Journal of Cotton Research

RESEARCH Open Access

Characters and structures of the nucleobase–ascorbate transporters (NAT) family genes in *Gossypium hirsutum* and their roles in responding to salt and drought stresses

GUO Lixue, ZHAO Lanjie, LU Xuke, CHEN Xiugui, WANG Shuai, WANG Junjuan, WANG Delong, YIN Zujun, CHEN Chao, FAN Yapeng, ZHANG Yuexin and YE Wuwei*

Abstract

Background: Nucleobase–ascorbate transporters (NAT), synonymously called nucleobase–cation symporter 2 (NCS2) proteins, were earlier reported to be involved in plant growth, development and resistance to stress. Previous studies concluded that s a polymorphic SNP associated with *NAT12* was significant different between salt-tolerant and salt-sensitive materials of upland cotton. In current study, a comprehensive analysis of *NAT* family genes was conducted for the first time in cotton.

Results: In this study, we discovered 32, 32, 18, and 16 *NAT* genes in *Gossypium hirsutum*, *G. barbadense*, *G. raimondii* and *G. arboreum*, respectively, which were classified into four groups (groups I–IV) based on the multiple sequence analysis. These *GhNAT* genes were unevenly distributed on At and Dt sub-genome in *G. hirsutum*. Most *GhNAT* members in the same group had similar gene structure characteristics and motif composition. The collinearity analysis revealed segmental duplication as well as tandem duplication contributing to the expansion of the *GhNATs*. The analysis of *cis*-acting regulatory elements of *GhNATs* showed that the function of *GhNAT* genes in cotton might be related to plant hormone and stress response. Under different conditions, the expression levels further suggested the *GhNAT* family genes were associated with plant response to various abiotic stresses. GhNAT12 was detected in the plasma membrane. And it was validated that the *GhNAT12* gene played an important role in regulating cotton resistance to salt and drought stress through the virus-induced gene silencing (VIGS) analysis.

Conclusions: A comprehensive analysis of *NAT* gene family was performed in cotton, including phylogenetic analysis, chromosomal location, collinearity analysis, motifs, gene structure and so on. Our results will further broaden the insight into the evolution and potential functions of *NAT* genes in cotton. Current findings could make significant contribution towards screening more candidate genes related to biotic and abiotic resistance for the improvement in cotton.

Keywords: *Gossypium hirsutum*, *NAT* gene family, VIGS, Salt and drought stresses

Background

Cotton is an important economic crop around the world, due to its most common consumption as the raw material in textile industry. The genus *Gossypium* contains two cultivated tetraploids of *G. hirsutum* and *G. barbadense*,

*Correspondence: yew158@163.com
State Key Laboratory of Cotton Biology/Institute of Cotton Research of Chinese Academy of Agricultural Sciences/Zhengzhou Research Base, School of Agricultural Sciences, Zhengzhou University/Key Laboratory for Cotton Genetic Improvement, MOARA, Anyang 455000, Henan, China

which originated from the transoceanic hybridization of the A-genome-like progenitor *G. arboreum* with the native D-genome-like ancestor *G. raimondii* in New World, followed by the chromosome doubling (Li et al. 2015). Subsequently, the nascent dicotyledonous allotetraploid cotton diverged into five *Gossypium* species. And *G. hirsutum* and *G. barbadense* independently evolved in diverse geographic regions (Hu et al. 2019). With the rise of global warming, extreme weather conditions such as rainstorm, flood, drought, typhoon, heat wave, cold wave, and sandstorm occur frequently. In addition, the excessive development and utilization of land resources led to scarcity of water and soil salinization, which have brought huge losses to agricultural production. Though, cotton is a drought and saline-alkali tolerant crop. But extreme stresse condition will cause huge losses to cotton production. So that, it is of great significance for the development of cotton industry to study the mechanism of cotton salt and drought tolerance, necessitating to mine the related functional genes, and pyramid such genetic factors in new developing varieties.

Abiotic stress responses are important mechanisms of sessile organisms such as plants which cannot survive unless capable to cope with environmental changes (Hirayama and Shinozaki 2010). The abiotic stresses mostly refer to extreme environmental conditions, are always the great threats to plant growth and yield (Grattan and Grieve 1998). Under salt stress conditions, physiological and metabolic activities are affected by ionic and osmotic stresses, nutritional imbalances, or a combination of these factors (Chen et al. 2014; Slama et al. 2015). Ionic stress mainly involves the excessive accumulation of Na^+ in plant leaves. During periods of high salt stress, the Na^+ uptake competes with the K^+ uptake which results in excess of cytoplasmic Na^+ sequestration instead of Cl^- within the cell (Muchate et al. 2016). This will cause osmotic imbalances, and a series of enzymatic and nonenzymatic antioxidants in plant cells, which finally lead to water deficits, stomatal closure, and reduces plant growth and the rate of photosynthesis (Jithesh et al. 2006; Gill and Tuteja 2010; Roy et al. 2014). Similarly, K^+ as a major inorganic osmoticum gets dramatically decreased under drought stress in *Malus hupehensis* (Qi et al. 2019). And drought stress inhibits photosynthesis by decreasing stomatal aperture (Cornic 2000). The endomembrane trafficking is tightly linked to stress signaling pathways to meet the demand of rapid changes in cellular processes and to ensure the correct delivery of stress-related cargo molecules (Wang et al. 2020a), which had already been proved as an important regulator in osmotic stress, such as drought, salinity, and cold stresses.

Nucleobase–ascorbate transporters (NATs), also known as nucleobase–cation symporter 2 (NCS2), is a group of transporter proteins being found in almost all domains of life, which contain transporting ascorbate, pyrimidines and purines (De Koning and Diallinas 2000). The *NAT* family has been identified in bacteria, archaea, diatoms, fungi, plants and animals (De Koning and Diallinas 2000; Bürzle et al. 2013). The *NAT* family members generally have common characters, containing about 400∼650 amino acid residues and 10∼14 transmembrane domains. Within these transmembrane domains, there are two conserved domains. Of them, one is highly conserved domain located at the upstream of transmembrane segment 9 (TMS9), also called "NAT signature motif" (Q/E/P) NXGXXXXT (R/K/G) a critical domain of purine transporters UapA and UapC to recognize substrate in *Aspergillus* (Diallinas et al. 1998; Koukaki et al. 2005). In principle, if the first amino acid of the motif is Q/E, the protein will transport nucleic acids, or P will transport ascorbic acid. And another highly conserved domain is a QH structure located in the middle of transmembrane segment 1 (TMS1) (Diallinas et al. 1998; Koukaki et al. 2005; Karatza et al. 2006; Gournas et al. 2008; Frillingos 2012; Kosti et al. 2012). Phylogenetic tree analysis of NAT proteins depicted five clades based on multiple sequence alignments in plant (Maurino et al. 2006; Cai et al. 2014). The NAT proteins in *Arabidopsis thaliana* split into five subfamilies, including I, II, III, IV and V, respectively. These classifications were closely associated with the specific expression of tissue and organ in the growth process of *Arabidopsis thaliana* (Maurino et al. 2006). The *NAT* family members are evolutionarily conserved and the nucleobase is the principal substrate of this family. Among the five known families of transporters, only the NAT family utilizes nucleobases as their principal substrate (Frillingos 2012; Kourkoulou et al. 2018). The most renowned characteristics of NAT proteins include transport uric acid, xanthine, uracil, ascorbic acid, adenine, guanine, hypoxanthine, and uracil (Gournas et al. 2008; Frillingos 2012; Bürzle et al. 2013; Niopek-Witz et al. 2014), such as AtNAT3 and AtNAT12 proteins have been proved to transport adenine, guanine, and uracil (Niopek-Witz et al. 2014). Eight Arabidopsis NATs (AtNAT1–AtNAT8) all transport xanthine (Hunt 2013).

The NAT proteins are important for plant growth and development. Some plants' NAT proteins have been functionally characterized. The maize leaf permease1 *Lpe1* was the first identified NAT family member in plant. Loss of LPE1 function results in a defective chloroplast phenotype and affects the integrity of plasma membrane (Schultes et al. 1996). LPE1 acts as a transporter for xanthine and uric acid, but not for ascorbic acid (Argyrou

et al. 2001). The biochemical characterization of NAT3 and NAT12 proteins from *Arabidopsis thaliana* have been reported to transport guanine, uracil, and adenine with high affinity (Niopek-Witz et al. 2014). The transient expression of *AtNAT3* and *AtNAT12* has indicated that the encoded proteins are localized in the plasma membrane (Niopek-Witz et al. 2014). Overexpression of *MdNAT7* in transgenic apple plants has showed enhanced xanthine and uric acid concentrations and improved tolerance to salinity stress compared with non-transgenic plants, while opposite phenotypes have been observed in *MdNAT7* RNAi plants (Sun et al. 2021).

In a previous study, SNP (Single nucleotide polymorphism) analysis has been performed in different salt-tolerant cotton materials, and 1 282 SNPs have been screened in association with salt-tolerance (Wang et al. 2016). A nucleobase–ascorbate transporter has been identified and named as *GhNAT12* (*Gh_A07G117800*). The SNP variation associated with *NAT12* was significantly different between salt-tolerant (AA) and salt-sensitive materials (AG), which suggests that the *NAT12* may be related to cotton salt-tolerance. In this study, the *NATs* were identified through a whole genome search in *Gossypium.* The gene phylogenetic relationship, gene structure, protein conserved motif, and chromosome localization were systematically analyzed in *G. hirsutum*. In addition, the relative expression of *GhNATs* under salt and drought stress was performed using qRT-PCR to validate the function of *NATs*. The silencing of *GhNAT12* in *G. hirsutum* under salt and drought stress demonstrated that GhNAT12 positively regulates cotton resistance against salt and drought. Further experiments revealed that *GhNAT12* is localized to the cell membrane. These results could make contribution to understand *NAT* family potential functions regarding resistance against abiotic stress in cotton.

Material and methods

Identification of *NAT* family members in cotton

Genomic sequence data and protein sequence of four cotton species *G. hirsutum*, *G. raimondii*, *G. arboreum*, and *G. barbadense* were acquired from the Cotton Functional Genomics Database (CottonFGD) (http://www.cottonfgd.org/) (Zhu et al. 2017). The Hidden Markov Mode (HMM) profile of PF00860 which most probably belongs to *NAT* gene family was downloaded from Pfam (https://pfam.xfam.org/). Protein sequence containing PF00860 were screened to identify the *NAT* candidate homologous genes containing target domain in four cotton species by HMMER software (Jones et al. 2014). Subsequently, the genes above the inclusion threshold were selected, the selected genes were stored temporarily. Using the Data Fetch & Enrichment interface on CottonFGD website, the selected genes were searched and the result in terms of the domain of PF00860 were downloaded. In order to further confirm the screened genes, the Batch Web CD-Search Tools website (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) was used to screen the genes associated domains. The genome sequence data of *A. thaliana* was retrieved from JGI Phytozome 13 (https://phytozome.jgi.doe.gov/). We obtained some features of *NAT* genes in cotton, as for amino acid length of the protein, isoelectric points (pI), molecular weights (MWs), grand average of hydropathy (GRAVY) and so on, by using CottonFGD (http://www.cottonfgd.org/).

Phylogentic analysis

To discover the evolutionary relationship of *NAT* family in different species, amino acid sequences of NATs identified in *G. hirsutum*, *G. raimondii*, *G. arboreum, G. barbadense*, and *A. thaliana* were applied to conduct multiple sequence alignment utilizing the Clustalw program (Edgar 2004). Subsequently, the phylogenetic tree was constructed using the neighbor joining (NJ) method with the Poisson substitution model by MEGA7.0 (Kumar et al. 2016).

Chromosomal locations and collinearity analysis of *NAT* genes in *Gossypium hirsutum*

The physical chromosomal location of all *NAT* genes in *G. hirsutum* were generated by TBtools software (Chen et al. 2018). Gene information was obtained from CottonFGD and CottonGen (https://www.cottongen.org/) (Yu et al. 2013). MCScanX software was used to analyze genomic collinearity blocks (Wang et al. 2012). The diagrammatical result of NAT from *G. hirsutum* was visualized by using Circos-0.69 software (Krzywinski et al. 2009).

Gene structure analysis and conserved motif identification

The conserved protein motifs were identified using Multiple Em for Motif Elicitation (MEME) website (http://meme-suite.org/) (Bailey et al. 2009). The MAST (Motif Alignment & Search Tool) file from the MEME website, the NWK (Newick) file from the phylogenetic tree analysis and the GFF3 (General Feature Format Version 3) genome file from *G. hirsutum* were used to visualize the phylogenetic tree, conservative protein motif and the gene structure map by the TBtools software.

Analysis of *GhNATs* promoter regions and differentially expressed genes

DNA sequences of the 2 000 bp upstream regions of *GhNATs* were extracted from CottonFGD database (http://www.cottonfgd.org/) as promoter regions. The PlantCARE website (http://bioinformatics.psb.ugent.be/

webtools/plantcare/html/) was used for the prediction of *cis*-acting elements. The RNA-seq data (PRJNA248163) was downloaded from National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/). The expression levels of *NAT* genes were analysized under salt, cold, hot, and PEG stress (Hu et al. 2019). And the localization of *cis*-acting elements were drawn using the TBtools software.

Plant materials and abiotic stresses

The upland cotton (*G. hirsutum*) cv. Zhong9807 is a salt tolerant cultivar, H177 is a drought tolerant cultivar (Wang et al. 2016; Lu et al. 2017). The cotton plants and *Nicotiana benthamiana w*ere sown in soil and seedlings were grown in greenhouse under 8 h (dark)/16 h (light), at 23 °C (dark)/28 °C (light) with a relative humidity of 60%.

The three-leaf stage seedlings of Zhong9807 were treated with 200 mmol·L^{-1} NaCl, and the drought stress was carried out on H177 with 12% PEG6000 (Wang et al. 2020b). At 0, 1, 3, 6, 9, and 12 h, the leaves of cotton were collected and immediately frozen with liquid nitrogen, then stored at − 80 °C before RNA extraction.

qRT-PCR analysis

Total RNA was extracted from cotton leaves using RNA prep Pure Plant Kit (TIANGEN) according to the protocol provided by the manufacturer. First-strand cDNA was generated using Prime ScriptTM II First Strand cDNA Synthesis Kit (Takara) according to the manufacturer's instructions. Quantitative real-time PCR (qRT-PCR) was performed following with the previously program (Zhang et al. 2021). The expression level of *NAT* genes was analyzed by the $2^{-\Delta\Delta C_t}$ method. And the cotton *histone3* (GenBank accession No. AF024716) gene was used as the internal reference gene. Three independent biological and technical repeats were performed. The primers used in qRT-PCR were listed in Additional file 1: Table S1.

GhNAT12 gene silenced

The fragments (Additional file 1: Table S1) of *GhNAT12* were integrated into the tobacco rattle virus vector pYL156 plasmid at the *Xba*I-*Sac*I site using the T4 DNA ligase (Takara) to form the pYL156-*GhNAT12* vector. And then the vector were transformed into *Agrobacterium tumefaciens* GV3101. The pYL156-*PDS* was used as a positive control, and the empty pYL156 plasmid was the negative control. *Agrobacterium* harboring the empty vector of pYL-156 or one of its derivatives (pYL15-*GhNAT12*, pYL156-*PDS*) were mixed with an equal volume of *Agrobacterium* harboring pYL192. When the cotton plants were 7 days old, each mixture was infiltrated into two fully expanded cotyledons of cotton plants until filled with the entire cotyledons as previously described (Fan et al. 2021). Seedlings were grown at 25 °C in dark treatment for 24 h, and then incubated with a 16 h (light) /8 h (dark) photoperiod. When the TRV:*PDS* plants displayed albino phenotype in their newly developed true leaves, the TRV:*00* and TRV:*GhNAT12* infiltrated Zhong9807 were treated with 200 mmol·L^{-1} NaCl and the TRV:00 and TRV:GhNAT12 infiltrated H177 were treated with 12% PEG6000 at the same time to observe their response to salt and drought stress. The expression level of the *GhNAT12* were examined, and the successfully silenced plants were used to evaluate the resistance against salt and drought stresses.

Subcellular localization

To observe the subcellular localization of GhNAT12, the open reading frame of *GhNAT12* was fused to GFP (Green fluorescent protein) under the control of 35S promoter in the expression vector pCAMBIA3300 to generate pCAMBIA3300-*GhNAT12*-GFP constructs. And the vectors were transiently expressed in tobacco (*Nicotiana benthamiana*) leaf cells by *A. tumefaciens* strain GV3101 infiltration method. The infiltrated plants were incubated for 2∼3 d at 23 °C under dark condition. The GFP fluorescence signals were detected using confocal microscopy (Olympus FV1200). Primers used in this study were listed in Additional file 1: Table S1.

For visualization of GhNAT12 in onion epidermal cells, recombinant plasmids of pCAMBIA3300-*GhNAT12*-GFP were transformed into onion epidermal cells by particle bombardment using the PDS-1000/He system (Bio-Rad, USA). After incubation on MS medium for 24∼36 h, GFP fluorescence was visualized by confocal microscopy. For plasmolysis experiment, cells were treated with 20% sucrose. Signals were visualized by confocal microscopy.

Results

Phylogenetic analysis of *NAT* genes

To clarify the evolutionary relationship of *NAT* genes in cotton, we identified 32, 32, 18, and 16 *NATs* in *G. hirsutum*, *G. barbadense*, *G. arboreum,* and *G. raimondii* (Additional file 2: Table S2), respectively. The phylogenetic tree of cotton and *A. thaliana NATs* had been constructed by MEGA 7 using the neighbor-joining (NJ) method. The results showed that all *NAT* genes in five species were classified into four different groups (I–IV) (Fig. 1). Group I included 46 cotton NATs and *AtNAT4–AtNAT10* in *Arabidopsis thaliana*. Group II consisted 12 cotton *NATs* and *AtNAT11*, *AtNAT12*. The second largest group was III, containing 25 cotton *NATs* and *AtNAT1*, *AtNAT2*. The group IV consisted of

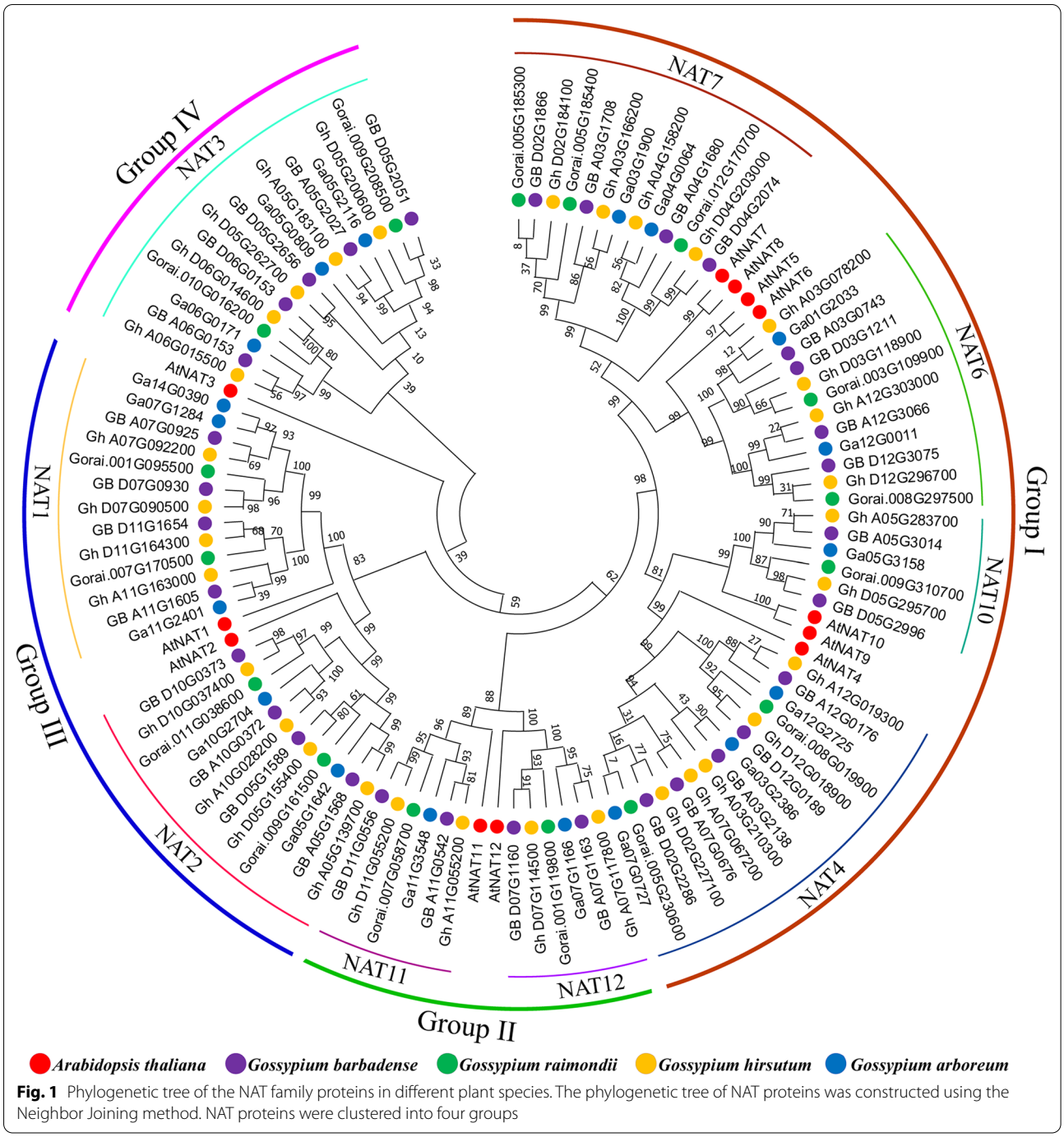

Fig. 1 Phylogenetic tree of the NAT family proteins in different plant species. The phylogenetic tree of NAT proteins was constructed using the Neighbor Joining method. NAT proteins were clustered into four groups

15 cotton *NATs* and *AtNAT3*. From the evolutionary distance on phylogenetic tree, homelogous genes were identified. A total of 9 pairs were higher similarly and identity. Interestingly, the diploid species of *G. raimondii* and *G. arboreum* were less as compared with the tetraploids of *G. hirsutum* and *G. barbadense*. This indicated that there was a presumed gene loss event during the evolutionary process.

Chromosome localization of *GhNATs* genes

To further investigate the genomic distribution of *NAT* family genes, the chromosome map of 32 *NAT* members from *G. hirsutum* were constructed (Fig. 2). A total of 16

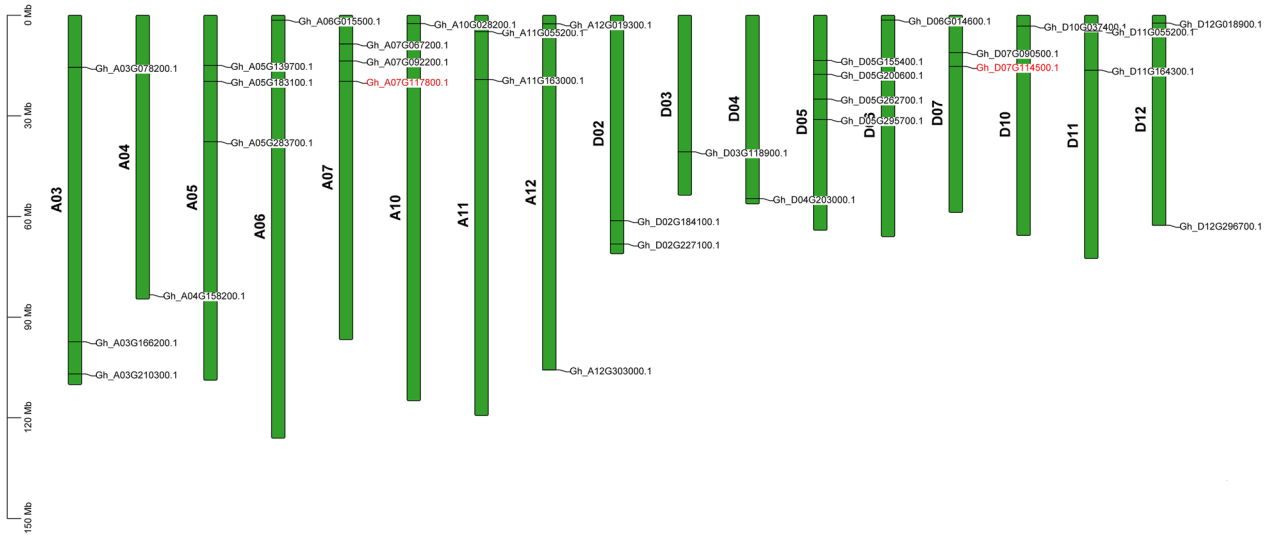

Fig. 2 Chromosomal locations of *NAT* genes in *Gossypium hirsutum*

genes were located on 8 chromosomes of At sub-genome and 16 genes were positioned on 9 chromosomes of Dt sub-genome. The distribution pattern of *NAT* genes was uneven on the At and Dt chromosomes. The maximum number of *NAT* members on the same chromosome were four, and those genes were located on the D05 chromosome. The minimum number of *NAT* members presented on the same chromosome was one, occuring at chromosome A04, A06, A10, D03, D04, D06, D10. The distribution of some NATs in terms of number and position on the chromosome of At were the same as the chromosome of Dt, such as A04–D04, A06–D06, A10–D10, A11–D11, A12–D12. The result indicated that the genes were conserved during evolution. Comparing with the homeologous chromosome of At, Dt sub-genome, *Gh_D05G262700* and *Gh_A07G067200* were distinctive, it might be the result of incomplete genome assembly. Some genes were distributed on different chromosomes, which implied that these genes might been added or lost during the evolution process. There were more NAT homologues in A03, A05 than D03, D05, suggested the possible gene loss event. Based on the primitive resistance genes, the gene loss event might arise before polyploid formation during the evolution. And this statement requries further verification.

Collinearity relationship of *GhNATs* genes

To identify paralogous gene pairs of At and Dt sub-genomes in *G. hirsutum*, the *NAT* genes locus on chromosome and collinearity analysis were investigated (Fig. 3). The collinearity analysis indicated that several *NAT* gene loci were highly conserved between the At and Dt sub-genomes. A total of 35 gene pairs were investigated, and 34 pairs were segmental duplication and 1 pair were tandem duplication (Additional file 3: Table S3). The duplication events resulted in the amplification of *GhNAT* in cotton genome.

Gene structure and conserved motifs analysis of *GhNATs*

Gene structure and protein motifs distribution have been analyzed for the evolutionary relationship among diverse gene family members in *G. hirsutum* (Fig. 4). Twenty conserved motifs were identified in NAT proteins. There were conserved protein motifs distributed in all GhNAT proteins. The number of motifs ranged from 1 to 17 in each protein with similar arrangements of protein motifs on same group. These distribution patterns illustrated that all GhNATs had at least one conserved protein motif . The composition of specific conserved motifs was found in particular groups, which suggested that the distribution of motifs might correlated with functional specificity. The motif 1–9 were found in the GhNAT members except *Gh_D05G262700* and *Gh_A07G067200*. As for these motifs, they were highly conserved.

The *NAT* genes contained 10∼15 exons in *G. hirsutum*, except *Gh_D05G262700* (3 exons) and *Gh_A07G067200* (6 exons). Intron and exon structure had similar permutation within the same group. The comprehensive illustration showed exon–intron different distribute among the *GhNAT* genes among various groups. Taken together, the gene structure of *NAT* genes had a strong relationship with gene evolution, but its mechanism was not clear, needing further study.

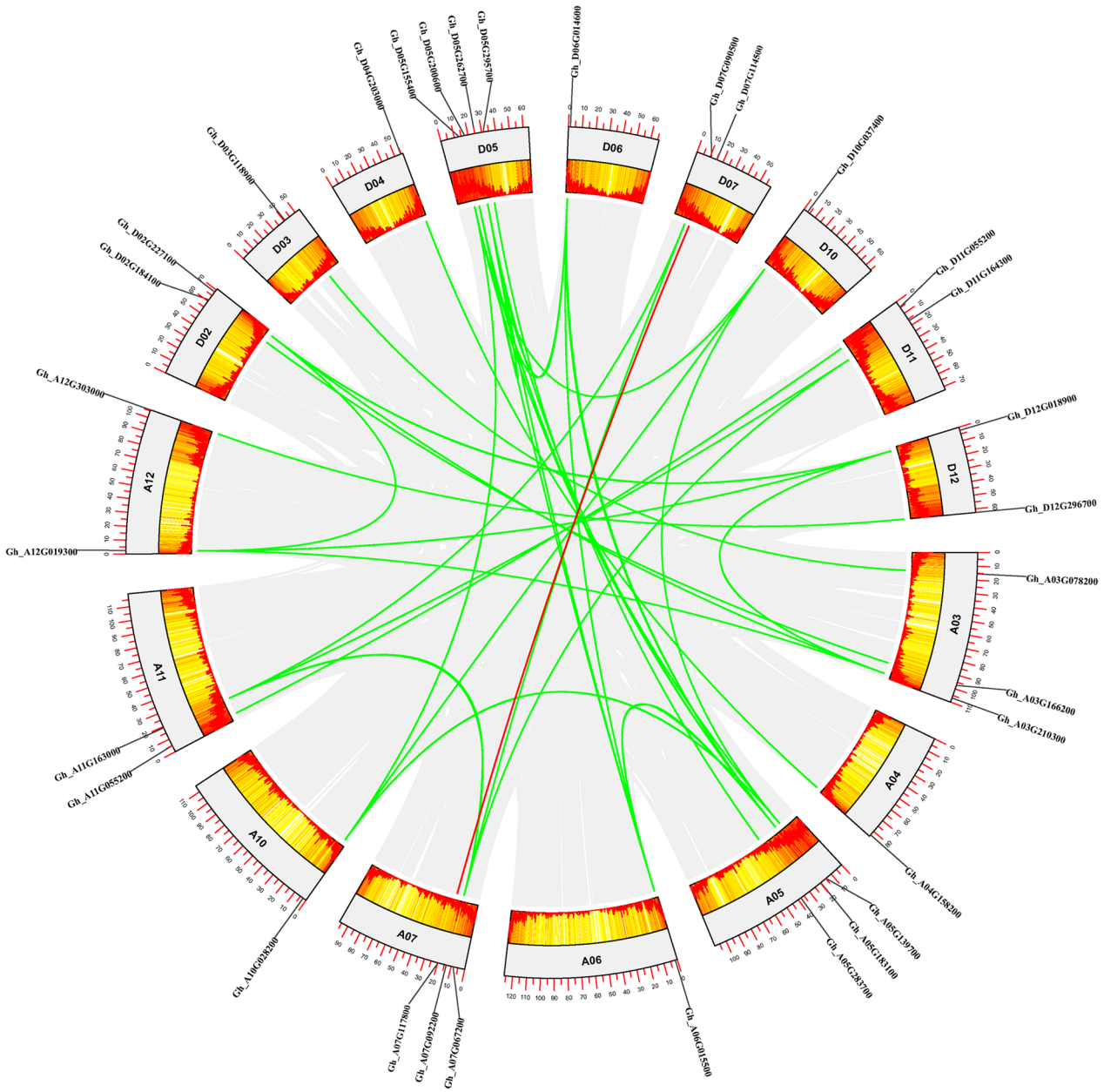

Fig. 3 Collinearity and gene duplication analysis of *NAT* genes in *Gossypium hirsutum*

Promoter analysis of *GhNATs* genes

Promoters were strongly associated with the gene expression patterns and gene function (Xue et al. 2008). In order to predict the function of *NAT* gene family, the upstream *cis*-acting element presented in promoter regions was analyzed, which was found to be related to plant hormones (Abscisic acid, gibberellin, salicylic acid, auxin, MeJA) and abiotic stress (drought, defense and stress, low-temperature responsiveness) (Fig. 5; Additional file 4: Table S4). The distribution of *cis*-acting elements was diversified among different *NAT* genes. Hormones responsive elements were inspected on all genes. For example, MeJA regulation was found in more than half of *NAT* genes, reported that MeJA existed extensively in plants and was closely connection with stress resistance (Yu et al. 2009). *GhNAT12* contained drought, abscisic acid, MeJA, gibberellin, salicylic acid, defense, and low-temperature responsive element . In general, majority of *NAT* promoters contain sequence features associated with drought, MeJA, abscisic acid, salicylic acid, defense,

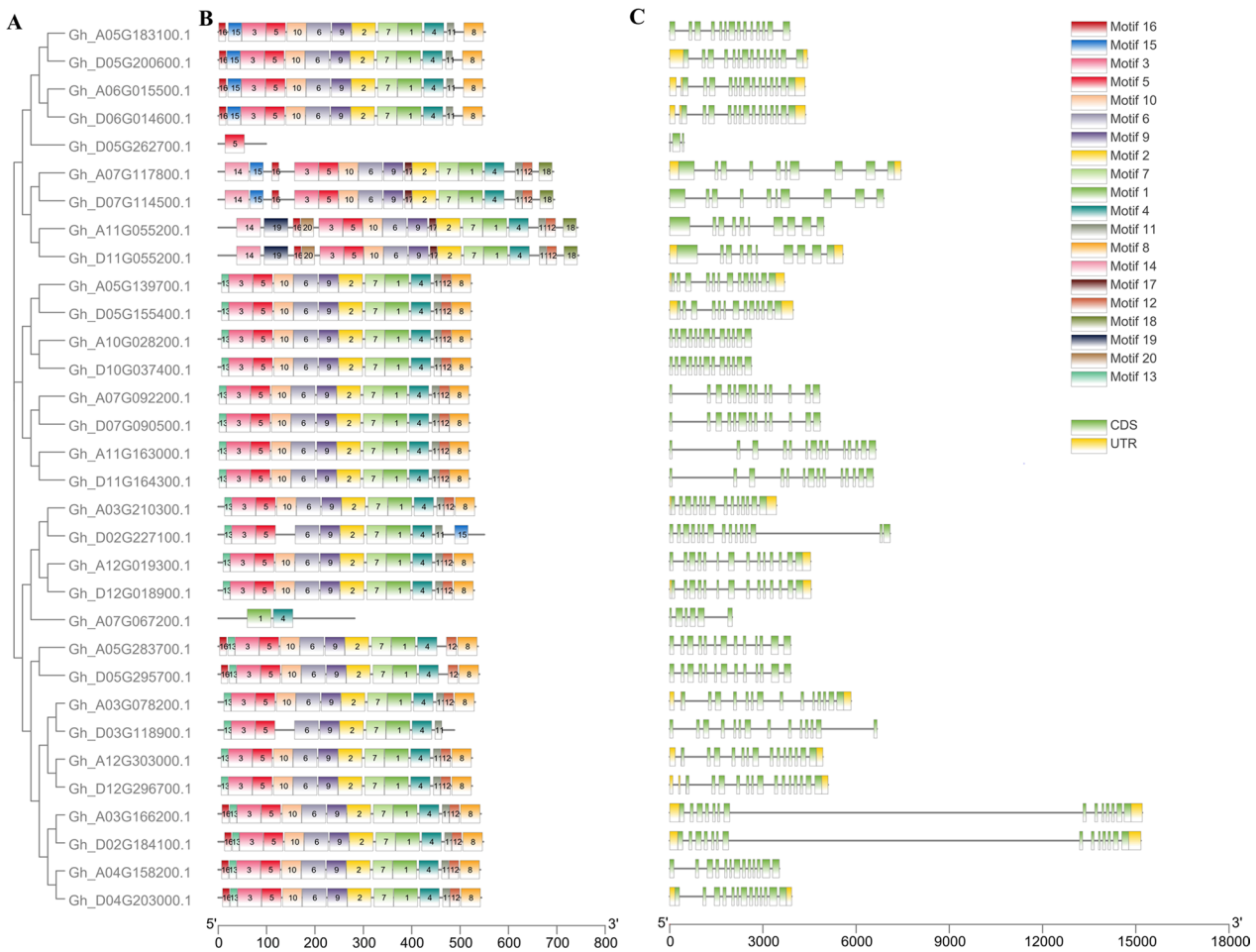

Fig. 4 Phylogenetic analysis, conserved motifs and gene structures of *GhNAT* genes. **A** Phylogenetic tree of *GhNAT* genes. **B** Conserved motifs of GhNAT proteins. Twenty conserved motifs are represented by different color boxes. **C** Exon–intron organization analyses of *GhNAT* genes. Yellow boxes mean Untranslated Regions (UTR), Exons and introns are distinguished by green boxes and black lines, respectively

and stress and low-temperature response. This confirmed our guess that various *cis*-acting elements played very important roles in response to abiotic stress.

Expression analysis of *GhNATs* genes

Through exploring the expression pattern, we detected expression level of *NAT* genes under salt, PEG, cold, and heat stress in different periods (1 h, 3 h, 6 h, 12 h) by RNA-seq data (Additional file 5: Table S5). The gene expression under different stress conditions showed that the *GhNAT* members were related in the regulation of abiotic stress. The results confirmed that different genes showed different expression trends. The expression levels of three genes (*Gh_A06G015500*, *Gh_D06G014600*, *Gh_D05G262700*) were low to high under salt, PEG, cold, and heat stresses, and the expression level peaked at 12 h (Fig. 6). The expression pattern was roughly similar on the same group of genes, but there were slight differences among gene expression under cold and heat stress. *Gh_A12G303000* and *Gh_D12G296700* were significantly highly expressed in all periods of cold, heat, salt, and drought. *Gh_A06G015500*, *Gh_D06G014600*, and *Gh_D05G262700* showed continuous high expression levels under four different stress conditions, and the expression level was high at 12 h. The similar trend of the *NAT* genes expression were detected under PEG and salt stress. Likewise, the expression pattern of *NAT12* genes (*Gh_A07G117800*, *Gh_D07G114500*) were gradually increased under salt stress, and peaked at 12 h. In brief, *GhNAT* genes played an important role in response to plant hormones and abiotic stress, and can provide important basis for gene function analysis.

To further confirm the expression levels of the *GhNATs* under drought and salt stresses, qRT-PCR was performed using the leaves of upland cotton Zhong9807 and H177, which were treated with 200 mmol·L^{-1} NaCl and 12%

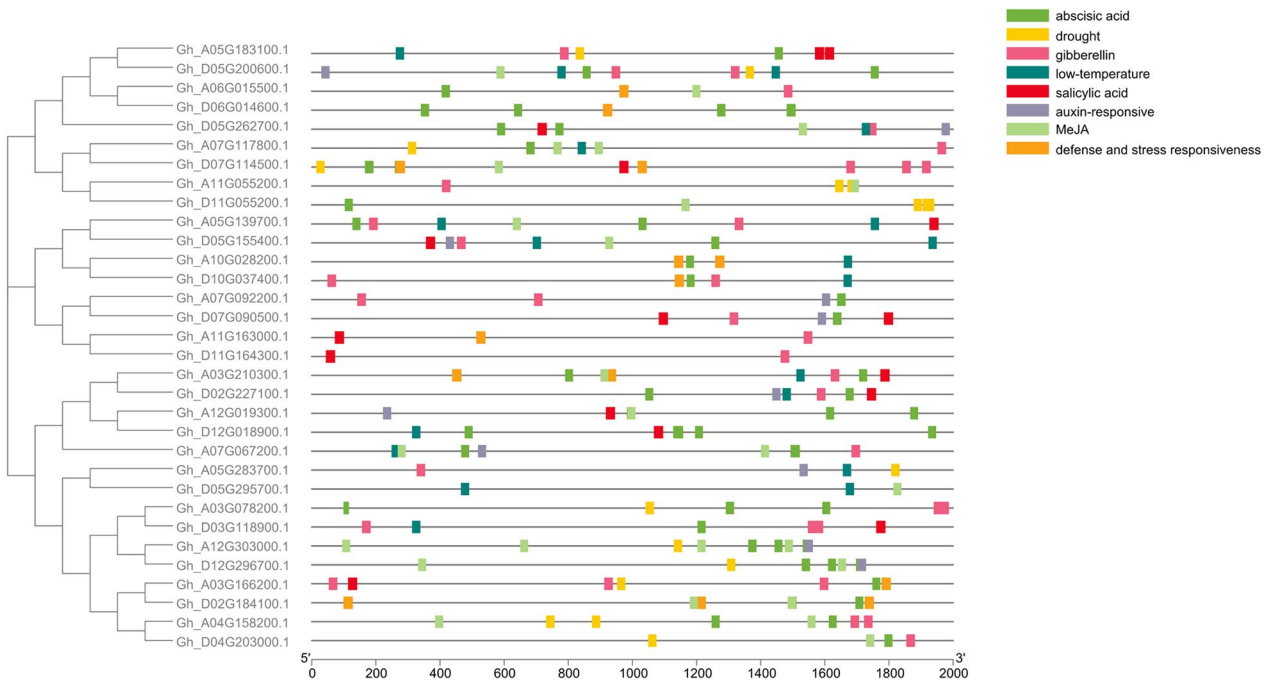

Fig. 5 Analysis *cis*-elements in promoters of *GhNAT* family

PEG6000, respectively (Fig. 7). According to the multiple sequence analysis, the *GhNAT* genes were classified into four groups, similarly, different expression levels of the genes also were divided into four clades. The result showed that most of *NAT* genes were upregulated after salt and PEG treatment, while only a few genes were downregulated in upland cotton. For example, the *GhNAT12* (*Gh_A07G117800*, *Gh_D07G114500*) expression level was generally up-regulated and peaked at 12 h under the salt and drought stresses. *Gh_A03G078200*, *Gh_A03G166200*, *Gh_A06G015500* and *Gh_D06G014600* showed continuously high expression level in generally and were upregulated at 12 h under salt and drought stress conditions.

Silencing *GhNAT12* reduced the resistance of cotton to abiotic stress

To further investigate the role of *GhNAT12* under abiotic stress, virus-induced gene silencing (VIGS) was performed in Zhong9807 and H177. When the albino phenotype appeared on the leaves at about 3 weeks on TRV:*PDS* plants (Fig. 8), the gene silencing efficiency of TRV:*00* and TRV:*GhNAT12* plants was confirmed by qRT-PCR. The results showed that *GhNAT12* expression was significantly decreased in the silenced plants of Zhong9807 compared with TRV:*00* plants (Fig. 8B), and *GhNAT12* expression was also significantly decreased in the TRV:*GhNAT12* plants of H177 than that in TRV:*00* plants (Fig. 8D), which indicated that the gene silencing of *GhNAT12* was successful in TRV:*GhNAT12* plants. Subsequently, the gene-silenced plants were treated with 200 mmol·L^{-1} NaCl and 12% PEG6000 at the three leaf-stage in Zhong9807 and H177, respectively. The result showed that the leaves were not wilting in CK plant treated with water, while, wilting phenotype of leaves began to appear after 24 h under treatment in TRV:*GhNAT12* and TRV:*00* plants. But the TRV:*GhNAT12* plants displayed more severe withered phenotype than the TRV:*00* plants, especially after 48 h. It suggested that the *GhNAT12* gene played an active role in regulating plant resistance to various stresses.

Subcellular localization of *GhNAT12*

To experimentally confirm the GhNAT12 protein localization, we constructed a *GhNAT12-GFP* vector. This vector was transiently expressed in tobacco leave cells and onion epidermal cells. According to the fluorescence signal, the GhNAT12 protein was present in the nucleus and cell membrane (Fig. 9). Meanwhile, GhNAT12-GFP protein was detected on the plasma membrane in tobacco cells. In addition, the experiment of plasmolysis was conducted on onion epidermal cells. The same result appeared that the GhNAT12-GFP fluorescence signal was located on the cell membrane before and after plasmolysis.

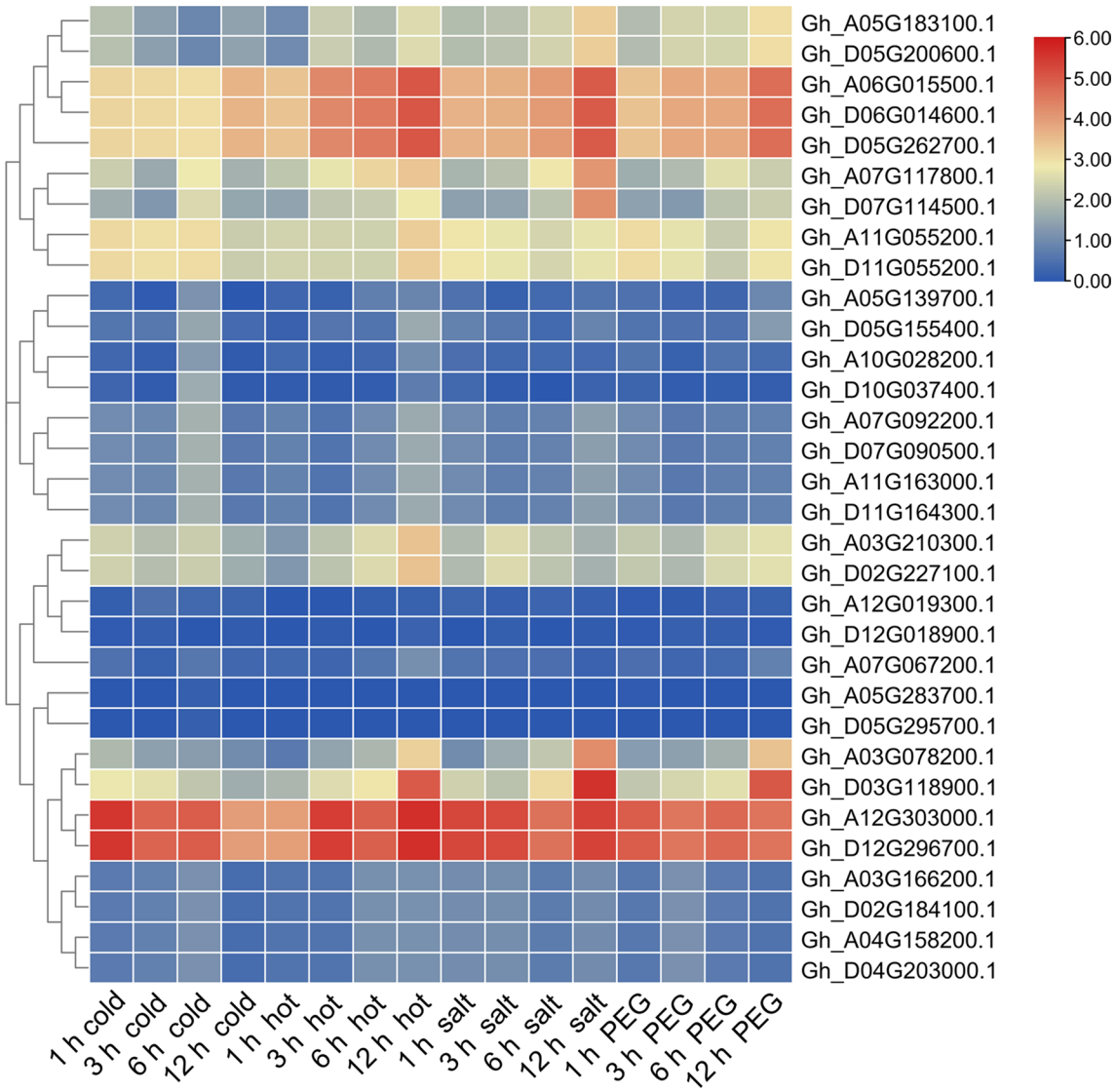

Fig. 6 Differentially expressed *GhNAT* genes under cold, heat, salt and PEG stress conditions.

Discussion

Cotton is cultivated worldwide, as an important cash crop, is facing severe biotic and abiotic stresses when they grow and develop under natural conditions. Although cotton is a relatively salt-tolerant, thermophilic and drought-resistant species, the growth and development of this plant can still be greatly affected by adverse salt and drought conditions (Zhang et al. 2016). So breeders focused on key molecular factors involved in salt and drought stress response and attempted to develop salt-tolerant cotton varieties. The nucleobase–ascorbate transporter gene family was named for its transport of ascorbic acid in mammals (Liang et al. 2001), which had critical functions in many biological processes involved in the plant growth and development. AtNAT3 and AtNAT12 are mainly involved in the transport of adenine, guanine, uracil, and hypoxanthine, which play an important role in maintaining normal physiological metabolism in *Arabidopsis thaliana* (Hunt 2013; Niopek-Witz et al. 2014). Overexpression of MdNAT7 increased the concentrations of xanthine and uric acid in apple, leading to enhanced tolerance to salinity stress (Sun et al. 2021).

The evolutionary process of the *NAT* gene family lead to intriguing functional diversity and broad expansion (De Koning and Diallinas 2000; Frillingos 2012). It is very intriguing that the mammalian L-ascorbate transporters seem to have evolved from uric acid/xanthine

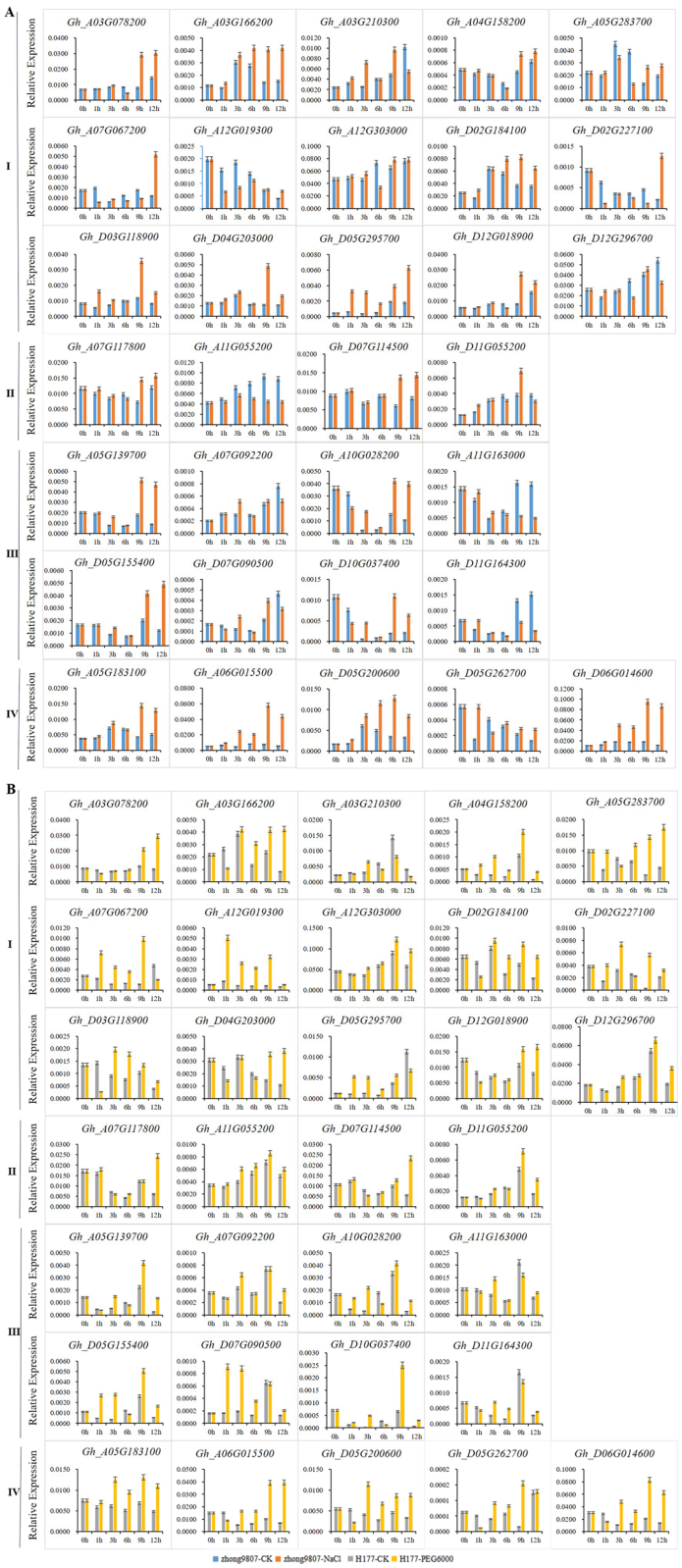

Fig. 7 Expression pattern of *GhNAT* genes in response to abiotic stresses by qRT-PCR. **A** The expression level of *GhNATs* in leaves of Zhong9807 under 0 and 200 mmol·L^{-1} NaCl treatment. **B** The expression level of *GhNATs* in the leaves of H177 under 0 and 12% PEG6000 treatment

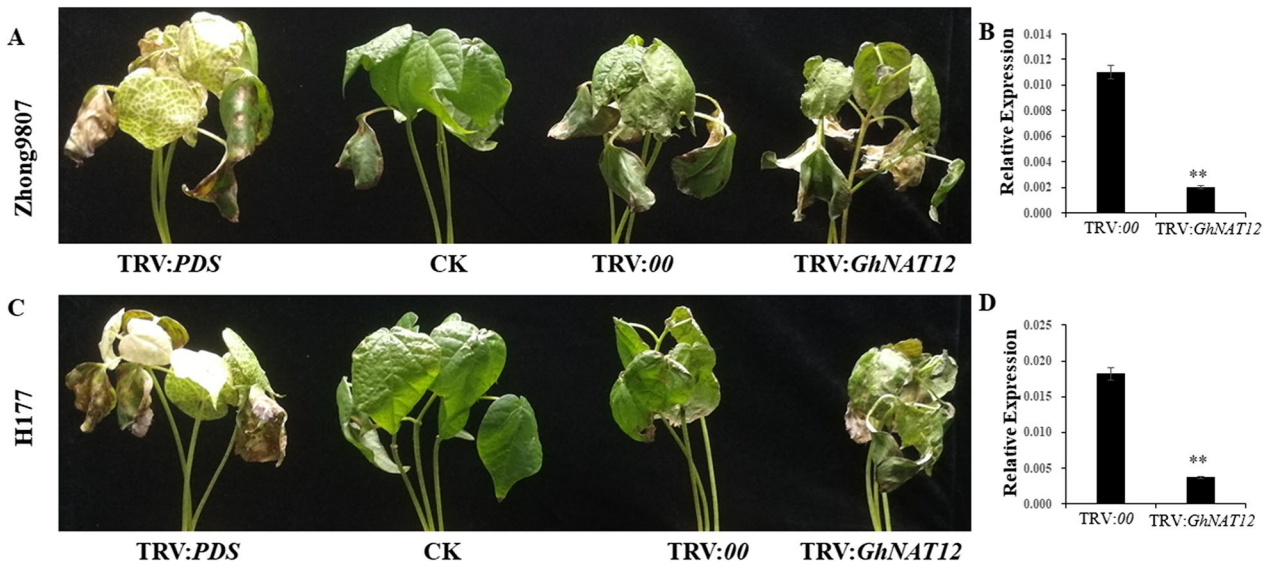

Fig. 8 The phenotype of cotton leaves of *GhNAT*12 silenced Zhong9807 and H177 under NaCl and PEG6000 treatment, respectively. **A** Phenotype of cotton leaves of *GhNAT*12 silenced Zhong9807, the CK was treated with 0 mmol·L^{-1} NaCl, and TRV:*00* and TRV:*GhNAT12* treated with 200 mmol·L^{-1} NaCl. **B** *GhNAT12* gene silence efficiency detection in Zhong9807. **C** Phenotype of cotton leaves of *GhNAT*12 silenced in H177, CK was treated with 0% PEG6000, and TRV: *00* and TRV: *GhNAT12* treated with 12% PEG6000. **D** *GhNAT12* gene silence efficiency detection in H177

transporters of lower eukaryotes, since uric acid and xanthine have entirely different formulas compared with ascorbic acid (Gournas et al. 2008). Several studies of *NAT* genes had been focused on *Arabidopsis*, tomato, apple, maize, and pepper. However, cotton was still lacking for researches about *NATs*. With the development of cotton genome sequencing, it is conceivable to investigate thoroughly on cotton *NAT* genes and study its potential functions. Based on the genome-wide characterization of the *NAT* gene family, 32, 32, 18, and 16 *NAT* genes were determined and characterized in four sequenced cotton species *G. hirsutum*, *G. barbadense*, *G. raimondii,* and *G. arboreum*, respectively. This result indicated that the loss of *NAT* genes appeared in the allotetraploid *G. hirsutum* and *G. barbadense*, and these *NAT* genes were conserved during the evolutionary process in cotton. The allotetraploid cotton evolved from two kinds of diploid cotton containing an A-genome (*G. arboreum*) and D-genome (*G. raimondii*). The number of genes in the allotetraploid cotton was twice that of diploid cotton implying that NAT homologs are evolutionarily conserved. And it was consisted with high rates of gene loss in the allotetraploid plants (Paterson et al. 2012; Zhang et al. 2015).

In present study, the phylogenetic tree has been constructed for 110 *NAT* genes from four cotton species and *Arabidopsis*. These *NATs* were divided into four distinct groups, i.e. group I, II, III and IV. It was similar to those already described for *Arabidopsis*, apple, maize, and tomato family members (Maurino et al. 2006; Cai et al. 2014; Sun et al. 2016). The *NAT* family is one of the five known groups that utilize nucleobases as their principal substrates. They belong to an evolutionarily widespread family of transport proteins in prokaryotes, fungi, plants, and mammals (De Koning and Diallinas 2000; Frillingos 2012). In addition, the number of genes in group I was the largest among different species, and was the most complex *NATs*. It has been previously reported that *ZmLpe1* derived from group I, and characterized as a unique transporter for xanthine and uric acid (Argyrou et al. 2001). This has indicated that the substrate of most plant *NATs* maybe the nucleobase.

In terms of the conserved protein motifs and gene structure, most *NAT* genes from the same group share the similar exon–intron structure and conserved motifs in *G. hirsutum*. It implies that both function and structure of NATs remain conserved during evolution in the same group. Obviously, exon–intron structure and conserved motifs are slightly different among *NAT* gene members, and it may be responsible for their respective functional specificity. In general, this phenomenon suggests that exon–intron structure and conserved motifs have a vital role in the distinctive function. And during evolution, gene fusions or chromosomal rearrangement occurr that enlarge the NAT domain, ultimately producing the typical domain (Chai et al. 2018). The gene structure of *GhD02G227100.1*, *GhA03G166200.1* and *GhD02G184100*.1 contain large intron, it worth further investigation of whether transposon insertion events

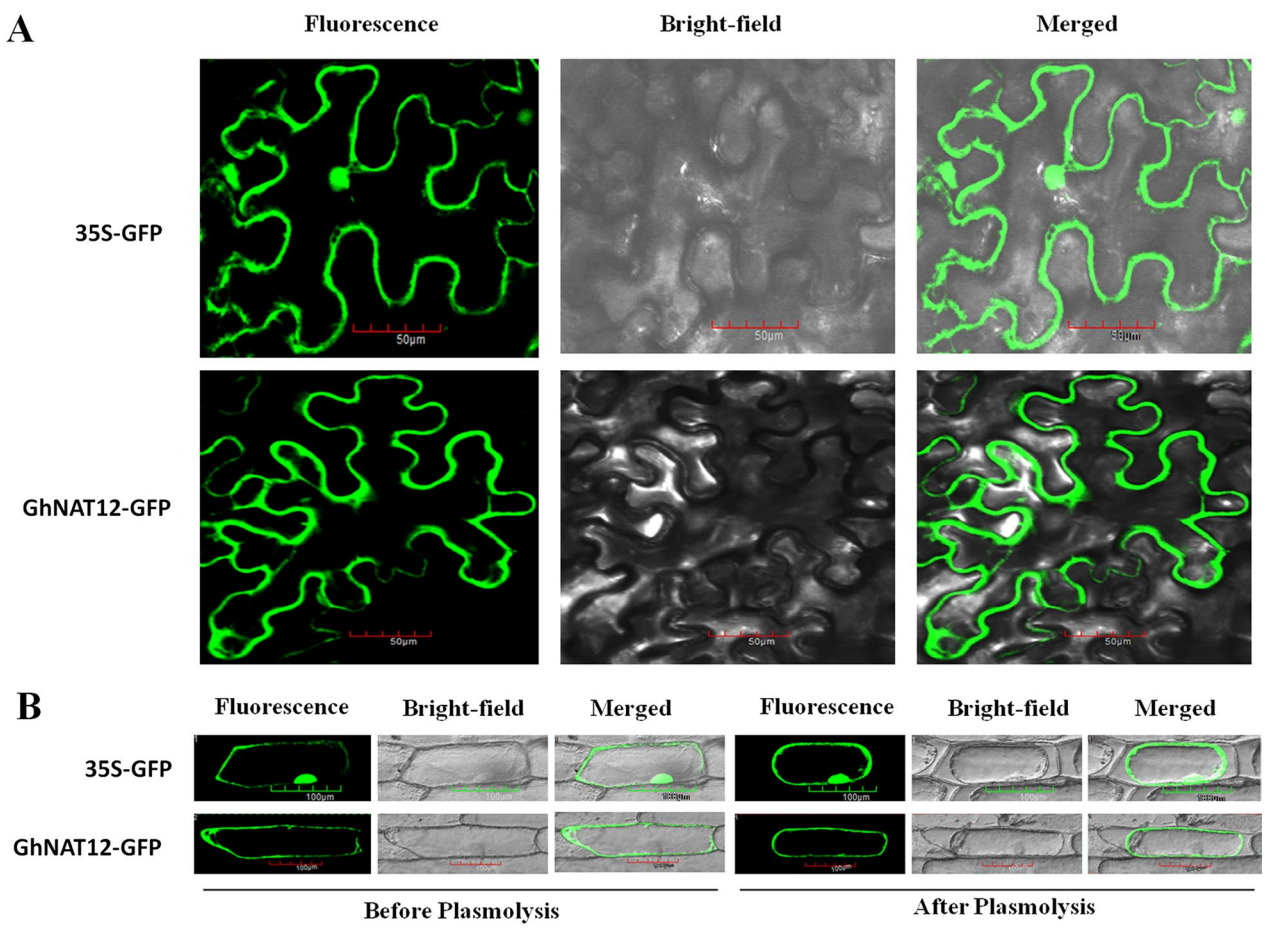

Fig. 9 Subcellular localization of GhNAT12. **A** The GhNAT12-GFP signals were localized to the plasma membrane. **B** GhNAT12-GFP and GFP signals on the plasmolysis of onion cells

have occurred. According to the gene expression profile, those three genes' expression level were relatively low, and the biological meaning needs further study. As for the chromosome position, the uneven distribution of *NAT* genes was displayed on specific chromosomes in At and Dt sub-genomes. The distribution pattern on the two sub-genomes tend to be the same. However, different distribution of some genes might be lost or gained compared with At and Dt sub-genomes. According to the collinearity analysis, there were 34 pairs of segmental duplication and 1 pair of tandem duplication in *G. hirsutum,* and they played a crucial role in the expansion of the gene family. According to the previously reported findings, one of the main drivers of the genome evolution is gene duplication (Moore and Purugganan 2003). Tandem duplication, segmental duplication, and whole-genome duplication played a crucial role in the expansion of gene family (Xu et al. 2012).

When plants are threatened by abiotic stresses, *cis*-acting elements played an important role (Nakashima et al. 2014). It was consistent with the previous research that similar *cis*-acting elements in promoter regions had relevant function (Fan et al. 2021), and the *NAT* genes might be associated with various physiological and biochemical processes refer to previous studies. The *cis*-acting elements of the *NAT* family contained stimulating phytohomones (abscisic acid, gibberellin, salicylic acid, auxin, MeJA) and various stresses (drought, defense and stress, low-temperature responsiveness), these results indicated that *NAT* genes not only participated in multiple signaling pathways but also took part in plant growth, development and defensive responses. In addition, qRT-PCR analysis was used to verify the results of the transcriptome, they were generally the same in the upregulation or downregulation. Some differences may be related to experimental conditions. The gene expression profile provides important clues for *GhNAT* genes in response to abiotic stress. For example, among all the 14 *NAT* genes in apple, only *MdNAT6* proved responsive to both drought and salinity (Sun et al. 2016).

In this study, it is proved that *GhNAT12* gene played an active role in regulating plant resistance to salt and drought stress. Also *GhNAT12* was detected in the plasma membrane, which was consistent with the discovered localization of *AtNAT12*, *AtNAT7* and *AtNAT8* in the plasma membrane (Maurino et al. 2006). The allocation of NATs in different clades correlates with their expression during the life cycle of Arabidopsis. Some of the members of this gene family show unique expression (*AtNAT12*), while the expression of other *AtNAT* genes is restricted to specific tissues (*AtNAT7*, *AtNAT8*, *AtNAT9*) (Maurino et al. 2006; Niopek-Witz et al. 2014). Interestingly, these genes were in cluster I. In addition, *MdNAT7* overexpression in transgenic apple plants improved their tolerance to salinity stress compared with non-transgenic plants, while opposite phenotypes were observed for *MdNAT7* RNAi plants (Sun et al. 2021).

Conclusions

In this study, a comprehensive analysis of *NAT* gene family was performed in cotton. A total of 98 *NAT* genes were identified in four cotton species, including 32 genes in *G. hirsutum*. These genes were classified into four groups in the phylogenetic tree. Chromosomal location and collinearity analysis illustrated segmental duplication and tandem duplication might play a vital role in the expansion of the *NAT* gene family. Conserved motifs and gene structure analysis demonstrated evolutionary difference and conservation. The expression patterns suggested that the *NAT* family genes were associated with plant response to various adversity. Further, it was revealed that the *GhNAT12* gene played an important role in regulating plant resistance to salt and drought stress. The present study lays a foundation for further study to explore the detailed present function of the *NAT* genes in cotton.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42397-022-00118-7.

Additional file 1: Table S1. List of primers used in this study.

Additional file 2: Table S2. Characteristics of *NAT* genes in cotton.

Additional file 3: Table S3. Duplication events of *NAT* genes in cotton.

Additional file 4: Table S4. *Cis*-element predicted in promotors of *GhNAT* gene family members.

Additional file 5: Table S5. RNA-seq data analysis of *GhNAT* genes.

Acknowledgements

We would like to thank the anonymous reviewers for their valuable comments and helpful suggestions which help to improve the manuscript. Meanwhile, we are grateful to Dr. Zareen Sarfraz giving help for this paper.

Authors' contributions

Formal analysis, Guo LX; investigation, Guo LX; software, Zhao LJ, Lu XK, Chen XG, Chen C, Fan YP and Zhang YX; validation, Guo LX, Wang S, Wang JJ, Wang DL, Yin ZJ, Zhang YX and Fan YP; visualization, Guo LX, Zhao LJ and Wang S; original draft, Guo LX; reviewing and editing, Ye WW. All authors read and approved the final manuscript.

Funding

This work was supported by China Agriculture Research System of MOF and MOARA, and Agricultural Science and Technology Innovation Program of Chinese Academy of Agricultural Science.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 17 September 2021 Accepted: 7 March 2022
Published online: 21 March 2022

References

Argyrou E, Sophianopoulou V, Schultes N, et al. Functional characterization of a maize purine transporter by expression in *Aspergillus nidulans*. Plant Cell. 2001;13(4):953–64. https://doi.org/10.2307/3871352.

Bailey TL, Boden M, Buske FA, et al. MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(suppl_2):W202–8. https://doi.org/10.1093/nar/gkp335.

Bürzle M, Suzuki Y, Ackermann D, et al. The sodium-dependent ascorbic acid transporter family SLC23. Mol Asp Med. 2013;34(2–3):436–54. https://doi.org/10.1016/j.mam.2012.12.002.

Cai XF, Ye J, Hu T, et al. Genome-wide classification and expression analysis of nucleobase–ascorbate transporter (NAT) gene family in tomato. Plant Growth Regul. 2014;73:19–30. https://doi.org/10.1007/s10725-013-9864-x.

Chai WB, Peng XJ, Liu B, et al. Comparative genomics, whole-genome re-sequencing and expression profile analysis of nucleobase:cation symporter 2 (NCS2) genes in maize. Front Plant Sci. 2018;9:856. https://doi.org/10.3389/fpls.2018.00856.

Chen MX, Lung SC, Du ZY, et al. Engineering plants to tolerate abiotic stresses. Biocatal Agric Biotechnol. 2014;3(1):81–7. https://doi.org/10.1016/j.bcab.2013.09.010.

Chen C, Chen H, He Y, et al. TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. BioRxiv. 2018;289660. https://doi.org/10.1016/j.molp.2020.06.009.

Cornic G. Drought stress inhibits photosynthesis by decreasing stomatal aperture-not by affecting ATP synthesis. Trends Plant Sci. 2000;5(5):187–8. https://doi.org/10.1016/S1360-1385(00)01625-3.

De Koning H, Diallinas G. Nucleobase transporters (review). Mol Membr Biol. 2000;17(2):75–94. https://doi.org/10.1080/09687680050117101.

Diallinas G, Valdez J, Sophianopoulou V, et al. Chimeric purine transporters of *Aspergillus nidulans* define a domain critical for function and specificity conserved in bacterial, plant and metazoan homologues. EMBO J. 1998;17(14):3827–37. https://doi.org/10.1093/emboj/17.14.3827.

Edgar RC. Muscle: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. https://doi.org/10.1093/nar/gkh340.

Fan YP, Zhang YX, Rui C, et al. Zinc finger transcription factor ZAT family genes confer multi-tolerances in *Gossypium hirsutum* L. J Cotton Res. 2021;4:24. https://doi.org/10.1186/s42397-021-00098-0.

Frillingos S. Insights to the evolution of nucleobase–ascorbate transporters (NAT/NCS2 family) from the Cys-scanning analysis of xanthine permease XanQ. Int J Biochem Mol Biol. 2012;3(3):250–72.
Gill SS, Tuteja N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol Biochem. 2010;48(12):909–30. https://doi.org/10.1016/j.plaphy.2010.08.016.
Gournas C, Papageorgiou I, Diallinas G. The nucleobase ascorbate transporter (NAT) family: genomics, evolution, structure, function relationships and physiological role. Mol Biosyst. 2008;4(5):404–16. https://doi.org/10.1039/b719777b.
Grattan SR, Grieve CM. Salinity–mineral nutrient relations in horticultural crops. Sci Hortic Amst. 1998;78(1–4):127–57. https://doi.org/10.1016/S0304-4238(98)00192-7.
Hirayama T, Shinozaki K. Research on plant abiotic stress responses in the post-genome era: past, present and future. Plant J. 2010;61:1041–52. https://doi.org/10.1111/j.1365-313X.2010.04124.x.
Hu Y, Chen J, Fang L, 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. https://doi.org/10.1038/s41588-019-0371-5.
Hunt KA. Functional characterization of the nucleobase–ascorbate transporter family of *Arabidopsis thaliana*. Hosp J. 2013;198:239–50. https://doi.org/10.1016/j.scienta.2015.09.034.
Jithesh MN, Prashanth SR, Sivaprakash KR, et al. Antioxidative response mechanisms in halophytes: their role in stress defence. J Genet. 2006;85(3):237–54. https://doi.org/10.1007/BF02935340.
Jones P, Binns D, Chang HY, et al. Interproscan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236–40. https://doi.org/10.1093/bioinformatics/btu031.
Karatza P, Panos P, Georgopoulou E, et al. Cysteine-scanning analysis of the nucleobase–ascorbate transporter signature motif in YgfO permease of *Escherichia coli*: Gln-324 and Asn-325 are essential and Ile-329-Val-339 form an α-helix. J Biol Chem. 2006;281(52):39881–90. https://doi.org/10.1074/jbc.M605748200.
Kosti V, Lambrinidis G, Myrianthopoulos V, et al. Identification of the substrate recognition and transport pathway in a eukaryotic member of the nucleobase–ascorbate transporter (NAT) family. PLoS ONE. 2012;7(7):e41939. https://doi.org/10.1371/journal.pone.0041939.
Kourkoulou A, Pittis AA, Diallinas G. Evolution of substrate specificity in the Nucleobase-Ascorbate Transporter (NAT) protein family. Microbial Cell. 2018;5:280–92. https://doi.org/10.15698/mic2018.06.636.
Koukaki M, Vlanti A, Goudela S, et al. The nucleobase–ascorbate transporter (NAT) signature motif in UapA defines the function of the purine translocation pathway. J Mol Biol. 2005;350(3):499–513. https://doi.org/10.1016/j.jmb.2005.04.076.
Krzywinski M, Schein J, Birol I, et al. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19(9):1639–45. https://doi.org/10.1101/gr.092759.109.
Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4. https://doi.org/10.1093/molbev/msw054.
Li FG, Fan GY, Lu CR, et al. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. Nat Biotechnol. 2015;33:524–30. https://doi.org/10.1038/nbt.3208.
Liang WJ, Johnson D, Jarvis SM. Vitamin C transport systems of mammalian cells. Mol Membr Biol. 2001;18(1):87–95. https://doi.org/10.1080/09687680110033774.
Lu XK, Wang XG, Chen XG, et al. Single-base resolution methylomes of upland cotton (*Gossypium hirsutum* L.) reveal epigenome modifications in response to drought stress. BMC Genomics. 2017;18:297. https://doi.org/10.1186/s12864-017-3681-y.
Maurino VG, Grube E, Zielinski J, et al. Identification and expression analysis of twelve members of the nucleobase–ascorbate transporter (NAT) gene family in *Arabidopsis thaliana*. Plant Cell Physiol. 2006;47(10):1381–93. https://doi.org/10.1093/pcp/pcl011.
Moore RC, Purugganan MD. The early stages of duplicate gene evolution. Proc Natl Acad Sci. 2003;100(26):15682–7. https://doi.org/10.1073/pnas.2535513100.
Muchate NS, Nikalje GC, Rajurkar NS, et al. Plant salt stress: adaptive responses, tolerance mechanism and bioengineering for salt tolerance. Bot Rev. 2016;82(4):371–460. https://doi.org/10.1007/s12229-016-9173-y.
Nakashima K, Yamaguchi-Shinozaki K, Shinozaki K. The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. Front Plant Sci. 2014;5:170. https://doi.org/10.3389/fpls.2014.00170.
Niopek-Witz S, Deppe J, Lemieux MJ, et al. Biochemical characterization and structure-function relationship of two plant NCS2 proteins, the nucleobase transporters NAT3 and NAT12 from *Arabidopsis thaliana*. Biochim Biophys Acta. 2014;1838(12):3025–35. https://doi.org/10.1016/j.bbamem.2014.08.013.
Paterson AH, Wendel JF, Gundlach H, et al. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature. 2012;492(7429):423–7. https://doi.org/10.1038/nature11798.
Qi JG, Sun SM, Yang L, et al. Potassium uptake and transport in apple roots under drought stress. Hortic Plant J. 2019;5(1):10–6. https://doi.org/10.1016/j.hpj.2018.10.001.
Roy SJ, Negrao S, Tester M. Salt resistant crop plants. Curr Opin Biotechnol. 2014;26:115–24. https://doi.org/10.1016/j.copbio.2013.12.004.
Schultes NP, Brutnell TP, Allen A, et al. Leaf permease1 gene of maize is required for chloroplast development. Plant Cell. 1996;8(3):463–70. https://doi.org/10.1105/tpc.8.3.463.
Slama I, Abdelly C, Bouchereau A, et al. Diversity, distribution and roles of osmoprotective compounds accumulated in halophytes under abiotic stress. Ann Bot. 2015;115(3):433–47. https://doi.org/10.1093/aob/mcu239.
Sun TT, Jia DF, Huang LL, et al. Comprehensive genomic identification and expression analysis of the nucleobase–ascorbate transporter (NAT) gene family in apple. Sci Hortic. 2016;198:473–81. https://doi.org/10.1016/j.scienta.2015.09.034.
Sun TT, Pei TT, Yang LL, et al. Exogenous application of xanthine and uric acid and nucleobase–ascorbate transporter *MdNAT7* expression regulate salinity tolerance in apple. BMC Plant Biol. 2021;21(1):52. https://doi.org/10.1186/s12870-021-02831-y.
Wang Y, Tang H, DeBarry JD, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49. https://doi.org/10.1093/nar/gkr1293.
Wang XG, Lu XK, Wang JJ, et al. Mining and analysis of SNP in response to salinity stress in upland cotton (*Gossypium hirsutum* L.). PLoS ONE. 2016;11(6):e0158142. https://doi.org/10.1371/journal.pone.0158142.
Wang XF, Xu M, Gao CJ, et al. The roles of endomembrane trafficking in plant abiotic stress responses. J Integr Plant Biol. 2020a;62(1):55–69. https://doi.org/10.1111/jipb.12895.
Wang XG, Lu XK, Waqar AM, et al. Differentially expressed bZIP transcription factors confer multi-tolerances in *Gossypium hirsutum* L. Int J Biol Macromol. 2020b;146:569–78. https://doi.org/10.1016/j.ijbiomac.2020.01.013.
Xu G, Guo C, Shan H, et al. Divergence of duplicate genes in exon-intron structure. Proc Natl Acad Sci. 2012;109(4):1187–92. https://doi.org/10.1073/pnas.1109047109.
Xue T, Wang D, Zhang S, et al. Genome-wide and expression analysis of protein phosphatase 2C in rice and *Arabidopsis*. BMC Genomics. 2008;9(1):550. https://doi.org/10.1186/1471-2164-9-550.
Yu M, Shen L, Fan B, et al. The effect of MeJA on ethylene biosynthesis and induced disease resistance to *Botrytis cinerea* in tomato. Postharvest Biol Technol. 2009;54(3):153–8. https://doi.org/10.1016/j.postharvbio.2009.07.001.
Yu J, Jung S, Cheng CH, et al. CottonGen: a genomics, genetics and breeding database for cotton research. Nucleic Acids Res. 2013;42:D1229–36. https://doi.org/10.1093/nar/gkt1064.
Zhang TZ, Hu Y, Jiang WK, et al. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol. 2015;33(5):531–7. https://doi.org/10.1038/nbt.3207.
Zhang F, Zhu G, Du L, et al. Genetic regulation of salt stress tolerance revealed by RNA-Seq in cotton diploid wild species *Gossypium Davidsonii*. Sci Rep. 2016;6:20582. https://doi.org/10.1038/srep20582.
Zhang YX, Fan YP, Rui C, et al. Melatonin improves cotton salt tolerance by regulating ROS scavenging system and Ca^{2+} signal transduction. Front Plant Sci. 2021;12:693690. https://doi.org/10.3389/fpls.2021.693690.
Zhu T, Liang C, Meng Z, et al. CottonFGD: an integrated functional genomics database for cotton. BMC Plant Biol. 2017;17(1):1–9. https://doi.org/10.1186/s12870-017-1039-x.