A peer-reviewed version of this preprint was published in PeerJ on 12 September 2019.

Qian W, Yang X, Li J, Luo R, Yan X, Pang Q. 2019. Genome-wide characterization and expression analysis of aquaporins in salt cress (*Eutrema salsugineum*) PeerJ 7:e7664
https://doi.org/10.7717/peerj.7664

Genome-wide identification and expression analysis of aquaporins in salt cress (*Eutrema salsugineum*)

Weiguo Qian [1], **Xiaomin Yang** [1], **Jiawen Li** [1], **Rui Luo** [1], **Xiufeng Yan** [1], **Qiuying Pang** [Corresp. 1]

[1] Alkali Soil Natural Environmental Science Center, Northeast Forestry University/Key Laboratory of Saline-alkali Vegetation Ecology Restoration in Oil Field, Ministry of Education, Harbin, China

Corresponding Author: Qiuying Pang
Email address: qiuying@nefu.edu.cn

Aquaporins (AQPs) serve as water channel proteins and belong to major intrinsic proteins (MIPs) family, functioned in rapidly and selectively transporting water and other small solutes across biological membranes. Importantly, AQPs have been shown to play critical roles in abiotic stress response of plants. *Eutrema salsugineum* is close to *Arabidopsis thaliana* and proposed as a model system for studying plant salt resistance. Here we identified 35 full-length *AQP* genes in *E. salsugineum*. Phylogenetic analysis showed EsAQPs were similar with AtAQPs and grouped into four subfamilies including 12 plasma membrane intrinsic proteins (PIPs), 11 tonoplast intrinsic proteins (TIPs), 9 NOD-like intrinsic proteins (NIPs), and 3 small basic intrinsic proteins (SIPs). Gene structure, also the conserved motifs (MEME) of EsAQPs in each subfamily shared high similarities. In detailed sequence analysis, EsAQPs comprised 237-323 amino acids, with a theoretical molecular weight (MW) of 24.31-31.80 kDa and an isoelectric point (pI) value of 4.73-10.49. Functional prediction based on the NPA motif, aromatic/arginine (ar/R) selectivity filter, Froger's position and specificity-determining position suggested there was a big difference in the specificity of substrate transport between EsAQPs. Gene expression profiles illustrated *EsAQP* genes could be detected in all organs and appear to play an important role in response salt, cold and drought signals. These results will bring a better understanding on the characterizations of AQPs in *E. salsugineum* and its complex transport networks in homeostasis control.

Genome-Wide Identification and Expression Analysis of Aquaporins in Salt cress (*Eutrema salsugineum*)

Weiguo Qian, Xiaomin Yang, Jiawen Li, Rui Luo, Xiufeng Yan, Qiuying Pang

Alkali Soil Natural Environmental Science Center, Northeast Forestry University/Key Laboratory of Saline-alkali Vegetation Ecology Restoration in Oil Field, Ministry of Education, Harbin, China

Corresponding Author:
Qiuying Pang
Hexing road, Harbin, Helongjiang, 150040, China
Email address: qiuying@nefu.edu.cn

Abstract

Aquaporins (AQPs) serve as water channel proteins and belong to major intrinsic proteins (MIPs) family, functioned in rapidly and selectively transporting water and other small solutes across biological membranes. Importantly, AQPs have been shown to play critical roles in abiotic stress response of plants. *Eutrema salsugineum* is close to *Arabidopsis thaliana* and proposed as a model system for studying plant salt resistance. Here we identified 35 full-length *AQP* genes in *E. salsugineum*. Phylogenetic analysis showed EsAQPs were similar with AtAQPs and grouped into four subfamilies including 12 plasma membrane intrinsic proteins (PIPs), 11 tonoplast intrinsic proteins (TIPs), 9 NOD-like intrinsic proteins (NIPs), and 3 small basic intrinsic proteins (SIPs). Gene structure, also the conserved motifs (MEME) of EsAQPs in each subfamily shared high similarities. In detailed sequence analysis, EsAQPs comprised 237-323 amino acids, with a theoretical molecular weight (MW) of 24.31-31.80 kDa and an isoelectric point (pI) value of 4.73-10.49. Functional prediction based on the NPA motif, aromatic/arginine (ar/R) selectivity filter, Froger's position and specificity-determining position suggested there was a big difference in the specificity of substrate transport between EsAQPs. Gene expression profiles illustrated *EsAQP* genes could be detected in all organs and appear to play an important role in response salt, cold and drought signals. These results will bring a better understanding on the characterizations of AQPs in *E. salsugineum* and its complex transport networks in homeostasis control.

Introduction

Water is the most abundant molecule in living cells, also the medium which all biochemical activities take place in (*Dev and Herbert, 2018*). Aquaporins (AQPs) belong to the major intrinsic proteins (MIPs) superfamily, which could efficiently and selectively transport water molecules across the cell membrane. In addition, AQPs can also transport many small molecules, such as glycerol, urea, carbon dioxide (CO_2), silicon, boron, ammonia (NH_3) and hydrogen peroxide (H_2O_2) (*Biela et al., 1999*; *Gerbeau et al., 1999*; *Uehlein et al., 2003*; *Ma et al., 2006*; *Takano et al., 2006*; *Loque et al., 2005*; *Dynowski et al., 2008*). AQPs were discovered in animals and subsequently found in almost all living organisms (*Gomes et al., 2009*). Compared with animals, plants have more robust and diverse AQPs. For instance, there are 35 AQPs in *Arabidopsis thaliana*, 33 in *Oryza sativa*, 40 in *Sorghum bicolor*, 72 in *Glycine max*, 47 in *Cicer arietinum* and 45 in *Manihot esculenta* (*Johanson et al., 2001*; *Sakurai et al., 2005*; *Kadam et al., 2017*; *Zhang et al., 2013*; *Deokar et al., 2013*; *Putpeerawit et al., 2017*).

Plant AQPs can be divided into seven subfamilies based on the protein sequence similarity analysis. Plasma membrane intrinsic proteins (PIPs) are the largest subfamily of plant AQPs. The most of the PIPs are commonly localized in the plasma membrane and are further divided into two phylogenetic groups PIP1 and PIP2. Tonoplast intrinsic proteins (TIPs) subfamily is usually localized in the tonoplast, which contain five classes TIP1, TIP2, TIP3, TIP4 and TIP5. NOD26-like intrinsic proteins (NIPs) named from NIP protein (Nodulin-26, GmNOD26), were discovered in the plasma membrane of soybean cells (Fortin *et al*., 1987). Small basic intrinsic proteins (SIPs) are typically localized in the endoplasmic reticulum. X intrinsic proteins (XIPs) are present in some dicots but absent in Brassicaceae and monocots (*Maurel et al., 2015*). GlpF-like intrinsic proteins (GIPs) are found in moss (*Physcomitrella patens*) and similar to bacterial glycerol channels (*Danielson and Johanson, 2008*; *Gustavsson et al., 2005*). Hybrid intrinsic proteins (HIPs) are found in fern (*Selaginella moellendorffii*) and moss (*Anderberg et al., 2012*; *Gustavsson et al., 2005*). Therefore, some classes (such as XIPs, HIPs, or GIPs) are considered to be lost during the evolution of certain plant lineages pointing to functional redundancies (*Maurel et al., 2015*).

AQPs are highly conserved in molecular structure, consisting of six transmembrane α-helical domains (TM1-TM6) linked by five loops (A-E), with both the N and C terminal having a cytoplasmic orientation. There are two highly conserved NPA (Asn-Pro-Ala) motifs in two half helices (HB and HE) of loopB and loopE at the center of the pore that have substrate selectivity (*Tajkhorshid et al., 2002*). The narrow aromatic/arginine (ar/R) selectivity filter is formed by four residues from TM helix 2 (H2), TM helix 5 (H5), and loop E (LE1 and LE2), which has been shown to provide a size barrier for solute permeability (*Bansal and Sankararamakrishnan, 2007*). Froger's position consists of five residues (P1-P5) that could transport two different types of molecules, water and glycerol (*Froger et al., 1998*). Moreover, a comprehensive analysis on functional characterization of AQPs, predicting nine specificity-determining positions (SDPs) for non-aqua substrates, such as ammonia, boron, carbon dioxide, hydrogen peroxide, silicon and urea, for each unique group (*Hove and Bhave, 2011*).

Salt cress previously named as *Thellungiella halophila* or *Thellungiella salsuginea,* recently was corrected to *Eutrema salsugineum* based on taxonomy and systematics, which is close to *A. thaliana* (Koch and German, 2013). *A. thaliana* is a salt-sensitive plant which has certain limits in studying the mechanism of salt and drought resistance. Importantly, *E. salsugineum* has a small genome, and also tolerant to salt, drought and low temperature stress, thus it is considered to be a halophyte model plant for investigating the mechanism of plant resistance to stress (*Zhu, 2001*; *Inan et al., 2004*). The *E. salsugineum* AQPs like TsTIP1;2, TsMIP6 and TsPIP1;1 have been found to play an important role in plant response to abiotic stress (*Wang et al., 2014*; *Sun et al., 2015*; *Li et al., 2018*). Since the *E. salsugineum* genome was sequenced in 2012 and 2013 at the chromosome level and scaffold level respectively (*Wu et al., 2012*; *Yang et al., 2013*), promoting the bioinformatics analysis of whole aquaporin family.

In this study, a genome-wide analysis of AQP genes was carried out in *E. salsugineum*, a total of 35 full-length AQP genes were identified. Based on the phylogenetic analysis, we found the identified EsAQPs were quite similar to AtAQPs. The EsAQPs could be grouped into four subfamilies, including PIPs, TIPs, NIPs and SIPs. Each of these members was analyzed to identify their protein sequences, chromosome distribution, gene structure and putative function. The expression level of EsAQPs in different organs and the RNA relative fold changes of EsAQPs in response to salt, drought and cold stress were also investigated.

Materials & Methods

Identification and chromosomal location of EsAQPs

The whole genome of *E. salsugineum* was downloaded from NCBI (https://www.ncbi.nlm.nih.gov/genome/12266, *Wu et al., 2012*; *Yang et al., 2013*). To identify *E. salsugineum* AQP candidate genes, a Hidden Markov Model (HMM) analysis was used. HMM profile of MIP (PF00230) was downloaded from Pfam protein family database (http://pfam.sanger.ac.uk/) and used as the query ($P<0.05$) to search for AQP proteins in the *E. salsugineum* genome. To avoid missing potential AQP members, the NCBI BLAST tool was used to search *Arabidopsis* AQP proteins, and the top five aligned sequences were considered as candidates. After removing all of the redundant sequences, the sequences of putative *EsAQP* genes were loaded on relative chromosomes of *E. salsugineum* using the SnapGene tool. The map of chromosome position of each *EsAQP* genes was drawn by MapInspect 1.0.

Classification, phylogenetic analysis and structural features

Multiple sequence alignments of putative AQP proteins were performed by ClustalW, and a phylogenetic tree was constructed using neighbor joining with MEGA 6.0 (*Tamura et al., 2013*). The transmembrane regions were detected using TOPCONS (http://topcons.cbr.su.se/pred/) and TMHMM (http://www.cbs.dtu.dk/services/TMHMM/). Protein subcellular localization of *E. salsugineum* AQPs was predicted in Plant-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/) and WoLF PSORT (http://www.genscript.com/wolf-psort.html). Functional predictions, such as NPA motifs, ar/R filters (H2, H5, LE1 and LE2), Froger's positions (P1-P5) and nine specificity-determining positions (SDP1-SDP9), were analyzed by the alignments with function known AQPs (*Quigley et al., 2001*; *Park et al., 2010*; *Hove and Bhave, 2011*). The gene structure

for each EsAQP was illustrated with the Gene Structure Display Server 2.0 (http://gsds.cbi.pku.edu.cn/). The conserved motifs of EsAQP proteins were analyzed by MEME suite (http://meme-suite.org/).

Plant materials and stress treatments

E. salsugineum seeds (ecotype Shandong, China) were provided by Prof. Hui Zhang (Shandong Normal University, Jinan, China). The seeds were plated on 1/2 MS medium and treated at 4°C in the dark for 7 days, then cultured in plant growth chamber with illumination of 150 μmol/m^2/s, photoperiod 16/8 h of light/darkness at 25°C and 60% relative humidity. After one week, transfer the seedlings into a mixed medium with soil and vermiculite (3:1). Vernalization treatment for bolting was conducted in 4-week old seedlings at 4°C for 4 weeks, and moved them back to growth chamber until getting flowers. Samples of roots, stems, leaves, flowers and siliques, were collected, immediately frozen in liquid nitrogen and stored at -80°C for further analysis.

For abiotic stress assays, the 4-week old seedlings were exposed to 300 mM NaCl for 24 h as salt stress condition, treated at 4 °C for 24 h as cold stress, and lack of irrigation until the soil moisture content was less than 20% for 7 days as drought stress. The aerial part of seedlings was collected for further analysis.

RNA extraction, cDNA synthesis and qRT-PCR

The total RNA was extracted using TRIzol reagent (Takara) following the manufacturer's protocol. The quality of the RNA was determined using an ultraviolet spectrophotometer (Thermo, BioMate 3S). After removing genomic DNA contamination with DNase I, cDNA was synthesized by using the PrimeScript™ RT Reagent Kit (Takara). Three biological replicates of cDNA samples were used for qRT-PCR analysis with three technical replicates.

All of primers were designed using Primer 3.0 (http://bioinfo.ut.ee/primer3-0.4.0/) and listed in Table S1. The qRT-PCR analysis was conducted in Applied Biosystems 7500 Real-Time PCR System (ABI, USA) by using SYBR Premix Ex TaqTM II (Takara). Reaction system contained 10 μL SYBR Premix Ex Taq II, 2 μL5-fold diluted cDNA, 0.8 μL of each primer (10 mM), and ddH_2O to a final volume of 20 μL. The PCR program was set as follows: 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s and 60 °C for 34 s. Then, a melting curve was generated to analyze the specificity of each primer with a temperature shift from 60 to 95 °C. The fold changes of the *EsAQPs* expression under abiotic stresses were calculated with the $2^{-\Delta\Delta}$Ct method, while the gene expressions level of *EsAQPs* in each organ were calculated with the $^{\Delta}$Ct method. The heat map of gene expression pattern was visualized using HemI software.

Results

Characters, classification and chromosome localization of EsAQPs

A total of 35 putative AQPs were identified in *E. salsugineum* at the scaffold level (GenBank assembly accession GCA_000478725.1) based on HMM analysis and BLAST searches against *Arabidopsis* AQPs. The *AQP* genes were aligned into *E. salsugineum* chromosomes (GenBank assembly accession GCA_000325905.2), along with their scaffold numbers, CDS numbers and protein IDs, were listed in Table 1. To classify the AQP members, a phylogenetic tree was constructed according to the similarity of AQP protein sequences of *E. salsugineum* and *A.*

thaliana through the neighbor-joining method (Fig. 1). Based on the phylogenetic analysis, we found the identified EsAQPs have very high similarity with AtAQPs and can be grouped into four subfamilies, including 12 PIPs, 11 TIPs, 9 NIPs and 3 SIPs. In addition, the EsPIP subfamily was further divided into two classes (5 EsPIP1s and 7EsPIP2s), the EsTIP subfamily into five classes (3 EsTIP1s, 4 EsTIP2s, 2 EsTIP3s, 1 EsTIP4s and 1 EsTIP5s), the EsNIP subfamily into seven classes (1 EsNIP1s, 1 EsNIP2s, 1 EsNIP3s, 3 EsNIP4s, 1 EsNIP5s, 1 EsNIP6s and 1 EsNIP7s), and the EsSIP subfamily into two classes (2 EsSIP1s and 1 EsSIP2s). The nomenclature of *E. salsugineum* AQPs was based on their corresponding homolog in AtAQPs (Fig. 1). According to the amino acid homology, XP_006410897.1 and XP_006392950.1, which were annotated as EsPIP2-2 and EsTIP2-1 in NCBI, were corrected into EsPIP2;3 and EsPIP2;4, respectively. Additionally, XP_006405831.1 and XP_006405829, both annotated as EsNIP4-1 in NCBI, were corrected into EsNIP4;2 and EsNIP4;3, respectively (Table 1). Based on the comparison with *Arabidopsis* aquaporins, PIP2;8 and NIP1;1 were not identified in *E. salsugineum* but were replaced by TIP2;4 and NIP4;3.

As shown in Table 1 and Figure 2, 34 *EsAQP* genes were randomly located at different chromosomes as sequenced by Wu *et al*. (*2012*). Chromosome 4 and 5 contained the maximum number of seven *EsAQP* genes, chromosome 7 contained six members. Chromosomes 6, 1, 3 and 2 contained five, four, three, and two *EsAQP* genes, respectively. In addition, all *EsAQPs* were found in 15 different scaffolds sequenced by Yang *et al*. (*2013*). It is notable that *EsAQPs* with same scaffold numbers were located at same chromosomes with neighbor positions, indicating that the two sequencing results were consistent (Table 1), except for *EsTIP2;2*, which was found at the scaffold level but not located at the chromosomes.

Gene structure and subcellular localization analysis of EsAQPs

Gene structure analysis of the 35 *EsAQPs* was performed in the Gene Structure Display Server of NCBI. Based on their mRNA and genomic DNA sequences, we found exon lengths were mostly conserved in each subfamily of *EsAQP* gene with same exon number, but introns varied in both length and position (Fig. 3). All members of EsPIP subfamily contained four exons with similar length (289-328, 296, 141 and 93-126 bp, respectively) and conserved sequences in the 2nd and 3rd exon, except for *EsPIP2;4*, which have a shorter 2nd and longer 3rd exon (307, 151, 286, and 111 bp). The majority members of EsTIP subfamily contained three exons with similar lengths, and the other members had two exons with similar lengths, except for *EsTIP1;3*, which had only one exon without intron. In the EsNIP subfamily, some members exhibited five exons with similar lengths, while others had four exons with varied lengths. All EsSIP subfamily genes were characterized by three exons with similar lengths. This description of exon-intron structure provides additional evidence to support the classification results (*Kong et al., 2017*).

The prediction of EsAQP subcellular localization in Plant-mPLoc showed that all EsPIP, EsNIP and EsSIP subfamilies were localized in plasma membrane, while EsPIP1;2 was localized in both plasma membrane and tonoplast membrane, all EsTIP subfamily members were localized in tonoplast membrane, and EsTIP5;1 was localized in both tonoplast membrane and plasma membrane. However, the prediction of EsAQP subcellular localization in WoLF PSORT showed

that most EsAQPs were localized in plasma membrane, except for four TIPs (EsTIP2;2, EsTIP2;3, and EsTIP2;4 in tonoplast membrane and EsTIP5;1 in chloroplast), two NIPs (EsNIP2;1 and EsNIP3;1 in tonoplast membrane) and two SIPs (EsSIP2;1 and EsSIP1;2 in tonoplast membrane). Combining the results of EsAQP subcellular localization predictions in Plant-mPLoc and WoLF PSORT, all EsPIP subfamily members were predicted to localize in the plasma membrane, and the other EsAQPs were localize in plasma membrane or tonoplast membrane.

Structure characteristics of EsAQPs

Sequence analysis showed that all EsAQPs contain six transmembrane domains (TMDs) comprising 237-323 amino acids ,had theoretical molecular weights (MW) of 24.31-31.80 kDa and isoelectric point (pI) values of 4.73-10.49 (Table 2). The EsPIP subfamily had a similar molecular weight of approximately 30.84 kDa. Most members of the EsNIP subfamily exhibited a similar molecular weight and isoelectric point of EsPIP subfamily. The EsTIP and EsSIP subfamilies had lower MW among the EsAQPs, and the isoelectric points of these two subfamilies were acidic and alkaline, respectively (Fig. S1).

NPA motifs, ar/R selectivity filters and Froger's positions of AQP protein sequences play critical roles in channel selectivity. Sequence alignment between AtAQPs and GhAQPs was carried out to analyze the conserved domains (*Quigley et al., 2001*; *Park et al., 2010*). The results in Table 2 showed that all EsPIP subfamily members had two typical NPA motifs in loop B and loop E, with a water transport ar/R filter with amino acid of F-H-T-R. Froger's position consists of Q-S-A-F-W in most cases, except for EsPIP2;7, which had an M at the P1 position. All EsTIP subfamily had two typical NPA motifs. The ar/R was composed of H-I-A-V in EsTIP1s, H-I-G-R in EsTIP2s and H-T/M/I-A-R in other EsTIP members, while in EsTIP5;1, it was composed of N-V-G-C. Froger's position consists of T-A/S-A-Y-W, except for EsTIP5;1 and EsTIP3;2, which had a V at the P1 position and a T at the P2 position respectively. Most members of EsNIP subfamily had two typical NPA motifs, not in EsNIP2;1 (with an NPG in LE), EsNIP5;1 and EsNIP7;1 (with an NPS in LB). The ar/R filter consists of residues like W/A-V/I-A/G-R, and Froger's position consists of F-S-A-Y-L, except for EsNIP7;1, which had a Y at the P1 position, and for EsNIP5;1 and EsNIP6;1 had a T at the P2 position. The EsSIP subfamily showed a variable site in the first NPA, the alanine (A) was replaced by threonine (T), cysteine (C) or leucine (L). The ar/R filter was also inconsistent with each other: I-V-P-I in EsSIP1;1, V-F-P-I in EsSIP1;2 and S-H-G-A in EsSIP2;1. The Forger's position was composed of I-A-A-Y-W in EsSIP1s, while it was F-V-A-Y-W in EsSIP2;1.

Conserved motifs of EsAQP proteins were predicted by MEME suite (Fig. 4). The results showed that motif 3 was found in all EsAQPs, and EsTIPs and EsNIPs having two motif 3 (except for EsNIP5;1 and EsNIP7;1). Motif 1 was absent only in EsSIPs and EsPIPs had three, while the others had two (except for EsNIP2;1 which had one). Motif 6 was present in all EsTIPs and EsNIPs, and EsNIPs had two (except for EsNIP3;1). Motif 8 was present in EsPIPs and EsSIPs (except for EsSIP2;1). However, some motifs were family-specific, such as motifs 2, 4, 7

and 10, which were present only in EsPIPs, and motif 5 was present only in EsTIPs (except for EsTIP5;1). In addition, motif 9 was present only in EsPIP1s.

Expression pattern of *EsAQPs*

The expression of *EsAQP* genes in different organs, including root, stem, leaf, flower and silique, was analyzed by RT-qPCR. The results showed that 35 *EsAQP* genes were detected in all the organs (Fig. 5A). Almost all *EsPIP* genes were highly expressed in all organs, except for *EsPIP2;5* in leaf. In addition, the *EsPIP* genes, *EsTIP1;1*, *EsTIP1;2*, *EsNIP1;2*, *EsNIP5;1*, *EsSIP1;1* and *EsSIP2;1* were also highly expressed in all organs. Some *EsAQP* genes, such as *EsTIP2;3*, *EsTIP2;4*, *EsNIP2;1* and *EsNIP3;1,* were specifically highly expressed in root. Two *EsTIPs* (*EsTIP2;2* and *EsTIP5;1*), three *EsNIPs* (*EsNIP4;1*, *EsNIP4;3* and *EsNIP7;1*) and *EsSIP1;2* were highly expressed only in flower. Two *EsTIPs* (*EsTIP3;1* and *EsTIP3;2*) were highly expressed in silique. Compared analysis of each *EsAQP* gene between different organs revealed that most *EsAQPs* showed higher expression level in flower than in other organs.

Abiotic stresses are the main limiting factors for plants during environmental conditions that induce osmotic stress and disturb water balance. AQPs play major roles in maintaining water homeostasis and responding to environmental stresses in plants. Therefore, we further investigated the expression patterns of *EsAQPs* under salt, drought and cold stress by qRT-PCR. The results showed that most of the *EsAQP* genes were upregulated under salt and cold stress but downregulated under drought stress (Fig. 5B). We found that five *EsAQP* genes were upregulated under all the types of abiotic stresses, including *EsPIP2;4*, *EsTIP1;2*, *EsNIP4;3*, *EsNIP5;1* and *EsSIP1;2*, while three *EsAQP* genes were downregulated under all the types of abiotic stresses, including *EsPIP1;5*, *EsTIP2;2* and *EsTIP2;4*. In addition, *EsPIP1;1* and *EsPIP2;2* were specifically upregulated under salt stress, and *EsPIP2;1*, *EsTIP2;1*, *EsTIP5;1*, *EsNIP4;1* and *EsNIP6;1* were upregulated only under cold stress.

Discussion

Gene duplication is a ubiquitous event that plays an important role in biological evolution, may also contribute to stress tolerance via gene dosage increasing, avoid some deleterious mutations and create the opportunity for immediate emergence of a new function (*Innan and Kondrashov, 2010*). AQPs are abundant, diverse and widely distributed in plants and involved in regulate plant growth and development. From algae (e.g., 2 in *Thalassiosira pseudonana* and 5 in *Phaeodactylum tricornutum*) (*Armbrust et al., 2004*; *Bowler et al., 2008*) to fern (19 in *Selaginella moellendorffii*) (*Danielson and Johanson, 2008*) and moss (23 in *Physcomitrella patens*) (*Anderberg et al., 2012*) to the higher plants (e.g., 35 AQPs in *Arabidopsis thaliana*, 33 in *Oryza sativa*, 72 in *Glycine max*) (*Johanson et al., 2001*; *Sakurai et al., 2005*; *Zhang et al., 2013*), the number of AQPs has largely increased with evolution. Here, we provide a genome-wide information of AQP family of *E. salsugineum*.

A total of 35 full-length AQPs were identified from *E. salsugineum* and grouped into four subfamilies, including twelve PIPs, eleven TIPs, nine NIPs and three SIPs (Fig. 1). The number of AQPs identified in *E. salsugineum* is same as *A. thaliana*, and their protein sequences have very high similarity. For instance, the similarity was even up to 99% between EsPIP1;1 and

AtPIP1;1. In previous studies, it was shown that more than 95% gene families are shared in *T. salsuginea* and *A. thaliana* (*Wu et al., 2012*) or that more than 80% *E. salsugineum* genes had high-homology orthologs in *A. thaliana* (*Yang et al., 2013*). In the AQP family, 33 of the 35 (over 94%) AQP genes from *E. salsugineum* could align with *A. thaliana* genes. Therefore, the nomenclature of *E. salsugineum* AQPs was based on their homologs in AtAQPs. Although they have very high similarity, many physiological characteristics differ from each other (*Pilarska et al., 2016*; *Prerostova et al., 2017*). The biological functions of AQPs need to be further investigated.

The comparison of EsAQPs with AtAQPs showed that *EsNIP1;2* shared 86% and 88% sequence similarities with *AtNIP1;1* and *AtNIP1;2* in nucleotide sequence and 83% and 91% sequence similarities with AtNIP1;1 and AtNIP1;2 in protein sequence, respectively; so it was named as EsNIP1;2. However, the position of AtNIP1;1 and AtNIP1;2 are very close at chromosome 4 and had the same ar/R filter (W-V-A-R) and P5 position (F-S-A-Y-L) (Quigley *et al*., 2001), it is same as EsNIP1;2 in our study (Table 1). This suggests that these genes may have same function. The four EsTIP2s members were named according to their homology of three AtTIP2s (Fig. 1). Moreover, EsTIP2;4 shared sequence similarities of 72%, 66% and 66% with EsTIP2;1, EsTIP2;2 and EsTIP2;3, respectively. This result implies that EsTIP2;4 may evolved from EsTIP2;1. *A. thaliana* has two NIP4s located closely at chromosome 5 (*Quigley et al., 2001*). The same phenomenon was also found in our study, which three EsNIP4s were very close at chromosome 7 (Table 1 and Fig. 2). Moreover, the gene structures of *EsNIP4;1*, *EsNIP4;2*, *AtNIP4;1* and *AtNIP4;2* were identical and had 5 exons, and the length of each exon (132, 225, 198, 62, and 235 bp) was consistent (*Tabata et al., 2000; Feng et al., 2017*).

Exon-intron structural divergences happened commonly in duplicate gene evolution and even in sibling paralogs; these changes occurred through the mechanisms of gain/loss, exonization/pseudoexonization and insertion/deletion (*Xu et al., 2012*). In common bean (*Phaseolus vulgaris* L.), each aquaporin subfamily are completely conserved in number, order and length of exons but varies in introns (*Ariani and Gepts, 2015*). The MEME motifs of the AQPs were conserved in all subfamilies, while a few were deleted, unique or family-specific, and a previous report also found this pattern in ZmPIPs (*Bari et al., 2018*). In our study, the exon-intron structure of *EsAQP* genes and the conserved MEME motifs of EsAQP protein sequences showed some common patterns (Fig. 3 and Fig. 4). All EsPIP subfamily members had four or three exon-intron structures, and the length of each exon was similar, except for *EsPIP2;4*, which had a shorter 2nd exon and a longer 3rd exon. Motif 1, 2, 3, 4, 7, 8, and 10 were same in all EsPIPs, and motif 2, 4, 7, and 10 were unique among EsAQPs. In addition, motif 9 was unique in EsPIP1s and may be used to distinguish EsPIP1s from EsPIP2s. This pattern of conserved motifs in the PIP subfamily also occurs in other plants and PIP1s contain one unique motif (*Tao et al., 2014*; *Yuan et al., 2017*). In the EsTIP subfamily, most genes contained 2 or 3 exons, and the length of each corresponding exon was similar (except for *EsTIP2;1*). The conserved motif analysis showed that almost all EsTIPs had two motif 1, two motif 3, one motif 5 and one motif 6. The exception was EsTIP1;3, which had no intron and motif 6. Motif 5 could

be an identifier of EsTIPs among the AQPs of *E. salsugineum* except for EsTIP5;1. The EsNIP subfamily contained 5 exons with similar length or 4 exons with various length (*EsNIP2;1*, *EsNIP3;1*, *EsNIP4;1* and *EsNIP5;1*). While most of the EsNIP genes with 4 exons were also different in MEME motifs among the NIP subfamily, most of members in NIP subfamily had two motif 1, two motif 3, and two motif 6, except for EsNIP2;1 (lose one motif 1), EsNIP3;1 (lose one motif 6) and EsNIP5;1 (lose one motif 3). The two motif 6 might be used to distinguish EsNIPs with other EsAQPs. All EsSIP subfamily had 3 exons with similar lengths and carried motif 3. Motif 8 appeared in EsSIP1s but not in EsSIP2;1, so it might be an specific trait of this group. This is a common phenomenon in plant SIP subfamily contains less motifs *(Tao et al., 2014*; *Reddy et al., 2015*; *Yuan et al., 2017*; *Kong et al., 2017*). These results indicated that the gene structure and the conserved motifs of EsAQPs shown subfamily-specific, these traits may provide new evidences to support the classification.

A high degree of conservation of signature sequences or residues was shown in plant PIP proteins. In our study (shown in Table 2), EsPIPs showed a typical NPA motif, a highly conserved ar/R selectivity filter and Froger's position of F-H-T-R and Q/M-S-A-F-W, these characteristics are correlated with water transport activity (*Quigley et al., 2001*). In addition to water transport, plant PIPs also could transfer carbon dioxide, hydrogen peroxide, boric acid, and urea (*Gaspar et al., 2003*; *Bienert et al., 2014*; *Heckwolf et al., 2011*). According to the SDP analysis proposed by Hove and Bhave (*2011*), all EsPIPs had H_2O_2-type and urea-type SDPs (Table 3, Fig. S2). In addition, all EsPIP1s and EsPIP2;5 had boric acid-type SDPs, and all EsPIP1s had CO_2-type SDPs, including two novel types of SDP showed in EsPIP1;3 and EsPIP1;4which have an M in place of I in SDP2, it also have been found in RcPIPs, JcPIPs and BvPIPs (*Zou et al., 2015*; *Zou et al., 2016*; *Kong et al., 2017*). In addition, EsPIP2;4 owned another novel CO_2-type SDPs (V-I-C-A-V-E-W-D-W), with E replaced by D in SDP6. These results showed the conservation of plant PIPs in the transportation of urea and hydrogen peroxide (*Gaspar et al., 2003*; *Bienert et al., 2014*), and PIP1s not PIP2s are the main CO_2 and boric acid channels (*Heckwolf et al., 2011*). Compared to PIPs, TIPs are more diverse and have a variety of selectivity filters. As shown in Table 2, typical NPA motifs were found in all the EsTIPs, and the ar/R filters and Froger's position were conserved in the EsTIP1s and EsTIP2s classes but different with other classes. All the EsTIPs showed urea-type SDPs, and most of them had H_2O_2-type SDPs (except for TIP3;1 and TIP5;1). EsTIP2;1 had an NH_3-type SDPs, as confirmed in *Arabidopsis* TIP2;1 (*Loque et al., 2005*), but EsTIP3;1 possessed a novel NH_3-type SDPs (T-L-G-T-A-S-H-P-A) with F/T replaced by G in SDP3. The NIP subfamily has low intrinsic water permeability and the ability to transport solutes like glycerol and ammonia (*Choi et al., 2007*). Most NIPs held two typical NPA motifs, but some varied at the third residue in the first or second NPA motif. All NIPs had urea-type SDPs, EsNIP1;2, EsNIP3;1 and EsNIP5;1 had H_2O_2-type SDPs. EsNIP5;1, EsNIP6;1 and EsNIP7;1had boric acid-type SDPs, which have been found in *Arabidopsis* (*Takano et al., 2006*). EsNIP1;2 possessed a novel NH_3-type SDPs with a substitution of G for A at SDP4. In addition, EsNIP4;1 and EsNIP4;3, which both had the substitution of T for K/L/N/V at SDP2. EsSIPs varied in the third residue of the first NPA motif,

with diverse ar/R filters and Froger's positions. However, the residues were consistent with the corresponding SIP in *Arabidopsis*. AtSIP1;1 and AtSIP1;2 could transport water in the ER. AtSIP2;1 might act as an ER channel for other small molecules or ions (*Ishikawa et al., 2005*), and their similarity in these motifs suggests that these EsSIPs may have similar functions. These results indicate that the diversity of AQPs in *E. salsugineum* may have crucial roles in response to environmental stress.

Plant *AQP* genes exhibit various expression patterns in different organs or under different stress conditions. The studies on *AQP* gene expression in different cells, tissues and organs exposed to different environmental conditions provided the first evidence on the biological function of AQPs in plants (*Kapilan et al., 2018*). PIPs and TIPs are highly abundant in all organs in many plant species (*Quigley et al., 2001*; *Venkatesh et al., 2013*; *Reuscher et al., 2013*). Our study showed the transcripts of *EsAQP* genes could be detected in all organs, and the most abundant transcripts were *EsPIPs* and a few *EsTIPs* (*EsTIP1;1* and *EsTIP1;2*; Fig. 5A). Regarding different organs, most *EsAQP* genes were highly expressed in root, implying their crucial roles in transporting of water and nutrient. We also found that the majority of *EsAQP* genes were highly expressed in flower and silique. The morphology of flowers in *Hydrangea macrophylla* owe much to AQPs (*Negishi et al., 2012*), and the loss of function of *NIP5;1* delayed flowering and also affected silique development under boron limitation in *Arabidopsis* (*Takano et al., 2006*). This implies that EsAQPs are involved in growth and development, but the underlying mechanisms need to be further investigated.

Environmental stress factors such as salt, drought and low temperature can quickly reduce water transport rates (*Javot and Maurel, 2003*), thus the maintenance of osmotic potential is a major challenge for plants. Because the leaf status represents a major marker for testing plant water transport potential (*Maurel et al., 2015*), we investigated the expression levels of *EsAQPs* in leaf under salt, drought and low temperature stress. Most *Arabidopsis PIPs* are downregulated in response to drought stress (*Surbanovski et al., 2013*), and the expression of most *PeTIPs* is downregulated under drought stress and upregulated under salt stress (*Sun et al., 2016*). In this study, the results were consistent with previous reports showing that most *EsAQP* genes were induced by salinity in contrast to drought condition (Fig. 5B), suggesting their potential roles in maintaining water balance under environmental stress. However, *EsTIP3;2* was significantly upregulated under drought stress, suggesting that *EsTIP3;2* may play a unique role in drought stress response. In rice, cold stress could induce the expression of *OsPIP2;5* and causes the enhancement of root hydraulic conductivity (Lpr) (*Ahamed et al., 2012*). Our study showed that most of the *EsAQP* genes were upregulated after 4°C treatment for 24 h in leaf, particularly in *EsPIP2;5*, *EsPIP2;6* and *EsTIP2;3*. The varied expression patterns of *EsAQP* genes (and even subfamilies) indicate that their roles in maintaining water homeostasis response to abiotic stress may be different although they shared a higher structural similarity.

Conclusions

In our study, a genome-wide information of *E. salsugineum* AQP gene family was provided. 35 EsAQPs were identified and divided into four subfamilies based on sequence similarity and

phylogenetic relationships according to their homologs in *Arabidopsis*. Furthermore, their structural and functional properties were investigated through the analysis of gene structures, chromosome distributions, ar/R filters, Froger's positions and SDPs, which all have potential outputs on the function of EsPIPs in water balance. Moreover, the expression analysis was performed by qRT-PCR, showing EsAQP genes could be detected in all organs and also when the plants subjected to abiotic stress. This study will provide important information for further analysis of *E. salsugineum* AQPs in abiotic stress response.

Acknowledgements

The authors appreciate those contributors who make the *Eutrema salsugineum* genome data accessible in public databases.

References

Ahamed A, Murai-Hatano M, Ishikawa-Sakurai J, Hayashi H, Kawamura Y, Uemura M. 2012. Cold stress-induced acclimation in rice is mediated by root-specific aquaporins. *Plant Cell Physiology* **53:**1445-1456 DOI 10.1093/pcp /pcs089

Anderberg HI, Kjellbom P, Johanson U. 2012. Annotation of *Selaginella moellendorffii* major intrinsic proteins and the evolution of the protein family in terrestrial plants. *Frontiers in Plant Science* **3:**33 DOI 10.3389/fpls.2012.00033

Ariani A, Gepts P. 2015. Genome-wide identification and characterization of aquaporin gene family in common bean (*Phaseolus vulgaris*, L.). *Molecular Genetics & Genomics* **290:**1771-1785 DOI 10.1007/s00438-015-1038-2

Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kröger N, Lau WW, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS. 2004. The Genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* **306:**79–86 DOI 10.1126/science.1101156

Bansal A, Sankararamakrishnan R. 2007. Homology modeling of major intrinsic proteins in rice, maize and arabidopsis: comparative analysis of transmembrane helix association and aromatic/arginine selectivity filters. *BMC Structural Biology* **7:**27-27 DOI 10.1186/1472-6807-7-27

Biela A, Grote K, Otto B, Hoth S, Hedrich R, Kaldenhoff R. 1999. The *Nicotiana tabacum* plasma membrane aquaporin NtAQP1 is mercury-insensitive and permeable for glycerol. *Plant Journal* **18:**565–570 DOI 10.1046/j.1365-313X.1999.00474.x

Bienert GP, Heinen RB, Berny MC, Chaumont F. 2014. Maize plasma membrane aquaporin ZmPIP2;5, but not ZmPIP1;2, facilitates transmembrane diffusion of hydrogen peroxide. *Biochimica et Biophysica Acta* **1838:**216–222 DOI 10.1016/j.bbamem.2013.08.011

Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, Fournet J, Haruta M, Huysman MJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert Y, Kaplan A, Kröger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-Jézéquel V, Lopez PJ, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A, Oudot-Le Secq MP, Napoli C, Obornik M, Parker MS, Petit JL, Porcel BM, Poulsen N, Robison M, Rychlewski L, Rynearson TA, Schmutz J, Shapiro H, Siaut M, Stanley M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van de Peer Y, Grigoriev IV. 2008. The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature* 456: 239–244 DOI 10.1038/nature07410

Choi WG, Roberts DM. 2007. *Arabidopsis* NIP2;1, a major intrinsic protein transporter of lactic acid induced by anoxic stress. *Journal of Biological Chemistry* **282:**24209-24218 DOI 10.1074/jbc.M700982200

Danielson JÅ, Johanson U. 2008. Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. *BMC Plant Biology* **8:**45 DOI 10.1186/1471-2229-8-45

Deokar AA, Tar'an B. 2016. Genome-wide analysis of the aquaporin gene family in Chickpea (*Cicer arietinum* L.). *Frontiers in Plant Science* **7:**1802 DOI 10.3389/fpls.2016.01802

Dev TB, Herbert JK. 2018. From aquaporin to ecosystem: plants in the water cycle. *Journal of Plant Physiology* **227:**1-2 DOI 10.1016/j.jplph.2018.06.008

Dynowski M, Schaaf G, Loque D, Moran O, Ludewig U. 2008. Plant plasma membrane water channels conduct the signaling molecule H_2O_2. *Biochemical Journal* **414:**53–61 DOI 10.1042/BJ20080287

Feng ZJ, Xu SC, Liu N, Zhang GW, Hu QZ, Xu ZS, Gong YM. 2017. Identification of the AQP members involved in abiotic stress responses from *Arabidopsis*. *Gene* **646:**64–73 DOI 10.1016/j.gene.2017.12.048

Fortin MG, Morrison NA, Verma DPS. 1987. Nodulin-26, a peribacteroid membrane nodulin is expressed independently of the development of the peribacteroid compartment. *Nucleic Acids Research* **15:**813-824 DOI 10.1093/nar/15.2.813

Froger A, Tallur B, Thomas D, Delamarche C. 1998. Prediction of functional residues in water channels and related proteins. *Protein Science* **7:**1458-1468 DOI 10.1002/pro.5560070623

Gaspar M, Bousser A, Sissoëff I, Roche O, Hoarau J, Mahé A. 2003. Cloning and characterization of ZmPIP1-5b, an aquaporin transporting water and urea. *Plant Science* **165:**21–31 DOI 10.1016/j.jinsphys.2013.08.013

Gerbeau P, Guclu J, Ripoche P, Maurel C. 1999. Aquaporin Nt-TIPa can account for the high permeability of tobacco cell vacuolar membrane to small neutral solutes. *Plant Journal* **18:**577–587 DOI 10.1046/j.1365-313x.1999.00481.x

Gomes D, Agasse A, Thiébaud P, Delrot S, Gerós H, Chaumont F. 2009. Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. *Biochimica et Biophysica Acta* **1788:**1213-1228 DOI 10.1016/j.bbamem.2009.03.009

Gustavsson S, Lebrun AS, Kristina N, François C, Johanson U. 2005. A novel plant major intrinsic protein in *Physcomitrella patens* most similar to bacterial glycerol channels. *Plant Physiology* **139:**287-295 DOI 10.1104/pp.105.063198

Heckwolf M, Pater D, Hanson DT, Kaldenhoff R. 2011. The *Arabidopsis thaliana* aquaporin AtPIP1;2 is a physiologically relevant CO_2 transport facilitator. *Plant Journal* **67:**795-804 DOI 10.1111/j. 1365-313X.2011.04634.x

Hove RM, Bhave M. 2011. Plant aquaporins with non-aqua functions: deciphering the signature sequences. *Plant Molecular Biology* **75:**413-430 DOI 10.1007/s11103-011-9737-5

Inan G, Zhang Q, Li P, Wang Z, Cao Z, Zhang H, Zhang C, Quist TM, Goodwin SM, Zhu J, Shi H, Damsz B, Charbaji T, Gong Q, Ma S, Fredricksen M, Galbraith DW, Jenks MA, Rhodes D, Hasegawa PM, Bohnert HJ, Joly RJ, Bressan RA, Zhu JK. 2004. Salt Cress. A halophyte and cryophyte *Arabidopsis* relative model system and its applicability to molecular genetic analyses of growth and development of extremophiles. *Plant Physiology* **135:**1718-37 DOI 10.1104/pp.104.041723

Ishikawa F, Suga S, Uemura T, Sato MH, Maeshima M. 2005. Novel type aquaporin SIPs are mainly localized to the er membrane and show cell-specific expression in *Arabidopsis thaliana*. *FEBS Letters* **579:**5814-5820 DOI 10.1016/j.febslet.2005.09.076

Javot H, Maurel C. 2002. The role of aquaporins in root water uptake. *Annals of Botany* **90:**301-313 DOI 10.1093/aob/mcf199

Johanson U, Karlsson M, Johansson I, Gustavsson S, Sjövall S, Fraysse L, Weig AR, Kjellbom P. 2001. 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 Physiology* **126:**1358-1369 DOI 10.1104/pp.126.4.1358

Kadam S, Abril A, Dhanapal AP, Koester RP, Vermerris W, Jose. S, Fritschi FB. 2017. Characterization and regulation of aquaporin genes of sorghum [*Sorghum bicolor* (L.) Moench] in response to waterlogging stress. *Frontiers in Plant Science* **8:**862 DOI 10.3389/fpls.2017.00862

Koch MA, German DA. 2013. Taxonomy and systematics are key to biological information: *Arabidopsis*, *Eutrema* (*Thellungiella*), *Noccaea* and *Schrenkiella* (Brassicaceae) as examples. *Frontiers in Plant Science* **4:**267 DOI 10.3389/fpls.2013.00267

Kong W, Yang S, Wang Y, Bendahmane M, Fu X. 2017. Genome-wide identification and characterization of aquaporin gene family in *Beta vulgaris*. *Peer J* **5:**e3747 DOI 10.7717/peerj.3747

Li W, Qiang XJ, Han XR, Jiang LL, Zhang SH, Han J, He R, Cheng XG. 2018. Ectopic expression of a *Thellungiella salsuginea* aquaporin gene, TsPIP1;1, increased the salt tolerance of Rice. *International Journal of Molecular Science* **19:**2229 DOI 10.3390/ijms19082229

Loque D, Ludewig U, Yuan LX, von Wiren N. 2005. Tonoplast intrinsic proteins AtTIP2;1 and AtTIP2;3 facilitate NH_3 transport into the vacuole. *Plant Physiology* **137:**671–680 DOI 10.1104/pp.104.051268

Ma JF, Tamai K, Yamaji N, Mitani N, Konishi S, Katsuhara M, Ishiguro M, Murata Y, Yano M. 2006. A silicon transporter in rice. *Nature* **440:**688–691 DOI 10.1038/nature04590

Maurel C, Boursiac Y, Luu DT, Santoni V, Shahzad Z, Verdoucq L. 2015. Aquaporins in plants. *Physiological Reviews* **95:**1321-1358 DOI 10.1152/physrev.00008.2015

Park W, Scheffler BE, Bauer PJ, Campbell BT. 2010. Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.). *BMC Plant Biology* **10:**142 DOI 10.1186/1471-2229-10-142

Pilarska M, Wiciarz M, Jajić I, Kozieradzka-Kiszkurno M, Dobrev P, Vanková R, Niewiadomska E. 2016. A different pattern of production and scavenging of reactive oxygen species in Halophytic *Eutrema salsugineum* (*Thellungiella salsuginea*) plants in comparison to *Arabidopsis thaliana* and its relation to salt stress signaling. *Frontiers in Plant Science* **7:**1179 DOI 10.3389/fpls.2016.01179

Prerostova S, Dobrev PI, Gaudinova A, Hosek P, Soudek P, Knirsch V, Vankova R. 2017. Hormonal dynamics during salt stress responses of salt-sensitive, *Arabidopsis thaliana*, and salt-tolerant, *Thellungiella salsuginea*. *Plant Science* **264:**188-198. DOI 10.1016/j.plantsci.2017.07.020

Putpeerawit P, Sojikul P, Thitamadee S, Narangaiavana J. 2017. Genome-wide analysis of aquaporin gene family and their responses to water-deficit stress conditions in cassava. *Plant Physiology and Biochemistry* **121:**118-127 DOI 10.1016/j. plaphy.2017.10.025

Quigley F, Rosenberg JM, Shacharhill Y, Bohnert HJ. 2001. From genome to function: the *Arabidopsis* aquaporins. *Genome Biology* **3:**1-17 DOI 10.1186/gb-2001-3-1-research0001

Reddy PS, Rao TSRB, Sharma KK, Vadez V. 2015. Genome-wide identification and characterization of the aquaporin gene family in *Sorghum bicolor* (L.). *Plant Gene* **1:**18-28 DOI 10.1016/j.plgene.2014.12.002

Reuscher S, Akiyama M, Mori C, Aoki K, Shibata D, Shiratake K. 2013. Genome-wide identification and expression analysis of aquaporins in tomato. *PLOS ONE* **8:**e79052 DOI 10.1371/journal.pone.0079052

Sakurai J, Ishikawa F, Yamaguchi T, Uemura M, Maeshima M. 2005. Identification of 33 rice aquaporin genes and analysis of their expression and function. *Plant Cell Physiology* **46:**1568-1577 DOI 10.1093/pcp/pci172

Sun HY, Li LC, Lou YF, Zhao HS, Gao ZM. 2016. Genome-wide identification and characterization of aquaporin gene family in moso bamboo (*Phyllostachys edulis*). *Molecular Biology Report* **43:**437-450 DOI 10.1007/s11033-016-3973-3

Sun LL, Yu GH, Han XR, Xin SC, Qiang XJ, Jiang LL, Zhang SH, Cheng XG. 2015. TsMIP6 enhances the tolerance of transgenic rice to salt stress and interacts with target proteins. *Journal of Plant Biology* **58:**285-292 DOI 10.1007/s12374-015-0069-x

Surbanovski N, Sargent DJ, Else MA, Simpson DW, Zhang H, Grant OM. 2013. Expression of fragaria vesca PIP aquaporins in response to drought stress: PIP down-regulation correlates with the decline in substrate moisture content. *PLOS ONE* **8:**e74945 DOI 10.1371/journal.pone.0074945

Tabata S, Kaneko T, Nakamura Y, Kotani H, Kato T, Asamizu E, Miyajima N, Sasamoto S, Kimura T, Hosouchi T, Kawashima K, Kohara M, Matsumoto M, Matsuno A, Muraki A, Nakayama S, Nakazaki N, Naruo K, Okumura S, Shinpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Sato S, de la Bastide M, Huang E, Spiegel L, Gnoj L, O'Shaughnessy A, Preston R, Habermann K, Murray J, Johnson D, Rohlfing T, Nelson J, Stoneking T, Pepin K, Spieth J, Sekhon M, Armstrong J, Becker M, Belter E, Cordum H, Cordes M, Courtney L, Courtney W, Dante M, Du H, Edwards J, Fryman J, Haakensen B, Lamar E, Latreille P, Leonard S, Meyer R, Mulvaney E, Ozersky P, Riley A, Strowmatt C, Wagner-McPherson C, Wollam A, Yoakum M, Bell M, Dedhia N, Parnell L, Shah R, Rodriguez M, See LH, Vil D, Baker J, Kirchoff K, Toth K, King L, Bahret A, Miller B, Marra M, Martienssen R, McCombie WR, Wilson RK, Murphy G, Bancroft I, Volckaert G, Wambutt R, Düsterhöft A, Stiekema W, Pohl T, Entian KD, Terryn N, Hartley N, Bent E, Johnson S, Langham SA, McCullagh B, Robben J, Grymonprez B, Zimmermann W, Ramsperger U, Wedler H, Balke K, Wedler E, Peters S, van Staveren M, Dirkse W, Mooijman P, Lankhorst RK, Weitzenegger T, Bothe G, Rose M, Hauf J, Berneiser S, Hempel S, Feldpausch M, Lamberth S, Villarroel R, Gielen J, Ardiles W, Bents O, Lemcke K, Kolesov G, Mayer K, Rudd S, Schoof H, Schueller C, Zaccaria P, Mewes HW, Bevan M, Fransz P. 2000. Sequence and analysis of chromosome 5 of the plant *Arabidopsis thaliana*. *Nature* **408:**823-826. DOI 10.1038/35048507

Tajkhorshid E, Nollert P, Jensen MØ, Miercke LJ, O'Connell J, Stroud RM, Schulten K. 2002. Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science* **296:**525-530 DOI 10.1126/science.1067778

Takano J, Wada M, Ludewig U, Schaaf G, von Wiren N, Fujiwara T. 2006. The *Arabidopsis* major intrinsic protein NIP5;1 is essential for efficient boron uptake and plant development under boron limitation. *Plant Cell* **18:**1498–1509 DOI 10.1105/tpc.106.041640

Negishi T, Oshima K, Hattori M, Kanai M, Mano S, Nishimura M, Yoshida K. 2012. Tonoplast- and plasma membrane-localized aquaporin-family transporters in blue hydrangea sepals of aluminum hyperaccumulating plant. *PLOS ONE* **7:**e43189 DOI 10.1371/journal.pone.0043189

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 30:2725–2729 DOI 10.1093/molbev/mst197

Tao P, Zhong X, Li B, Wang W, Yue Z, Lei J, Guo W, Huang X. 2014. Genome-wide identification and characterization of aquaporin genes (AQPs) in Chinese cabbage (*Brassica rapa* ssp. pekinensis). *Molecular Genetics and Genomics* **289:**1131-1145 DOI 10.1007/s00438-014-0874-9

Uehlein N, Lovisolo C, Siefritz F, Kaldenhoff R. 2003. The tobacco aquaporin NtAQP1 is a membrane CO_2 pore with physiological functions. *Nature* **425:**734–737 DOI 10.1038/nature02027

Venkatesh J, Yu JW, Park SW. 2013. Genome-wide analysis and expression profiling of the *Solanum tuberosum* aquaporins. *Plant Physiology & Biochemistry* **73:**392-404 DOI 10.1016/j.plaphy.2013.10.025

Wang LL, Chen AP, Zhong NQ, Liu N, Wu XM, Wang F, Yang CL, Romero MF, Xia GX. 2014. The *Thellungiella salsuginea* tonoplast aquaporin TsTIP1;2 functions in protection against multiple abiotic stresses. *Plant & Cell Physiology* **55:**148-161 DOI 10.1093/pcp/pct166

Wu HJ, Zhang Z , Wang JY , Oh DH, Dassanayake M, Liu B, Huang Q, Sun HX, Xia R, Wu Y, Wang YN, Yang Z, Liu Y, Zhang W, Zhang H, Chu J, Yan C, Fang S, Zhang J, Wang Y, Zhang F, Wang G, Lee SY, Cheeseman JM, Yang B, Li B, Min J, Yang L, Wang J, Chu C, Chen SY, Bohnert HJ, Zhu JK , Wang XJ, Xie Q. 2012. Insights into salt tolerance from the genome of *Thellungiella salsuginea*. *Proceedings of the National Academy of Sciences* **109:**12219-12224 DOI 10.1073/pnas.1209954109

Xu GX, Guo CC, Shan HY, Kong HZ. 2012. Divergence of duplicate genes in exon-intron structure. *Proceedings of the National Academy of Sciences* **109:**1187-1192 DOI 10.1073/pnas.1109047109

Yang R, Jarvis DE, Chen H, Beilstein MA, Grimwood J, Jenkins J, Shu S, Prochnik S, Xin M, Ma C, Schmutz J, Wing RA, Mitchell-Olds T, Schumaker KS, Wang X. 2013. The Reference Genome of the Halophytic Plant *Eutrema salsugineum*. *Frontiers of Plant Science* **4:**46 DOI 10.3389/fpls.2013.00046

Yuan D, Li W, Hua YP, King GJ, Xu FS, Shi L. 2017. Genome-wide identification and characterization of the aquaporin gene family and transcriptional responses to boron deficiency in *Brassica napus*. *Frontiers of Plant Science* **8:**1336 DOI 10.3389/fpls.2017.01336

Zhang DY, Ali Z, Wang CB, Xu L, Yi JX, Xu ZL, Liu XQ, He XL, Huang YH, Khan IA, Trethowan RM, Ma HX. 2013. Genome-wide sequence characterization and expression analysis of major intrinsic proteins in soybean (*Glycine max* L.). *PLOS ONE* **8:**e56312 DOI 10.1371/journal.pone.0056312

Zhu JK. 2001. Plant salt tolerance. *Trends in Plant Science* **6:**66-71 DOI 10.1016/S1360-1385(00)01838-0

Zou Z, Gong J, Huang QX, Mo YY, Yang LF, Xie GS. 2015. Gene structures, evolution, classification and expression profiles of the aquaporin gene family in castor bean (*Ricinus communis* L.). *PLOS ONE* **10:**e0141022 DOI 10.1371/journal.pone.0141022

Zou Z, Yang L, Gong J, Mo Y, Wang J, Cao J, An F, Xie G. 2016. Genome-wide identification of *Jatropha curcas* aquaporin genes and the comparative analysis provides insights

628 into the gene family expansion and evolution in *Hevea brasiliensis*. *Frontiers of Plant Science*
629 **7:**395 DOI 10.3389/fpls.2016.00395

Figure 1

Phylogenetic tree of AQP amino acid sequences from *Eutrema salsugineum* and *Arabidopsis thaliana*.

Alignments were performed using the default parameter of ClustalW and the phylogenetic tree was constructed using Neighbor-Joining tree method with 1000 bootstrap replicates in MEGA6.0 software. Each subfamily of AQPs was well separated in different clades and represented by different colors. The solid circle represents EsAQPs and the hollow circle represents AtAQPs.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27563v1 | CC BY 4.0 Open Access | rec: 3 Mar 2019, publ: 3 Mar 2019

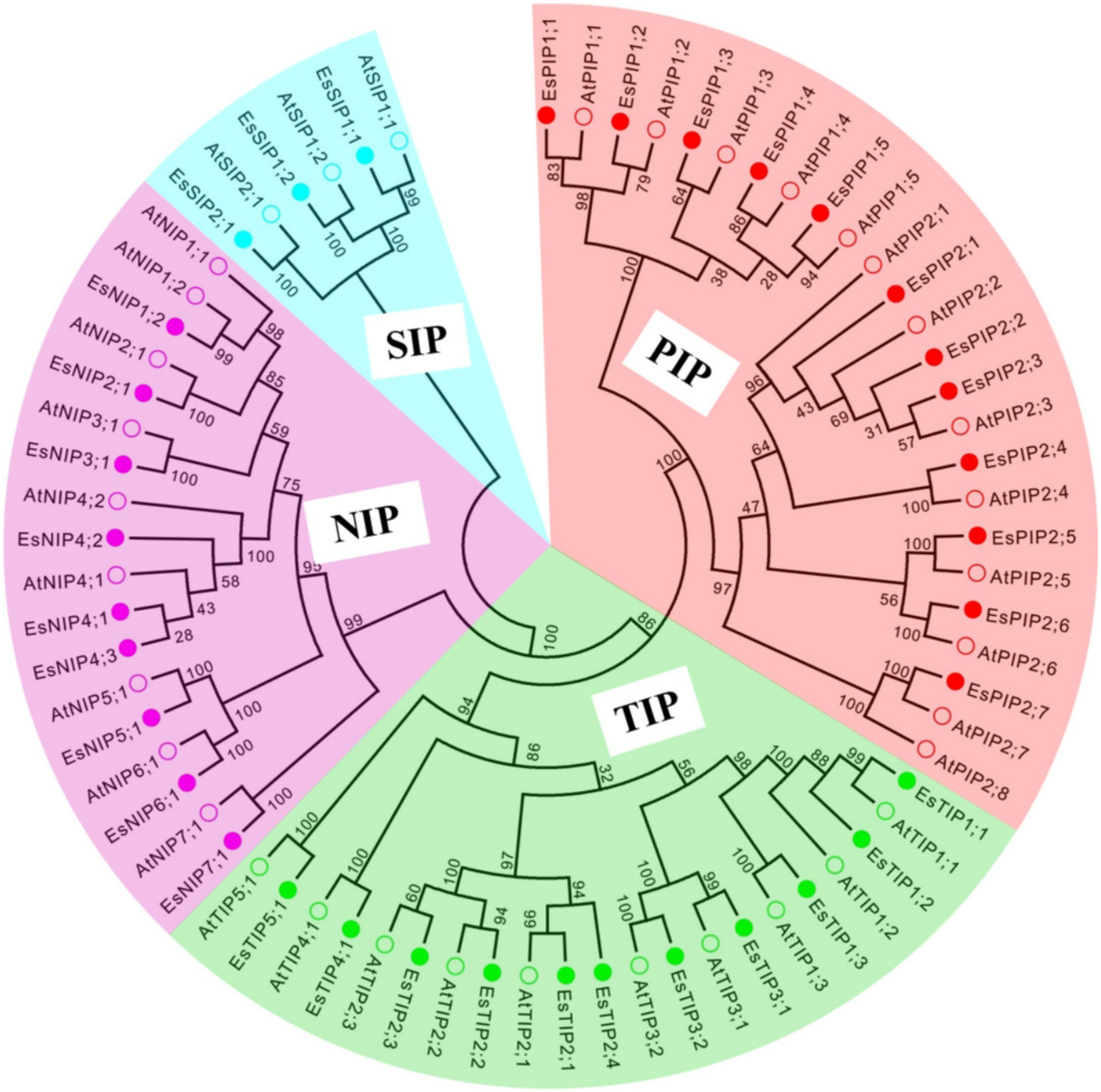

Figure 2

Chromosomal localization of the EsAQP genes.

The diagram was drawn using the MapInspect software, and 34 of 35 EsAQPs were located on 7 chromosomes (except *EsTIP2;2*).

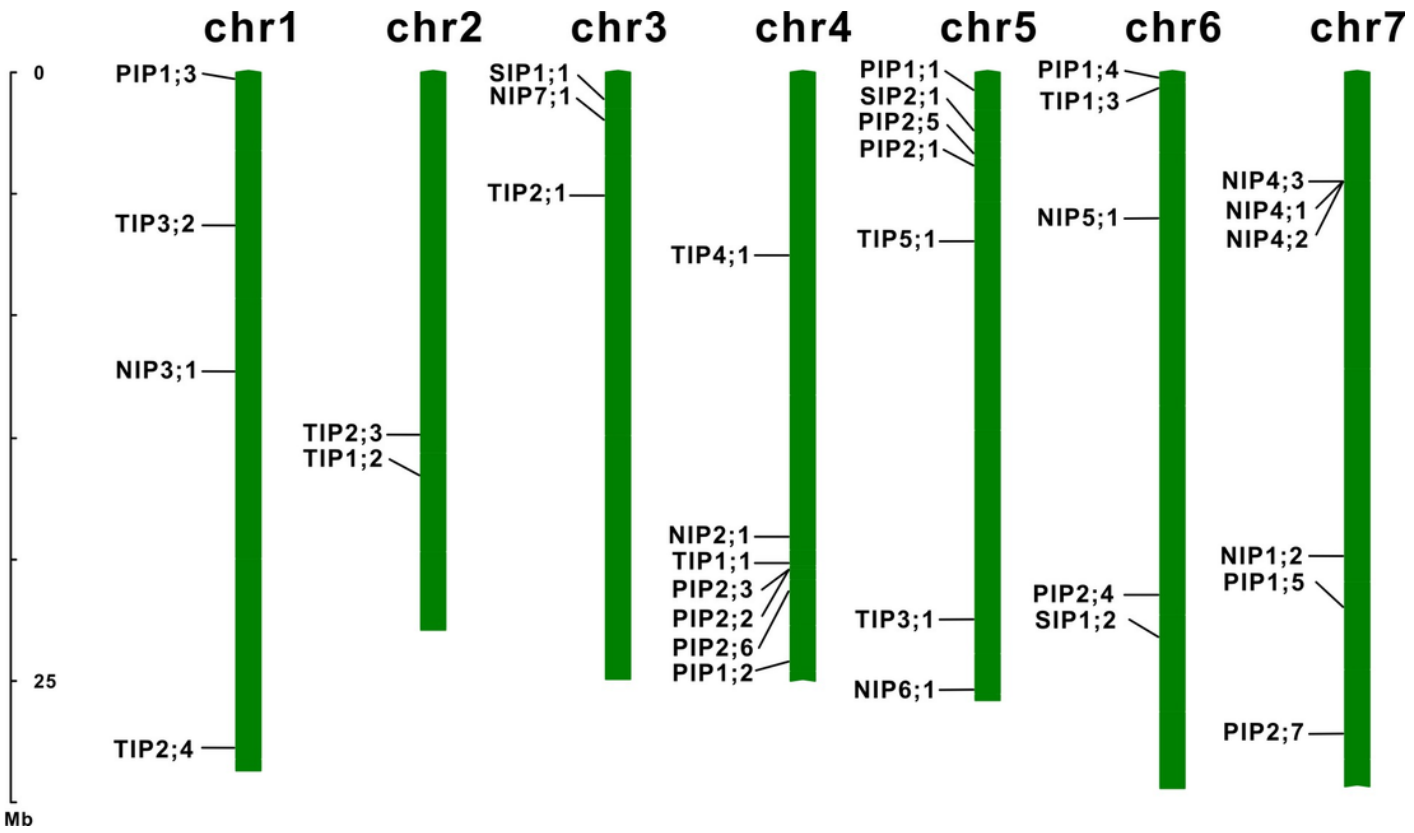

Figure 3

Gene structures of the 35 EsAQP genes.

The blue rectangle, yellow rectangle and black line represent UTR, exon and intron, respectively.

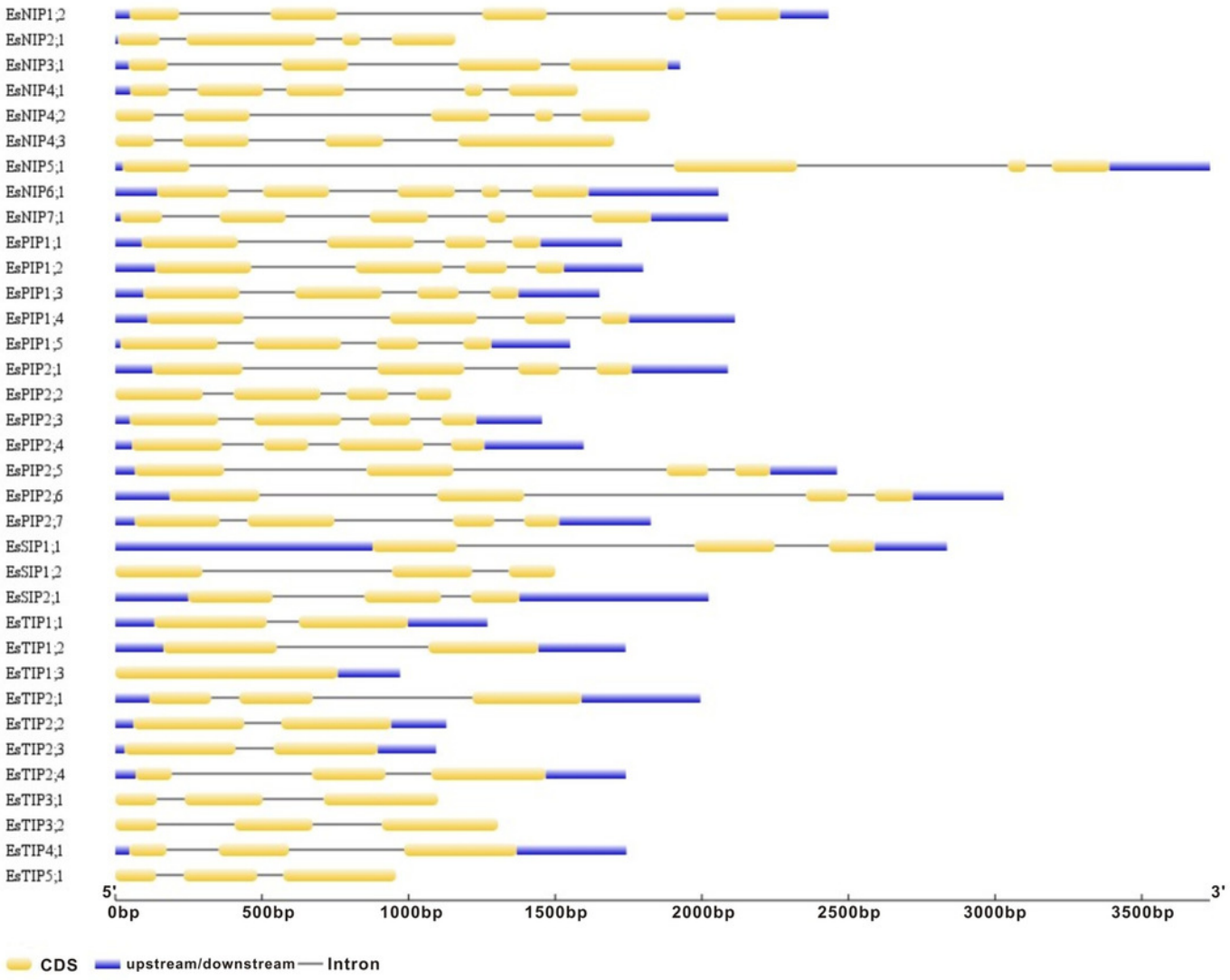

Figure 4

Conversed motif analysis in EsAQPs.

The conversed motif prediction was identified using MEME motif search analysis, and the maximum number parameter was set to 10. Different motifs were represented by different colors. (A) Conversed motifs of 35 EsAQP proteins correspond to *p*-values. (B) Motif consensus sequences.

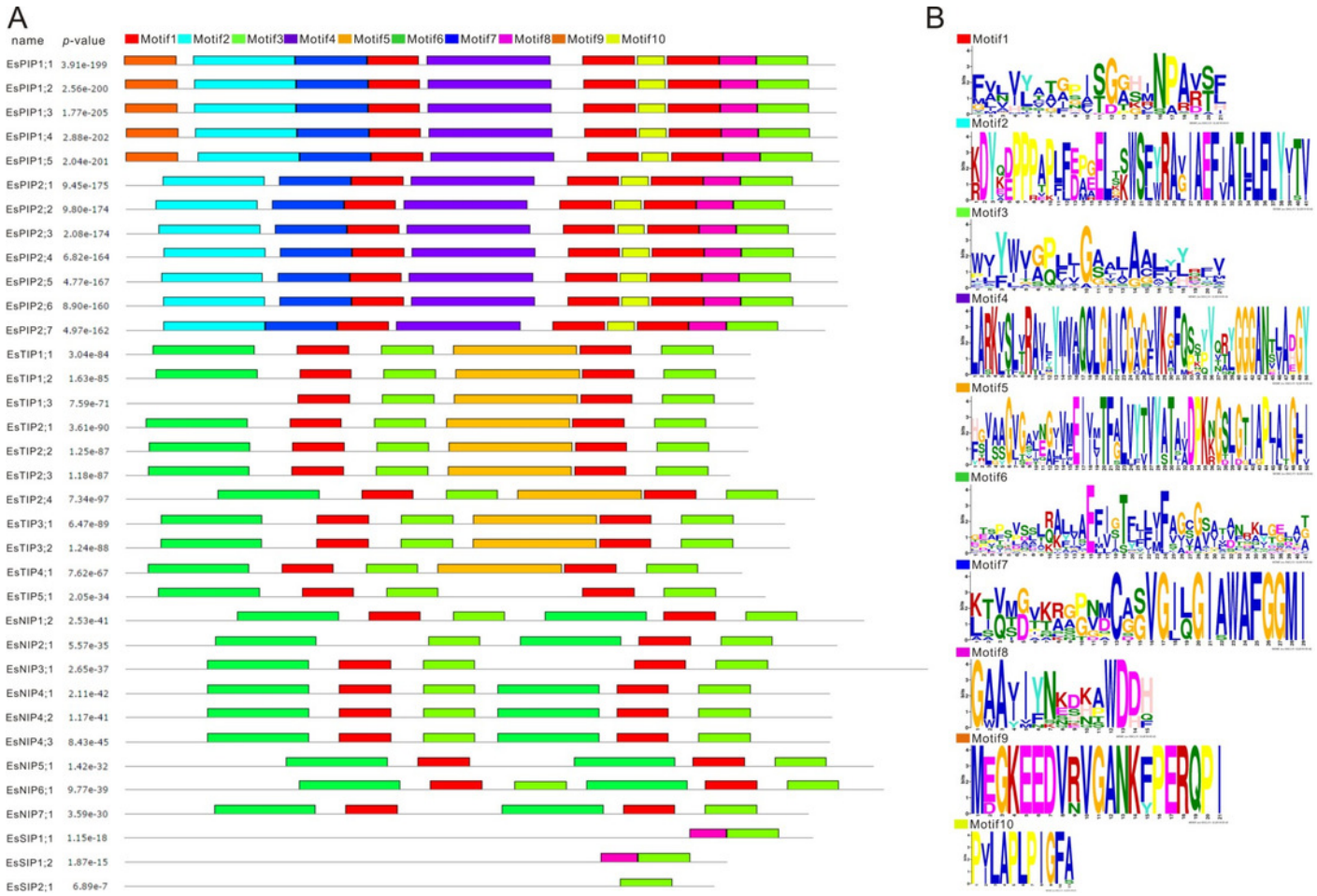

Figure 5

Expression profiles of the EsAQP genes.

(A) EsAQPs expression in response to abiotic stress. The color scale represents the $2^{-\Delta\Delta Ct}$ value normalized to untreated controls and $\log_2$ transformed counts, where green indicates downregulated expression and red indicates upregulated expression. (B) Expression of EsAQPs in various organs of *E. salsugineum*. Color scales represent $2^{\Delta Ct}$ values normalized to actin and $\log_2$ transformed counts, where green indicates low expression and red indicates high expression.

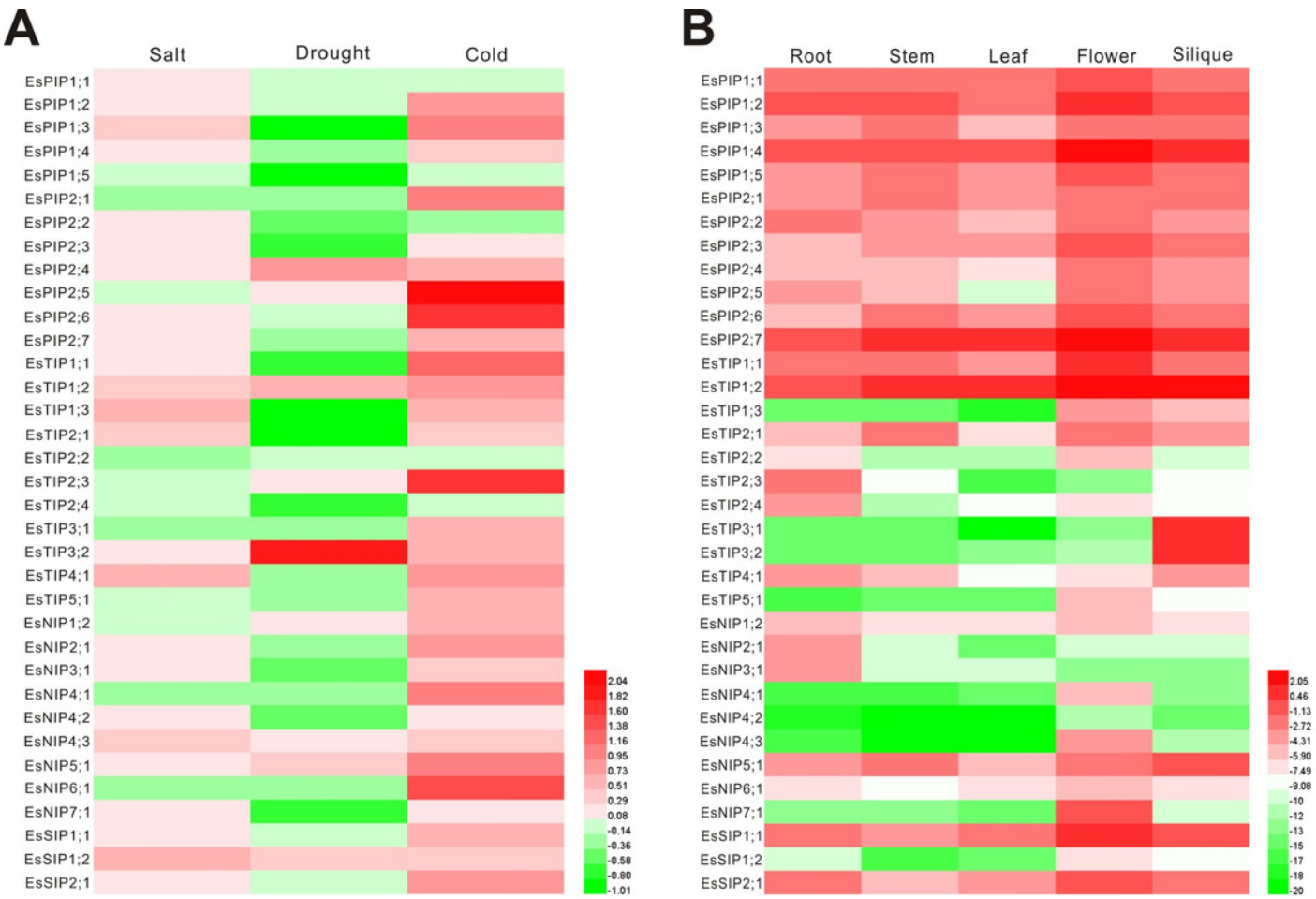

Table 1(on next page)

List of identified EsAQP genes in *Eutrema salaugineum* along with subcellular localization.

1**TABLE1** List of identified EsAQP genes in *Eutrema salaugineum* along with subcellular localization.

	Name	Chromosomal Localization	Scaffold	CDS[a]	Protein ID	Plant-mPLoc[b]	WoLF PSORT[c]
PIPs	EsPIP1;1	Chr5;748,014~746,287	NW_006256838.1	XM_006402419.1	XP_006402482.1	PM	PM
	EsPIP1;2	Chr4;24,198,933~24,200,732	NW_006256812.1	XM_006397718.1	XP_006397781.1	PM	PM
	EsPIP1;3	Chr1;227,418~229,068	NW_006256612.1	XM_006418376.1	XP_006418439.1	PM	PM
	EsPIP1;4	Chr6;182,520~180,408	NW_006256756.1	XM_006396178.1	XP_006396241.1	PM	PM
	EsPIP1;5	Chr7;21,955,256~21,956,964	NW_006256909.1	XM_006413496.1	XP_006413559.1	PM	PM
	EsPIP2;1	Chr5;3,815,044~3,817,131	NW_006256858.1	XM_006403628.1	XP_006403691.1	PM	PM
	EsPIP2;2	Chr4;20,408,518~20,407,373	NW_006256908.1	XM_006410833.1	XP_006410896.1	PM	PM
	EsPIP2;3	Chr4;20,411,864~20,413,318	NW_006256908.1	XM_006410834.1	XP_006410897.1	PM	PM
	EsPIP2;4	Chr6;21,418,342~21,416,629	NW_006256829.1	XM_006400761.1	XP_006400824.1	PM	PM
	EsPIP2;5	Chr5;3,318,416~3,315,956	NW_006256858.1	XM_006403468.1	XP_006403531.1	PM	PM
	EsPIP2;6	Chr4;21,319,556~21,322,584	NW_006256908.1	XM_006411061.1	XP_006411124.1	PM	PM
	EsPIP2;7	Chr7;27,180,960~27,182,785	NW_006256909.1	XM_006412089.1	XP_006412152.1	PM	PM
TIPs	EsTIP1;1	Chr4;20,182,942~20,184,210	NW_006256908.1	XM_006410791.1	XP_006410854.1	V	PM
	EsTIP1;2	Chr2;16,508,526~16,506,789	NW_006256547.1	XM_006395487.1	XP_006395549.1	V	PM
	EsTIP1;3	Chr6;663,103~662,130	NW_006256756.1	XM_006396285.1	XP_006396348.1	V	PM
	EsTIP2;1	Chr3;5,624,419~5,626,413	NW_006256885.1	XM_006406794.1	XP_006406857.1	V	PM
	EsTIP2;2	ND	NW_006256909.1	XM_006414179.1	XP_006414242.1	V	V
	EsTIP2;3	Chr2;14,894,399~14,893,306	NW_006256828.1	XM_006398375.1	XP_006398438.1	V	V
	EsTIP2;4	Chr1;27,709,976~27,708,236	NW_006256486.1	XM_006392888.1	XP_006392950.1	V	V
	EsTIP3;1	Chr5;22,490,388~22,491,488	NW_006256342.1	XM_006390520.1	XP_006390582.1	V	PM
	EsTIP3;2	Chr1;6,309,744~6,311,048	NW_006256612.1	XM_006416602.1	XP_006416665.1	V	PM
	EsTIP4;1	Chr4;7,484,947~7,486,691	NW_006256895.1	XM_006408738.1	XP_006408801.1	V	PM
	EsTIP5;1	Chr5;6,934,814~6,933,858	NW_006256858.1	XM_006404316.1	XP_006404379.1	V/PM	Chl
NIPs	EsNIP1;2	Chr7;19,890,089~19,892,520	NW_006256909.1	XM_006413978.1	XP_006414041.1	PM	PM
	EsNIP2;1	Chr4;19,043,681~19,042,522	NW_006256908.1	XM_006410521.1	XP_006410584.1	PM	V
	EsNIP3;1	Chr1;12,292,410~12,294,335	NW_006256612.1	XM_006415218.1	XP_006415281.1	PM	V
	EsNIP4;1	Chr7;4,484,562~4,482,986	NW_006256877.1	XM_006405767.1	XP_006405830.1	PM	PM
	EsNIP4;2	Chr7;4,513,301~4,511,485	NW_006256877.1	XM_006405768.1	XP_006405831.1	PM	PM
	EsNIP4;3	Chr7;4,481,446~4,479,745	NW_006256877.1	XM_006405766.1	XP_006405829.1	PM	PM
	EsNIP5;1	Chr6;6,005,178~6,008,910	NW_006256756.1	XM_006397006.1	XP_006397069.1	PM	PM
	EsNIP6;1	Chr5;25,383,958~25,386,014	NW_006256342.1	XM_006389768.1	XP_006389830.1	PM	PM
	EsNIP7;1	Chr3;1,929,290~1,927,201	NW_006256885.1	XM_006407920.1	XP_006407983.1	PM	PM
SIPs	EsSIP1;1	Chr3;1,105,251~1,102,416	NW_006256885.1	XM_024159977.1	XP_024015745.1	PM	PM
	EsSIP1;2	Chr6;23,161,081~23,162,581	NW_006256829.1	XM_006400314.1	XP_006400377.1	V/PM	V
	EsSIP2;1	Chr5;2,401,441~2,403,463	NW_006256838.1	XM_006402867.1	XP_006402930.1	PM	V

2 [a] *Coding sequence*

3 [b] *Prediction of subcellular localization using Plant-mPLoc: PM, Plasma membrane; V, tonoplast membrane;*

4 *Chl, chloroplast thylakoid membrane.*

[c] *Prediction of subcellular localization using WoLF PSORT*

Table 2(on next page)

Structural characteristics of the EsAQPs.

1**TABLE 2** Structural characteristics of the EsAQPs.

Name	AA	TM	MW (KD)	pI	NPA motif		ar/R selectivity filter				Froger's positions				
					LB	LE	H2	H5	LE1	LE2	P1	P2	P3	P4	P5
PIPs															
EsPIP1;1	286	6	30.77	9.14	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP1;2	286	6	30.60	9.16	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP1;3	286	6	30.62	9.02	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP1;4	286	6	30.56	9.02	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP1;5	287	6	30.61	9.00	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP2;1	287	6	30.48	6.95	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP2;2	284	6	30.21	6.50	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP2;3	285	6	30.31	6.51	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP2;4	285	6	30.12	7.62	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP2;5	286	6	30.57	8.82	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP2;6	290	6	31.11	7.69	NPA	NPA	F	H	T	R	Q	S	A	F	W
EsPIP2;7	281	6	29.82	9.11	NPA	NPA	F	H	T	R	M	S	A	F	W
TIPs															
EsTIP1;1	251	6	25.62	6.03	NPA	NPA	H	I	A	V	T	A	A	Y	W
EsTIP1;2	253	6	25.70	5.32	NPA	NPA	H	I	A	V	T	A	A	Y	W
EsTIP1;3	252	6	25.85	5.10	NPA	NPA	H	I	A	V	T	S	A	Y	W
EsTIP2;1	277	6	28.32	7.80	NPA	NPA	H	I	G	R	T	S	A	Y	W
EsTIP2;2	250	6	25.02	4.87	NPA	NPA	H	I	G	R	T	S	A	Y	W
EsTIP2;3	243	6	24.31	4.73	NPA	NPA	H	I	G	R	T	S	A	Y	W
EsTIP2;4	254	6	25.85	5.43	NPA	NPA	H	I	G	R	T	S	A	Y	W
EsTIP3;1	265	6	27.94	7.17	NPA	NPA	H	T	A	R	T	A	A	Y	W
EsTIP3;2	267	6	28.29	6.58	NPA	NPA	H	M	A	R	T	T	A	Y	W
EsTIP4;1	249	6	26.16	5.49	NPA	NPA	H	I	A	R	T	S	A	Y	W
EsTIP5;1	257	6	26.70	7.72	NPA	NPA	N	V	G	C	V	A	A	Y	W
NIPs															
EsNIP1;2	297	6	31.80	8.83	NPA	NPA	W	V	A	R	F	S	A	Y	L
EsNIP2;1	286	6	30.56	6.78	NPA	NPG	W	V	A	R	F	S	A	Y	L
EsNIP3;1	323	6	34.46	5.94	NPA	NPA	W	I	A	R	F	S	A	Y	L
EsNIP4;1	283	6	30.49	8.73	NPA	NPA	W	V	A	R	F	S	A	Y	L
EsNIP4;2	284	6	30.34	8.80	NPA	NPA	W	V	A	R	F	S	A	Y	L
EsNIP4;3	283	6	30.30	8.98	NPA	NPA	W	V	A	R	F	S	A	Y	L
EsNIP5;1	301	6	31.20	8.31	NPS	NPA	A	I	G	R	F	T	A	Y	L
EsNIP6;1	305	6	31.78	8.57	NPA	NPA	A	I	A	R	F	T	A	Y	L
EsNIP7;1	275	6	28.62	6.12	NPS	NPA	A	V	G	R	Y	S	A	Y	L
SIPs															
EsSIP1;1	238	6	25.41	9.89	NPT	NPA	I	V	P	I	I	A	A	Y	W
EsSIP1;2	242	6	25.96	9.83	NPC	NPA	V	F	P	I	I	A	A	Y	W

EsSIP2;1	237	6	25.85	9.64	NPL	NPA	S	H	G	A	F	V	A	Y	W

Abbreviation: AA ,amino acids length; TM, transmembrane domain; MW, molecular weight; pI, isoelectricpoint, NPA Asn-Pro-Ala motif; ar/R, aromatic/arginine.

Table 3(on next page)

Identified typical SDPs in EsAQPs.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27563v1 | CC BY 4.0 Open Access | rec: 3 Mar 2019, publ: 3 Mar 2019

TABLE 3 Identified typical SDPs in EsAQPs.

Aquaporin	Specificity-determining positions								
	SDP1	**SDP2**	**SDP3**	**SDP4**	**SDP5**	**SDP6**	**SDP7**	**SDP8**	**SDP9**
Ammonia Transporters	**F/T**	**K/L/N/V**	**F/T**	**V/L/T**	**A**	**D/S**	**A/H/L**	**E/P/S**	**A/R/T**
EsTIP2;1	T	L	T	V	A	S	H	P	A
EsTIP3;1	T	L	G	T	A	S	H	P	A
EsNIP1;2	F	K	F	T	G	D	L	E	T
EsNIP4;1	F	T	F	T	A	D	L	E	T
EsNIP4;3	F	T	F	T	A	D	L	E	T
Boric Acid transporter	**T/V**	**I/V**	**H/I**	**P**	**E**	**I/L**	**I/L/T**	**A/T**	**A/G/P/K**
EsPIP1;1	T	I	H	P	E	L	L	T	P
EsPIP1;2	T	I	H	P	E	L	L	T	P
EsPIP1;3	T	I	H	P	E	L	L	T	P
EsPIP1;4	T	I	H	P	E	L	L	T	P
EsPIP1;5	T	I	H	P	E	L	L	T	P
EsPIP2;5	T	I	H	P	E	L	L	T	P
EsNIP5;1	T	I	H	P	E	L	L	A	P
EsNIP6;1	T	I	H	P	E	L	L	A	P
EsNIP7;1	V	I	H	P	E	L	L	T	P
CO_2 transporter	**I/L/V**	**I**	**C**	**A**	**I/V**	**D**	**W**	**D**	**W**
EsPIP1;1	L	I	C	A	I	D	W	D	W
EsPIP1;2	V	I	C	A	I	D	W	D	W
EsPIP1;3	V	M	C	A	I	D	W	D	W
EsPIP1;4	V	M	C	A	I	D	W	D	W
EsPIP1;5	V	I	C	A	I	D	W	D	W
EsPIP2;4	V	I	C	A	V	E	W	D	W
H_2O_2 transporters	**A/S**	**A/G**	**L/V**	**A/F/L/V/T**	**I/L/V**	**H/I/L/Q**	**F/Y**	**A/V**	**P**
EsPIP1;1	A	G	V	F	I	H	F	V	P
EsPIP1;2	A	G	V	F	I	H	F	V	P
EsPIP1;3	A	G	V	F	I	H	F	V	P
EsPIP1;4	A	G	V	F	I	H	F	V	P
EsPIP1;5	A	G	V	F	I	H	F	V	P
EsPIP2;1	A	G	V	F	I	H	F	V	P
EsPIP2;2	A	G	V	F	I	H	F	V	P
EsPIP2;3	A	G	V	F	I	H	F	V	P
EsPIP2;4	A	G	V	F	I	Q	F	V	P
EsPIP2;5	A	G	V	F	I	H	F	V	P
EsPIP2;6	A	G	V	F	I	Q	F	V	P
EsPIP2;7	A	G	V	F	I	H	F	V	P
EsTIP1;1	S	A	L	A	I	H	Y	A	P
EsTIP1;2	S	A	L	A	I	H	Y	A	P

EsTIP1;3	A	A	L	S	I	H	Y	V	P
EsTIP2;1	S	A	L	V	I	H	Y	V	P
EsTIP2;2	S	A	L	V	I	I	Y	V	P
EsTIP2;3	S	A	L	V	I	I	Y	V	P
EsTIP3;2	A	A	L	A	I	H	Y	V	P
EsTIP4;1	S	A	L	L	T	H	Y	V	P
EsNIP1;2	S	A	L	L	V	I	Y	V	P
EsNIP3;1	S	A	L	V	I	L	Y	V	P
EsNIP5;1	S	A	L	V	V	L	Y	V	P
Silicic acid transporters	**C/S**	**F/Y**	**A/E/L**	**H/R/Y**	**G**	**K/N/T**	**R**	**E/S/T**	**A/K/P/T**
Not found									
Urea Transporters	**H**	**P**	**F/I/L/T**	**A/C/F/L**	**L/M**	**A/G/P**	**G/S**	**G/S**	**N**
EsPIP1;1	H	P	F	F	L	P	G	G	N
EsPIP1;2	H	P	F	F	L	P	G	G	N
EsPIP1;3	H	P	F	F	L	P	G	G	N
EsPIP1;4	H	P	F	F	L	P	G	G	N
EsPIP1;5	H	P	F	F	L	P	G	G	N
EsPIP2;1	H	P	F	F	L	P	G	G	N
EsPIP2;2	H	P	F	F	L	P	G	G	N
EsPIP2;3	H	P	F	F	L	P	G	G	N
EsPIP2;4	H	P	F	F	L	P	G	G	N
EsPIP2;5	H	P	F	F	L	P	G	G	N
EsPIP2;6	H	P	F	F	L	P	G	G	N
EsPIP2;7	H	P	F	F	L	P	G	G	N
EsTIP1;1	H	P	F	F	L	A	G	S	N
EsTIP1;2	H	P	F	F	L	A	G	S	N
EsTIP1;3	H	P	F	F	L	A	G	S	N
EsTIP2;1	H	P	F	A	L	P	G	S	N
EsTIP2;2	H	P	L	A	L	P	G	S	N
EsTIP2;3	H	P	L	A	L	P	G	S	N
EsTIP2;4	H	P	F	V	L	P	G	S	N
EsTIP3;1	H	P	F	L	L	P	G	S	N
EsTIP3;2	H	P	L	L	L	P	G	S	N
EsTIP4;1	H	P	I	L	L	A	G	S	N
EsTIP5;1	H	P	F	A	L	P	G	S	N
EsNIP1;2	H	P	I	A	L	P	G	S	N
EsNIP2;1	H	P	I	A	L	E	G	S	N
EsNIP3;1	H	P	I	A	L	P	G	S	N
EsNIP4;1	H	P	V	A	L	P	G	S	N
EsNIP4;2	H	P	F	A	L	P	G	S	N
EsNIP4;3	H	P	I	A	L	P	G	S	N

EsNIP5;1	H	P	I	A	L	P	G	S	N
EsNIP6;1	H	P	I	A	L	P	S	S	N
EsNIP7;1	H	P	I	A	V	P	G	S	N

2