duplication on chromosomes A05 and D04 of *G. hirsutum*, 1 tandem-duplication on chromosome 12 of *G. raimondii*, and no tandem duplication in *G. arboreum* and *G. barbadense*, indicating that there was a special evolutionary pattern in the evolution of different cotton species. The lack of tandem duplication may be the main reason for the low number of GbFLSs and GhFLSs. The collinearity analysis showed that whole gene replication and fragment replication played important roles in the development of FLSs. The *Ka*/*Ks* values of 164 pairs of genes were calculated, among which $Ka/Ks < 1$ for 65 pairs of duplicate genes, indicating purifying selection. It was speculated that the cotton *FLS* gene family underwent strong purifying selection after fragment replication, tandem replication, and WGD, but the functional difference was limited.

Motif prediction can provide a basis for researchers to analyze the functional and structural classification of families. A motif is a short sequence of relatively conservative characteristics shared by a group of genes. This may be a recognition sequence or a functional protein [64]. In this study, it was found that most of the Motif 9 in branch I and branch II of the 24 family members was located at the N-terminus, while the Motif 9 in branch III was located at the C-terminus and the number of these 24 genes was also different, which may be the main reason why *FLS* had different functions.

GhFLS1 plays an important role in salt response

Under NaCl stress, the expression of *GhFLS1*, *GhFLS2*, *GhFLS4*, *GhFLS6*, *GhFLS8*, and *GhFLS9* changed

significantly. While the *GhFLS6* gene was significantly downregulated, the other five genes were upregulated. After silencing the *GhFLS1* in cotton, cotton seedlings were more sensitive to NaCl, and the expression of the *GhFLS1* gene decreased significantly. Some studies showed that the overexpression of *EkFLS* significantly increased the flavonol content in guard cells, and the accumulation of flavonol reduced ABA-induced stomatal closure through inhibiting the accumulation of hydrogen peroxide in guard cells. In addition, the overexpression of *EkFLS* increased the content of superoxide dismutase and peroxidase enzymes in *A. thaliana*, thus effectively eliminating the ROS caused by drought [65]. In this study, it is speculated that *GhFLS1* can resist salt stress through increasing the ROS produced by flavonol accumulation under salt stress. When salt stress is applied to plants, the *FLS* gene in the plant responds positively and converts dihydroflavonols into flavonols, which can cope with ROS and thus resist salt stress (Fig. 11). The structural should be changed, taking into account the circularity, the contribution of compounds and genes, and the reaction of plant and metabolism. This model needs to be further validated and improved in future experiments.

Conclusion

This study transcriptome and metabolome data showed that flavonols play an important role under salt stress. This study analyzed the phenotype and total flavonoids content of cotton leaves under salt stress. *FLS* was identified in cotton for the first time, and four, seven, nine, and four *FLS* genes were identified in *G. arboreum*, *G. barbadense*, *G. hirsutum*, and *G. raimondii*, respectively. FLSs were divided into three branches according to the composition of the phylogenetic tree, gene structure, and motifs. Fragment replication and WGD were the main evolution modes of the *FLS* gene family. Silencing *GhFLS1* resulted in a more serious phenotype in cotton under NaCl stress, indicating that *GhFLS1* was involved in the response of cotton to NaCl stress. This study provides a reference and knowledge basis for further exploring the relationship between *GhFLS1* and NaCl stress.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12302-023-00743-2.

Additional file 1: Fig. S1. *GhFLS* genes displayed expression patters under cold, heat, salt and drought stress. **Fig. S2.** Interaction network of FLS protein. The FLS represented the protein AtFLS1 with the highest homology to GhFLS1. **Table S1.** Primers used in the experiment. **Table S2.** Basic information of 24 genes.

Acknowledgements
Not applicable.

Author contributions
MH: conceived and designed the experiments, writing original draft, data curation, software. RC, YC, TJ, HH, YL, XL, CR: data curation. review and editing. YF, YZ, KN, JW, SW, LS: methodology. XC, XL, DW, ZY, CC, LG and LZ: data curation. QC and WY: concept of study, supervision and revised manuscript. All authors have read and approved the final manuscript.

Funding
This work was supported by National Key Research and Development Program of China (No. 2022YFD1200300), China Agriculture Research System of MOF and MARA, Agricultural Science and Technology Innovation Program of Chinese Academy of Agricultural Sciences.

Availability of data and materials
The original contributions presented in the study are included in the article/Additional files, further inquiries can be directed to the corresponding author.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors agreed to publish the paper.

Competing interests
The authors declare no competing or financial interests.

Received: 16 February 2023 Accepted: 7 May 2023
Published online: 19 May 2023

References

1. Radanielson AM, Angeles O, Tao L, Ismail AM, Gaydon DS (2018) Describing the physiological responses of different rice genotypes to salt stress using sigmoid and piecewise linear functions. Field Crop Res 220:46–56
2. Zheng N, Guo M, Yue W, Teng Y, Zhai Y, Yang J, Zuo R (2021) Evaluating the impact of flood irrigation on spatial variabilities of soil salinity and groundwater quality in an arid irrigated region. Hydrol Res 52(1):229–240
3. Deinlein U, Stephan AB, Horie T, Luo W, Xu G, Schroeder JI (2014) Plant salt-tolerance mechanisms. Trends Plant Sci 19(6):371–379
4. Hossain MS, Dietz K-J (2016) Tuning of redox regulatory mechanisms, reactive oxygen species and redox homeostasis under salinity stress. Front Plant Sci 7:548
5. Chapman JM, Muhlemann JK, Gayomba SR, Muday GK (2019) RBOH-dependent ROS synthesis and ROS scavenging by plant specialized metabolites to modulate plant development and stress responses. Chem Res Toxicol 32(3):370–396
6. Maggioni D, Biffi L, Nicolini G, Garavello W (2015) Flavonoids in oral cancer prevention and therapy. Eur J Cancer Prev 24(6):517–528
7. Kang H-K, Ecklund D, Liu M, Datta SK (2009) Apigenin, a non-mutagenic dietary flavonoid, suppresses lupus by inhibiting autoantigen presentation for expansion of autoreactive Th1 and Th17 cells. Arthritis Res Ther 11(2):1–13
8. Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol 126(2):485–493
9. Gayomba SR, Watkins JM, Muday GK (2017) Flavonols regulate plant growth and development through regulation of auxin transport and cellular redox status. Recent advances in polyphenol research. Wiley, UK, pp 143–170
10. Pi E, Zhu C, Fan W, Huang Y, Qu L, Li Y, Zhao Q, Ding F, Qiu L, Wang H (2018) Quantitative phosphoproteomic and metabolomic analyses reveal *GmMYB173* optimizes flavonoid metabolism in soybean under salt stress. Mol Cell Proteomics 17(6):1209–1224
11. Nakabayashi R, Yonekura-Sakakibara K, Urano K, Suzuki M, Yamada Y, Nishizawa T, Matsuda F, Kojima M, Sakakibara H, Shinozaki K (2014)

Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. Plant J 77(3):367–379
12. Williams RJ, Spencer JP, Rice-Evans C (2004) Flavonoids: antioxidants or signalling molecules? Free Radical Biol Med 36(7):838–849
13. Bellocco R (2011) Dietary quercetin intake and risk of gastric cancer: results from a population-based study in Sweden. Ann Oncol 22(2):438–443
14. Xu M, Pirtskhalava T, Farr JN, Weigand BM, Palmer AK, Weivoda MM, Inman CL, Ogrodnik MB, Hachfeld CM, Fraser DG (2018) Senolytics improve physical function and increase lifespan in old age. Nat Med 24(8):1246–1256
15. Francisco P-V, Juan D (2010) Flavonols and cardiovascular disease. Mol Aspects Med 31(6):478–494
16. Jia X, Zhu Y, Zhang R, Zhu Z, Zhao T, Cheng L, Gao L, Liu B, Zhang X, Wang Y (2020) Ionomic and metabolomic analyses reveal the resistance response mechanism to saline-alkali stress in *Malus halliana* seedlings. Plant Physiol Biochem 147:77–90
17. Veit M, Pauli GF (1999) Major flavonoids from *Arabidopsis thaliana* leaves. J Nat Prod 62(9):1301–1303
18. Yang J, Zhang L, Jiang L, Zhan YG, Fan GZ (2021) Quercetin alleviates seed germination and growth inhibition in *Apocynum venetum* and *Apocynum pictum* under mannitol-induced osmotic stress. Plant Physiol Biochem 159:268–276
19. Peer WA, Brown DE, Tague BW, Muday GK, Taiz L, Murphy AS (2001) Flavonoid accumulation patterns of transparent testa mutants of *Arabidopsis*. Plant Physiol 126(2):536–548
20. Alexei S, Michaela S (2003) Significance of skin flavonoids for UV-B-protection in apple fruits. J Exp Bot 389:1977–1984
21. Vogt T, Wollenweber E, Taylor LP (1995) The structural requirements of flavonols that induce pollen germination of conditionally male fertile *Petunia*. Phytochemistry 38(3):589–592
22. Yoshitama K, Ishikura N, Fuleki T, Nakamura S (1992) Effect of anthocyanin, flavonol co-pigmentation and pH on the color of the berries of *Ampelopsis brevipedunculata*. J Plant Physiol 139(5):513–518
23. Buer CS, Djordjevic MA (2009) Architectural phenotypes in the *transparent testa* mutants of *Arabidopsis thaliana*. J Exp Bot 60(3):751–763
24. Buer CS, Kordbacheh F, Truong TT, Hocart CH, Djordjevic MA (2013) Alteration of flavonoid accumulation patterns in transparent testa mutants disturbs auxin transport, gravity responses, and imparts long-term effects on root and shoot architecture. Planta 238:171–189
25. Ringli C, Bigler L, Kuhn BM, Leiber R-M, Diet A, Santelia D, Frey B, Pollmann S, Klein M (2008) The modified flavonol glycosylation profile in the *Arabidopsis rol1* mutants results in alterations in plant growth and cell shape formation. Plant Cell 20(6):1470–1481
26. Chapman JM, Muday GK (2021) Flavonols modulate lateral root emergence by scavenging reactive oxygen species in *Arabidopsis thaliana*. J Biol Chem 296:100222
27. Rejeb KB, Abdelly C, Savouré A (2014) How reactive oxygen species and proline face stress together. Plant Physiol Biochem 80:278–284
28. Costa RCLD, Lobato AKDS, Silveira JAGD (2011) ABA-mediated proline synthesis in cowpea leaves exposed to water deficiency and rehydration. Turk J Agric For 35(3):309–317
29. Watkins JM, Chapman JM, Muday GK (2017) Abscisic acid-induced reactive oxygen species are modulated by flavonols to control stomata aperture. Plant Physiol 175(4):1807–1825
30. Martens S, Teeri T, Forkmann G (2002) Heterologous expression of dihydroflavonol 4-reductases from various plants. FEBS Lett 531(3):453–458
31. Davies KM, Schwinn KE, Deroles SC, Manson DG, Bradley JM (2003) Enhancing anthocyanin production by altering competition for substrate between flavonol synthase and dihydroflavonol 4-reductase. Euphytica 131(3):259–268
32. Lou Q, Liu Y, Qi Y, Jiao S, Tian F, Jiang L, Wang Y (2014) Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. J Exp Bot 12:3157–3164
33. Forkmann G, De Vlaming P, Spribille R, Wiering H, Schram A (1986) Genetic and biochemical studies on the conversion of dihydroflavonols to flavonols in flowers of *Petunia hybrida*. Z Naturforsch C 41(1–2):179–186
34. Wisman E, Hartmann U, Sagasser M, Baumann E, Palme K, Hahlbrock K, Saedler H, Weisshaar B (1998) Knock-out mutants from an *En-1* mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. Proc Natl Acad Sci 95(21):12432–12437
35. Holton TA, Brugliera F, Tanaka Y (1993) Cloning and expression of flavonol synthase from *Petunia hybrida*. Plant J 4(6):1003–1010
36. Preu A, Stracke R, Weisshaar B, Hillebrecht A, Matern U, Martens S (2009) *Arabidopsis thaliana* expresses a second functional flavonol synthase. FEBS Lett 583(12):1981–1986
37. Pelletier MK, Murrell JR, Shirley BW (1997) Characterization of flavonol synthase and leucoanthocyanidin dioxygenase genes in *Arabidopsis* (further evidence for differential regulation of" early" and" late" genes). Plant Physiol 113(4):1437–1445
38. Moriguchi T, Kita M, Ogawa K, Tomono Y, Omura M (2002) Flavonol synthase gene expression during citrus fruit development. Physiol Plant 114(2):251–258
39. Takahashi R, Githiri SM, Hatayama K, Dubouzet EG, Shimada N, Aoki T, Ayabe S-I, Iwashina T, Toda K, Matsumura H (2007) A single-base deletion in soybean flavonol synthase gene is associated with magenta flower color. Plant Mol Biol 63:125–135
40. Falcone Ferreyra ML, Rius S, Emiliani J, Pourcel L, Feller A, Morohashi K, Casati P, Grotewold E (2010) Cloning and characterization of a UV-B-inducible maize flavonol synthase. Plant J Cell Mol Biol 62(1):77–91
41. Zhou X-W, Fan Z-Q, Chen Y, Zhu Y-L, Li J-Y, Yin H-F (2013) Functional analyses of a flavonol synthase-like gene from *Camellia nitidissima* reveal its roles in flavonoid metabolism during floral pigmentation. J Biosci 38:593–604
42. Liu W, Xiao Z, Fan C, Jiang N, Meng X, Xiang X (2018) Cloning and characterization of a flavonol synthase gene from *Litchi chinensis* and its variation among litchi cultivars with different fruit maturation periods. Front Plant Sci 9:567
43. Park S, Kim D-H, Park B-R, Lee J-Y, Lim S-H (2019) Molecular and functional characterization of *Oryza sativa* flavonol synthase (*OsFLS*), a bifunctional dioxygenase. J Agric Food Chem 67(26):7399–7409
44. Zhang X, Yang H, Schaufelberger M, Li X, Cao Q, Xiao H, Ren Z (2020) Role of flavonol synthesized by nucleus *FLS1* in *Arabidopsis* resistance to Pb stress. J Agric Food Chem 68(36):9646–9653
45. Yu Z, Dong W, Teixeira da Silva JA, He C, Si C, Duan J (2021) Ectopic expression of *DoFLS1* from *Dendrobium officinale* enhances flavonol accumulation and abiotic stress tolerance in *Arabidopsis thaliana*. Protoplasma 258:803–815
46. Muhlemann JK, Younts TL, Muday GK (2018) Flavonols control pollen tube growth and integrity by regulating ROS homeostasis during high-temperature stress. Proc Natl Acad Sci 115(47):E11188–E11197
47. Wei Y, Xu Y, Lu P, Wang X, Li Z, Cai X, Zhou Z, Wang Y, Zhang Z, Lin Z (2017) Salt stress responsiveness of a wild cotton species (*Gossypium klotzschianum*) based on transcriptomic analysis. PLoS ONE 12(5):e0178313
48. Wang D, Lu X, Chen X, Wang S, Wang J, Guo L, Yin Z, Chen Q, Ye W (2020) Temporal salt stress-induced transcriptome alterations and regulatory mechanisms revealed by PacBio long-reads RNA sequencing in *Gossypium hirsutum*. BMC Genomics 21(1):1–15
49. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102(43):15545–15550
50. Owens DK, Alerding AB, Crosby KC, Bandara AB, Westwood JH, Winkel BS (2008) Functional analysis of a predicted flavonol synthase gene family in *Arabidopsis*. Plant Physiol 147(3):1046–1061
51. Kumar S, Stecher G, Tamura K (2015) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874
52. Zhang H, Gao S, Lercher MJ, Hu S, Chen WH (2012) EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. Nucleic Acids Res 40(W1):569–572
53. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020) TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 13(8):1194–1202
54. Hu Y, Chen J, Fang L, Zhang Z, Ma W, Niu Y, Ju L, Deng J, Zhao T, Lian J (2019) *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. Nat Genet 51(4):739–748
55. Schilbert HM, Schöne M, Baier T, Busche M, Viehöver P, Weisshaar B, Holtgräwe D (2021) Characterization of the *Brassica napus* flavonol

synthase gene family reveals bifunctional flavonol synthases. Front Plant Sci 12:733762
56. Buer CS, Kordbacheh F, Truong TT, Hocart CH, Djordjevic MA (2013) Alteration of flavonoid accumulation patterns in transparent testa mutants disturbs auxin transport, gravity responses, and imparts long-term effects on root and shoot architecture. Planta 238(1):171–189
57. Yin R, Han K, Heller W, Albert A, Dobrev PI, Zažímalová E, Schäffner AR (2014) Kaempferol 3-*O*-rhamnoside-7-*O*-rhamnoside is an endogenous flavonol inhibitor of polar auxin transport in *Arabidopsis* shoots. New Phytol 201(2):466–475
58. Kawai Y, Ono E, Mizutani M (2014) Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants. Plant J 78(2):328–343
59. Nadi R, Mateo-Bonmatí E, Juan-Vicente L, Micol JL (2018) The 2OGD superfamily: emerging functions in plant epigenetics and hormone metabolism. Mol Plant 11(10):1222–1224
60. Hagel JM, Facchini PJ (2010) Dioxygenases catalyze the *O*-demethylation steps of morphine biosynthesis in opium poppy. Nat Chem Biol 6(4):273–275
61. Gu Z, Steinmetz LM, Gu X, Scharfe C, Davis RW, Li W-H (2003) Role of duplicate genes in genetic robustness against null mutations. Nature 421(6918):63–66
62. Li F, Fan G, Lu C, Xiao G, Zou C, Kohel RJ, Ma Z, Shang H, Ma X, Wu J (2015) Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. Nat Biotechnol 33(5):524–530
63. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature 492(7429):423–427
64. Morello L, Breviario D (2008) Plant spliceosomal introns: not only cut and paste. Curr Genomics 9(4):227–238
65. Wang M, Zhang Y, Zhu C, Yao X, Zheng Z, Tian Z, Cai X (2021) *EkFLS* overexpression promotes flavonoid accumulation and abiotic stress tolerance in plant. Physiol Plant 172(4):1966–1982

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

► Convenient online submission
► Rigorous peer review
► Open access: articles freely available online
► High visibility within the field
► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com