Published April 15, 2016

RESEARCH

Genome-wide Identification and Structural Analysis of Pyrophosphatase Gene Family in Cotton

Xiaojie Zhao, Xuke Lu, Zujun Yin, Delong Wang, Junjuan Wang, Weili Fan, Shuai Wang, Tianbao Zhang, Wuwei Ye*

ABSTRACT

Pyrophosphatase is a hydrolytic enzyme that utilizes pyrophosphate as a substrate. H^+-PPase, an important proton pump on the vacuolar membrane, plays an important role in regulating cell expansion, H^+ electrochemical gradient, and secondary active transport of inorganic ions, organic acids and saccharides. Under low temperature, drought, high NaCl concentration, and hypoxia, H^+-PPase gene expression is promoted to maintain the intracellular balance and enhance the ability of stress resistance in plants. The *GhVP* gene family is a class of H^+-PPase genes in *Gossypium hirsutum* L. The present study identified 13 H^+-PPase genes from diploid *Gossypium raimondii* Ulbrich (genome D) and *Gossypium arboreum* L. (genome A), and analyzed their genetic structures and systematic evolution. The 13 H^+-PPase genes, including six and seven from diploid *G. arboreum* and *G. raimondii*, respectively, were obtained with *GhVP* as a reference. The H^+-PPase genes belong to a relatively conserved family in plant evolution. Phylogenetic analysis indicated that the genes of the same species showed high sequence similarities and relatively close genetic relationships. Gene structural analysis indicated that most of the H^+-PPase genes had 7–8 exons, with a few having a variable number (1–14) of exons. By analyzing the conserved domain, the H^+-PPase genes showed several relatively conserved structural domains, such as a myristyl site, casein kinase site II, and protein kinase phosphorylation site. These findings provide a basis for the further characterization of the H^+-PPase gene function in cotton. Analysis by qRT-PCR showed that a different expression level of gene members in the V-PPase family existed in different tissues; for example, gene expression was highest in leaves and lowest in fibers to suggest that V-PPase expression is tissue-specific. Moreover, with the salt stress intensity increasing, the changes of expression level of the V-PPase genes were different.

X. Zhao, X. Lu, Z. Yin, D. Wang, J. Wang, W. Fan, S. Wang, T. Zhang, and W. Ye, State Key Laboratory of Cotton Biology and Institute of Cotton Research of Chinese Academy of Agricultural Sciences, Anyang 455000, Henan, China. X. Zhao and X. Lu contributed equally to this work. Received 16 Dec. 2014. Accepted 6 Jan. 2016. Assigned to Associate Editor Peng Chee. *Corresponding author (yew158@163.com).

Abbreviations: PPase, pyrophosphatase; JGI, Joint Genome Institute.

Pyrophosphatase (PPase), a hydrolytic enzyme utilizing pyrophosphate as substrate, takes part in a variety of metabolic pathways such as synthesis of saccharides, nucleic acids and proteins. At present, H^+-PPase and its analogues have been described in plants, algae, protozoa and photosynthetic bacteria, but not in fungi and animals (Maeshima, 2000). In higher plants, H^+-PPase is usually located on the vacuolar membrane; thus, it is also named V-PPase. The enzyme consists of a single 80 kDa polypeptide, utilizes the phosphate with a high-energy phosphate bond, and coexists with H^+-ATPase on the vacuolar membrane. Pyrophosphatases are mainly divided into three classes: soluble PPases, membrane-bound PPases, and H^+-PPases (Maeshima, 2000). Of the three classes, soluble PPases have been well described, while H^+-PPases differ in other structural domains except for the catalytic site (Cooperman et al., 1992). In addition, only H^+-PPases have the ability to transport H^+. There are two types of H^+-PPases in plant cells: type I and II. Type I H^+-PPase is mainly distributed on the vacuolar membrane, is K^+ sensitive, and could be activated by Ca^{2+}. An example is *AVP1* of *Arabidopsis thaliana* (L.) Heynh. Type II H^+-PPase is mainly distributed in the Golgi apparatus and lysosomes, and is insensitive to K^+ but hypersensitive to Ca^{2+}; the *AVP2* of *A. thaliana* is an example, which only has 36% homology with *AVP1* (Mitsuda et al., 2005). Vacuoles occupy most of plant cell volume, especially in mature plant cells. The vacuole, a multifunctional organelle,

Published in Crop Sci. 56:1831–1840 (2016).
doi: 10.2135/cropsci2014.12.0832

keeps the intracellular ion balance and recycles degrading intracellular proteins (Marty, 1999). H^+-PPase is an important proton pump on the vacuolar membrane and plays an important role in regulating cell expansion, H^+ electrochemical gradient and secondary active transport of inorganic ions, organic acids, and saccharides. Therefore, gene identification and structural analysis of H^+-PPases in cotton is significant for studying the salt tolerance in cotton and the H^+-PPase function.

As previously reported, one of the *A. thaliana* V-H^+-PPase family genes, *AVP1*, was shown to generate a H^+ electrochemical gradient through the vacuolar membrane (Shin and Podskarbi, 2004). The overexpression of the *AVP1* gene in *A. thaliana* enhanced salt tolerance and drought resistance (Gaxiolaet al., 1999). Overexpression of the *Thellungiella salsuginea* V-H^+-PPase gene in cotton resulted in increased salt tolerance in the latter plant (Lv et al., 2008). Numerous studies have shown that V-H^+-PPase is a regulatory factor that can improve plant stress resistance. Amino acid sequences coded by V-PPase genes cloned from *A. thaliana*, beet (*Beta vulgaris*), mung bean (*Vigna radiate*), tobacco (*Nicotiana tabacum)*, rice (*Oryza sativa*), bunge (*Suaeda glauca*), *Thellungiella salsuginea*, barley (*Hordeum vulgare)*, and cotton (*Gossypium* spp.) are highly conserved with >80% pairwise similarities. There are at least six homologous genes in rice that encode V-PPase, five in corn, four in tobacco, three in wheat, and two each in *A. thaliana*, barley, grape, and nectarine (Fukuda et al., 2004; Fonseca-de-Souza et al., 2014). We have obtained the *GhVP* gene by RACE and RT-PCR techniques, and demonstrated a strong hydrophobicity for *GhVP,* which showed 14 transmembrane regions as assessed by bioinformatics analysis.

At present, studies assessing plant PPase genes mainly concerned the cloning, identification, or functional analysis of a single gene; fewer analyses have been performed for this gene family at the whole genome level. In 2012, the genome of diploid *G. raimondii* D5 (genome D) was completed; the following year, the genome sequence of diploid *G. arboreum* A2-8 (genome A) was made available (Wang et al., 2012; Li et al., 2014). The whole genome sequence of *G. raimondii* (genome D) was completed by the US Department of Energy Joint Genome Institute (JGI) in 2012 (Paterson et al., 2012).The whole genome completion of diploid cotton provides opportunities for comprehensive analysis and comparisons within the PPase gene family. At present, this gene has been cloned from rice, barley, beet, and corn, and has been demonstrated to participate in the response of plants to adversity stress by analysis of stress-induced expression, tissue-specific expression, and its promoter sequence (Tanaka et al., 1993; Kim et al., 1994; Gaxiola et al., 2001; Brini et al., 2005; Ozolina et al., 2009). Since the PPase genes of cotton have homology with rice, barley, beet, and corn, it is speculated that PPase genes in cotton are related to adversity stress, such as chilling stress, drought and salt stress, and the secondary active transport of iron for example.

The *GhVP* gene and its homologous genes were analyzed by the means of bioinformatics on the cotton genome sequence. Results showed that sequence pairwise similarities were higher than 80%, and the open reading frames of these homologous genes were highly conserved. However, the non-transcribed regions were less conserved to suggest that the pyrophosphatase on the vacuole membrane belongs to a membrane protein family coded by polygenes. This study analyzed the *GhVP* gene based on the diploid cotton whole genome to evaluate the distribution of V-PPase genes in cotton. The findings provided here are significant in studying the salt-tolerance mechanism of cotton to improve salt tolerance in this species.

MATERIALS AND METHODS

Recognition and Identification of V-PPase Gene Family in Cotton

Taking *GhVP* sequences (GeneBank No. HM370494.1) as reference, local blast analysis against genomic data of diploid *G. raimondii* D5 and diploid *G. arboreum* A2-8 was performed. In combination of V-PPase gene homology analysis with species of *A. thaliana* among other plants (US National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov), the prediction of cotton V-PPase gene family was undertaken with Basic Local Alignment Search Tool (BLAST, US National Center for Biotechnology Information, http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Analysis of V-PPase Gene Family

Motif analysis was performed through online tools (MEME Suite, http://meme-suite.org/). By using Perl program the data for genomes D and A were parsed and information was selected for chromosome locations and structures of the V-PPase genes. Sequence alignment was done by use of DNAMAN 5.2.9 Demo version software. Finally the phylogenetic trees were generated using the neighbor joining methods and MEGA 6.0 software (Tamura et al., 2013). Functional domain analysis was performed with online scan prosite tool (http://prosite.expasy.org/scanprosite/).

Materials and Processing Methods

The materials used were the *G. hirsutum* lines Zhong 9807 and Zhong s9612, *G. arboreum* Shixiya I, and *G. raimondii* D5-3. All materials were provided from the Institute of Cotton Research, Chinese Academy of Agricultural Sciences (CAAS, Anyang, China). Seedlings were raised by sand culture (Olcott, 1939).

RNA Isolation and qRT-PCR Analysis

Samples of cotton roots, stems, leaves, flowers, and fibers were used to extract RNA, respectively. Afterward the extracted RNA was analyzed by qRT-PCR analysis. Tissue from *G. arboreum* Shixiya I and *G. raimondii* D5-3 were treated with 0.4% NaCl (0.4% by weight of the sand used in the pots) at the trefoil stage. We kept the seedlings in a humid environment for 24 h

after the treatment. Plant roots, stems, leaves were harvested at 0, 6, 12, and 24 h after the treatment. The samples were quickly frozen in liquid nitrogen for use. RNA was extracted (Carra et al., 2007), and the concentration and quality of total RNA were determined with Nanodrop2000 nucleic acid. The A260/280 ratio for each RNA sample was 2.0. Then, RNA was reversely transcribed to first strand cDNA by use of PrimeScript RT reagent kit with gDNA eraser (TaKaRa, China). The fluorescent quantitative primers were designed for all family members using Premier 5 (Supplemental Table S1). Actin gene served as a reference. The synthesized cDNA was pre-incubated at 95°C for 15 s, followed by 40 cycles of denaturation at 95°C for 5 s and extension at 60°C for 34 s. The fluorescence quantitative assay was used to analyze expression level of the V-PPase genes in different tissues of cotton, and expression changes in *G. raimondii* and *G. arboreum* under salt stress. The assay was designed with three replicates and the results were analyzed with the double delta Ct method.

RESULTS

Identification of Family Members and Analysis of Structural Domains at the Whole Genome Level

Plant PPase is a unique proton pump in plant cells, which plays an important role in maintaining the vacuole function and adjusting the osmotic pressure balance. As shown in Table 1, six *GhVP*-like genes named *GaVP1* through *GaVP6* and seven named *GrVP*7 through *GrVP13* were found in genomes A (*G. arboreum* A2-8) and D (*G. raimondii* D5), respectively. The number of amino acid residues in the V-PPase protein ranged from 750 to 790, with the majority between 760 and 770. The pI range was 5.11 to 6.13, and the protein molecular weight ranged from 79.32 to 84.51 kDa. All V-PPase proteins were located in the plasma membrane according to a subcellular localization prediction. However, a further experimental verification is required.

Table 1. Basic characteristics of V-PPase genes in the cotton genome.

Gene	Accession	Amino acids	pI	Mol. wt.
		No.		kDa
GhVP	HM370494.1	760	5.38	80.44
GaVP1	Cotton_A_07299	770	5.32	80.49
GaVP2	Cotton_A_25209	760	5.38	80.48
GaVP3	Cotton_A_17055	750	5.32	79.32
GaVP4	Cotton_A_18486	760	5.49	80.03
GaVP5	Cotton_A_10266	790	5.93	84.51
GaVP6	Cotton_A_40017	770	6.13	81.76
GrVP7	Cotton_D_gene_10010970	760	5.30	80.41
GrVP8	Cotton_D_gene_10031919	760	5.26	80.42
GrVP9	Cotton_D_gene_10016525	760	5.24	80.11
GrVP10	Cotton_D_gene_10004549	770	5.11	80.63
GrVP11	Cotton_D_gene_10016108	770	5.24	80.52
GrVP12	Cotton_D_gene_10001961	760	5.40	80.23
GrVP13	Cotton_D_gene_10020039	760	5.49	79.97

Functional domain analysis (Fig. 1) showed that all of the sequences contained myristyl site, casein kinase site II, protein kinase phosphorylation site, and N-glycosylation site, but site locations and numbers differed. Therefore, it is speculated that during long-term evolutionary process, V-PPase genes maintained similarities of structure and function and might exhibit individual differences in gene transcription regulation. Further analysis with Scan Prosite tool found a cAMP and cGMP dependent protein phosphorylation site in *GaVP6* and *GrVP13*, but not in other genes. Thus, it is speculated that the two genes resulted from structural and functional changes of the V-PPase gene in long-term evolutionary process.

As shown in Figure 2, each gene contained three motifs (motif 1, motif 2, motif 3), and motif 2 was the most conserved. Except for *GaVP4*, *GaVP5* and *GaVP13* V-PPase genes might all contain two motif 1, while *GaVP5* and *GaVP6* likely contain an additional motif 3 at 71 to 120 amino acid residues. Therefore, it is speculated that *GaVP4*, *GaVP5* and *GaVP6* in the cotton genome A might have changed in conserved domains, which resulted in a directional evolution of the V-PPase gene function.

Distribution of V-PPase Gene Family Members in the Cotton Whole Genome

The information about gene location in the chromosome provided important evidence in studying the evolution and function of the gene family. As shown in Figure 3, ten V-PPase genes were mapped on eight chromosomes of the cotton genomes A and D. However, *GaVP6*, *GrVP7,* and *GrVP10* were mapped on scaffold320, scaffold121, and scaffold295, respectively, but not on any chromosome. Of the eight chromosomes, CA10 of genome A and chromosome 9 of genome D both had two V-PPase genes, while the other six chromosomes had one V-PPase gene each. These results were not the same according to JGI genome data. Using the JGI *G. raimondii* assembly (www.phytozome.net), *GrVP7* can be mapped to chromosome 6 and *GrVP10* can be placed on chromosome 4. The JGI genome data also places *GrVP13* on chromosome 4 instead of chromosome 2. The JGI genome data corroborates the placement of the other four *GrVP* genes. It is known that chromosome fragments might result in scattered distribution of gene family members on given chromosomes (Schauser et al., 2005).

Evolution Analysis of V-PPase Gene Family Members

The phylogenetic analysis indicated that the homology of V-PPase genes in a same species was closer than other members of their family (Fig. 4), e.g., *GaVP5* and *GaVP6*, *GrVP13* and *GaVP4*, *GrVP11* and *GrVP12, GaVP1* and *GaVP9*, *GaVP2* and *GrVP7*, *GaVP3* and *GrVP8*. The K^+ ion is important for plants in adversity stress, and *AVP2* in

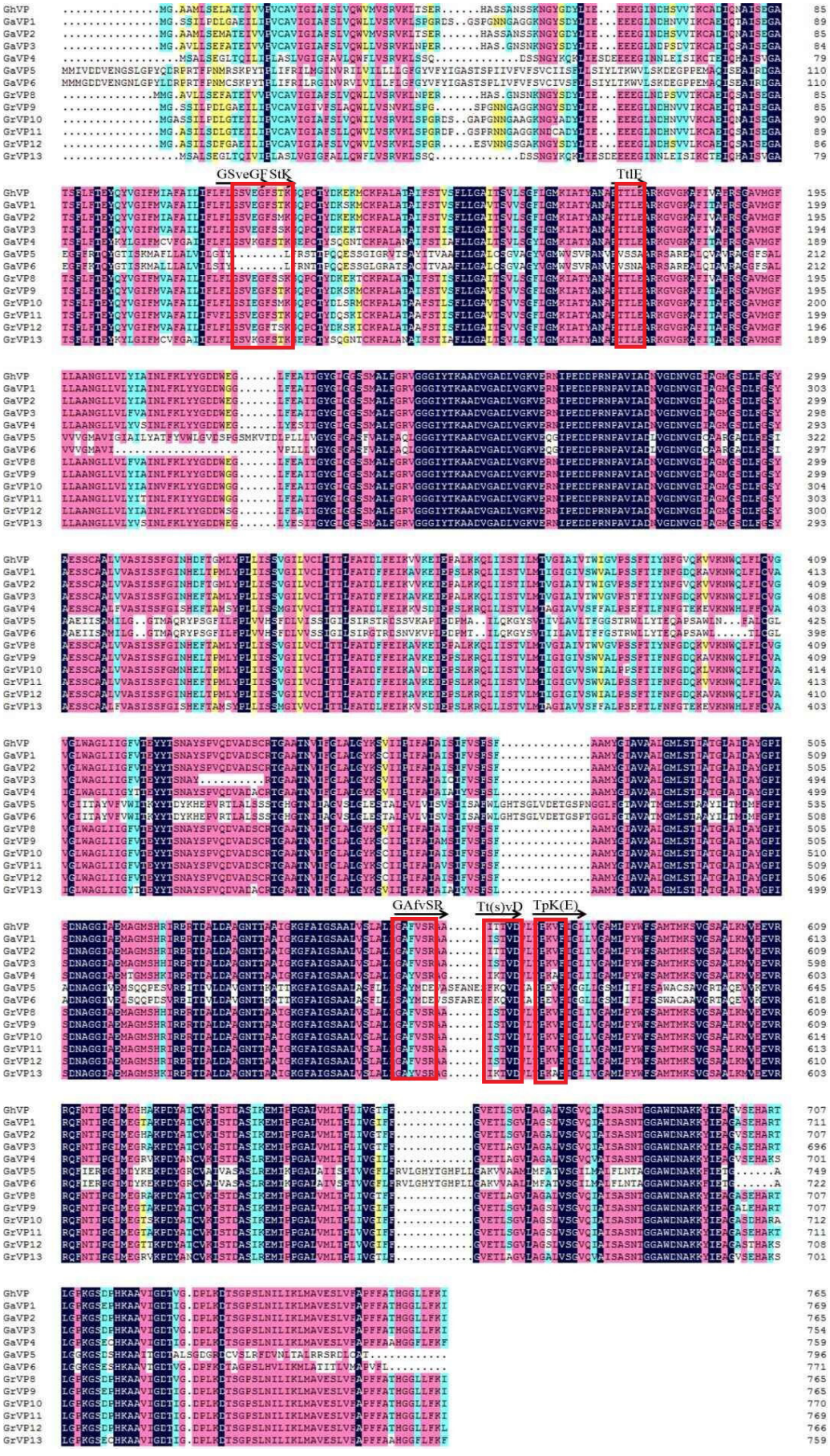

Fig. 1. Alignment of V-PPase gene family members and analysis of conserved functional domains in cotton. Many domains in the V-PPase family were identified. The five sequences marked with arrows represent some conserved domains in different gene members.

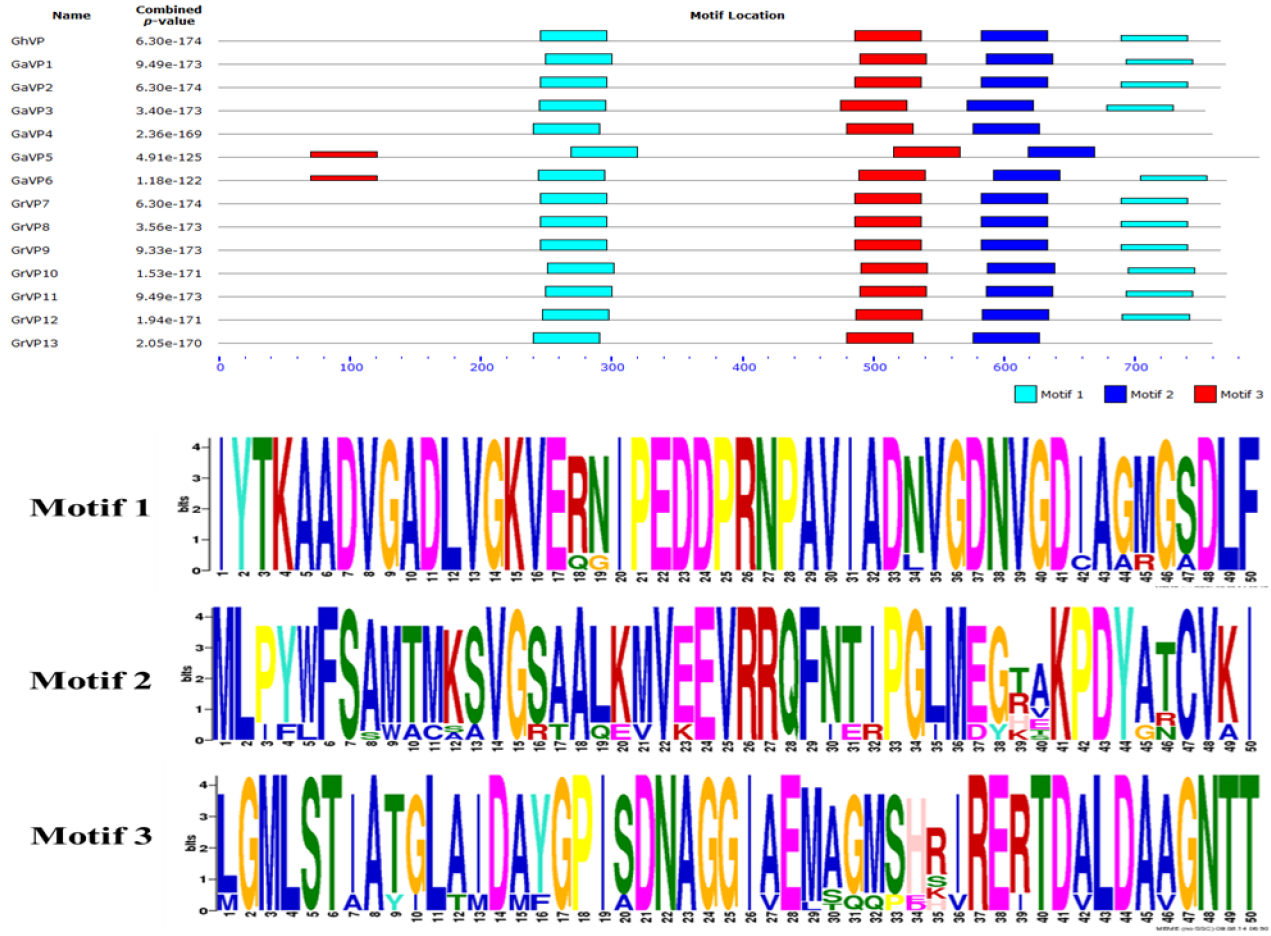

Fig. 2. Motif information of V-PPase gene family in cotton. (a) Combined block diagrams of three Motifs. (b) Sequence logos of three Motifs in expanded form.

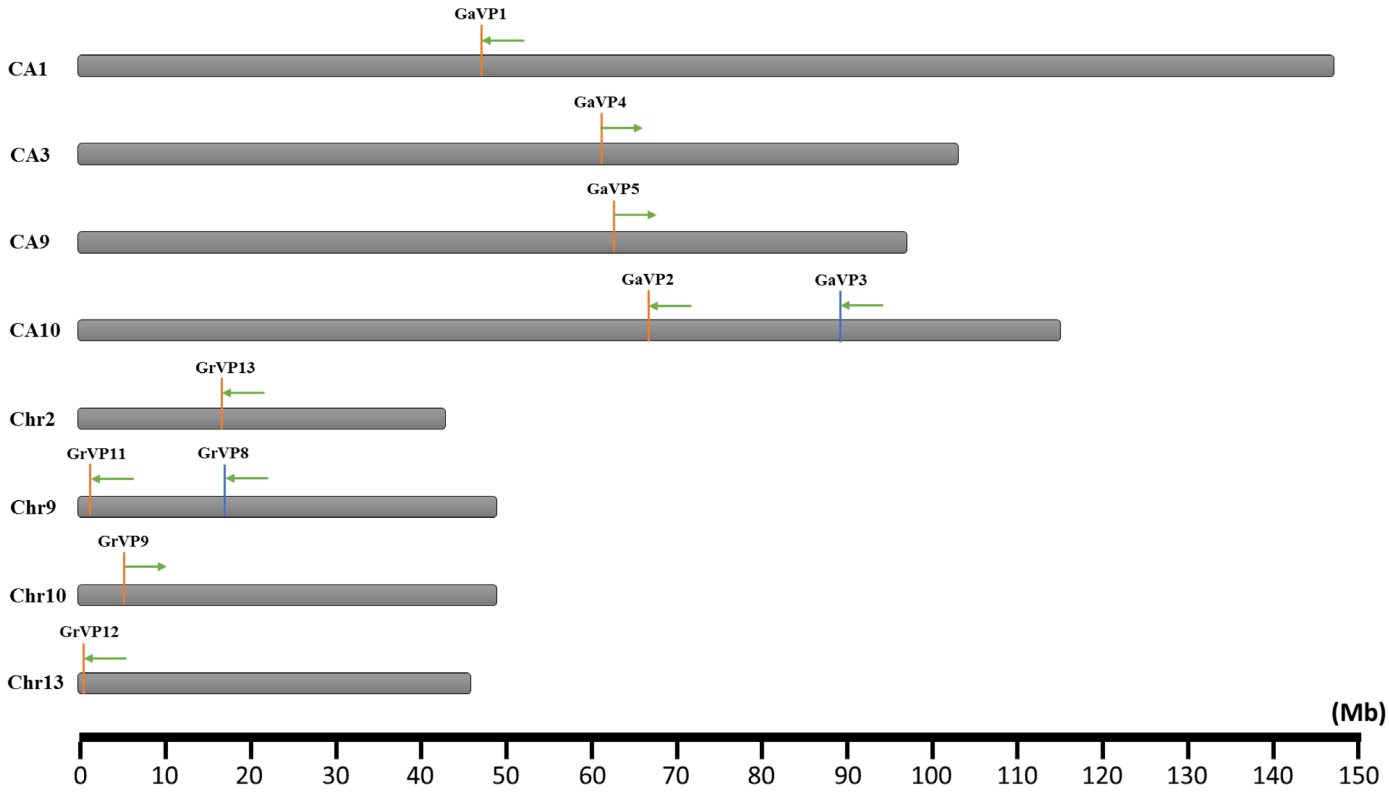

Fig. 3. Chromosome distribution of V-PPase genes in the cotton genome.

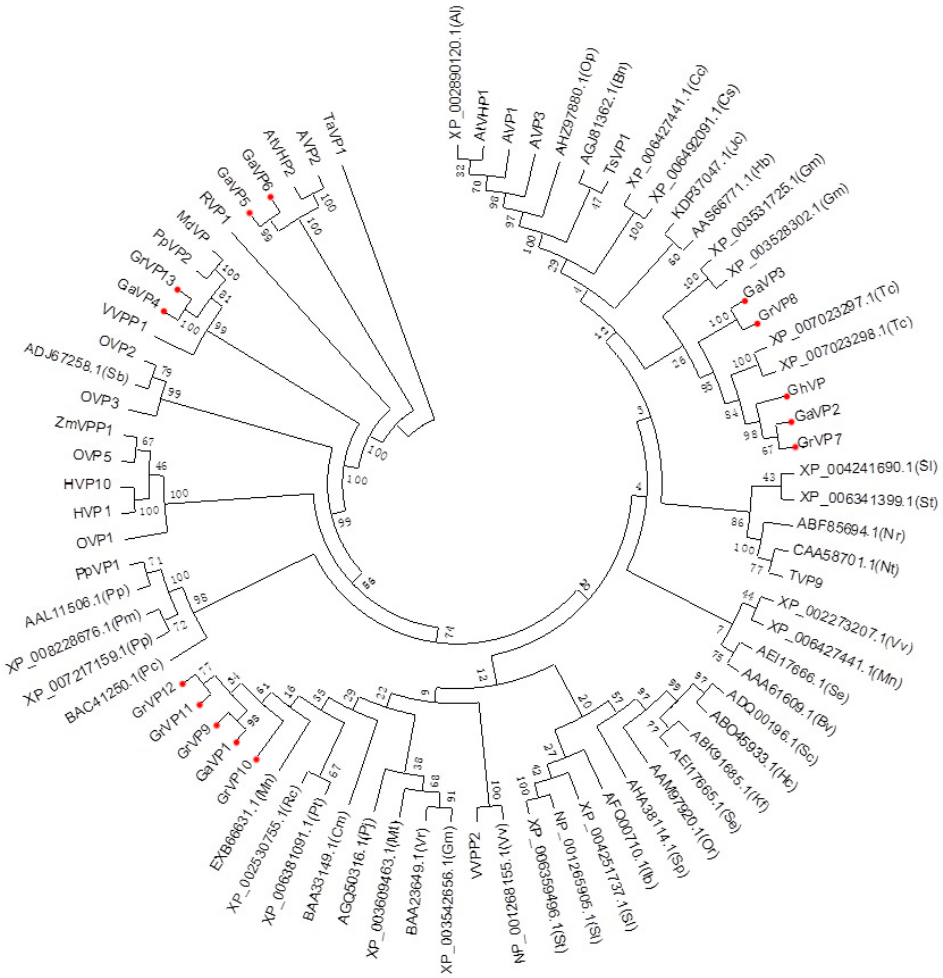

Fig. 4. Phylogenetic trees of the V-PPase genes in cotton and other species.

A. thaliana was shown to promote K^+ assimilation, therefore conferring resistance to adversity (Drozdowicz et al., 2000). *ZmVPP1* in corn might participate in segregating Na^+ and contributing to salt resistance. Thus, V-PPase members in cotton might play important roles in resistance to adversity. As all V-PPase genes sequences were aligned, the similarity of V-PPase gene sequences between cotton genomes A and D reached 89.12%, while a similarity level of only 76.15% was obtained between cotton and other species included in the phylogenetic tree. These findings suggested that during the evolutionary process, the V-PPase gene has evolved in different directions for various species, with some conserved structural domains remaining unchanged.

Genetic Structure of V-PPase Gene Family Members in Different Species

Gene structure is important in gene evolution studies. As shown in Figure 5, there were large differences among various members of the cotton V-PPase gene family. The H^+-PPase genes had 1 to 14 exons, with most genes (57.14%) having 8 exons, to suggest that exons and introns are relatively conserved in the evolutionary process of these genes. Three members contained 7 exons, *GhVP* and *GaVP6* each only contained 1 exon, and *GaVP5* had the largest number of exons (14 exons). V-PPase genes of genome A had wider range in exon number (1–14 exons), while those of genome D mainly displayed 7 and 8 exons. Thus, V-PPase members of cotton genome D were similar in structure in the genetic evolutionary process and relatively conserved, which suggested some consistency for the function of these genes. A previous study showed that small DNA segment inserts could change gene function and cause disappearance of the gene by natural selection (Long et al., 1995). *GaVP5* in the cotton genome A contained the most exons and introns and had large changes in length. This has probably resulted from large changes in structure or function of the gene during evolution.

Fluorescent Quantitative Assay

We selected two V-PPase genes, *GhVP* and *GrVP7*, as representatives to investigate the expression change in roots, stems, leaves, flowers, and fibers, and the two V-PPase genes demostrated similar expression level in all tissues. Moreover, the expression level of the V-PPase genes was found to be highest in leaves, which is threefold greater

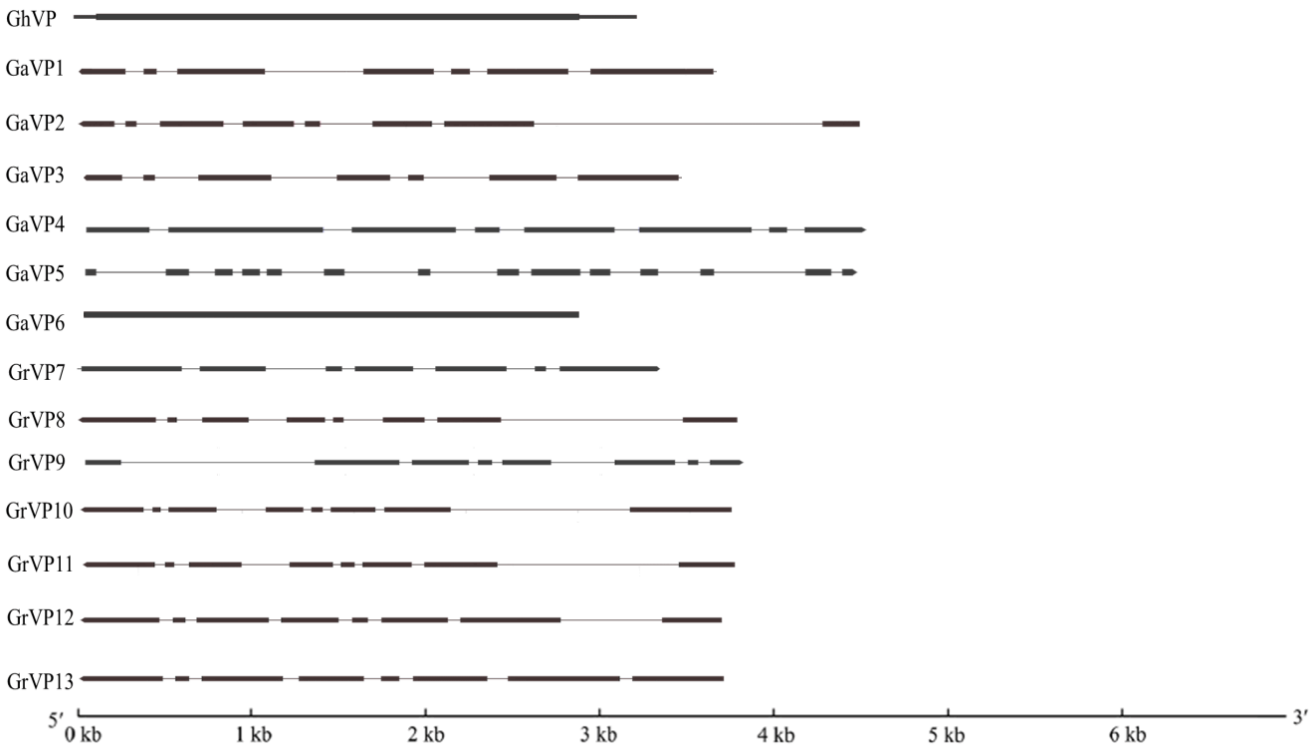

Fig. 5. Intron-exon structures of V-PPase gene family in cotton.

than in roots. Also we found that expression level of the V-PPase genes in stems was twofold greater than in roots. However, it was relatively lower in flowers and fibers. Our data indicate that V-PPase genes are expressed mainly in the leaves of cotton. Figure 6a shows fluorescent quantitative analysis of genes *GhVP* and *GrVP7* in tissues of upland cotton. The highest expression level was detected in leaves and the lowest expression level was in fibers. Other V-PPase genes had similar expression patterns.

As shown in Figure 6b, taking expression level of V-PPase genes in roots at 0 h after salt treatment as a control, *GaVP1* and *GaVP6* showed maximal expression level at 6 h after salt treatment. However, both genes displayed decreased level at 12 and 24 h after salt treatment, but still higher than in the control. In contrast, expression of *GrVP7*, *GrVP9* and *GrVP11* started to rise at 6 h after salt treatment and reached the maximum at 12 h. However, it was very low at 24 h, especially the expression level of *GrVP11* was lower than control. The data above indicate that expression of the V-PPase genes was upregulated after a short time of salt stress. However, it decreased as salt stress intensity increased. It can be concluded that expression of the V-PPase genes is associated with the salt tolerance mechanism. Figure 6c shows that expression level of the V-PPase genes in leaves was much higher than in stems, suggesting that expression of the V-PPase genes is tissue specific. The findings also indicated the change of the expression of these genes is different in two diploid species, but the overall change pattern was the same, showing the function of these genes in two species were consistent. This study was focused on expression changes of the V-PPase genes in stems and leaves when facing salt stress. As shown in Figure 6c, expression level of the V-PPase genes in leaves was downregulated at 24 h after salt treatment. However, it was upregulated in stems at 24 h. It indicated that the leaves express these stress response genes faster, but for a shorter duration than the stems. The expression pattern in different tissues was different.

DISCUSSION

By bioinformatics analysis, we identified 13 *GhVP*-like genes, with six and seven from genomes A and D, respectively, using *GhVP* as the reference. These genes constitute a relatively conserved family in plant evolution. The phylogenetic analysis indicated that some V-PPase genes are more closely related to their homolog in a closely related species than another member of their family within the same species. Gene structural analysis revealed that most V-PPase genes had 7 or 8 exons, with a minority presenting a variable number (1–14) of exons; this might result from a directional evolution of structure and function of V-PPase genes. By analyzing the conserved domain, all V-PPase genes showed several relatively conserved structural domains, e.g., myristyl site, casein kinase site II and protein kinase phosphorylation site. These domains might be important guarantee of gene function and the reason for changes in gene structure and function. However, metabolic pathways and specific functions of these domains require further analyses. The results indicated that V-PPase genes located in multiple chromosomes. Compared with other eukaryotes, the copying rate has been shown to

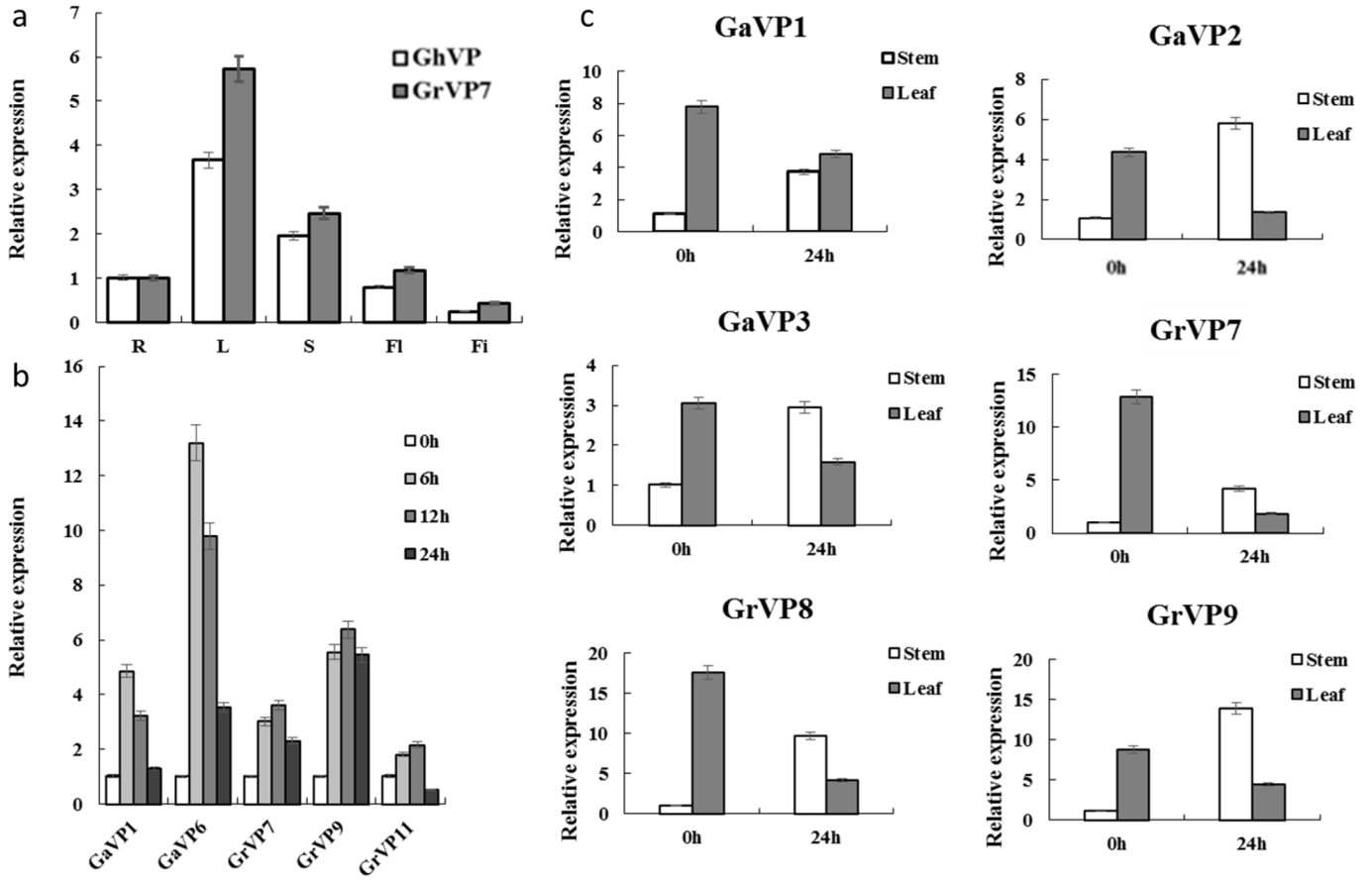

Fig. 6. Expression analysis of V-PPase genes in different tissues and different treatments. (a) Fluorescent quantitative analysis of genes GhVP and GrVP7 in tissues of cotton; Relative expression level of mRNA in tissue types: root (R), stem (S), leaf (L), flower (Fl), and fiber (Fi); Each sample was analyzed in triplicate. (b) Relative expression level of the V-PPase genes in roots after salt treatment for 0, 6, 12, and 24 h. (c) Expression level of the V-PPase genes after salt treatment in stems and leaves; mRNA relative expression level for 0 and 24 h of salt stress.

be higher for plants (Wei et al., 2014). Indeed, a recent study showed the whole genome of *G. raimondii* was copied at least twice (Wang et al., 2012). Chromosomes CA10 in genome A and Chr9 in genome D both had two V-PPase genes, which might result from gene duplication or partial gene duplication in the evolutionary history of the cotton genome. Previous studies have shown that gene duplication and later divergence phenomenon were the two major impetuses in evolution (Ohno et al., 1968; Chothia et al., 2003).

With advances in genome characterization, the study of extensive genome families using comparative genomics has become the hotspot in species analyses. To date, several gene families in several species have been identified and studied, including the LEA family in soybean (Shih et al., 2012), LBD and MAPK families in tomato (Popescu et al., 2009; Feng et al., 2012) and MAPKKK family in cotton (Yin et al., 2013). Cotton is a pioneer crop on saline-alkaline soils, and the V-PPase gene is important in stress resistance. With the successful completion of the cotton genomic sequencing, the study of the V-PPase gene from cotton whole genome became of significance. This study identified V-PPase gene family members from the cotton whole genomes A and D, for the first time, and researched V-PPase genes as for chromosomal location, evolution and structure, laying a foundation for future study of V-PPase genes.

Two PPase genes, i.e., *AVP1* (Yang et al., 2014) and *AVP2* (Drozdowicz et al., 2000), have been cloned from *A. thaliana*. The *AVP1* gene was shown to improve salt tolerance in yeast, while *AVP2* enhanced stress resistance by assimilating K^+. The *GhVP* gene was first cloned from cotton in our laboratory (Song et al., 2010) and the analysis of semiquantitative RT-PCR indicated the expression of *GhVP* in root system increased after salt treatment, which might be due to improved V-PPase activity under salt stress or related to increased V-PPase number (Fig. 7). Gene sequence alignment showed that similarities between *GhVP* and other V-PPase gene members in the cotton genome were as high as 89.12%, and their structures were also similar. Thus, other V-PPase gene members in cotton might also be related to stress resistance. The high V-PPase gene expression in corn was shown to improve the growth and development of the corn root system as well as assimilation and accumulation of phosphorus, to achieve stress

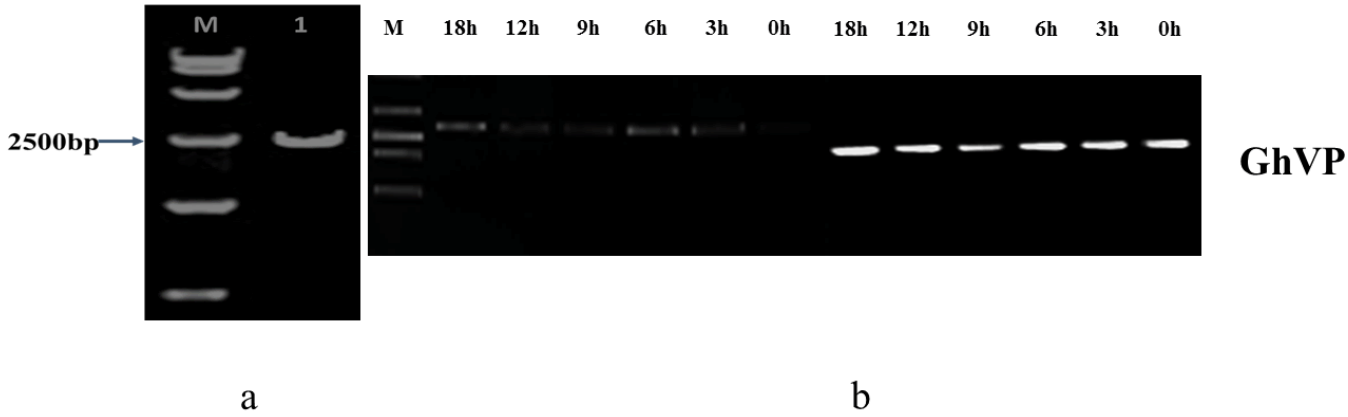

Fig. 7. PCR amplification and semiquantitative RT-PCR analysis of V-PPase gene. (a) Product of PCR amplification of V-PPase gene. (b) Expression analysis of V-PPase by semiquantitative RT-PCR.

resistance (Pei et al., 2012). Salt and drought resistance in *A. thaliana* was improved by transferring the *TVP1* gene from wheat into *A. thaliana* (Brini et al., 2007). These findings all demonstrated that H^+-PPase is an important gene that could directly or indirectly enhance the salt tolerance of plants. Therefore, the H^+-PPase gene members in cotton are likely to be related to stress resistance.

The fluorescent quantitative analysis demonstrates that expression level of the V-PPase genes was significantly different from that after salt treatment. It is in line with the previous study by Wang et al. (2001). Our results exhibit that the expression level of the V-PPase genes was highest in leaves and lowest in fibers, indicating that it is tissue specific. The stems are expressing certain V-PPase genes at significantly higher levels than leaves 24 h after salt stress, which may be related with the different function of different tissues. It was reported that expression level of the V-PPase genes increased along with drought, strong sunlight and heavy metal stress (Kumari and Sharma, 2010; Kabala et al., 2014).

Studies have found that H^+-PPase is distributed in almost all vacuolated plant cells. Therefore, PPase is an essential part of higher plant cells. The enzyme decomposes pyrophosphate and plays an important role in the expansion process of plant cells and vacuoles. Other studies have shown that PPase expression changes under different physiological and adverse conditions. However, the molecular evolution, gene expression regulation, and post-transcription of H^+-PPase remain to be elucidated. To preferably assess gene structure and function, this study identified H^+-PPase genes in cotton whole genome and analyzed their genetic structures and evolution, laying the foundation for further study of the gene function of H^+-PPase in cotton.

Supplemental Materials Available

A supplemental table of the primer sequences for qRT-PCR is available with the online verison of this article.

Acknowledgments

All materials were provided by the Institute of Cotton Research of Chinese Academy of Agricultural Sciences. This work was supported by the National High-tech R&D Program (Grant no. 2013AA102601-01-13).

References

Brini, F., R.A. Gaxiola, G.A. Berkowitz, and K. Masmoudi. 2005. Cloning and characterization of a wheat vacuolar cation/proton antiporter and pyrophosphatase proton pump. Plant Physiol. Biochem. 43:347–354. doi:10.1016/j.plaphy.2005.02.010

Brini, F., M. Hanin, I. Mezghani, G.A. Berkowitz, and K. Masmoudi. 2007. Overexpression of wheat Na+/H+ antiporter TNHX1 and H^+-pyrophosphatase TVP1 improve salt- and drought-stress tolerance in *Arabidopsis thaliana* plants. J. Exp. Bot. 58:301–308. doi:10.1093/jxb/erl251

Carra, A., G. Gambino, and A. Schubert. 2007. A cetyltrimethylammonium bromide-based method to extract low-molecular-weight RNA from polysaccharide-rich plant tissues. Anal. Biochem. 360:318–320. doi:10.1016/j.ab.2006.09.022

Chothia, C., J. Gough, C. Vogel, and S.A. Teichmann. 2003. Evolution of the protein repertoire. Science 300:1701–1703. doi:10.1126/science.1085371

Cooperman, B.S., A.A. Baykov, and R. Lahti. 1992. Evolutionary conservation of the active site of soluble inorganic pyrophosphatase. Trends Biochem. Sci. 17:262–266.

Drozdowicz, Y.M., J.C. Kissinger, and P.A. Rea. 2000. AVP2, a sequence-divergent, K(+)-insensitive H(+)-translocating inorganic pyrophosphatase from Arabidopsis. Plant Physiol. 123:353–362. doi:10.1104/pp.123.1.353

Feng, Z., J. Zhu, X. Du, and X. Cui. 2012. Effects of three auxin-inducible LBD members on lateral root formation in *Arabidopsis thaliana*. Planta 236:1227–1237. doi:10.1007/s00425-012-1673-3

Fonseca-de-Souza, A.L., A.L. Freitas-Mesquita, L.P. Vieira, D. Majerowicz, N. Daflon-Yunes, L.C. Soares-de-Medeiros, et al. 2014. Identification and characterization of an ecto-pyrophosphatase activity in intact epimastigotes of *Trypanosoma rangeli*. PLoS ONE 9:E106852. doi:10.1371/journal.pone.0106852

Fukuda, A., K. Chiba, M. Maeda, A. Nakamura, M. Maeshima, and Y. Tanaka. 2004. Effect of salt and osmotic stresses on the expression of genes for the vacuolar H+-pyrophosphatase, H^+-ATPase subunit A, and Na+/H+ antiporter from barley. J. Exp. Bot. 55:585–594. doi:10.1093/jxb/erh070

Gaxiola, R.A., J. Li, S. Undurraga, L.M. Dang, G.J. Allen, S.L. Alper, et al. 2001. Drought- and salt-tolerant plants result from overexpression of the AVP1 H+-pump. Proc. Natl. Acad. Sci. USA 98:11444–11449. doi:10.1073/pnas.191389398

Gaxiola, R.A., R. Rao, A. Sherman, P. Grisafi, S.L. Alper, and G.R. Fink. 1999. The *Arabidopsis thaliana* proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast. Proc. Natl. Acad. Sci. USA 96:1480–1485. doi:10.1073/pnas.96.4.1480

Kabala, K., M. Janicka-Russak, M. Reda, and M. Migocka. 2014. Transcriptional regulation of the V-ATPase subunit c and V-PPase isoforms in *Cucumis sativus* under heavy metal stress. Physiol. Plant. 150:32–45. doi:10.1111/ppl.12064

Kim, Y., E.J. Kim, and P.A. Rea. 1994. Isolation and characterization of cDNAs encoding the vacuolar H(+)-pyrophosphatase of *Beta vulgaris*. Plant Physiol. 106:375–382. doi:10.1104/pp.106.1.375

Kumari, N., and V. Sharma. 2010. Stress-mediated alteration in V-ATPase and V-PPase of *Butea monosperma*. Protoplasma 245:125–132. doi:10.1007/s00709-010-0153-5

Li, F., G. Fan, K. Wang, F. Sun, Y. Yuan, G. Song, et al. 2014. Genome sequence of the cultivated cotton Gossypium arboreum. Nat. Genet. 46:567–572. doi:10.1038/ng.2987

Long, M., C. Rosenberg, and W. Gilbert. 1995. Intron phase correlations and the evolution of the intron/exon structure of genes. Proc. Natl. Acad. Sci. USA 92:12495–12499. doi:10.1073/pnas.92.26.12495

Lv, S., K. Zhang, Q. Gao, L. Lian, Y. Song, and J. Zhang. 2008. Overexpression of an H+-PPase gene from *Thellungiella halophila* in cotton enhances salt tolerance and improves growth and photosynthetic performance. Plant Cell Physiol. 49:1150–1164. doi:10.1093/pcp/pcn090

Maeshima, M. 2000. Vacuolar H(+)-pyrophosphatase. Biochim. Biophys. Acta 1465:37–51. doi:10.1016/S0005-2736(00)00130-9

Marty, F. 1999. Plant vacuoles. Plant Cell 11:587–600. doi:10.1105/tpc.11.4.587

Mitsuda, N., M. Seki, K. Shinozaki, and M. Ohme-Takagi. 2005. The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. Plant Cell 17:2993–3006. doi:10.1105/tpc.105.036004

Ohno, S., U. Wolf, and N.B. Atkin. 1968. Evolution from fish to mammals by gene duplication. Hereditas 59:169–187. doi:10.1111/j.1601-5223.1968.tb02169.x

Olcott, H.S. 1939. Sand culture of cotton plants. Science 89:608–609. doi:10.1126/science.89.2322.608

Ozolina, N.V., E.V. Kolesnikova, I.S. Nesterkina, V.N. Nurminsky, J.V. Boyarkin, L.A. Sitneva, et al. 2009. Interaction of signal systems (nitric oxide and calcium) in regulation of hydrolytic activity of tonoplast H(+)-pyrophosphatase under normal conditions and stress. Dokl. Biochem. Biophys. 428:242–244. doi:10.1134/S1607672909050056

Paterson, A.H., J.F. Wendel, H. Gundlach, et al. 2012. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibers. Nature 492(7429):423–427. doi:10.1038/nature11798

Pei, L., J. Wang, K. Li, Y. Li, B. Li, F. Gao, et al. 2012. Overexpression of Thellungiella halophila H+-pyrophosphatase gene improves low phosphate tolerance in maize. PLoS ONE 7:E43501. doi:10.1371/journal.pone.0043501

Popescu, S.C., G.V. Popescu, S. Bachan, Z. Zhang, M. Gerstein, M. Snyder, et al. 2009. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. Genes Dev. 23:80–92. doi:10.1101/gad.1740009

Schauser, L., W. Wieloch, and J. Stougaard. 2005. Evolution of NIN-like proteins in Arabidopsis, rice, and *Lotus japonicus*. J. Mol. Evol. 60:229–237. doi:10.1007/s00239-004-0144-2

Shih, M.D., T.Y. Hsieh, W.T. Jian, M.T. Wu, S.J. Yang, F.A. Hoekstra, et al. 2012. Functional studies of soybean (*Glycine max* L.) seed LEA proteins GmPM6, GmPM11, and GmPM30 by CD and FTIR spectroscopy. Plant Sci. 196:152–159. doi:10.1016/j.plantsci.2012.07.012

Shin, Y.S., and T. Podskarbi. 2004. Molecular and biochemical basis for variants and deficiency of GALT: Report of 4 novel mutations. Bratisl. Lek. Listy (Bratislava Medical Journal) 105:315–317.

Song, L.Y., W.W. Ye, Y.L. Zhao, J.J. Wang, B.X. Fan, and D.L. Wang. 2010. Isolation and analysis of aalt tolerance related gene (*GhVP*) from *Gossypium hirsutum* L. Cotton Science 22(3):285–288.

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar. 2013. MEGA6: Molecular evolutionary genetics analysis, version 6.0. Mol. Biol. Evol. 30:2725–2729. doi:10.1093/molbev/mst197

Tanaka, Y., K. Chiba, M. Maeda, and M. Maeshima. 1993. Molecular cloning of cDNA for vacuolar membrane proton-translocating inorganic pyrophosphatase in Hordeum vulgare. Biochem. Biophys. Res. Commun. 190:1110–1114. doi:10.1006/bbrc.1993.1164

Wang, B., U. Luttge, and R. Ratajczak. 2001. Effects of salt treatment and osmotic stress on V-ATPase and V-PPase in leaves of the halophyte *Suaeda salsa*. J. Exp. Bot. 52:2355–2365. doi:10.1093/jexbot/52.365.2355

Wang, K., Z. Wang, F. Li, W. Ye, J. Wang, G. Song, et al. 2012. The draft genome of a diploid cotton *Gossypium raimondii*. Nat. Genet. 44:1098–1103. doi:10.1038/ng.2371

Wei, K., Y. Wang, and D. Xie. 2014. Identification and expression profile analysis of the protein kinase gene superfamily in maize development. Mol. Breed. 33:155–172. doi:10.1007/s11032-013-9941-x

Yang, H., X. Zhang, R.A. Gaxiola, G. Xu, W.A. Peer, and A.S. Murphy. 2014. Over-expression of the Arabidopsis proton-pyrophosphatase AVP1 enhances transplant survival, root mass, and fruit development under limiting phosphorus conditions. J. Exp. Bot. 65:3045–3053. doi:10.1093/jxb/eru149

Yin, Z., J. Wang, D. Wang, W. Fan, S. Wang, and W. Ye. 2013. The MAPKKK gene family in *Gossypium raimondii*: Genome-wide identification, classification and expression analysis. Int. J. Mol. Sci. 14:18740–18757. doi:10.3390/ijms140918740