

Genome-Wide Bioinformatics Analysis of Aquaporin Gene Family in Maize (*Zea mays* L.)

Amna Bari¹, Muhammad Farooq², Athar Hussain^{2*}, Muhammad Tahir ul Qamar³, Malik Waseem Abbas², Ghulam Mustafa², Asad Karim⁴, Imtiaz Ahmed⁵ and Tahir Hussain¹

¹Government College University Faisalabad, Faisalabad, Pakistan

²National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan

³College of Informatics, Huazhong Agricultural University, Wuhan, China

⁴Jamil-ur-Rahman Center for Genome Research, International Center for Chemical and Biological Sciences, University of Karachi, Pakistan

⁵School of Chemistry and Environment, Beihang University, Beijing 100191, China

*Corresponding author: Athar Hussain, National Institute for Biotechnology and Genetic Engineering (NIBGE), Jang Road, Faisalabad, Pakistan, Tel: +923469565692; E-mail: atharmutahari@gmail.com

Received date: March 08, 2018; Accepted date: March 28, 2018; Published date: April 06, 2018

Copyright: © 2018 Bari A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Aquaporins are a super family of major intrinsic proteins, which facilitate the fast and passive movement of water across the cell membrane. This study presented genome-wide identification, characterization and functional prediction of aquaporins in maize using bioinformatics. A total of 41 non-redundant putative aquaporins were identified and were classified into four subfamilies: 18 TIPs, 12 PIPs, 8 NIPs and 3 SIPs. The finding reveals that exon-intron organization were conserved within subfamilies. Several transmembrane domains(TM1-TM6) were predicted by analyzing of conserved domains and motifs, along with various selectivity filters(ar/R). The functional prediction demonstrated ZmAQPs roles in regulation of multiple compounds i.e. water, glycerol, carbohydrates, metal ions and others small solutes. Furthermore, ZmAQPs were the crucial constituent of membranous structure such as plasma membrane and vacuolar membrane etc. These results deliver valuable information to address function of ZmAQPs as well as provide basic data for the improvement of plant growth and development.

Keywords: Bioinformatics; Aquaporins; Conserved motifs; Phylogenetic tree; Gene structure; Trans-membrane

Introduction

Plants are the most adversely affected entities by water scarcity due to exponential depletion of underground water level. The movement of water from underground to plants is mainly carried out through their roots by three parallel pathways i.e. symplastic, trans-cellular or apoplastic [1]. Aquaporins (AQPs) are an ancient channel protein family embedded in membranous structures of plants that transport water and other neutral metabolites across membranes [2]. Most of them are involved in hydraulic conductivity and have potential to increase 10 to 20-fold of water transportation across plasma membrane [3]. This property of AQPs is very important in plants to hold different activities like maintenance and regulation of water [4], cell elongation [5], soil-water relations [6,7], plant cell osmo-regulation [8], seed germination [9] and even in plant reproduction [10,11]. Aquaporins also influence leaf movements and their physiology [12], salt tolerance, fruit ripening [13] and drought resistance in plants [14].

As these proteins cover major portion of membranous structure so named as major intrinsic protein (MIP). The phylogenetic analysis revealed that the MIP encoding genes can be broadly classified into four different sub-families i.e. plasma membrane intrinsic proteins (PIPs), tonoplast intrinsic proteins (TIPs), nodulin-26 intrinsic proteins (NIPs) and small basic intrinsic proteins (SIPs) [3,15-17]. Additionally, three more subfamilies including glycerol intrinsic proteins (GIPs), hybrid intrinsic proteins (HIPs), and unrecognized X intrinsic proteins (XIPs) have also been reported in non-vascular

mosses and in which GIPs were homologous to bacterial glycerol channels [18,19]. The first aquaporin gene was identified from human erythrocytes CHIP28 as human AQP1 [20,21], followed by *Escherichia coli* (aqua-glyceroporin GLP-F) [22] and NOD26, a nitrogen-fixating symbiosomes in root nodules of soybean plants [23]. Currently, aquaporin genes have been identified in large numbers in many plants species e.g. 38 in *Arabidopsis thaliana* [24], 66 in *Glycine max* [25], 35 in *Physcomitrella patens* [19], 33 in *Oryza sativa* [26], 71 in *Gossypium hirsutum* [27], 54 in *Populus trichocarpa* [28] and 47 in *Solanum lycopersicum* [29].

Zea mays (maize or corn) is an important cash crop, which belongs to the grass family i.e. Poaceae. It is the 3rd most significant crop grown globally for food (<http://par.com.pk/>). In addition to human food, maize has major contribution in animal feeds and many other purposes like bioethanol production and secondary metabolites. But maize yield is adversely affected due to environmental stresses especially shortage of water, thus solution of this problem is a major concern of world-wide researches.

Many studies have been carried out on maize aquaporins, showing its crucial role in biological regulation. The first study was performed by Chaumont [3] based on expressed sequence tags (ESTs). Later on, functional aspects of different aquaporins were studied under different conditions; for example, expression of some ZmPIPs (PIPs of *Zea mays*) brought change in stomata opening and closing regulation, and also induced tolerance under high concentration of boron (B) and salinity [30]. So, a comprehensive study is required to identify the important functions of aquaporins in maize. In order to fulfill these requirements, we performed a genome wide identification,

characterization and evolutionary analysis of aquaporins of Arabidopsis and chickpea. Furthermore, we also predicted many important biological features of maize aquaporins, including protein sequence analysis, identification of trans-membrane domains, conserved motif analysis with respect to phylogenetic tree, gene structure analysis with evolutionary tree, chromosomal distribution, gene ontology includes biological process, molecular function and sub-cellular localization. These bioinformatics results might be helpful for further experimental analysis of aquaporins in maize genome.

Materials and Methods

Database searching and identification of ZmAQP genes

Aquaporin protein sequences from *Arabidopsis thaliana* (35 sequences) [24] and *Cicer aritinum* (40 sequences) [31] were used as a query in different database search engines including NCBI-Blastp, Phytozome-Blast and MaizDB-Blast. Furthermore, the position specific iteration (PSI) was also used to make it more specific [32]. The proteins having query similarity more than 75% identified through blastp were selected. After removal of redundant sequences, the identified aquaporins were validated using CD-Search program [33] for presence of MIPs domain. Finally, the data related to these selected ZmAQPs was collected including amino acid sequences, cDNA and genomic DNA sequences from MaizeDB [34].

Multiple sequence alignment and phylogenetic tree analysis

The full length amino acid sequences of 41 identified ZmAQP genes were aligned using MUSCLE program [35] and the aligned sequences used for phylogenetic tree construction through maximum likelihood method with 1000 bootstrap value in PhyML3.0 [36]. The phylogenetic tree was visualized in MEGA 6.0 [37]. In order to authenticate the evolutionary tree another program, Mr Bayes v3.2.6 [38] was also used to construct phylogenetic tree with parameters Markov Chain Sampling over the space of all possible reversible substitution models and prior for the amino acid model to mixed. The ZmAQPs were classified into four subfamilies NIPs, TIPs, PIPs and SIPs, based on known nomenclature of AQPs that were used as query in initial BLAST search. Thirty-one ZmAQPs were already annotated with AQP names by Chaumont [3] and the 10 ZmAQP genes were named according to their cumulative phylogenetic tree of ZmAQPs with Arabidopsis [24] and chickpea AQPs [31].

Protein characterization and identification of conserved domain, conserved motifs, trans membrane domain and ar/R selectivity filter of ZmAQPs

In order to determine the characteristics of ZmAQPs, the amino acid length (a.a), molecular weight (kDa) and iso-electric point (pI) were predicted through the ExPASy server (<http://web.expasy.org>). Conserved domains in ZmAQPs identifies using conserved domain database, NCBI Batch CD search [33]. Transmembrane domains were

predicted using TMHMM Server v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). The NPA motifs, ar/R selectivity filters (H2, H5, LE1, LE2) and Froger's positions (P1–P5) were identified by careful comparison of multiple sequence alignment of ZmAQPs with AQP protein structure as identified in literature [39–42]. Beside NPA motifs other conserved motifs also predicted using MEME motif discovery server 4 [43] with parameters as described by Hussain [44]. The cellular localization of ZmAQPs were predicted using two other algorithms; Plant-mPLoc (<http://www.csbio.sjtu.edu.cn>) and WoLF PSORT (<http://www.genscript.com/wolf-psort.html>).

Gene structure analysis, chromosomal distribution and Gene Ontology (GO) of ZmAQPs

The genomic data including DNA and cDNA sequences, chromosome number, start base pair, end base pair, number of exons as well as introns were retrieved from MaizeGDB (<http://www.maizegdb.org>) for gene structure analysis and chromosomal distribution. Exon-intron organization of ZmAQPs were analyzed by using Gene Display Server 2 (<http://gsds.cbi.pku.edu.cn/>) [45] and the physical location graph was manually created using MS excel sheet. Gene Ontology of ZmAQP proteins was predicted by Blast2GO program [46] using amino acid sequences with default parameters and using different databases like Swiss-Prot protein, NCBI non-redundant protein (nr), Gene ontology (GO), Kyoto Encyclopedia of Genes (KEGG) protein family and Cluster of Orthologs Groups (COGs).

Results

Identification and prediction of *Zea mays* Aquaporin's gene family

The database searches resulted in 46 putative AQP genes showing strong matches in maize (*Zea mays* L.) genome. After removal of duplication in gene locus and validation of MIP superfamily domain in amino acid sequences, 41 putative ZmAQPs were proposed in maize genome. A literature survey of AQPs superfamily along with their subfamilies in other plant species were also included for a comparative study, given in Table 1. Generic names (ZmAQPs: ZmPIPs, ZmNIPs, ZmTIPs and ZmSIPs) were assigned for *Zea mays* aquaporins. The individual gene data and their predicted characteristics are listed in Table 2, including gene name, NCBI accession, chromosome number (Chr#), amino acid residues (aa), Trans-membrane domain (TMD), isoelectric point (Ip), molecular weight (MWkDa), start base pair, end base pair and sub-cellular localization (C.L) (Table 2). Our results demonstrated that all ZmAQPs were unevenly distributed on all 10 chromosomes of maize genome. The amino acid sequences length ranged from 116 bp (ZmTIP4;3a) to 311 bp (ZmTIP4;2a) and Ip ranged 5.30 pH (ZmTIP2;3) to 9.79pH (ZmTIP3;3). Similarly, molecular weight (MW) ranged 24.988kDa (ZmTIP2;2) to 31.760 (ZmNIP2;1) (Table 2).

Species	Total AQPs	PIP	TIP	NIP	SIP	XIP	Mono/dicot	References
(<i>Oryza sativa</i>)	33	11	10	10	2	0	nonocot	[48]
Moso bamboo (<i>Phyllostachya puberula</i>)	26	10	6	8	2	0	nonocot	[47]

Sorghum(<i>Sorghum bicoler</i>)	35	9	13	10	3	0	nonocot	[46]
Banana (<i>Musa acuminata</i>)	47	18	17	9	3	0	nonocot	[51]
Maize (<i>Zea mays</i> L.)	31	12	12	4	3	0	nonocot	[3]
Maize (<i>Zea mays</i> L.)	41	12	18	8	3	0	nonocot	This study
Barley(<i>Hordeum vulgare</i>)	40	20	11	8	1	0	nonocot	[44]
Chickpea(<i>Cicer arietinum</i> L.)	39	9	11	15	4	0	dicot	[65]
Wild cabbage (<i>Brassica oleracea</i>)	41	23	8	6	4	0	dicot	[9]
Chinese cabbage (<i>Brassica rapa</i>)	29	16	5	3	5	0	dicot	[81]
<i>Arabidopsis thaliana</i>	35	13	10	9	3	0	dicot	[28]
Physic nut (<i>Jatropha curcas</i> L.)	32	9	9	8	4	2	dicot	[82]
Sweet Orange (<i>Citrus Sinensis</i>)	34	8	11	9	3	3	dicot	[83]
Rubber tree (<i>Hevea brasiliensis</i>)	51	15	17	9	4	6	dicot	[84]
Common bean (<i>Phaseolus vulgaris</i> L.)	41	12	13	10	4	2	dicot	[85]
Tomato (<i>Solanum lycopersicum</i>)	47	14	11	12	4	6	dicot	[33]
Soybean (<i>Glycine max</i>)	66	22	23	13	6	2	dicot	[29]
Potato (<i>Solanum tuberosum</i>)	47	15	11	10	3	8	dicot	[86]
Tree plant (<i>Populus trichocarpa</i>)	55	15	17	11	6	6	dicot	[32]

Table 1: Genome-wide studies of aquaporins. Literature survey on genome wide studies of aquaporins in plant species.

Gene name	NCBI accession	Chr#	AA	TMDd	Ip (pH)	MW (kDa)	Start bp	End bp	C.La	C.Lb
ZmNIP1;1	NP_001105721	5	282	6	8.58	29.55	148343777	148345925	C.M	plas: 7
ZmNIP1;2	NP_0011151947	6	284	5	7.67	29.36	131635199	131636926	C.M	plas: 8
ZmNIP2;1	NP_001105637	5	295	6	6.79	31.76	206972623	206976131	C.M	plas: 4
ZmNIP2;1a	NP_0011131324	4	303	6	6.39	32.33	182093856	182096521	C.M	plas: 5
ZmNIP2;2	NP_001105020	6	294	6	7.71	31.39	113136429	113140203	C.M	plas: 6
ZmNIP2;3	NP_001105517	9	301	6	7.04	31.65	4504397	4507849	C.M	E.R.: 5
ZmNIP3;1	XP_008660914	9	284	6	6.07	30.14	77428241	77429525	C.M	cyto: 5
ZmNIP7;1	NP_001150784	4	288	6	9.44	29.26	34365878	34367308	C.M	vacu: 9
ZmPIP1;1	NP_001105466	2	288	6	9	30.67	18757697	18759478	C.M	plas: 7
ZmPIP1;2	NP_001104934	5	289	6	9	30.79	193359281	193362537	C.M	plas: 7
ZmPIP1;3	NP_001105022	4	292	6	8.83	30.99	153654785	153658151	C.M	plas: 7
ZmPIP1;5	NP_001105131	4	288	6	8.3	30.72	170004670	170006094	C.M	plas: 8
ZmPIP1;6	NP_001105023	9	296	6	6.7	31.02	6467916	6469385	C.M	cyto: 4
ZmPIP2;1	NP_001105024	7	290	6	7.69	30.21	41435404	41438423	C.M	plas: 8
ZmPIP2;2	NP_001105638	2	292	6	8.29	30.25	169186930	169190239	C.M	plas: 6
ZmPIP2;3	NP_001105025	4	289	6	6.95	30.42	143891383	143893926	C.M	plas: 1

ZmPIP2;4	NP_001105026	5	288	6	6.5	30.32	195239919	195242604	C.M	plas:11
ZmPIP2;5	NP_001105616	2	285	6	7.7	29.83	28493234	28495453	C.M	plas: 9
ZmPIP2;6	NP_001105027	7	288	7	8.38	30.19	41553964	41555424	C.M	plas: 7
ZmPIP2;7	XP_008670063	2	286	6	8.38	29.82	168928036	168929344	C.M	plas: 8
ZmTIP1;1	NP_001105029	8	254	7	5.87	25.35	156113368	156114463	vacu	vacu: 6
ZmTIP1;2	NP_001104896	1	250	6	6.02	25.82	10068811	10070364	vacu	cyto: 4
ZmTIP2;1	NP_001104907	10	248	6	6.16	25.13	139149979	139151283	vacu	vacu: 7
ZmTIP2;2	XP_008669054	2	248	6	5.87	24.98	20452237	20453860	vacu	plas: 8
ZmTIP2;3	NP_001105030	4	249	6	5.3	24.86	152458474	152459746	vacu	plas: 7,
ZmTIP2;4	NP_001105031	5	250	6	5.59	25.04	191963086	191964272	vacu	cyto: 5
ZmTIP3;1	NP_001105032	5	262	6	8.12	27.2	25531209	25532312	Vacu	mito: 9
ZmTIP3;2	NP_001105045	1	266	6	8.13	27.39	228464363	228465624	Vacu	mito: 9
ZmTIP3;3	NP_001146930	9	267	5	9.79	27.28	154814247	154815523	vacu	cyto: 5
ZmTIP3;4	XP_008669132	2	265	6	9.05	27.46	27233154	27235078	Vacu	chlo: 5
ZmTIP4;1	XP_008648279	6	274	6	8.86	28.8	133537380	133539724	vacu	chlo:11
ZmTIP4;1a	NP_001105033	6	255	6	6.7	26.52	133537380	133539724	vacu	vacu: 5
ZmTIP4;2	NP_001105034	5	257	6	7.88	26.6	93150897	93151178	vacu	vacu: 5
ZmTIP4;2a	XP_008656321	8	311	6	7.9	32.27	110236609	110239084	vacu	chlo: 9
ZmTIP4;3	NP_001105035	3	249	6	6.2	25.26	13536805	13539433	vacu	cyto: 4
ZmTIP4;2b	XP_008653758	7	116	2	8.29	12.37	48621286	48621731	Vacu	chlo: 7
ZmTIP4;4	NP_001105641	3	252	6	6.42	25.26	1847759	1848777	vacu	cyto: 5,
ZmTIP5;1	NP_001105036	10	260	6	7.74	26.49	139151482	139151482	vacu	chlo:12
ZmSIP1;1	NP_001105514	4	245	5	8.43	25.62	9896605	0.990258	C.M	vacu: 1
ZmSIP1;2	NP_001105028	8	243	6	9.35	25.7	106815373	106819980	C.M, vacu	cyto: 5
ZmSIP2;1	AF326499	1	249		9.86	26.6	52645832	52648286	C.M	chlo, nucl

Table 2: Bioinformatics analysis of ZmAQPs. Bioinformatics predicted features of ZmAQPs, including gene name, accession, chromosome number, number of amino acids, trans-membrane domain, iso-electric point molecular weight, start and end base pair and sub-cellular location.

Phylogenetic analysis

Phylogenetic analysis revealed 41 ZmAQPs, which were divided into 4 subfamilies: ZmTIPs, ZmNIPs, ZmPIPs and ZmSIPs. ZmTIPs contained 18 ZmAQPs genes forming five subgroups: ZmTIP1 (two genes), ZmTIP2 (four genes), ZmTIP3 (four genes), ZmTIP4 (seven genes), and ZmTIP5 (one gene). Eight genes ZmNIPs were found and these genes were dividing into four subgroups i.e. ZmNIP1 (2 genes), ZmNIP2 (4 genes), ZmNIP3 (1 genes) and ZmNIP7 (1 gene). Twelve ZmPIPs were found and were divided into two subgroups, ZmPIP1 (5 genes) and ZmPIP2 (7 genes). Similarly, ZmSIPs possessed only 3 genes: ZmSIP1(two genes) and ZmSIP2 (one gene). All ZmAQPs distributed into 14 sisters branches and 13 single genes (Figure 1). In order to understand the evolutionary relationship of AQP's orthologs

between monocots and dicots plants, a phylogenetic analysis was performed among maize (41 AQP's genes), Arabidopsis (35 AQP's genes) and chickpea (40 AQP's genes) using their amino acid sequences. Result depicted that all AQP's were distributed into 38 sister groups including 13 sisters of ZmAQP's-ZmAQP's, 11 of AtAQP's-AtAQP's, 9 of CaAQP's-CaAQP's, 5 of CaAQP's-AtAQP's and 36 single genes (Figure 2). Interestingly there is no any sister branch of either ZmAQP's-AtAQP or ZmAQP-CaAQP's. Moreover most of the maize aquaporin gene formed separated branches in comparison to chickpea and Arabidopsis. For example, in NIP subfamily all ZmNIPs formed two distinct branches. Similarly, in PIPs, ZmPIP1 and ZmPIP2 showed separate branches except ZmPIP1;6. Within TIPs, ZmTIPs formed three separated branches e.g all ZmTIP3, ZmTIP2 and ZmTIP4.

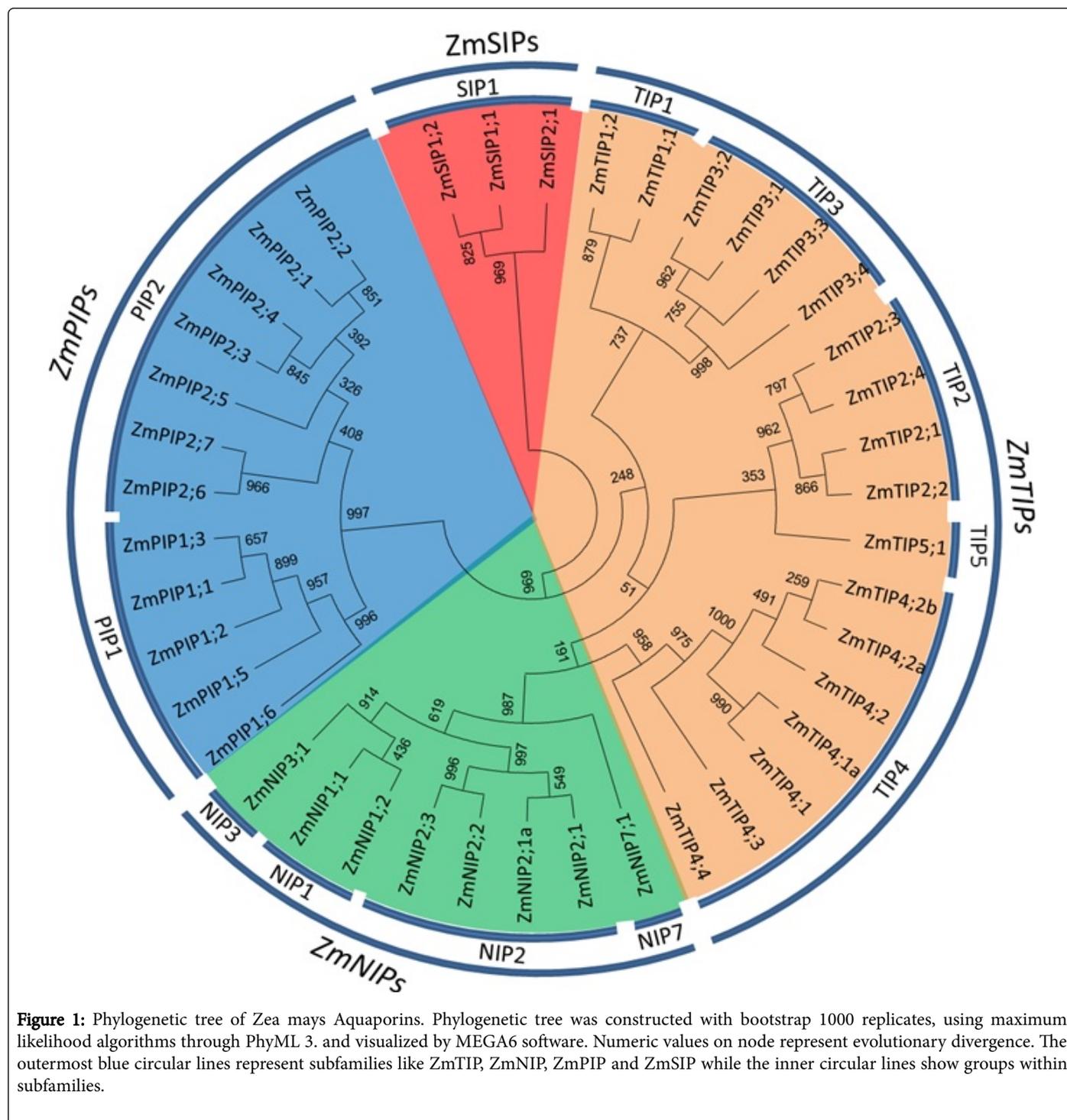
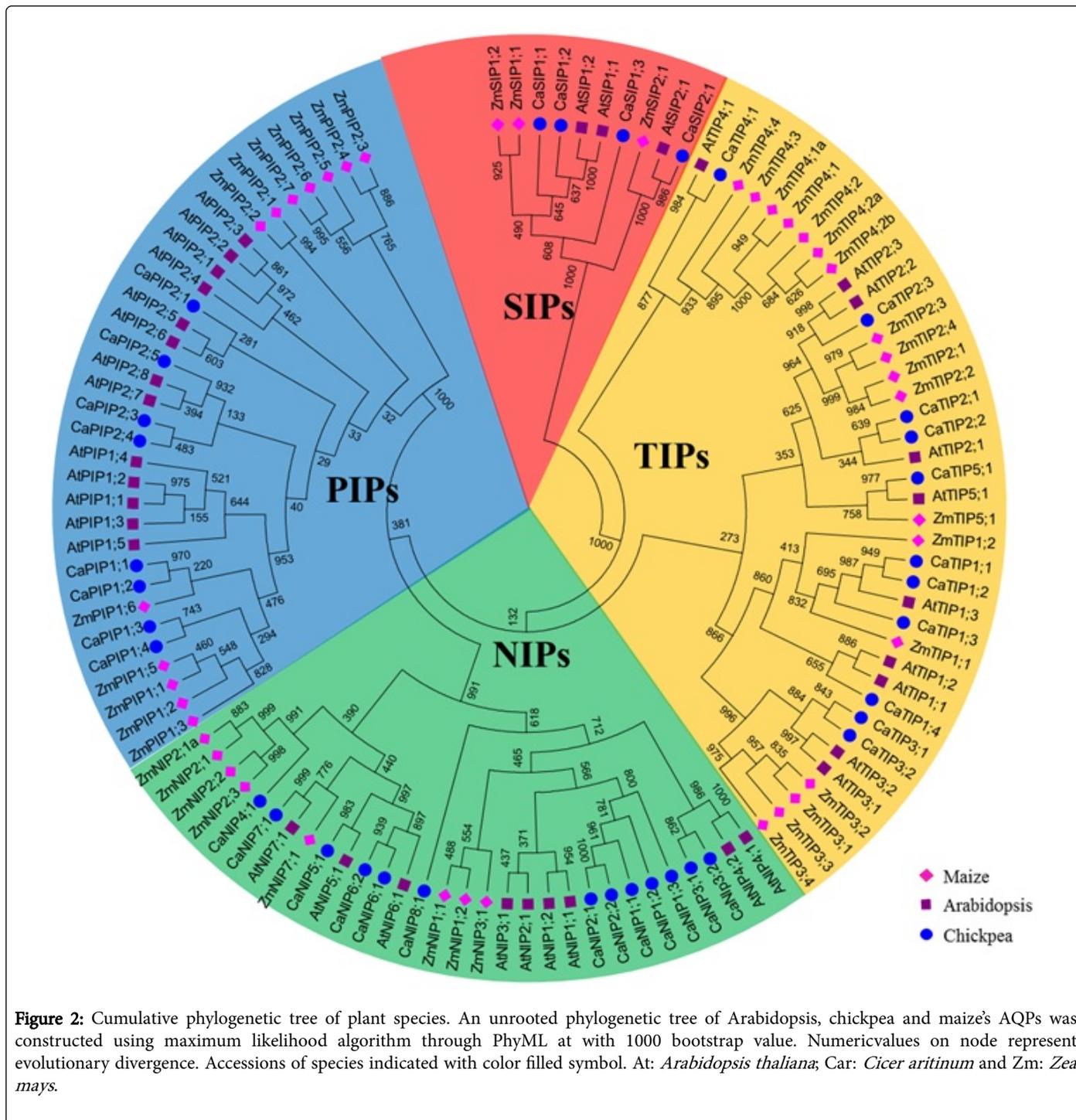


Figure 1: Phylogenetic tree of *Zea mays* Aquaporins. Phylogenetic tree was constructed with bootstrap 1000 replicates, using maximum likelihood algorithms through PhyML 3. and visualized by MEGA6 software. Numeric values on node represent evolutionary divergence. The outermost blue circular lines represent subfamilies like ZmTIP, ZmNIP, ZmPIP and ZmSIP while the inner circular lines show groups within subfamilies.



Trans-membrane domains (TMDs) and sub-cellular localization of ZmAQPs

The Trans-membrane domain (TMDs) prediction demonstrated that majority of ZmAQPs consist of six TMDs (34 of 41, 83%), two ZmAQPs (ZmPIP2;6 and ZmTIP1;1) showed seven TMDs, three ZmAQPs (ZmTIP3;3, ZmNIP1;2 and ZmSIP1;1) carried five TMDs and only one ZmAQPs (ZmTIP4;2) has two TMDs. The sub-cellular localization prediction revealed that 21 aquaporins were localized in cell membrane, whereas, 18 of them demonstrated multiple locations

including vacuole membrane, plasma membrane, mitochondrial membrane, rough endoplasmic reticulum and chloroplast membrane (Table 2, Figure S1).

Sequence alignment and identification of NPA motifs and ar/R selectivity filter residues and Frogger's positions

Multiple sequence alignment of ZmAQPs exhibited that most of the amino acid residues within domains were conserved among all 41 ZmAQPs. In NPA motifs, ar/R selectivity filter and Frogger's positions

some residues were family specific. A complete full length multiple sequence alignment of all ZmAQPs amino acid sequences has been shown in (Figure S2), indicating NPA motifs, ar/R filter, foreigner position and trans-membrane domain. The 1st NPA motif pattern (Asp-Pro-Ala) was conserved in all subfamilies except ZmSIPs which showed replacement of alanine (A) with threonine (T), while ZmTIP4;2 depicted deletion of aforementioned motif. Similarly, the 2nd NPA motif was also conserved in all 41 ZmAPQs excluding ZmTIP4;2 i.e. A(alanine) replaced with T(tyrosine) (Table 3). In case of ar/R selectivity filter regions amino acid residues were conserved intra-subfamilies but it varied inter-subfamilies like in H2 helix, histidine (H) remained intact in all of the ZmTIPs except ZmTIP4;3 and ZmTIP5;1(histidine was replaced with glutamine). In ZmPIP subfamily, histidine (H) was substituted with phenylalanine (F), in case of ZmNIPs; histidine was changed with either glycine (G) or tryptophan (W). Similarly, ZmSIPs possessed leucine (L) in place of histidine. The H5 helix exhibited bit different amino acids placements e.g. ZmTIP subfamily has isoleucine/valine/serine residue, in ZmNIPs; alanine/serine/valine, and in ZmSIPs; valine/isoleucine. However, in all ZmPIPs; histidine(H) was conserved in H5 helix. In case of loops, (G) in all ZmAQPs, while in LE2 most of the ZmAQPs found arginine (R) except ZmSIPs (R replace with N) and two gene of ZmTIPs i.e ZmTIP1;1 and ZmTIP1;2 which depicted valine (V). Similarly, the amino acid residues at P1, P2, P3, P4 and P5 are displayed in Table 3 and Figure S2.

Gene name	NPA motifs		Ar/R selectivity filter				Froger's positions				
	LB	LE	H2	H5	LE1	LE2	P1	P2	P3	P4	P5
ZmTIP2;2	NPA	NPA	H	I	G	R	T	S	A	Y	W
ZmTIP2;1	NPA	NPA	H	I	G	R	T	S	A	Y	W
ZmTIP2;4	NPA	NPA	H	I	G	R	T	S	A	Y	W
ZmTIP2;3	NPA	NPA	H	I	G	R	T	S	A	Y	W
ZmTIP1;1	NPA	NPA	H	I	G	V	T	S	A	Y	W
ZmTIP1;2	NPA	NPA	H	I	G	V	T	S	A	Y	W
ZmTIP3;1	NPA	NPA	H	V	G	R	T	V	A	Y	W
ZmTIP3;2	NPA	NPA	H	V	G	R	T	V	A	Y	W
ZmTIP3;3	NPA	NPA	H	V	G	R	T	A	A	Y	W
ZmTIP3;4	NPA	NPA	H	I	G	R	S	A	A	Y	W
ZmTIP4;2 ^a	NPA	NPA	H	S	G	R	S	S	A	Y	W
ZmTIP4;2 ^b	-	NPT	-	S	G	R	-	S	A	Y	W
ZmTIP4;2	NPA	NPA	H	S	G	R	S	S	A	Y	W
ZmTIP4;1 ^a	NPA	NPA	H	S	G	R	S	S	A	Y	W
ZmTIP4;1	NPA	NPA	H	S	G	R	S	S	A	Y	W
ZmTIP4;4	NPA	NPA	H	V	G	R	A	S	A	Y	W
ZmTIP4;3	NPA	NPA	Q	S	G	R	T	S	A	Y	W
ZmTIP5;1	NPA	NPA	Q	V	G	R	;	S	A	Y	W

ZmPIP2;3	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP2;4	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP2;5	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP2;1	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP2;2	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP2;6	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP2;7	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP1;2	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP1;3	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP1;1	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP1;5	NPA	NPA	F	H	G	R	Q	S	A	F	W
ZmPIP1;6	NPA	NPA	F	H	G	R	G	S	A	F	W
ZmNIP7;1	NPA	NPA	G	A	G	R	A	T	A	Y	M
ZmNIP2;3	NPA	NPA	G	S	G	R	L	T	A	Y	F
ZmNIP2;2	NPA	NPA	G	S	G	R	L	T	A	Y	F
ZmNIP2;1	NPA	NPA	G	S	G	R	L	T	A	Y	F
ZmNIP2;1 ^a	NPA	NPA	G	S	G	R	L	T	A	Y	F
ZmNIP1;1	NPA	NPA	W	V	G	R	F	S	A	Y	V
ZmNIP1;2	NPA	NPA	W	V	G	R	F	T	A	Y	F
ZmNIP3;1	NPA	NPA	W	A	G	R	F	S	A	Y	I
ZmSIP1;2	NPT	NPA	L	V	G	N	R	A	A	Y	W
ZmSIP1;1	NPT	NPA	L	I	G	N	K	A	A	Y	W
ZmSIP2;1	NPT	NPA	L	I	G	N	R	A	A	Y	W

Table 3: ar/R selectivity regions. Specificity determining conserved residues including NPA, ar/R filter region and P1-P5 in *Zea mays* aquaporins.

Conserved motif within ZmAQPs

Besides NPA motifs and hexa-helical domain, there were other motifs too which might play important role in regulation of ZmAQPs. MEME motif discovery resulted that most of the motifs remained conserved in all subfamilies, but some showed different pattern i.e. few were deleted out and few were unique and family specific. Motifs 1-5 (1: NPA, 2: NPA, 3: Phosphoserine, 4: Amidationsite, 5 : Casein kinase II) were conserved among all four subfamilies except ZmSIPs which had only 2nd (NPA) and 4th (Amidation site) motif. Some motifs were family specific like motif 6th, 7th, 8th and 9th (6: Novel motif, 7: N-myristoylation site, 8: Novel , 9: N-myristoylation site) were present only in ZmPIPs while its subgroups ZmPIP1 and ZmPIP2 exhibited some differences, ZmPIP1 has 11th (PK_Phospho site) motif, which was deleted in ZmPIP2 protein. Motif 10th (: N-myristoylation site) and motif 12th (N-myristoylation site) were present only in ZmTIPs and ZmSIPs, while motif 13th (Novel motif) was present only in ZmNIPs. Motif 14th (Phosphothreonine kinase) was only present in ZmTIPs subfamily members except the ZmTIP3;3, ZmTIP3;4,

ZmTIP4;4, ZmTIP4;3 and ZmTIP5;1. Motif 15th (N-myristoylation site) was present in ZmNIPs and ZmPIPs subfamily members.

However, few genes were found as outlier to aforementioned pattern (Figure 3, supplementary file S1).

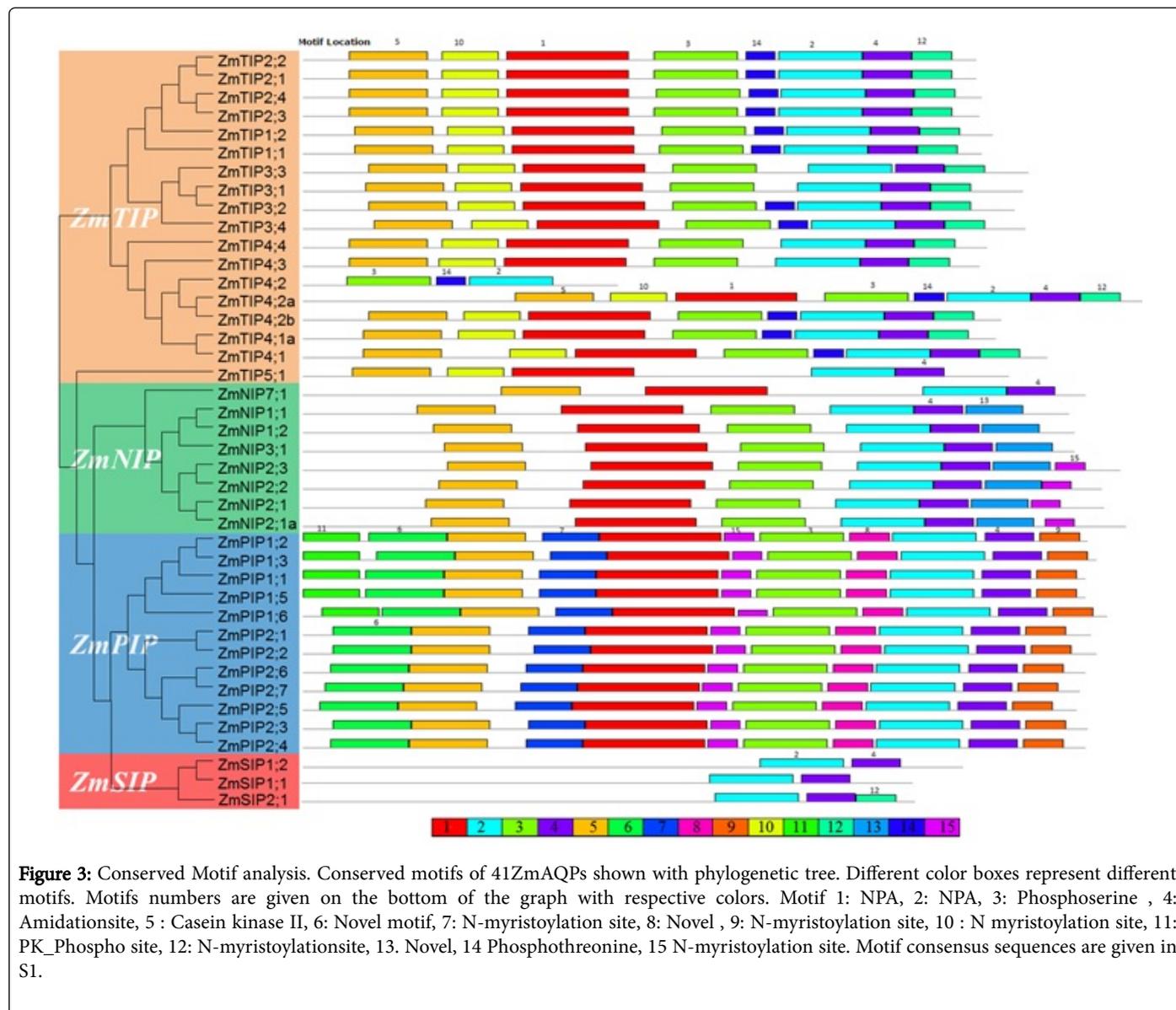


Figure 3: Conserved Motif analysis. Conserved motifs of 41ZmAQPs shown with phylogenetic tree. Different color boxes represent different motifs. Motifs numbers are given on the bottom of the graph with respective colors. Motif 1: NPA, 2: NPA, 3: Phosphoserine , 4: Amidationsite, 5 : Casein kinase II, 6: Novel motif, 7: N-myristoylation site, 8: Novel , 9: N-myristoylation site, 10 : N myristoylation site, 11: PK_Phospho site, 12: N-myristoylationsite, 13. Novel, 14 Phosphothreonine, 15 N-myristoylation site. Motif consensus sequences are given in S1.

Gene ontology (GO)

Gene Ontology analysis revealed a critical role of ZmAQPs in distinct biological, cellular and molecular processes. The biological process involved transportation of ions, glycerol and other small solute (Figure 4a, Figure S3). The molecular function revealed that ZmAQPs have crucial role in substrate-specific transmembrane transporter activity, active and passive transmembrane transporter activity, carbohydrate and organic hydroxyl compound transporter activity, receptor activity, metal binding activity, heterocyclic compound binding activity and structural molecular activity (Figure 4a, Figure

S4). In summary, significant activity of ZmAQPs was observed in transportation of compounds such as glycerol, ion, water and alcohol across plasma membrane. These activities may be due to substrate-specific channel forming activity (Figure 4a). Sub-cellular localization exhibited that all ZmAQPs were incorporated into membrane of various cellular components like plasma membrane, vacuolar membrane, membrane bounded organelles and other intrinsic component of membrane. ZmAQPs were also integrated into cell periphery, intracellular organelle, plasmodesmata and cell-cell junction (Figure 4b, Figure S5).

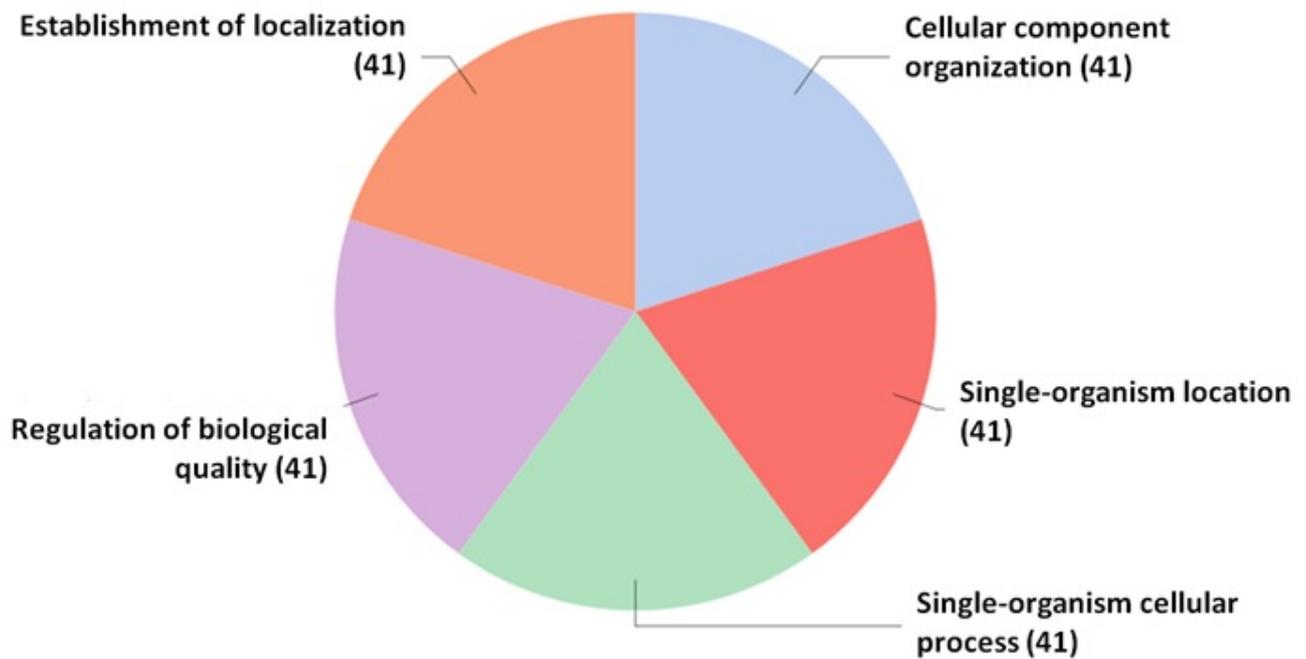


Figure 4a: Gene Ontology. Gene ontology of ZmAQPs proteins Biological process of 41 ZmAQPs.

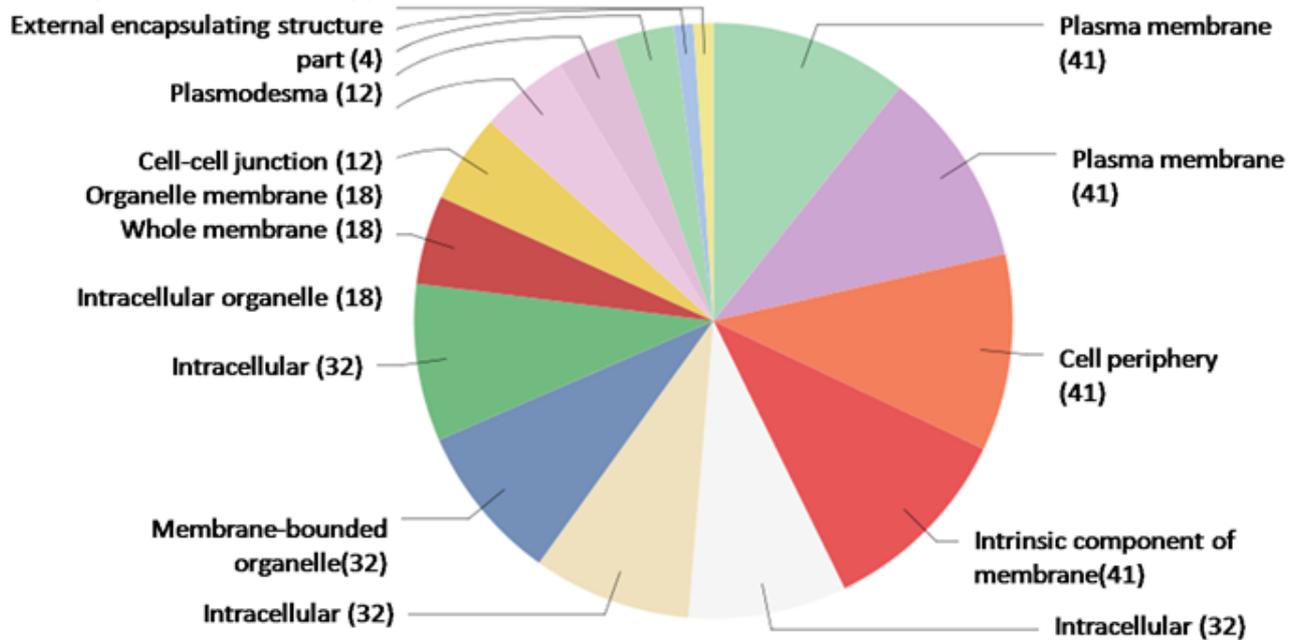


Figure 4b: Gene Ontology. Gene ontology of ZmAQPs proteins Cellular components of 41 ZmAQPs.

Gene structure analysis of ZmAQPs gene family / Genomic organization of ZmMIPs

The comprehensive transcriptomic data of *Z. mays* made it possible to analyze the gene structural components of aquaporins within *Z. mays* genome. The gene displays server (GDS) resulted that all four families showed differences in number of exon and introns. Among ZmAQPs families, ZmNIPs has highest number of exons (five exons) followed by ZmPIPs (four exons), ZmTIPs (three exons) and ZmSIPs

(three exons). For more depth study, in ZmTIPs; nine genes contain two exons and remaining eight genes carried three exons while ZmTIP4;2 showed no intron. In ZmNIPs all genes carried five exons except ZmNIP1;1 and ZmNIP1;2, which showed loss of 2nd exon. In ZmPIPs; five genes had four exons, three genes (ZmPIP2;5, ZmPIP2;6 and ZmPIP2;7) has three exons and only two genes (ZmPIP1;5 and ZmPIP1;5) has two exons. However, all ZmSIPs genes carried three exons (Figure 5).

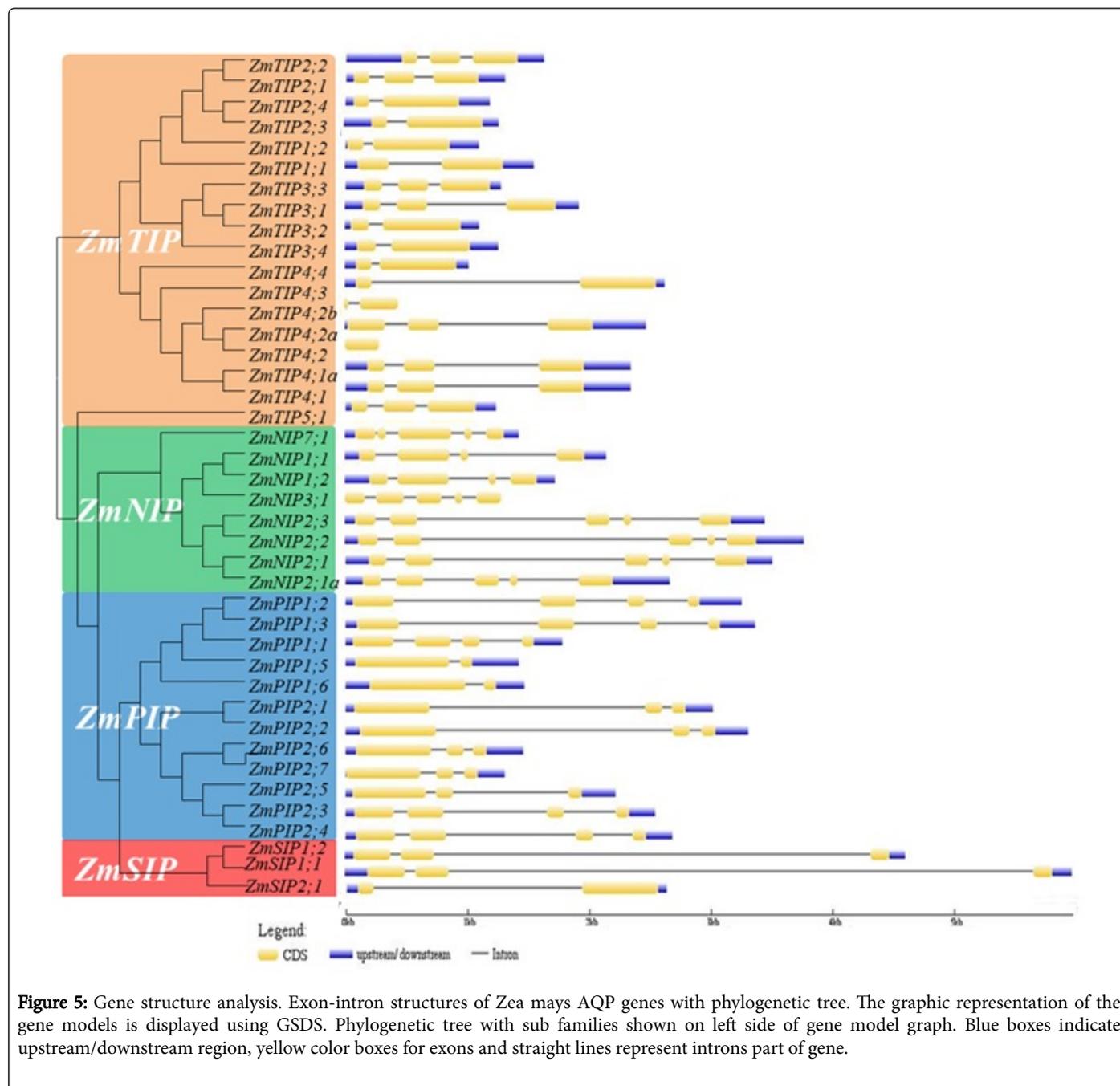


Figure 5: Gene structure analysis. Exon-intron structures of *Zea mays* AQP genes with phylogenetic tree. The graphic representation of the gene models is displayed using GSDS. Phylogenetic tree with sub families shown on left side of gene model graph. Blue boxes indicate upstream/downstream region, yellow color boxes for exons and straight lines represent introns part of gene.

Chromosomal location of ZmAQP gene

The chromosomal locations of 41 ZmAQPs were graphically presented in Figure 6. The ZmAQPs distribution in genome showed that 4th and 5th chromosome has highest number of ZmAQPs genes

(7 genes), followed by chr2 (6 genes), chr6 (4 genes) and chr9 (4 genes). Chromosome 1, 7 and 8 carried three ZmAQPs genes, while chromosome 3 and 10 has only two. Furthermore, chromosomal location identification disclosed that most of the ZmAQPs genes were

found in cluster form and located either on top or on bottom of the chromosome. Interestingly, most of the genes belonging to same family were located on same chromosome e.g. ZmPIP2;2, ZmPIP2;7 and ZmPIP2;5 formed a cluster, within a~500kbp segment on chromosome

2 while ZmPIP1;5 and ZmPIP1;3 were clustered on chromosome 4. Similarly, ZmNIP1;2 and ZmNIP2;2 were clustered on chromosome 6 (Figure 6).

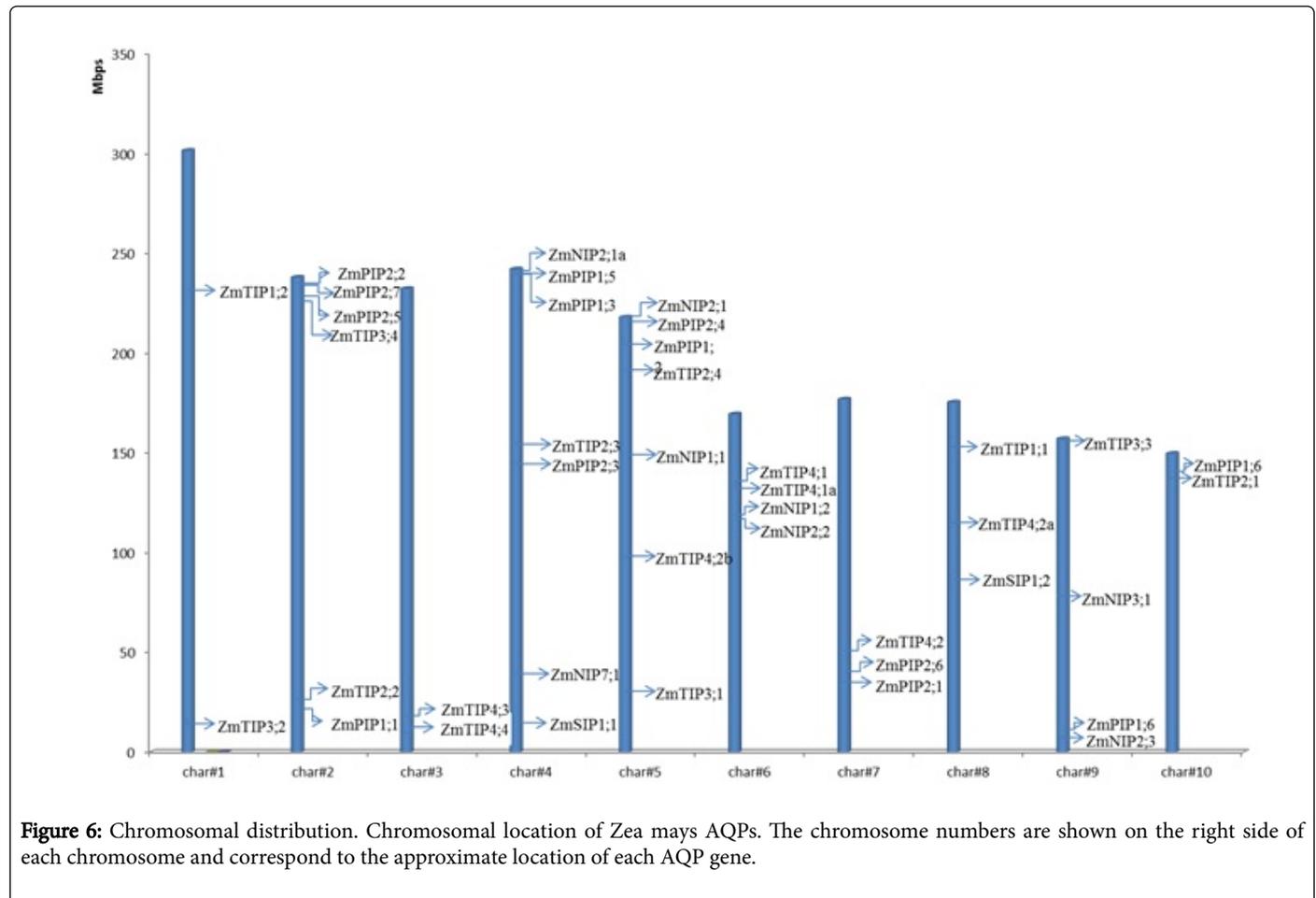


Figure 6: Chromosomal distribution. Chromosomal location of *Zea mays* AQPs. The chromosome numbers are shown on the right side of each chromosome and correspond to the approximate location of each AQP gene.

Discussion

Robust advancements in the computational biology helped to sequence large number of plant genomes, which certainly improved the identification and characterization of physiologically vital gene families in plants. The computational analysis of aquaporin proteins will definitely help in hypothesis generation and subsequent experimental validations and ultimately lead towards genetically engineered improved crop. Researchers have always been concerned to study different genes and their expression in order to understand their role in the context of a particular plant. Here in this study, AQPs family was selected due to its significant input to growth and development of a plant [47]. The prime function performed by AQPs is the regulation of water and some solutes across cell membrane [47,48]. Considering AQPs importance in plant growth and development, its genome wide analysis has been conducted in many plant species including monocots, such as rice, wheat, barley etc, as well as in dicots i.e. potato, tomato, cabbage, carrot, celery and Arabidopsis (see Table 1). Current study on ZmAQPs has its own worth because to the best of our knowledge few studies conducted on genome wide identification, characterization and functional prediction of aquaporins in maize [49].

However, Chaumont [3] carried a study on ZmAQPs based on EST/cDNA sequence and identified 31 AQPs, while current study is based on whole genome identification of aquaporins gene and proposed 41 AQPs gene. The difference in number of AQPs between the two studies might be due to varying number of expressed transcripts at a given time. The limitation in Chaumont study [3] was addressed in this study by exploiting the whole genome sequence of maize. In contrast to Chaumont study, we also performed a genome-wide comparative study of ZmAQPs with Arabidopsis and chickpea that gave the evolutionary insight among the monocots and dicots. Chaumont Study was restricted to expressed sequence tags and its phylogenetic tree, while in this study, we provided many other biological parameters like ZmAQPs protein sequences analysis, conserved motif prediction, gene ontology, gene structure analysis and chromosomal and sub-cellular localization prediction in details. The genome wide study of aquaporins in maize suggested that the total number of AQP genes in maize were higher than in Arabidopsis (35 AQPs), sweet orange (34 AQPs), rice (34 AQPs) and physic nut (32 AQPs), but lesser than other few species like banana (47 AQPs), tomato (47 AQPs), cottonwood (55 AQPs), soybean (66 AQPs), and chinese cabbage (53 AQPs) (see Table 1 for the comparison chart). These changes in number of gene in

different species may be due to size of their genome or due to evolutionary process for adaptation in natural environment [50].

The phylogenetic tree demonstrated that all ZmAQPs were classified into four sub-families viz ZmTIPs, ZmPIPs, ZmNIPs and ZmSIPs (Figure 1), which is in agreement with the Chaumont [3] and other studies listed in Table 1. Additionally, we also indicated groups and sub-groups of subfamilies; for instance, ZmTIPs was divided into five subgroups (TIP1, TIP2, TIP3, TIP4 and TIP5). Such findings were similar to other monocots AQPs i.e rice, banana, sorghum, and barley (Table 1). As TIPs proteins involve in the transport of various small solutes like NH_4^+ , H_2O_2 , and urea [51-53], so the information about TIPs in maize genome may leads to improvement of transportation mechanism. The ZmPIPs were divided into two subgroups; ZmPIP1 and ZmPIP2 which are seeming to be conserved in all other plant species (listed in Table 1). Experimental analysis of PIPs conformed its role in water absorption inside roots as well as turgor pressure in leaf [54-56]. Furthermore, PIPs also facilitate CO_2 diffusion in mesophyll that enhances photosynthesis process [57,58]. So the study of PIPs individual genes in maize will help in understanding of C_4 photosynthesis mechanism that are crucial in engineering of C_4 features into C_3 plants, such as rice, wheat and potato [59]. In case of ZmNIPs, it was divided into four subgroups like ZmNIP1, ZmNIP2, ZmNIP3 and ZmNIP7;1. Conversely, Chaumont proposed three subgroups of ZmNIPs (not identified ZmNIP7). The NIP7 subgroup has its unique sequence and was studied in all genome wide studies of AQPs in plant species like Arabidopsis, chickpea, rice, banana, barley, sweet orange, tomato, common bean and sorghum (Table 1). NIPs were found a bit more diverse than other sub-families, as it is more specific to species. It was observed that different species of monocots has different number of subgroups of NIPs subfamily like sorghum and rice have four sub-groups; while moso bamboo has three, and banana has five sub-groups [60-62]. Generally, NIPs reported as transporter of water and various small solutes like glycerol, silicon, lactic acid and urea transport facilitator [63,64]. These differences in crops might be due to crops potency to uptake glycerol, silicon and other small solutes. The ZmSIPs sub-family has been divided in two sub-groups (ZmSIP1 and ZmSIP2) which is quite similar to Chaumont [3] study and in other plants as listed in Table 1.

The comparative phylogenetic analysis of AQPs among maize, Arabidopsis and chickpea indicated that AQPs of Arabidopsis and chickpea were closer than maize. It concluded that AQPs within dicots are more similar as compare to dicots monocots relation. For example, ZmAQPs showed close relation with itself rather than other two dicots. However further deep study is required to understand the diversity of AQPs related to its functions. Furthermore, the 13 sister branches of ZmAQPs-ZmAQPs revealed segmental duplication events within maize genome that may have some key role in evolutionary adaptation against various environmental stress. The other important sequence features of ZmAQPs including molecular weight (Mw), iso-electric point (Ip), and amino acid length were similar as reported for MIP proteins in Arabidopsis, banana, moso bamboo and other species [17,61,65]. These predicted features would be helpful for the functional characterization of ZmAQPs. The identification of trans-membrane domains in ZmAQPs gives information about structural association with various functions. The identified trans-membrane domains in ZmAQPs were same as reported in other crops like TM1 to TM6 [17,61,65].

The transportation characteristics of AQPs are due to NPA motifs and ar/R selectivity filters which form water channels [66,67]. They

have high specific substrate binding capacity and are essential for selective transport of water and small solutes[66,67]. Chaumont study [3]only presented the NPA motifs, while current study included the ar/R selectivity filter and Frogger's positions that are most important for selection of molecule across biological membrane [66,67]. All ZmAQPs showed two representative NPA motifs as reported by Chaumont [3]. The ZmPIPs contain highly conserved ar/R selectivity filter region (F-H-G-R) and same pattern was also identified in other PIPs, like in Arabidopsis, tomato, wheat, barley and poplar (Table 1). Additionally, the presence of the S, A, F, W residues at P2-P5 positions in PIPs, has been reported as a signature of CO_2 transporter [68]. Any mutation in these conserved amino acids in ZmPIPs can alter the capacity of the protein for CO_2 diffusion. If the mutation leads increment in CO_2 diffusion, then it will also helpful for establishment of C_4 crops. Similarly ZmTIPs carried H, I, G, R or H, I, A, V residues in the ar/R selectivity filter region, and T, S, A, Y, W or T/S, A, A, Y, W amino acid residues at P1-P5 positions are reported to transport urea and H_2O_2 across membrane [68]. The ZmNIPs ar/R filters in maize were identical to sweet orange and soybean where these genes act as water facilitator [69]. The ZmNIPs sub-group showed G, S, G and R, at ar/R filter that are involved in transportation of water, silicon (Si) and boron (B) as identified in Arabidopsis (AtNIP5 with A, I, G and R in ar/R filter region involved in the transport of arsenic and boron but not silicon) and sweet orange [70]. The SIPs protein's basic structure and function are still under characterization but in Arabidopsis, AtSIPs are involved in functional water channel [71]. The differences in ar/R selectivity filter residues might alter the selectivity behavior for substrate transport. Thus, point mutations in these patterns either increase or decrease the transportation capacity of aquaporin proteins [70,72,73]. The variation in these conserved segments described the evolutionary diversity among ZmAQPs. Besides NPA and ar/R filter regions, there were some other important motifs, predicted through MEME discovery server (Figure 3) including phosphoserine, amidation site, casein kinase II, N-myristoylation site, PK_Phospho, phosphothreonine and some novel motifs. Such motifs have also essential role in organization and regulation of MIP domain and sometime these motifs can be targeted for regulation and interaction of AQPs protein with other cellular components. The prediction on sub-cellular localization as plasma membrane, vacuole membrane and mitochondrial membrane for ZmAQPs that were corresponds to the reports in other plant species including soybean, sorghum, rubber tree, sweet orange, common bean, soybean and moso bamboo (see Table 1). The ZmTIPs mainly express in vacuole that may help to control osmotic potential and ZmPIPs integrated within plasma membrane, while ZmNIPs are expressed in various membrane such finding were similar to literature [74,75].

We have presented schematic representation of gene structures of all ZmAQPs with their evolutionary relationship (Figure 5) that were not demonstrated in Chaumont [3]. Our results were in agreement to previous studies in other plant species[17,61,65]. Most of ZmAQPs showed similar gene structure to their orthologs in other plants like in TIPs; ZmTIP1, SbTIP1(sorghum) [60], MaTIP1(banana) [65], PeTIP1 (moso bamboo) [61] has two exons; ZmTIP2 has also covenant with its orthologs in sorghum, banana and moso bamboo except ZmTIP2;1 which demonstrated an additional exon as compare to aforementioned species[17,61,65]. The ZmTIP3 showed similarity to sorghum, while somehow diverges to moso bamboo and banana. Similarly, other sub-families also demonstrated similar patterns. Thus, most of our findings were similar to previous studies (Table 1). So, we suggest that the loss or gain of exons in ZmAQP genes might have ensued under natural

selection [28,76]. Previous studies have depicted that lost or gain of exons are the common feature of evolutionary process in plants genomes [76,77].

The chromosomal location of genes tells about the expression capacity of a gene [78]. Chaumont study [3] on ZmAQPs did not give any clue about gene location within genome of maize. However, we demonstrated chromosomal locations of ZmAQPs on whole genome sequence. All ZmAQPs genes were distributed on all 10 chromosomes forming groups. The genes belonging to same sub-family mostly form clusters within the window size of 500kb. The clustering of ZmAQPs depicted segmental duplication that is one of a key mechanisms for gene expansion that increase genetic diversity [79]. Moreover, this mechanism may also be responsible for the functional divergence by increasing total members of a given gene subfamily [80].

Conclusion

In this study, we used different bioinformatics tools for the genome-wide identification, analysis and characterization of aquaporins in maize genome. Moreover, we have predicted several physiological structures with their biological functions. The aquaporin genes family widely studied in many important crops demonstrating its structural and functional diversity. The availability of genome sequences made this study possible. Current study identified 41 aquaporins in *Zea mays* L. and these genes assigned nomenclature as well as classified into four subfamilies. Moreover, the structural and functional features of ZmAQPs have been predicted, and a comparative phylogenetic study of ZmAQPs, CaAQPs, and AtAQPs was also conducted, which provided insights about the evolution of AQPs within plant species. The results achieved in this study not only provide valuable information for future functional analysis of ZmAQP genes but also make a suitable reference to survey the gene family expansion in *Zea mays* and other crops from evolutionary perspective.

Competing Interests

The Authors declare no competing interest.

Acknowledgements

We are thankful to our colleagues and teachers at National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan

Ethical Approval and Consent to Participate

Not applicable.

References

1. Steudle E (2000) Water uptake by roots: Effects of water deficit. J Exp Bot 51: 1531-1542.
2. Noronha H, Agasse A, Martins AP, Berny MC, Gomes D, et al. (2014) The grape aquaporin VvSIP1 transports water across the ER membrane. J Exp Bot 65: 981-993.
3. Chaumont F, Barrieu F, Wojcik E, Chrispeels MJ, Jung R (2001) Aquaporins constitute a large and highly divergent protein family in maize. Plant Physiol 125: 1206-1215.
4. Maurel C, Verdoucq L, Luu DT, Santoni V (2008) Plant aquaporins: Membrane channels with multiple integrated functions. Annu Rev Plant Biol 59: 595-624.
5. Higuchi T, Suga S, Tsuchiya T, Hisada H, Morishima S (1998) Molecular cloning, water channel activity and tissue specific expression of two isoforms of radish vacuolar aquaporin I. Plant Cell Physiol 39: 905-913.
6. Maurel C (2007) Plant aquaporins: Novel functions and regulation properties. FEBS Lett 581: 2227-2236.
7. Siefritz F, Tyree MT, Lovisolo C, Schubert A, Kaldenhoff R (2002) PIP1 Plasma membrane aquaporins in Tobacco: From cellular effects to function in plants. Plant Cell 14: 869-876.
8. Kjellbom P (1999) Aquaporins and water homeostasis in plants. Trends Plant Sci 4: 308-314.
9. Gao YP, Young L, Bonham-Smith P, Gusta LV (1999) Characterization and expression of plasma and tonoplast membrane aquaporins in primed seed of Brassica napus during germination under stress conditions. Plant Mol Biol 40: 635-644.
10. Bots M, Vergeldt F, Wolters-Arts M, Weterings K, van As H, et al. (2005) Aquaporins of the PIP2 class are required for efficient anther dehiscence in tobacco. Plant Physiol 137: 1049-1056.
11. Kaldenhoff R, Fischer M (2006) Functional aquaporin diversity in plants. Biochim Biophys Acta 1758: 1134-1141.
12. Uehlein N, Kaldenhoff R (2008) Aquaporins and plant leaf movements. Ann Bot 101: 1-4.
13. Mut P, Bustamante C, Martínez G, Allea K, Sutka M, et al. (2008) A fruit-specific plasma membrane aquaporin subtype PIP1; 1 is regulated during strawberry (*Fragaria × ananassa*) fruit ripening. Physiologia Plantarum 132: 538-551.
14. Lian HL, Yu X, Ye Q, Ding XS, Kitagawa Y, et al. (2004) The role of aquaporin RWC3 in drought avoidance in rice. Plant and Cell Physiol 45: 481-489.
15. Bansal A, Sankaramakrishnan 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 Struct Biol 7: 27.
16. Johanson U, Gustavsson S (2002) A new subfamily of major intrinsic proteins in plants. Mol Biol Evol 19: 456-461.
17. Johanson U, Karlsson M, Johansson I, Gustavsson S, Sjoval S, et al. (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 Physiol 126: 1358-1369.
18. Gustavsson S, Lebrun AS, Nordén K, Chaumont F, Johanson U (2005) A novel plant major intrinsic protein in Physcomitrella patens most similar to bacterial glycerol channels. Plant Physiol 139: 287-295.
19. Danielson JÁ, Johanson U (2008) Unexpected complexity of the Aquaporin gene family in the moss Physcomitrella patens. BMC Plant Biol 8: 45.
20. Denker BM, Smith BL, Kuhajda FP, Agre P (1988) Identification, purification, and partial characterization of a novel Mr 28,000 integral membrane protein from erythrocytes and renal tubules. J Biol Chem 263: 15634-15642.
21. Preston GM, Agre P (1991) Isolation of the cDNA for erythrocyte integral membrane protein of 28 kilodaltons: Member of an ancient channel family. Proc Natl Acad Sci USA 88: 11110-11114.
22. Hénin J, Tajkhorshid E, Schulten K, Chipot C (2008) Diffusion of Glycerol through Escherichia coli Aquaglyceroporin GlpF. Biophys J 94: 832-839.
23. 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 Res 15: 813-824.
24. Quigley F, Rosenberg JM, Shachar-hill Y, Bohnert HJ (2002) From genome to function: The Arabidopsis aquaporins. Genome Biol 3: 1-17.
25. Zhang DY, Ali Z, Wang CB, Xu L, Yi JX, et al. (2013) Genome-Wide Sequence Characterization and Expression Analysis of Major Intrinsic Proteins in Soybean (*Glycine max* L.). PLoS One 8: e56312.

26. 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 Physiol* 46: 1568-1577.
27. 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 Biol* 10: 142.
28. Gupta AB, Sankararamkrishnan R (2009) Genome-wide analysis of major intrinsic proteins in the tree plant *Populus trichocarpa*: Characterization of XIP subfamily of aquaporins from evolutionary perspective. *BMC Plant Biology* 9: 134.
29. Reuscher S, Akiyama M, Mori C, Aoki K, Shibata D, et al. (2013) Genome-wide identification and expression analysis of aquaporins in Tomato. *PLoS One* 8: e79052-e79052.
30. Heinen RB, Bienert GP, Cohen D, Chevalier AS, Uehlein N, et al. (2014) Expression and characterization of plasma membrane aquaporins in stomatal complexes of *Zea mays*. *Plant Mol Biol* 86: 335-350.
31. Deokar AA, Tar'an B (2016) Genome-wide analysis of the aquaporin gene family in Chickpea (*Cicer arietinum* L.). *Front Plant Sci* 7: 1802.
32. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.
33. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, et al. (2016) CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45: D200-D203.
34. Polacco M, Coe E (2002) MaizeDB: The maize genome database. *Bioinformatics: Databases and Systems* 151-162.
35. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792-1797.
36. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, et al. (2010) New algorithms and methods to estimate maximum-Likelihood phylogenies: Assessing the performance of PhyML 2.0. *Syst Biol* 59: 307-321.
37. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evolution* 30: 2725-2729.
38. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61: 539-542.
39. Gattolin S, Sorieul M, Frigerio L (2011) Mapping of tonoplast intrinsic proteins in maturing and germinating *Arabidopsis* seeds reveals dual localization of embryonic TIPs to the tonoplast and plasma membrane. *Mol Plant* 4: 180-189.
40. Wang L, Li Q, Lei Q, Feng C, Gao Y, et al. (2015) MzPIP2; 1: An aquaporin involved in radial water movement in both water uptake and transportation, altered the drought and salt tolerance of transgenic *Arabidopsis*. *PLoS One* 10: e0142446.
41. Jain D, Chattopadhyay D (2010) Analysis of gene expression in response to water deficit of chickpea (*Cicer arietinum* L.) varieties differing in drought tolerance. *BMC Plant Biol* 10: 24.
42. Zhou S, Hu W, Deng X, Ma Z, Chen L, et al. (2012) Overexpression of the wheat aquaporin gene, TaAQP7, enhances drought tolerance in transgenic tobacco. *PLoS One* 7: e52439.
43. Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* 34: W369-W373.
44. Hussain A, Mustafa G, Ahmed I, Azeem M, Shaheen H, et al. (2016) Genome-wide analysis of ethylene responsive factor in Maize : An *in silico* Approach. *Appl Ecol Env Research* 14: 177-200.
45. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, et al. (2015) GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* 31: 1296-1297.
46. Gotz S, Arnold R, Sebastian-Leon P, Martin-Rodriguez S, Tischler P, et al. (2011) B2G-FAR, a species-centered GO annotation repository. *Bioinformatics* 27: 919-924.
47. Maurel C, Boursiac Y, Luu DT, Santoni V, Shahzad Z, et al. (2015) Aquaporins in plants. *Physiological Reviews* 95: 1321-1358.
48. Chaumont F, Moshelion M, Daniels MJ (2005) Regulation of plant aquaporin activity. *Biol Cell* 97: 749-764.
49. Xiao Y, Liu H, Wu L, Warburton M, Yan J (2017) Genome-wide association studies in maize: Praise and stargaze. *Molecular Plant* 10: 359-374.
50. Narayan R (1998) The role of genomic constraints upon evolutionary changes in genome size and chromosome organization. *Annals of Botany* 82: 57-66.
51. Bienert GP, Moller ALB, Kristiansen KA, Schulz A, Moller IM, et al. (2007) Specific aquaporins facilitate the diffusion of hydrogen peroxide across membranes. *J Biol Chem* 282: 1183-1192.
52. Holm LM, Jahn TP, Møller ALB, Schjoerring JK, Ferri D, et al. (2005) NH₃ and NH₄⁺ permeability in aquaporin-expressing *Xenopus* oocytes. *Pflügers Archiv: Eur J Physiol* 450: 415-428.
53. Loque D, Ludewig U, Yuan L, von Wiren N (2005) Tonoplast Intrinsic Proteins AtTIP2;1 and AtTIP2; 3 Facilitate NH₃ Transport into the Vacuole. *Plant Physiol* 137: 671-680.
54. Hu W, Yuan Q, Wang Y, Cai R, Deng X, et al. (2012) Overexpression of a wheat aquaporin gene, TaAQP8, enhances salt stress tolerance in transgenic tobacco. *Plant Cell Physiol* 53: 2127-2141.
55. Li G, Santoni V, Maurel C (2014) Plant aquaporins: Roles in plant physiology. *Biochimica et Biophysica Acta (BBA): General Subjects* 1840: 1574-1582.
56. Xu Y, Hu W, Liu J, Zhang J, Jia C, et al. (2014) A banana aquaporin gene, MaPIP1;1, is involved in tolerance to drought and salt stresses. *BMC Plant Biol* 14: 59.
57. Heckwolf M, Pater D, Hanson DT, Kaldenhoff R (2011) The *Arabidopsis thaliana* aquaporin AtPIP1;2 is a physiologically relevant CO₂ transport facilitator. *Plant J* 67: 795-804.
58. Mori IC, Rhee J, Shibasaki M, Sasano S, Kaneko T, et al. (2014) CO₂ Transport by PIP2 Aquaporins of Barley. *Plant Cell Physiol* 55: 251-257.
59. Wang C, Guo L, Li Y, Wang Z (2012) Systematic comparison of C3 and C4 plants based on metabolic network analysis. *BMC Systems Biology* 6: S9.
60. Gene P, Reddy PS, Santosh T, Bhadra R, Sharma KK, et al. (2015) Genome-wide identification and characterization of the aquaporin gene family in *Sorghum bicolor* (L). *PLGENE* 1: 18-28.
61. Sun H, Li L, Lou Y, Zhao H (2016) Genome-wide identification and characterization of aquaporin gene family in moso bamboo (*Phyllostachys edulis*). *Mol Biol Rep*.
62. Nguyen MX, Moon S, Jung KH (2013) Genome-wide expression analysis of rice aquaporin genes and development of a functional gene network mediated by aquaporin expression in roots. *Planta* 238: 669-681.
63. Choi WG, Roberts DM (2007) *Arabidopsis* NIP2;1, a major intrinsic protein transporter of lactic acid induced by anoxic stress. *J Biol Chem* 282: 24209-24218.
64. Takano J, Wada M, Ludewig U, Schaaf G, Von Wirén N, et al. (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.
65. Hu W, Hou X, Huang C, Yan Y, Tie W, et al. (2015) Genome-wide identification and expression analyses of aquaporin gene family during development and abiotic stress in banana. *Int J Mol Sci* 16: 19728-19751.
66. Park JH, Jr MHS (1996) Phylogenetic characterization of the MIP family of transmembrane channel proteins. *J Membrane Biol* 153: 171-180.
67. Reizer J, Reizer A, Saier MH (1993) The MIP family of integral membrane channel proteins: Sequence comparisons, evolutionary relationships, reconstructed pathway of evolution, and proposed functional differentiation of the two repeated halves of the proteins. *Crit Rev Biochem Mol Biol* 28: 235-257.
68. Hove RM, Bhavne M (2011) Plant aquaporins with non-aqua functions: Deciphering the signature sequences. *Plant Mol Biol* 75: 413-430.
69. Wallace IS, Roberts DM (2004) Homology modeling of representative subfamilies of *Arabidopsis* major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter. *Plant Physiol* 135: 1059-1068.

70. Mitani-Ueno N, Yamaji N, Zhao FJ, Ma JF (2011) The aromatic/arginine selectivity filter of NIP aquaporins plays a critical role in substrate selectivity for silicon, boron, and arsenic. J Exp Bot 62: 4391-4398.
71. 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 Lett 579: 5814-5820.
72. Beitz E, Wu B, Holm LM, Schultz JE, Zeuthen T (2006) Point mutations in the aromatic/arginine region in aquaporin 1 allow passage of urea, glycerol, ammonia, and protons. Proc Natl Acad Sci USA 103: 269-274.
73. Hub JS, de Groot BL (2008) Mechanism of selectivity in aquaporins and aquaglyceroporins. Proc Natl Acad Sci USA 105: 1198-1203.
74. Wudick MM, Luu DT, Maurel C (2009) A look inside: Localization patterns and functions of intracellular plant aquaporins. New Phytol 184: 289-302.
75. Zelazny E, Borst JW, Muylaert M, Batoko H, Hemminga MA, et al. (2007) FRET imaging in living maize cells reveals that plasma membrane aquaporins interact to regulate their subcellular localization. Proc Natl Acad Sci USA 104: 12359-12364.
76. Song XM, Huang ZN, Duan WK, Ren J, Liu TK, et al. (2014) Genome-wide analysis of the bHLH transcription factor family in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Mol Genet Genomics 289: 77-91.
77. Roy SW, Penny D (2006) Patterns of intron loss and gain in plants: Intron loss-dominated evolution and genome-wide comparison of *O. sativa* and *A. thaliana*. Mol Biol Evol 24: 171-181.
78. Bryant JA, Sellars LE, Busby SJ, Lee DJ (2014) Chromosome position effects on gene expression in *Escherichia coli* K-12. Nucleic Acids Res 42: 11383-11392.
79. Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R (2013) Gene duplication as a major force in evolution. Journal of Genetics 92: 155-161.
80. Shi FX, Li MR, Li YL, Jiang P, Zhang C, et al. (2015) The impacts of polyploidy, geographic and ecological isolations on the diversification of *Panax* (Araliaceae). BMC Plant Biol 15: 297.