Article

Meta-Analysis of Aquaporin Gene Family in *Triticum turgidum* and Its Expression Profile in Response to Salt Stress

Mahnaz Yaghobi [1], Parviz Heidari [1,*]

[1] Faculty of Agriculture, Shahrood University of Technology, Shahrood 3619995161, Iran; yaghobi.univer@gmail.com (M.Y.)

* Correspondence: heidarip@shahroodut.ac.ir

Abstract: During the response of plants to adverse stresses, aquaporin (AQP) plays a prominent role in membrane water transport based on received upstream signals. In addition, they have various physical parts for dealing with environmental stresses. Due to the importance of the AQP gene family, studies have been conducted investigating the function and regulatory system of these genes. However, many of their molecular aspects are still unknown. This study aims to carry out a genomic-wide investigation of the AQP gene family in durum wheat using bioinformatics tools and to investigate the expression patterns of some members in response to salt stress. Our results showed that there are 80 *TtAQP* genes in durum wheat, which are classified into four main groups based on phylogenetic analysis. Many duplications were observed between the members of the TtAQP gene family, and high diversity in response to post-translational modifications was observed between TtAQP family members. The expression pattern of *TtAQP* genes disclosed that these genes are primarily upregulated in response to salt stress. Besides, qPCR data revealed that *TtAQPs* are more induced in delayed responses to salinity stress. Overall, our findings illustrate that TtAQP members are diverse in terms of their structure, regulatory systems and expression levels.

Keywords: Water channel; Abiotic stresses; Wheat genes; Post-translation modification; Gene expression

1. Introduction

Aquaporin proteins (AQPs) include a group of the major intrinsic protein (MIP) that facilitate the transport of water and other small neutral molecules through the cell membrane by forming channels in the membrane and play an important role in plant growth and response to abiotic stress [1,2]. AQP channels facilitate bidirectional flow across a concentration gradient in all cells [3]. In addition, some AQP isoforms play the role of peroxyporins and contribute to cellular redox signaling by transporting hydrogen peroxide [4]. Moreover, AQPs in plants have multiple functions, such as hydraulic control of plant tissue as well as seed germination and the emergence of lateral roots [2]. Based on the structural analysis, AQP family genes can be divided into five main evolutionary subfamilies in plants, including plasma membrane intrinsic proteins (PIP), tonoplast internal proteins (TIP), Nodulin-26 intrinsic proteins such as NIPs are small major intrinsic proteins (SIPs) and X intrinsic proteins (XIPs) [5]. Two fundamental aspects of plant AQPs include: first, their extraordinary diversity in plants [6], and second, some AQPs are multifunctional channel proteins that allow some small neutral solutes across the cell membrane, such as glycerol, CO 2, ammonia (NH 3), urea, boron and hydrogen peroxide [7].

In recent years, the function of AQPs has been investigated relating to plant resistance in response to biotic and abiotic stresses [8–11]. It has been shown that AQP can improve plant tolerance to abiotic stresses (such as drought, osmosis, cold, salt, and high temperature) [12]. Furthermore, previous studies show that AQP plays a positive role in response to biotic stresses. It has also been found that the expression of *AQP* genes differs depending on the organs and tissues, as well as hormones and abiotic stress treatment

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

such as drought, heat, and cold [13,14]. Between AQP subfamilies, the function of PIPs and TIPs, which have high transporter activity, is investigated more under adverse conditions [15]. The downregulation of PIPs and TIPs was reported in response to drought and salt stress [1,16–18]. Besides, it was stated that the NIP subfamily is the only AQP able to transport Si (silicon) that can improve plant tolerance in response to various environmental stresses [5,19].

Today, cereals are considered the most important crop, and wheat is considered a strategic product among cereals in the world. Wheat varieties exist in three ploidy levels, which include diploid, tetraploid, and hexaploid [20]. Durum wheat, *Triticum turgidum*, is tetraploid wheat (2n=4x=28). Durum wheat is best adapted to areas with a relatively dry climate with warm days and cool nights during the growing season (corresponding to a Mediterranean and temperate climate) [21]. As mentioned above, AQPs are critical in transporting water and maintaining cell balance in stresses such as drought and salinity. However, the sequence structure and function of AQP gene family members in response to salinity are largely unknown in durum wheat. In the current study, we aimed to identify and characterize the AQPs of durum wheat (TtAQPs), and the expression levels of several *TtAQPs* were evaluated in different salinity concentrations. As the first report, new perspectives on the structure and process of evolution and function of members of this gene family were presented, which can be considered in future studies related to durum wheat breeding.

2. Materials and Methods

2.1. Identification and sequence analysis of TtAQPs

To identify the *AQP* genes in *Triticum turgidum* (*TtAQPs*), the proteins sequence of putative AQPs in *Arabidopsis* (AtAQPs) and rice (OsAQPs) were applied as the query sequences against the complete protein sequence of *Triticum turgidum* using the BLASTP tool of the EnsemblePlants database [22]. In the next step, the extracted sequences are confirmed by the Conserved Domains Database (CDD) [23], and Pfam [24]. Then, the short sequences (less than 150 amino acids in length) were removed from the list. The identified sequences were analyzed using the Expasy site and their physicochemical properties, such as molecular weight (MW), isoelectric point (pI), instability index ,and GRAVY index were estimated using the ProtParam tool [25].

2.2. Investigating the evolutionary process of TtAQPs

To study the evolutionary relationships of TtAQP proteins in durum wheat, a phylogenetic tree was constructed with complete protein sequences of TtAQP gene family members in durum wheat along with their orthologues from Arabidopsis, rice, barley, corn and sorghum plants. *AQP* genes were identified in Arabidopsis (*Arabidopsis thaliana*), barley (*Hordeum vulgare*), corn (*Zea mays*) and rice (*Oryza sativa* japonica) using the EnsemblPlants database as well as durum wheat. In the first step, the sequence of AQP proteins was aligned using the ClustalW program using an online tool, Clustal Omega [26]. Then, the phylogenetic tree of AQPs was constructed based on the maximum likelihood (ML) method using the IQ tree website [27]. The iTOL tool [28] was used to illustrated the phylogenetic tree. Besides, ten conserved motifs into TtAQPs were identified by the MEME motif finder tool [29] based on the default setting.

2.3. Determining the duplication genes and estimating Ka and Ks values

The duplication events between *TtAQP* genes were identified based on similarity, more than 80% between pairs of *TtAQP* genes [30]. In addition, synonymous (Ks) and non-synonymous (Ka) values at each site among pairs of duplicated genes were calculated using TBtools software [31]. Then the Ka/Ks ratio was estimated to identify the effect of evolution pressure on the function of duplicated genes. The division time of pairs of duplicated *TtAQP* genes was estimated using the synonymous mutation rate of substitution λ per synonymous site per year, as $T = (Ks/2\lambda\ (\lambda = 6.5 \times 10^{-9})) \times 10^{-6}$ [32].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

2.4. Prediction of phosphorylation site into TtAQP proteins

Phosphorylation is one of the important types of post-translational modifications. The site of phosphorylation is serine, tyrosine and threonine residues on proteins. The potential phosphorylation regions of TtAQP proteins were predicted through the NetPhos 3-1 site, with a potential value higher than 0.80 [33].

2.5. Prediction of 3D structures and pocket sites of AQP proteins

The three-dimensional structure of TtAQP proteins was predicted using the Phyre2 database [34]. The validity of the predicted protein model was evaluated through Ramachandran plot analysis [35]. Analysis of ligand binding regions (pocket sites) in predicted protein models was also performed through the Phyre investigator tool of the Phyre2 server.

2.6. TtAQP gene promoter analysis

To identify the regulatory regions in the promoter region, a region 1500 nucleotides upstream of *TtAQP* genes was investigated as the promoter region. The sequence of the promoter region has been analyzed to recognize the putative cis-regulatory elements using the PlantCARE database [36]. Finally, the identified cis-regulatory elements were grouped based on their functions.

2.7. Plant materials and treatments

Seeds of *Triticum turgidum* cultivar Yavaros were sterilized with 3% sodium hypochlorite for 10 minutes. Then, they were cultivated in pots containing sterilized perlite and peat moss (2:1). Five seeds were planted in each pot, and after germination, three of the best seedlings were left in each pot. The cultivation conditions were kept with a photoperiod of 16 hours of light and 8 hours of darkness and a temperature of 25 ºC. 40-day-old seedlings were subjected to salt treatment by irrigation. In the current study, two salt concentrations, 150 and 250 mM of NaCl, were applied twice at 24-hour intervals. Some pots were also considered as controls (without using salt stress). After applying salt stress, the leaves of seedlings were sampled at different times of 6, 24 and 72 hours. The collected samples were immediately placed in liquid nitrogen, and then were transferred to a -80 freezer for other analyses.

2.8. Primer design to study gene expression

To investigate gene expression, six *TtAQP* genes, including *TtAQP18, TtAQP29, TtAQP34, TtAQP79 TtAQP58*, and *TtAQP42*, were selected based on the phylogenetic result. The primers of selected *TtAQP* genes were designed based on the coding region (CDS) by Primer3 online software [37]. In this study, *Actin7* was used as a reference gene (Table S1).

2.9. RNA extraction and real-time PCR

RNX plus kit (Sinaclon, Iran) was used for RNA extraction based on manufacturer instructions. The quality and quantity of extracted RNA were checked using a Nano Photometer (Implen N50). Complementary DNA (cDNA) was synthesized by reverse transcriptase (Roche, Germany) based on the manufacturer's protocols. The expression levels of *TtAQP* genes were investigated using Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher, France) based on the manufacturer's protocol and the ABI Step One. The relative expression of each gene was calculated by the delta delta ct method [38]. The expression difference between the treatments and the control sample was calculated based on the t-test. All experiments were performed in three biological replicates.

2.10. Analysis of RNA-seq data

To compare the expression pattern between tetraploid and hexaploid wheat, the available expression data for common wheat (*Triticum aestivum*) in response to abiotic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

stresses (heat, drought, and salt stress) were extracted from Wheat eFP Browser (https://bar.utoronto.ca). Finally, the expression of *AQPs* was illustrated based on log2 fold-change.

3. Results

3.1. Identification of AQPs

In the current study, 80 putative *TtAQP* genes were identified in the genome of *Triticum turgidum*. Detailed information on the 80 *TtAQP*s is provided in Table S2. In addition, TtAQPs were compared with their orthologues in *Arabidopsis thaliana*, *Hordeum vulgar*, *Oryza sativa*, and *Zea mays* based on physio-chemical properties (Table 1). It was observed that AQPs in the studied plants varied based on physio-chemical properties. In *Triticum turgidum*, TtAQPs varied between 157 (TtAQP61) to 392 (TtAQP27) amino acids, and their pI ranged between 4.65 (TtAQP42) and 11.05 (TtAQP09). The exon number of the deduced proteins varied from 1 (most members of the PIP2 subfamily) to 6 and their MWs ranged from 16.7 (TtAQP42) to 41.5 (TtAQP27) kDa. In addition, the Gravy value of TtAQPs varied between -0.46 and 0.94, and 96% of the deduced proteins were predicted as sable proteins (Table 1 and Table S2).

Table 1. Summary of AQPs properties in *Triticum turgidum*, *Arabidopsis thaliana*, *Hordeum vulgar*, *Oryza sativa*, and *Zea mays*. Full details of physio-chemical properties of TtAQPs are shown in Table S2.

Plant species	Peptide (aa)	Exon #	pI	MW (kDa)	GRAVY	Stability
Triticum turgidum	157-392	1-6	4.65-11.05	16.7-41.5	-0.46, 0.94	%96
Arabidopsis thaliana	220-328	1-5	4.50-10.60	23.9-35.4	0.28, 1.02	%100
Hordeum vulgar	150-623	1-19	5.48-12.80	16.3-69.2	-0.88, 0.90	%71
Oryza sativa	165-314	1-5	5.91-12.35	17.5-33.5	0.14, 0.92	%88
Zea mays	151-564	1-7	5.42-12.11	18.8-62.8	-0.25, 0.95	%86

3.2. Evolutionary analysis

To investigate the evolutionary relationship of TtAQPs, 237 AQP proteins, including 80 *T. turgidum* AQPs, 39 *H. vulgar* AQPs (HvAQPs), 35 *Arabidopsis* AQPs (AtAQPs), 39 rice AQPs (OsAQPs) and 44 maize AQPs (ZmAQPs), were used to design phylogenetic tree. According to the phylogenetic tree, all AQPs were classified into four main groups (Figure 1). According to the phylogenetic tree, all PIP1 proteins were located in group 1, PIP2 proteins were placed in group 2, and TIPs were in group 3. In addition, SIPs were located in group 4-a, and NIPs were placed in group 4-b. Ten TtAQPs were present in group 1 with five TtAQPs, four HvAQPs, three OsAQPs, and three ZmAQPs. Besides, 25 TtAQPs with 8 AtAQPs, 10 HvAQPs, 8 OsAQPs, and 9 ZmAQPs were located in group 2. In addition, 19 TtAQPs with 10 AtAQPs, 7 HvAQPs, 10 OsAQPs, and 14 ZmAQPs were present in group 4. Finally, 26 TtAQPs with 12 AtAQPs, 19 HvAQPs, 14 OsAQPs, and 18 ZmAQPs were located in group 4. Based on phylogenetic analysis, TtAQPs showed a close relationships with their orthologous in barley. Moreover, AQPs from group 1 showed more genetic distance from other members. Ten conserved domains were identified in TtAQPs (Figure 2). Motifs 3, 4, 5, and 8 were mainly present in proteins from groups 1 and 2, including PIP1 and PIP2 proteins. Besides, motifs 10, 7, and 9 were observed in proteins groups 3 and 4. These conserved motifs can be used to identify the subfamily of TtAQPs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

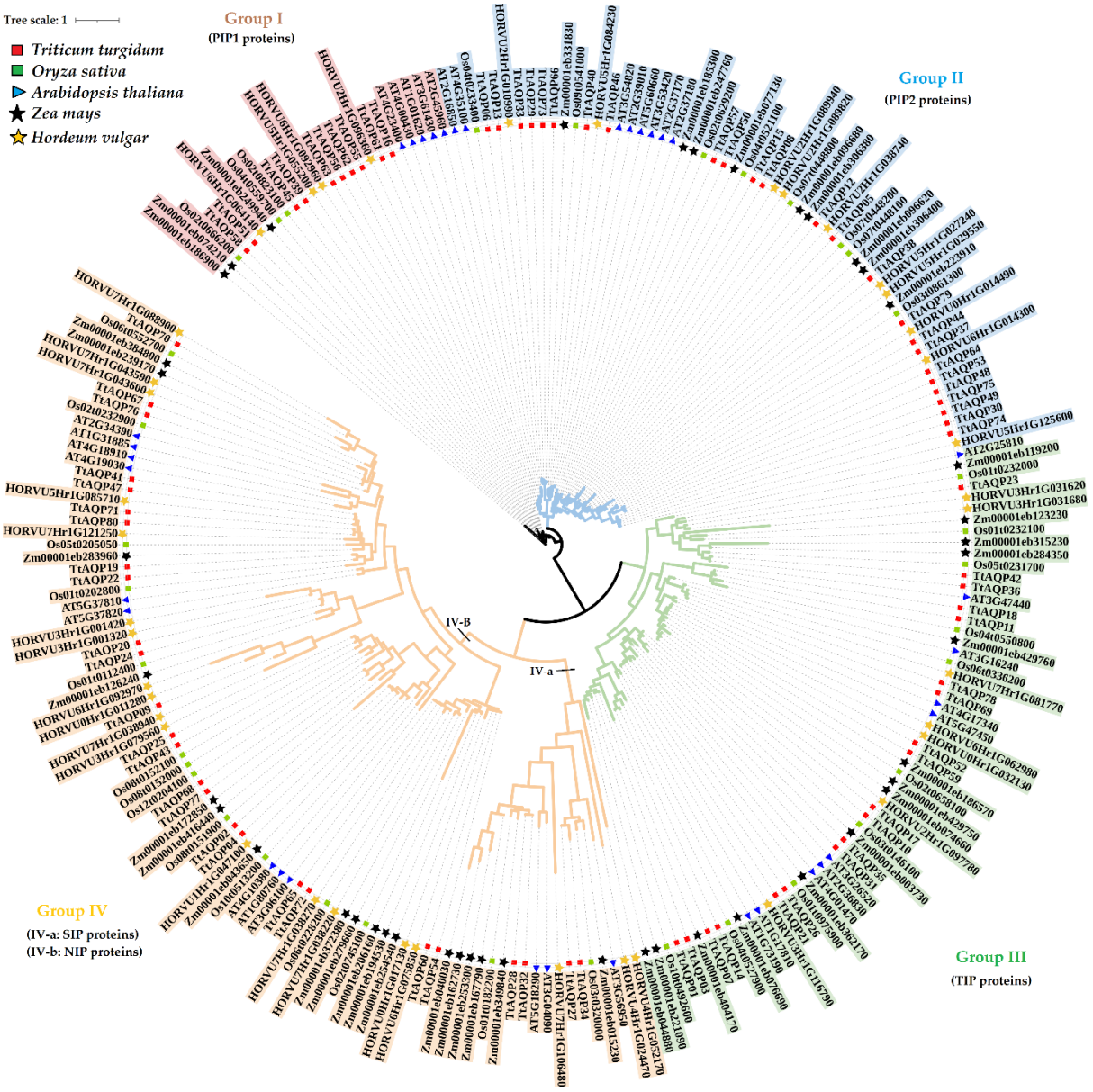

Figure 1. Phylogenetic analysis of AQP gene family in *Triticum turgidum* (TtAQPs), *Oryza sativa* (started with Os), *Zea mays* (started with Zm), *Hordeum vulgar* (started with HOR), *Arabidopsis thaliana* (started with AT).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

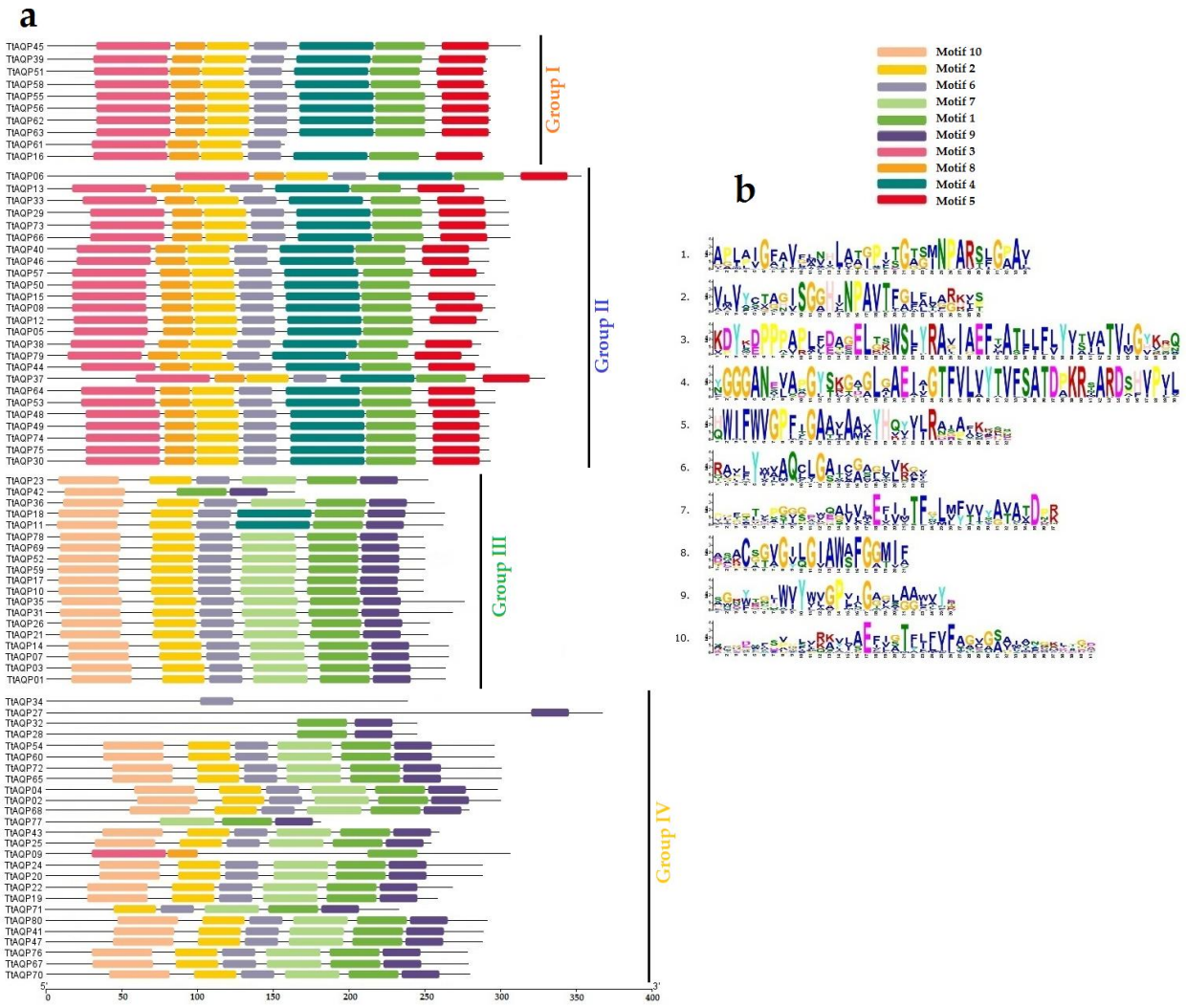

Figure 2. Conserved motifs into TtAQP proteins. Distribution of 10 conserved motifs in TtAQPs in four groups based on phylogenetic analysis (a), and conserved logos of identified motifs (b).

3.3. Duplication events in TtAQPs

The genomic distribution of each *TtAQP* was investigated in the durum wheat genome (Figure 3). Eighty *TtAQP* genes were located on 14 chromosomes and one chromosome was unknown (UN). In terms of the distribution of genes on chromosomes, Chr6B and Chr7B had the highest distribution and frequency of genes on them, with eight *TtAQP* genes on each. After these two chromosomes, Chr2A, Chr2B, Chr5A, Chr6A and Chr7A were more abundant, with seven *TtAQP* genes on each. In addition, Chr1A and Chr1B, each having two genes, included the least number of genes on the chromosomes. Finally, a *TtAQP* was located in unknown chromosomal positions. These results determined that *TtAQPs* are not uniformly distributed on durum wheat chromosomes, and probably during the evolution and polyploidy processes, these genes have increased randomly (Faraji et al, 2021). Besides, the investigation of the duplication process between the *TtAQP* genes showed that many duplications occurred between the members of the TtAQP family during evolution (Figure 3 and Table S3). Accordingly, 5% tandem duplication was observed for the TtAQP family (Table S3), and the most duplications were segmental type, which indicates gene transfer and change of chromosome set. Also, based on ka/ks ratio, the duplicated events between the TtAQP family members were under a purifying selection, which caused a decrease in diversity (Figure 4a). Moreover, the first duplication probably occurred around 59 million years ago between *TtAQP09* and *TtAQP51* (Figure 4b).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

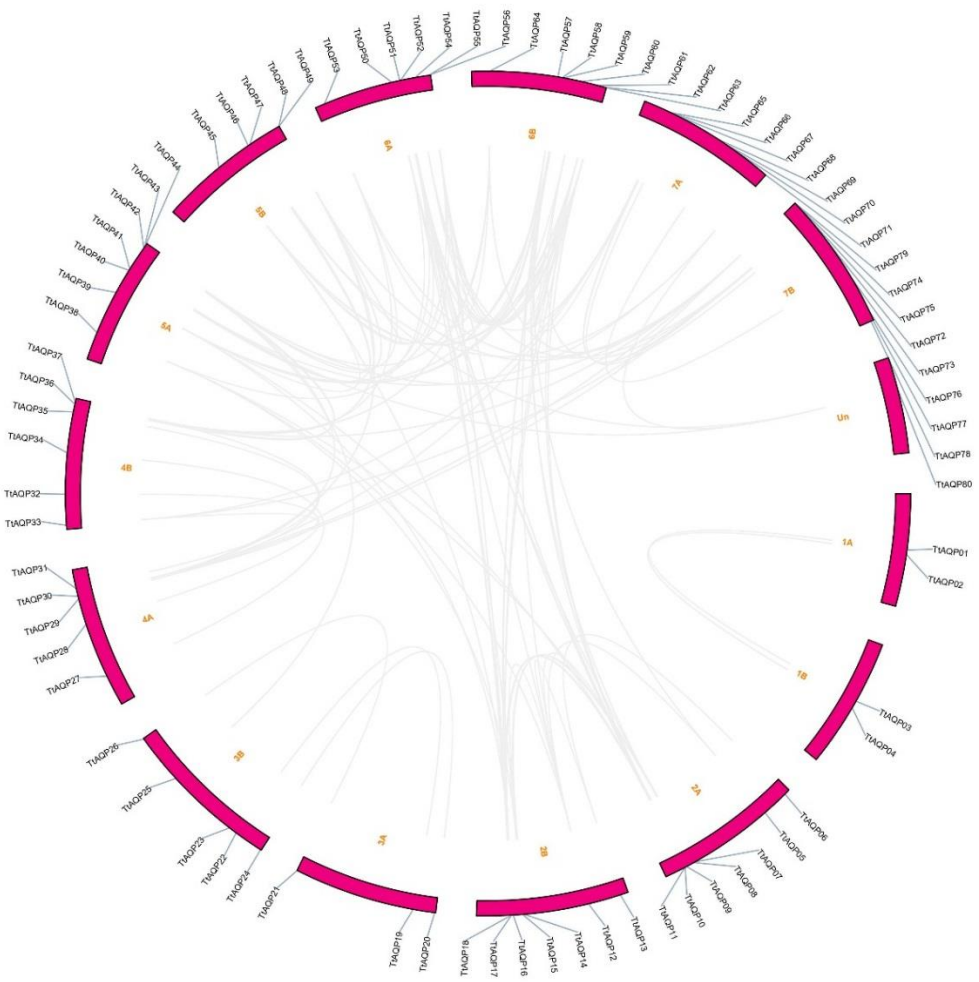

Figure 3. The location of *TtAQP* genes into the genome of *Triticum turgidum*. Duplicated genes are connected with gray lines

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

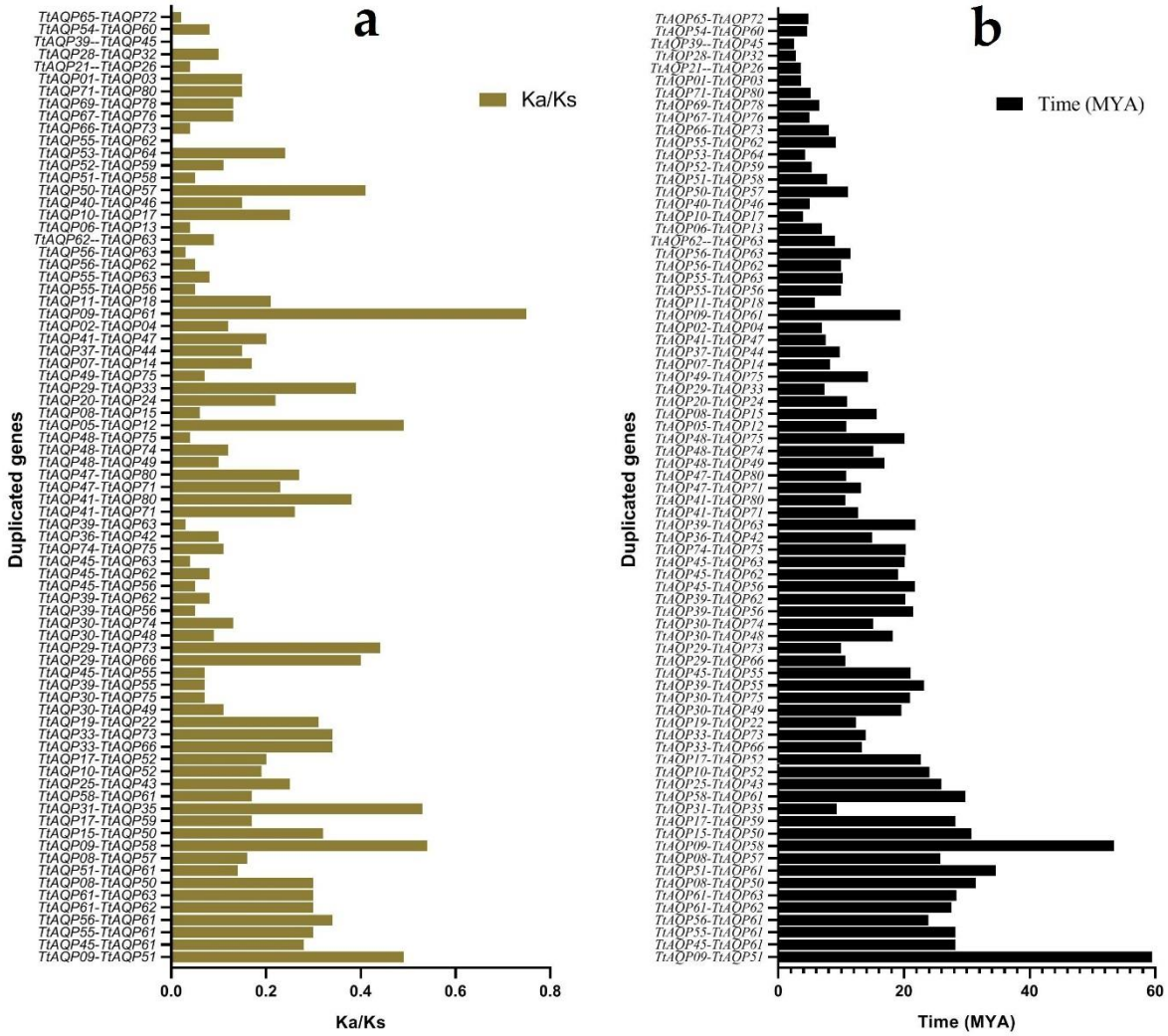

Figure 4. The Ka/Ks ratio for duplicated *TtAQP* genes (a), and time of divergence (MYA; million years ago) of the duplicated *TtAQP* genes (b).

3.4. Protein structure analysis of TtAQPs

The prediction results of the three-dimensional structure of TtAQP proteins showed that these proteins have different interaction areas (Figure 5a). Based on the three-dimensional structure, TtAQPs from group 1 and group 2, including PIP proteins, had almost similar structures. However, they differed in the location and type of binding region (pocket sites) (Figures 5a and b). The abundance of amino acids in the interaction and binding region of TtAQP proteins was also determined (Figure 5b). The amino acids, including glycine, alanine and valine had the highest frequency in the binding region of TtAQP proteins. These areas can be considered more in functional genomic studies.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

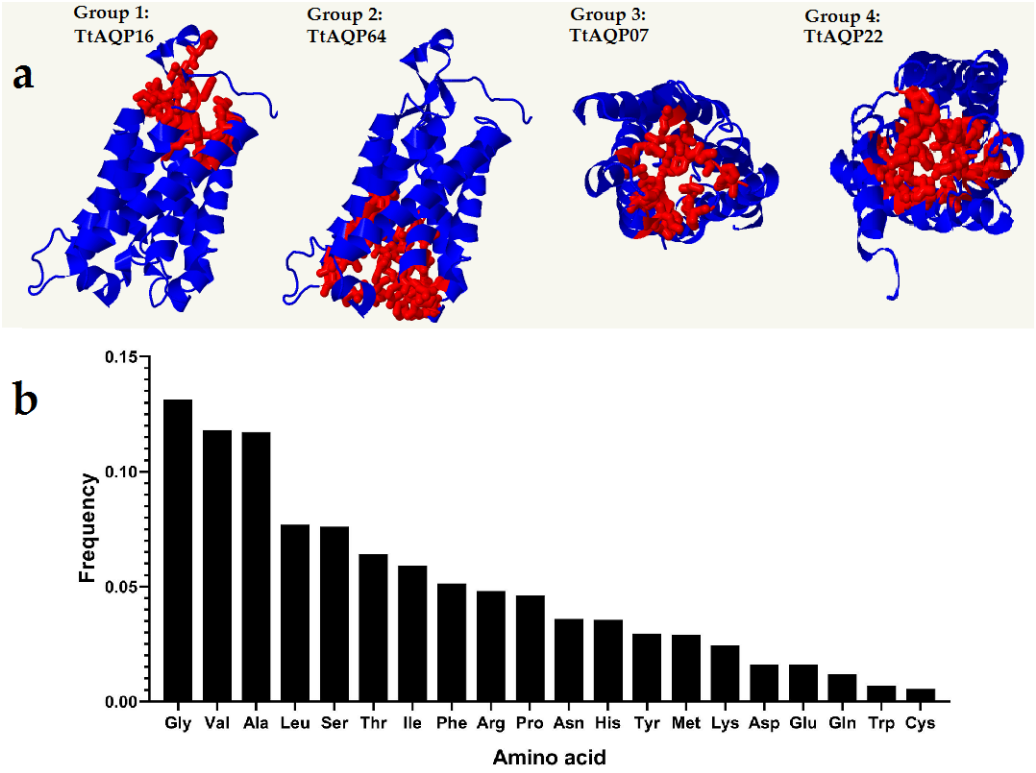

Figure 5. 3D structure analysis of TtAQP proteins (a), and frequency of amino acids present in binding sites of TtAQPs (b). Red balls show the binding site in the structure of selected proteins.

3.5. Phosphorylation analysis into TtAQPs

Phosphorylation regions in terms of three amino acids serine, tyrosine and threonine in TtAQP proteins were predicted. According to the analysis, it was found that the amount of serine compared to tyrosine and threonine is more subjected to post-translational phosphorylation modifications (Figure 6). The number of predicted phosphorylation areas in TtAQPs from group 2 showed a higher potential for phosphorylation. Phosphorylation is one of the important post-translational modifications that affect protein function, durability, and interaction [39,40]. It seems that TtAQP proteins of the second group (PIP2 subfamily) are more involved in cell signaling pathways.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

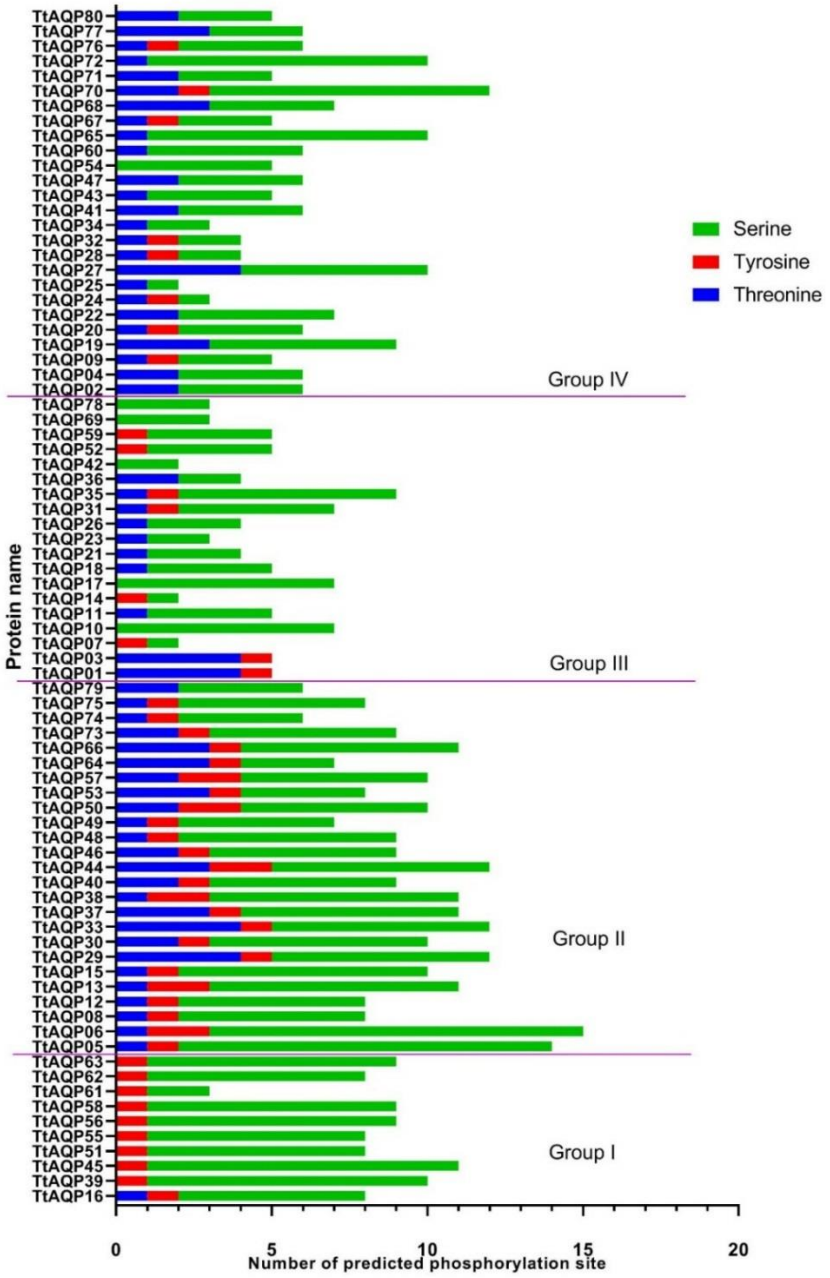

Figure 6. Prediction of phosphorylation site into TtAQPs based on three amino acids, including serine, tyrosine, and threonine. The potential value was > 0.80.

3.6. Expression profile of TtAQPs in response to salinity

The expression pattern of selected genes showed that *TtAQP* genes have different degrees of expression in response to salt stress (Figure 7). Most of the *TtAQP* genes were induced by salinity. In response to 150 mM NaCl treatment, *TtAQP58*, as a *PIP1* gene from group 1, *TtAQP18* and *TtAQP42*, as the *TIP* genes from group 3, and *TtAQP34*, as a *SIP* gene from group 4, were sharply upregulated after 72 h, while two *PIP2* genes from group 2, including *TtAQP29* and *TtAQP79* were not induced. However, *TtAQP29* showed a down-regulation after 72h of 150 mM NaCl. Interestingly, all selected *TtAQPs* were significantly upregulated in response to a high salinity concentration, 250 mM of NaCl, and the most expression levels were observed after 72h of salt stress. Overall, the expression patterns of six selected *TtAQP* genes in durum wheat showed that *TtAQPs* are mostly involved in the group of late responses to salinity stress.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

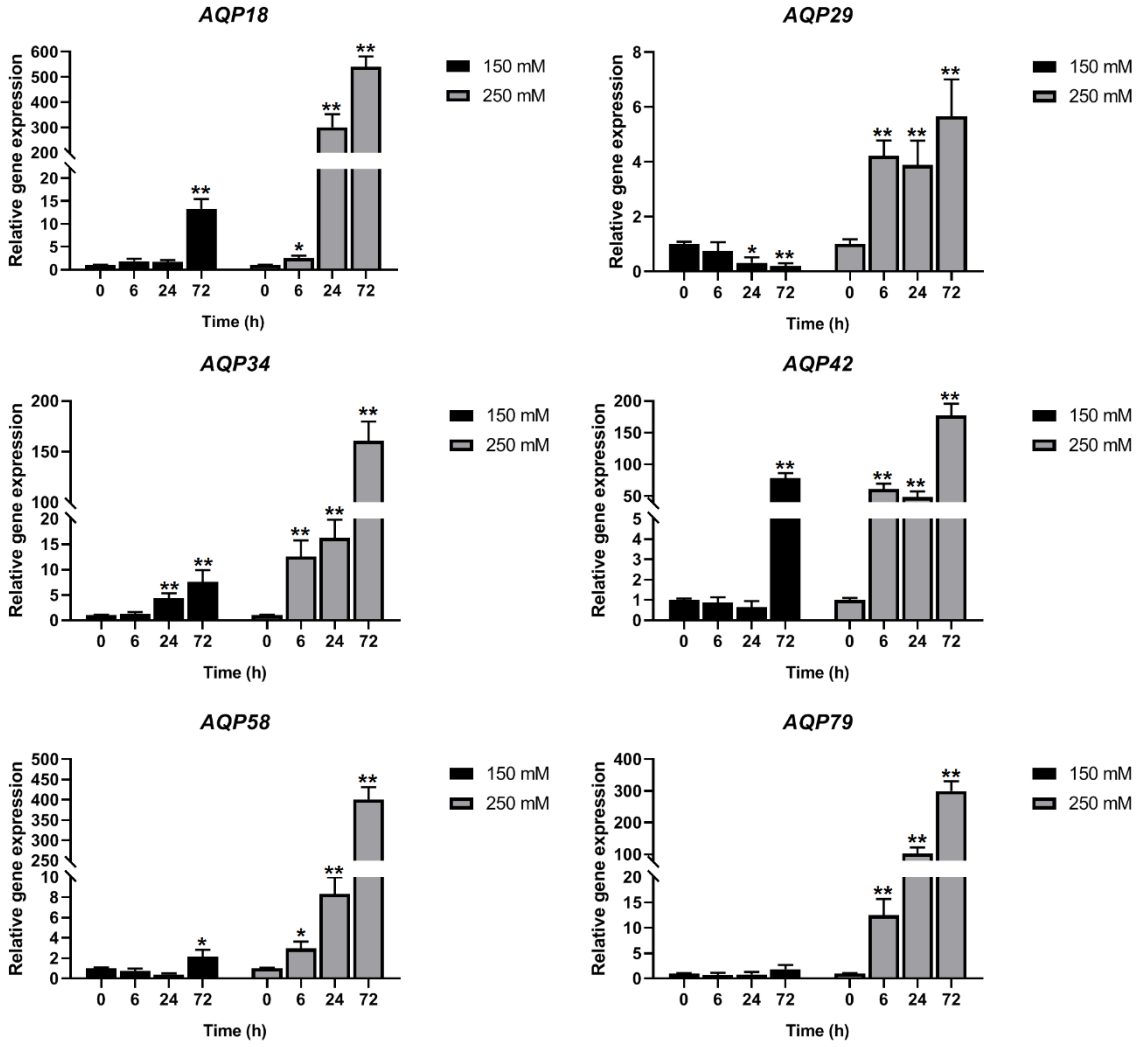

Figure 7. Expression profile of selected *TtAQP* genes under different concentrations of salinity, 150 and 250 mM of NaCl, and four time-points, including 0 (as a normal condition), 6, 24, and 72h after salt stress. * and ** indicate significant difference between the experimental treatments and control treatment (according to Student's t-test) at $p < 0.05$ and $p < 0.01$, respectively..

3.7. Promotor analysis of TtAQP genes

The promoter region of *TtAQP* genes was screened to identify the putative *cis*-regulatory elements. Results disclosed that the important key regions involved in response to biotic and abiotic stresses, as well as hormones, are located in the upstream region of *TtAQP* genes (Figure 8). Moreover, it was found that MYB elements, an element involved in response to stresses had the highest frequency in the promoter of *TtAQP* genes (Table S4). Then, CAT-box elements, dependent on meristem expression, and ABRE elements, which are important regulatory elements involved in response to abscisic acid (ABA) hormone, had the highest frequency in the promoter of *TtAQP* genes. In addition, MRE elements, which are elements present in the MYB binding site and play a role in responding to light, had the lowest frequency of presence in the promoter of *TtAQP* genes. Following that, WUN-motif, which are elements involved in response to abiotic stresses, and AuxRR-core, which include important regulatory elements in response to auxin hormone,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

had the lowest frequency of presence in the promoter of *TtAQP* genes. In general, the identified regulatory elements can be classified into four groups, including hormone-responsive elements, stress, light and growth. The highest of regulatory elements the upstream of *TtAQPs* was predicted in the field of stress and hormone-related functions. The presence of these important regulatory elements in the promoter sites indicates that *TtAQP* genes are involved in the response of plants to stress conditions. Overall, *TtAQP* genes showed high potential to participate in various cellular processes, so it is recommended to focus on this gene family in molecular works related to durum wheat breeding.

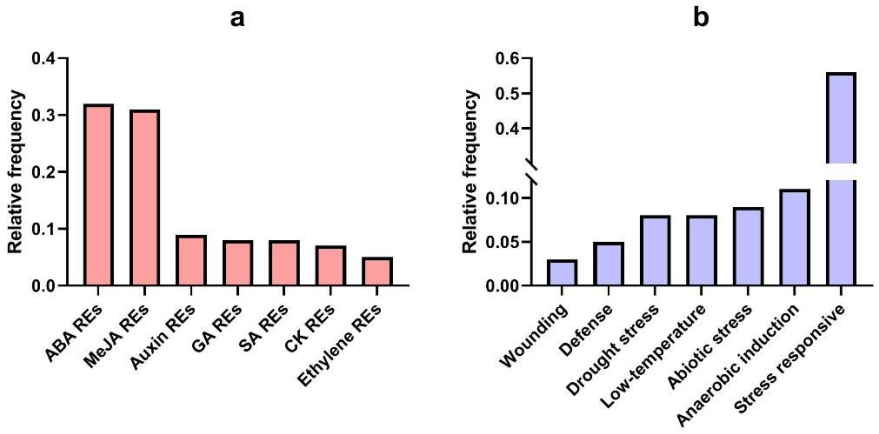

Figure 8. Grouping the cis-regulatory elements related to responsive-elements (REs) to phytohormones (a), and REs to stress (b). Full details are provided in Table S4.

3.8. Protein-protein interaction of TtAQPs

The interaction network of TtAQPs was constructed based on their orthologues in *Triticum aestivum* using the STRING database (Figure 9a). According to the predicted network, two PIP1 proteins, TtAQP51 and TtAQP45, and two SIP proteins, TtAQP27 and TtAQP34, and two NIP proteins, including TtAQP67 and TtAQP47, showed high interaction with other elements in the network (Figure 9a). Moreover, all proteins present in the network were analyzed to identify the significant (FDR < 0.001) gene ontology (GO) terms (Figure 9b). Cellular process in biological process terms, and channel activity and water channel activity from molecular function terms, were significantly enriched. The cellular component GO terms, including membrane, plasma membrane, vacuolar membrane, and cellular anatomical entity, were significantly enriched. These results revealed that TtAQPs are located in cell membranes and vacuoles and are more involved in the process of transferring water and ions. Results suggest that AQPs interact with each other, which probably affects their activity level.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0281.v1

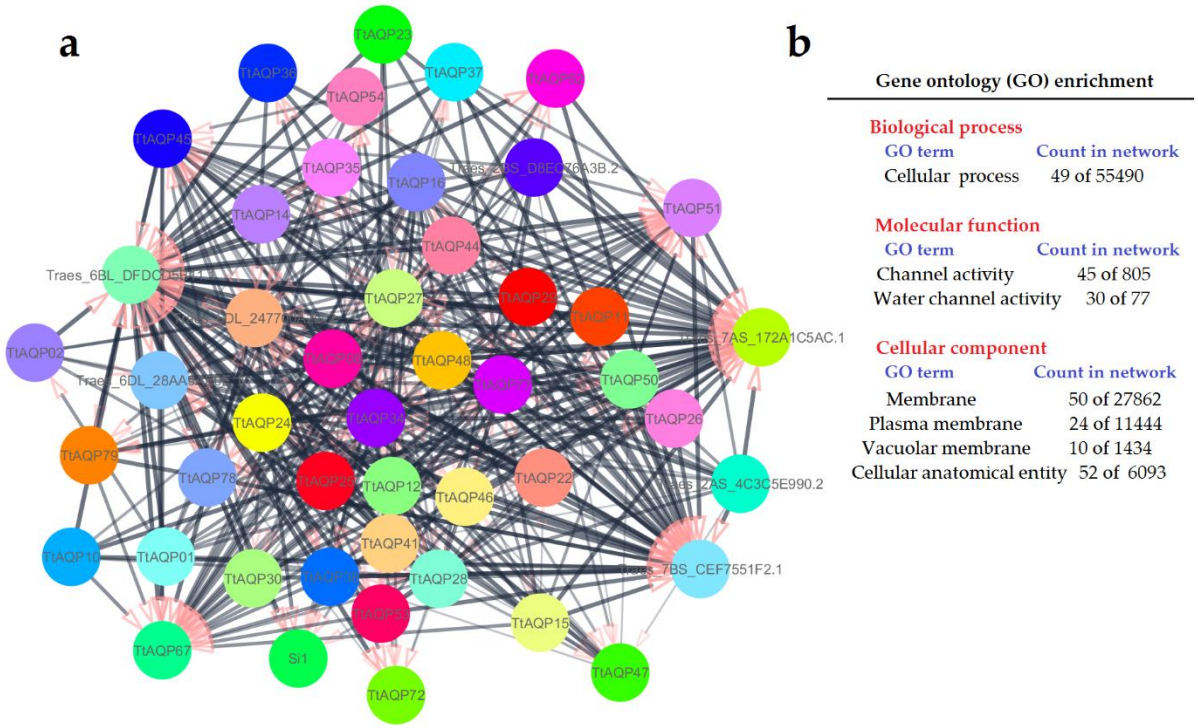

Figure 9. Interaction network of TtAQP proteins (a), and list of GO enrichment of proteins presented in the network (b). The network was constructed based on their homologs of TtAQPs in *Triticum aestivum* using the String database.

4. Discussion

The role of AQPs in regulating cell homeostasis by controlling the flow of water and some ions in the cell membrane is well known [1]. The function of the members of AQP gene family has been studied in several plants, although the structure, evolutionary process and function of this gene family have not been investigated in *Triticum turgidum*, so far. In the present study, 80 putative *TtAQP* genes were identified and characterized in the genome of *Triticum turgidum*, as the first report. Based on the results of previous studies, the number of AQP family members has been variable. For example, 32 *AQPs* in Physic nut [41], 35 in *Arabidopsis* [42], 41 in potato [43], 26 in bamboo [44], 45 in cassava [45], 71 in cotton [46], 67 in *Brassica oleracea* [47], and 51 in flax [48] were identified. The number of *AQP* was higher in durum wheat; there is a hypothesis that this gene family is more affected by polyploidy and duplication events under the evolution process [49,50]. Besides, a large number of segmental duplications was observed between TtAQP family members, indicating that segmental duplication has been the main responsible for the expansion of the TtAQP family under evolution. However, we know that durum wheat had undergone an allopolyploid event, which led to extending the homolog genes in the genome of durum wheat. Moreover, according to the Ka/Ks index, the duplicated members of *TtAQP*s have been under a negative selection; as a result, the functional diversity among *TtAQPs* has decreased. The exon/intron number was varied between TtAQP subfamilies, suggesting that each subfamily has had a different evolutionary process and each AQP subfamily was created before the derivation of monocotyledonous and dicotyledonous plants [51]. Furthermore, it has been reported that the intron/exon number can affect the gene expression in plant species [50]. According to phylogenetic analysis, AQP proteins were separated into four main groups, and subfamilies PIP1 and PIP2 showed more genetic distance than subfamily TIP, SIP, and NIP. Previously, it was reported that NIP subfamily originated from the bacteria genome [52,53]. It seems that this gene family has been subjected to evolutionary pressure and extended by duplication, polyploidy, and horizontal transmission.

 doi:10.20944/preprints202212.0281.v1

The post-translational modifications tightly regulated the activity of aquaporin [54]. In the current study, PIP2 subfamily proteins showed a high potential for phosphorylation events. PIP2 proteins have high water channel efficiency than PIP1 proteins [55] It is hypothesized that phosphorylation affects the channel activity of this subfamily. Phosphorylation is one of the important post-translational modifications that affect protein function, durability and interaction [49,56]. Besides, previous studies revealed that the channel opening of AQPs is influenced by phosphorylation at C terminal sites [14,57,58]. In addition, it was stated that ethylene can regulate the C terminal phosphorylation of Arabidopsis PIP2;1 (AtPIP2;1) [59]. Moreover, abiotic stresses such as salinity affect the phosphorylation of AtPIP2;1 [60]. Identifying more potential phosphorylation sites in TtAQP proteins could suggest that this subfamily has more channel activity and is more affected by upstream signaling pathways related to phosphorylation events.

The candidate *TtAQP* genes showed diverse expression patterns in response to salinity. In 150 mM NaCl condition, candidate genes from the PIP2 subfamily showed a downregulation unlike other selected genes from PIP1, TIP, and SIP subfamily. However, with increasing salinity concentration and duration of treatment, all *TtAQPs* were induced, and they showed the highest expression level. Results disclosed that *TtAQPs* are a part of durum wheat responses to salt stress. TIP and PIP subfamily, which are more studied than others [15], are mostly downregulated in response to abiotic stresses such as drought and salt stress [1,16,17]. The expression pattern of aquaporin genes in bread wheat (*Triticum aestivum*) showed decreased expression in response to salt stress (Figure S1). Bread wheat is more resistant to salt than durum wheat. According to the chromosomal set of bread wheat, AABBDD, it seems that genes relating to salt tolerance are probably located on DD set chromosomes, and affect the expression of AQPs. Furthermore, we speculated that *TtAQPs* are involved in the group of late responses of durum wheat to salinity stress.

5. Conclusions

In the present study, 80 putative *TtAQP* genes from the genome of *Triticum turgidum* were characterized. Results disclosed that subfamilies of TtAQP, including PIP1, PIP2, TIP, NIP, and SIP, have high genetic distances relative to each other. Furthermore, it was found that segmental duplications have played a major role in the extension of TtAQP family. In addition, we predicted that PIP2 subfamily members have more potential to be influenced by phosphorylation modification and they are probably involved in signaling pathways related to kinases. According to expression profile, we conclude that TtAQPs are associated with late responses to salt stress. Overall, we conclude that TtAQPs are diverse proteins, based on their structure, regulatory systems and expression. The results of this research can be used in further studies related to salt tolerance of durum wheat.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Table S1: List of *TtAQP* genes primers used in real-time PCR; Table S2: Accession number and Physio-chemical properties of TtAQP gene family members in *Triticum turgidum*; Table S3: Sequence and logo of conserved motifs distributed in AlLAC proteins; Table S4: List of the duplicated gene pairs in the TtAQP gene family; Table S5: Functional category of the putative cis-regulatory elements in *TtAQP* promoter; Figure S1: Expression heatmaps of AQP genes in bread wheat (*Triticum aestivum*) in response to abiotic stresses.

Author Contributions: Conceptualization, P.H.; methodology, M.Y. and P.H.; software, M.Y. and P.H.; validation, P.H.; formal analysis, M.Y.; investigation, P.H.; writing—original draft preparation, M.Y. and P.H.; writing—review and editing, P.H. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: Not applicable.

.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Afzal, Z.; Howton, T.C.; Sun, Y.; Mukhtar, M.S. The roles of aquaporins in plant stress responses. *J. Dev. Biol.* **2016**, *4*, 9.
2. Maurel, C.; Boursiac, Y.; Luu, D.-T.; Santoni, V.; Shahzad, Z.; Verdoucq, L. Aquaporins in plants. *Physiol. Rev.* **2015**, *95*, 1321–1358.
3. Kruse, E.; Uehlein, N.; Kaldenhoff, R. The aquaporins. *Genome Biol.* **2006**, *7*, 1–6.
4. Nordzieke, D.E.; Medraño-Fernandez, I. The plasma membrane: a platform for intra-and intercellular redox signaling. *Antioxidants* **2018**, *7*, 168.
5. Zhu, Y.-X.; Yang, L.; Liu, N.; Yang, J.; Zhou, X.-K.; Xia, Y.-C.; He, Y.; He, Y.-Q.; Gong, H.-J.; Ma, D.-F. Genome-wide identification, structure characterization, and expression pattern profiling of aquaporin gene family in cucumber. *BMC Plant Biol.* **2019**, *19*, 1–23.
6. Maurel, C.; Chrispeels, M.J. Aquaporins. A molecular entry into plant water relations. *Plant Physiol.* **2001**, *125*, 135–138.
7. Hachez, C.; Zelazny, E.; Chaumont, F. Modulating the expression of aquaporin genes in planta: a key to understand their physiological functions? *Biochim. Biophys. Acta (BBA)-Biomembranes* **2006**, *1758*, 1142–1156.
8. Hu, W.; Yuan, Q.; Wang, Y.; Cai, R.; Deng, X.; Wang, J.; Zhou, S.; Chen, M.; Chen, L.; Huang, C. Overexpression of a wheat aquaporin gene, TaAQP8, enhances salt stress tolerance in transgenic tobacco. *Plant Cell Physiol.* **2012**, *53*, 2127–2141.
9. Hu, W.; Hou, X.; Huang, C.; Yan, Y.; Tie, W.; Ding, Z.; Wei, Y.; Liu, J.; Miao, H.; Lu, Z. Genome-wide identification and expression analyses of aquaporin gene family during development and abiotic stress in banana. *Int. J. Mol. Sci.* **2015**, *16*, 19728–19751.
10. Li, G.; Chen, T.; Zhang, Z.; Li, B.; Tian, S. Roles of aquaporins in plant-pathogen interaction. *Plants* **2020**, *9*, 1134.
11. Xu, Y.; Hu, W.; Liu, J.; Zhang, J.; Jia, C.; Miao, H.; Xu, B.; Jin, Z. A banana aquaporin gene, MaPIP1; 1, is involved in tolerance to drought and salt stresses. *BMC Plant Biol.* **2014**, *14*, 1–14.
12. Madrid-Espinoza, J.; Brunel-Saldias, N.; Guerra, F.P.; Gutiérrez, A.; Del Pozo, A. Genome-wide identification and transcriptional regulation of aquaporin genes in bread wheat (Triticum aestivum L.) under water stress. *Genes (Basel).* **2018**, *9*, 497.
13. Kapilan, R.; Vaziri, M.; Zwiazek, J.J. Regulation of aquaporins in plants under stress. *Biol. Res.* **2018**, *51*, 1–11.
14. Maurel, C.; Verdoucq, L.; Luu, D.-T.; Santoni, V. Plant aquaporins: membrane channels with multiple integrated functions. *Annu. Rev. Plant Biol.* **2008**, *59*, 595–624.
15. Javot, H.; Maurel, C. The role of aquaporins in root water uptake. *Ann. Bot.* **2002**, *90*, 301–313.
16. Zhu, Y.-X.; Xu, X.-B.; Hu, Y.-H.; Han, W.-H.; Yin, J.-L.; Li, H.-L.; Gong, H.-J. Silicon improves salt tolerance by increasing root water uptake in Cucumis sativus L. *Plant Cell Rep.* **2015**, *34*, 1629–1646.
17. Sutka, M.; Li, G.; Boudet, J.; Boursiac, Y.; Doumas, P.; Maurel, C. Natural variation of root hydraulics in Arabidopsis grown in normal and salt-stressed conditions. *Plant Physiol.* **2011**, *155*, 1264–1276.
18. Boursiac, Y.; Chen, S.; Luu, D.-T.; Sorieul, M.; van den Dries, N.; Maurel, C. Early effects of salinity on water transport in Arabidopsis roots. Molecular and cellular features of aquaporin expression. *Plant Physiol.* **2005**, *139*, 790–805.
19. Yin, J.; Jia, J.; Lian, Z.; Hu, Y.; Guo, J.; Huo, H.; Zhu, Y.; Gong, H. Silicon enhances the salt tolerance of cucumber through increasing polyamine accumulation and decreasing oxidative damage. *Ecotoxicol. Environ. Saf.* **2019**, *169*,

8–17.

20. Dubcovsky, J.; Dvorak, J. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science (80-.).* **2007**, *316*, 1862–1866.
21. Bozzini, A. Origin, distribution, and production of durum wheat in the world. *Durum wheat Chem. Technol.* **1988**, 1–16.
22. Bolser, D.M.; Staines, D.M.; Perry, E.; Kersey, P.J. Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomic data. In *Plant Genomics Databases*; Springer, 2017; pp. 1–31.
23. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **2015**, *43*, D222–D226.
24. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J. Pfam: the protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230.
25. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, 2005; pp. 571–607.
26. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.
27. Nguyen, L.-T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating Maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274.
28. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296.
29. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208.
30. Musavizadeh, Z.; Najafi-Zarrini, H.; Kazemitabar, S.K.; Hashemi, S.H.; Faraji, S.; Barcaccia, G.; Heidari, P. Genome-Wide Analysis of Potassium Channel Genes in Rice: Expression of the OsAKT and OsKAT Genes under Salt Stress. *Genes (Basel).* **2021**, *12*, 784.
31. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202.
32. Yang, S.; Zhang, X.; Yue, J.-X.; Tian, D.; Chen, J.-Q. Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol. Genet. Genomics* **2008**, *280*, 187–198.
33. Blom, N.; Gammeltoft, S.; Brunak, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* **1999**, *294*, 1351–1362.
34. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858.
35. Lovell, S.C.; Davis, I.W.; Arendall III, W.B.; De Bakker, P.I.W.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins Struct. Funct. Bioinforma.* **2003**, *50*, 437–450.
36. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327.

37. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **2012**, *40*, e115–e115.

38. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods* **2001**, *25*, 402–408.

39. Rezaee, S.; Ahmadizadeh, M.; Heidari, P. Genome-wide characterization, expression profiling, and post-transcriptional study of GASA gene family. *Gene Reports* **2020**, *20*, 100795.

40. Ahmadizadeh, M.; Chen, J.-T.; Hasanzadeh, S.; Ahmar, S.; Heidari, P. Insights into the genes involved in the ethylene biosynthesis pathway in Arabidopsis thaliana and Oryza sativa. *J. Genet. Eng. Biotechnol.* **2020**, *18*, 1–20.

41. Zou, Z.; Yang, L.; Gong, J.; Mo, Y.; Wang, J.; Cao, J.; An, F.; Xie, G. Genome-wide identification of Jatropha curcas aquaporin genes and the comparative analysis provides insights into the gene family expansion and evolution in Hevea brasiliensis. *Front. Plant Sci.* **2016**, *7*, 395.

42. Johanson, U.; Karlsson, M.; Johansson, I.; Gustavsson, S.; Sjovall, S.; Fraysse, L.; Weig, A.R.; Kjellbom, P. The complete set of genes encoding major intrinsic proteins in Arabidopsis provides a framework for a new nomenclature for major intrinsic proteins in plants. *Plant Physiol.* **2001**, *126*, 1358–1369.

43. Venkatesh, J.; Yu, J.-W.; Park, S.W. Genome-wide analysis and expression profiling of the Solanum tuberosum aquaporins. *Plant Physiol. Biochem.* **2013**, *73*, 392–404.

44. Sun, H.; Li, L.; Lou, Y.; Zhao, H.; Gao, Z. Genome-wide identification and characterization of aquaporin gene family in moso bamboo (Phyllostachys edulis). *Mol. Biol. Rep.* **2016**, *43*, 437–450.

45. Putpeerawit, P.; Sojikul, P.; Thitamadee, S.; Narangajavana, J. Genome-wide analysis of aquaporin gene family and their responses to water-deficit stress conditions in cassava. *Plant Physiol. Biochem.* **2017**, *121*, 118–127.

46. Park, W.; Scheffler, B.E.; Bauer, P.J.; Campbell, B.T. Identification of the family of aquaporin genes and their expression in upland cotton (Gossypium hirsutumL.). *BMC Plant Biol.* **2010**, *10*, 1–17.

47. Diehn, T.A.; Pommerrenig, B.; Bernhardt, N.; Hartmann, A.; Bienert, G.P. Genome-wide identification of aquaporin encoding genes in Brassica oleracea and their phylogenetic sequence comparison to Brassica crops and Arabidopsis. *Front. Plant Sci.* **2015**, *6*, 166.

48. Shivaraj, S.M.; Deshmukh, R.K.; Rai, R.; Bélanger, R.; Agrawal, P.K.; Dash, P.K. Genome-wide identification, characterization, and expression profile of aquaporin gene family in flax (Linum usitatissimum). *Sci. Rep.* **2017**, *7*, 1–17.

49. Faraji, S.; Heidari, P.; Amouei, H.; Filiz, E.; Poczai, P. Investigation and Computational Analysis of the Sulfotransferase (SOT) Gene Family in Potato (Solanum tuberosum): Insights into Sulfur Adjustment for Proper Development and Stimuli Responses. *Plants* **2021**, *10*, 2597.

50. Heidari, P.; Puresmaeli, F.; Mora-Poblete, F. Genome-Wide Identification and Molecular Evolution of the Magnesium Transporter (MGT) Gene Family in Citrullus lanatus and Cucumis sativus. *Agronomy* **2022**, *12*, 2253.

51. Faraji, S.; Filiz, E.; Kazemitabar, S.K.; Vannozzi, A.; Palumbo, F.; Barcaccia, G.; Heidari, P. The AP2/ERF Gene Family in Triticum durum: Genome-Wide Identification and Expression Analysis under Drought and Salinity Stresses. *Genes (Basel).* **2020**, *11*, 1464.

52. Rivers, R.L.; Dean, R.M.; Chandy, G.; Hall, J.E.; Roberts, D.M.; Zeidel, M.L. Functional analysis of nodulin 26, an aquaporin in soybean root nodule symbiosomes. *J. Biol. Chem.* **1997**, *272*, 16256–16261.

53. Dean, R.M.; Rivers, R.L.; Zeidel, M.L.; Roberts, D.M. Purification and functional reconstitution of soybean nodulin 26. An aquaporin with water and glycerol transport properties. *Biochemistry* **1999**, *38*, 347–353.

 doi:10.20944/preprints202212.0281.v1

54. Singh, R.K.; Deshmukh, R.; Muthamilarasan, M.; Rani, R.; Prasad, M. Versatile roles of aquaporin in physiological processes and stress tolerance in plants. *Plant Physiol. Biochem.* **2020**, *149*, 178–189.

55. Chaumont, F.; Barrieu, F.; Jung, R.; Chrispeels, M.J. Plasma membrane intrinsic proteins from maize cluster in two sequence subgroups with differential aquaporin activity. *Plant Physiol.* **2000**, *122*, 1025–1034.

56. Abdullah; Faraji, S.; Mehmood, F.; Malik, H.M.T.; Ahmed, I.; Heidari, P.; Poczai, P. The GASA Gene Family in Cacao (Theobroma cacao, Malvaceae): Genome Wide Identification and Expression Analysis. *Agronomy* **2021**, *11*, 1425.

57. Törnroth-Horsefield, S.; Wang, Y.; Hedfalk, K.; Johanson, U.; Karlsson, M.; Tajkhorshid, E.; Neutze, R.; Kjellbom, P. Structural mechanism of plant aquaporin gating. *Nature* **2006**, *439*, 688–694.

58. Frick, A.; Järvå, M.; Ekvall, M.; Uzdavinys, P.; Nyblom, M.; Törnroth-Horsefield, S. Mercury increases water permeability of a plant aquaporin through a non-cysteine-related mechanism. *Biochem. J.* **2013**, *454*, 491–499.

59. Qing, D.; Yang, Z.; Li, M.; Wong, W.S.; Guo, G.; Liu, S.; Guo, H.; Li, N. Quantitative and functional phosphoproteomic analysis reveals that ethylene regulates water transport via the C-terminal phosphorylation of aquaporin PIP2; 1 in Arabidopsis. *Mol. Plant* **2016**, *9*, 158–174.

60. Prak, S.; Hem, S.; Boudet, J.; Viennois, G.; Sommerer, N.; Rossignol, M.; Maurel, C.; Santoni, V. Multiple phosphorylations in the C-terminal tail of plant plasma membrane aquaporins: role in subcellular trafficking of AtPIP2; 1 in response to salt stress. *Mol. Cell. Proteomics* **2008**, *7*, 1019–1030.