TÜBİTAK

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Research Article

Turk J Biol
(2016) 40: 600-611
© TÜBİTAK
doi:10.3906/biy-1501-91

Comparative and phylogenetic analysis of zinc transporter genes/proteins in plants

Recep VATANSEVER[1], İbrahim İlker ÖZYİĞİT[1], Ertuğrul FİLİZ[2,*]
[1]Department of Biology, Faculty of Science and Arts, Marmara University, İstanbul, Turkey
[2]Department of Crop and Animal Production, Çilimli Vocational School, Düzce University, Düzce, Turkey

Received: 28.01.2015 • **Accepted/Published Online:** 04.08.2015 • **Final Version:** 18.05.2016

Abstract: Zinc is an important catalytic element for more than 300 enzymes and plays a structural role in the stabilization of many proteins. Protein domain analysis showed that identified Zn transporter proteins belong to the ZIP protein family (PF02535). Zn transporter sequences were found to have similar molecular weights (33.1–51.4 kDa) and amino acid lengths (306–478 amino acids) with 5.31–8.92 *pI* values. Subcellular localization of Zn transporters was predicted as the plasma membrane. They had 6–9 putative transmembrane domains with a variable region between TM-3 and TM-4, which could contain a potential histidine-rich metal-binding domain. Moreover, alignment analysis showed that the TM-2, -4, and -5 domains could be potential metal-binding sites because they contain highly conserved His residues. Based on a homology search, the retrieved sequences were identified as corresponding homologs of either *Arabidopsis thaliana* or *Oryza sativa*. Phylogenetic analysis also supported that *A. thaliana* and *O. sativa* sequences could be used as a reference/benchmark to identify Zn transporter homologs in various plant species. The findings of this study will be valuable theoretical knowledge for feature studies in terms of understanding the gene and protein features of Zn transporters in various plants.

Key words: ZIP, transmembrane domain, histidine residue, metal-binding site, conserved motifs

1. Introduction

Zinc is an important catalytic element for more than 300 enzymes, including alcohol dehydrogenase, alkaline phosphatase, Cu-Zn superoxide dismutase, and carbonic anhydrase. It also plays structural roles in the stabilization of many protein motifs such as the Zn finger, Zn cluster, and RING finger domains (Fox and Guerinot, 1998). Zn is absorbed from soil as divalent cations and its cellular role depends on its behavior because it is not cellularly oxidized or reduced later (Marschner, 1995; Fox and Guerinot, 1998). Many studies have showed that Zn homeostasis in plants greatly requires a coordinated regulation of ZIP family metal transporters (Guerinot, 2000). Thus, the ZIP family has a major role in Zn transport. Additionally, the ZIP family is also involved in the transport of various metals such as iron (Fe^{2+}), manganese (Mn^{2+}), and cadmium (Cd^{2+}) (Guerinot, 2000; Maser et al., 2001). *Arabidopsis* ZIP family members of *ZIP1*, *ZIP2*, *ZIP3*, *ZIP4*, *IRT1*, *IRT2*, and *IRT3* genes have been functionally characterized as zinc uptake transporters with different affinities (Grotz et al., 1998; Maser et al., 2001; Lin et al., 2009; Assunção et al., 2010). *ZIP1*, *ZIP2*, and *ZIP3* genes were demonstrated to complement Zn uptake in yeast and they were also upregulated in roots of zinc-deficient plants. *ZIP4* is expressed in roots and shoots of Zn-deficient plants, suggesting its role in intracellular Zn transport (Grotz et al., 1998). IRT1 is a high-affinity Fe uptake transporter in roots (Vert et al., 2002); however, overexpressed *AtIRT1* was also reported to accumulate high levels of Cd and Zn under Fe deficiency (Connolly et al., 2002). AtIRT2 is also expressed under Fe deficiency and can transport Zn and Fe in yeast (Vert et al., 2001). *Arabidopsis* IRT3, which is a close homolog of *AtZIP4* (Grotz et al., 1998), functionally complemented Zn and Fe uptake in yeast, indicating a possible role in Zn and Fe transport (Lin et al., 2009). Most *ZIP* family members encode a polypeptide of 309–476 residues with eight potential TM domains, of which N- and C-terminal regions are located extracellularly. A cytoplasmic variable region between TM- 3 and -4 were also reported to contain a potential metal-binding domain rich in His residues (Guerinot, 2000). Moreover, TM-2, -4, and -5 contained a fully conserved histidine residue in all family members (Grotz et al., 1998; Guerinot, 2000). However, TM-4 was demonstrated to be the most conserved site of ZIP proteins, which, with an adjacent (semi)polar residue, may form part of the intramembranous metal-binding site in the transport pathway (Eng et al., 1998). Zn transporter gene homologs have also been characterized in some plant species, including soybean (Moreau et al., 2002), rice (Ramesh et al., 2003; Ishimaru et al., 2005), *Medicago truncatula* (Lopez-Millan et al., 2004), and maize (Li et al., 2013). Although

---

* Correspondence: ertugrulfiliz@gmail.com

Zn transporters have been identified in some plants, we still lack knowledge in many plant species. Thus, in the present study, we aimed to identify and comparatively analyze Zn transporter genes and proteins in 18 different plant species. In this context, potential Zn transporter genes and proteins in 18 plant species were identified, physicochemical properties and TM topologies of proteins were determined, ZIP family specific signature motifs were analyzed, gene organization was determined, and a phylogenetic tree, using protein sequences, was constructed.

2. Materials and methods

2.1. Sequence retrieval of zinc transporters

Eighteen plant species, including *Arabidopsis thaliana*, *Brachypodium distachyon*, *Brassica rapa*, *Cucumis sativus*, *Eucalyptus grandis*, *Glycine max*, *Gossypium raimondii*, *Medicago truncatula*, *Oryza sativa*, *Phaseolus vulgaris*, *Populus trichocarpa*, *Prunus persica*, *Solanum lycopersicum*, *Sorghum bicolor*, *Vitis vinifera*, *Zea mays*, *Chlamydomonas reinhardtii* (green alga), and *Physcomitrella patens* (moss) were selected from the Phytozome database for this study. The selected species are representatives of main plant groups such as monocots, dicots, and lower plants. Seven functionally characterized *Arabidopsis* Zn transporter sequences including ZIP1 (O81123.1), ZIP2 (Q9LTH9.1), ZIP3 (Q9SLG3.1), ZIP4 (O04089.1), IRT1 (Q38856.2), IRT2 (O81850.1), and IRT3 (Q8LE59.3) were then obtained from the NCBI protein database (http://www.ncbi.nlm.nih.gov/protein) (Romiti, 2010). These sequences were queried as references in proteome datasets of the 18 selected plant species in the Phytozome database for a threshold value of E-value $\leq 1 \times 10^{-10}$ (Goodstein et al., 2012). Redundant sequences were removed from the obtained genes and remaining sequences were further analyzed for their TM topologies by using the TMHMM server (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh et al., 2001).

2.2. Analysis of gene and protein features

Physicochemical properties such as sequence length, molecular weight, and isoelectric point of Zn transporter proteins were determined by using the ProtParam tool on the ExPASy server (http://web.expasy.org/protparam/) (Gasteiger et al., 2005). The subcellular localization of proteins was predicted by the CELLO server (http://cello.life.nctu.edu.tw/) (Yu et al., 2006). Protein domain families were searched in the Pfam database (http://pfam.xfam.org/) by using a sequence search tool (Sonnhammer et al., 1997). Exon/intron organizations of Zn transporter genes were analyzed by the Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/) (Guo et al., 2007).

2.3. Primary sequence analysis

Conserved motifs in protein sequences were analyzed by using the Multiple Em for Motif Elicitation tool (MEME; http://meme.nbcr.net/meme/) with the following parameters: maximum number of motifs to find, 5; minimum width of motifs, 6; and maximum width of motifs, 50 (Timothy et al., 2009).

2.4. Phylogenetic analysis

Protein sequences were aligned by ClustalW (Thompson et al., 1994). A phylogenetic tree was constructed by MEGA5 with the maximum likelihood (ML) method for 1000 bootstrap values (Tamura et al., 2011) and visualized by the FigTree tool (http://tree.bio.ed.ac.uk/software/figtree/).

3. Results and discussion

3.1. Analysis of gene and protein features

We have obtained a total of 112 Zn transporter genes from 18 plant species. These included 12 genes from *A. thaliana*, 11 genes from *B. distachyon*, 14 genes from *B. rapa*, 2 genes from *C. reinhardtii*, 5 genes from *C. sativus*, 7 genes from *E. grandis*, 7 genes from *G. max*, 7 genes from *G. raimondii*, 5 genes from *M. truncatula*, 8 genes from *O. sativa*, 4 genes from *P. vulgaris*, 2 genes from *P. patens*, 3 genes from *P. trichocarpa*, 5 genes from *P. persica*, 7 genes from *S. lycopersicum*, 2 genes from *S. bicolor*, 6 genes from *V. vinifera*, and 5 genes from *Z. mays* (Table). Protein domain analysis showed that the retrieved Zn transporter sequences belonged to the ZIP protein family (PF02535). They had similar molecular weights (33.1–51.4 kDa) and amino acid lengths (306–478 amino acids) with 5.31–8.92 *pI* values (Table). They included 6-9 potential TM domains. The subcellular localization of these proteins was predicted as the plasma membrane. About 85% of identified sequences demonstrated a slightly acidic character while the remaining 15% were basic in character with a *pI* of ≥8. Many studies have reported that ZIP family members are mainly predicted to be 309–476 amino acid residues in length and have similar topology with eight potential TM domains (Guerinot, 2000). cDNA analysis of soybean Zn transporter gene *GmZIP1* showed that it encodes a polypeptide of 354 amino acid residues with eight putative TM domains rich in His residues at loop regions (Moreau, et al., 2002). Rice OsZIP4 protein was reported to show 58% identity with AtZIP1. It was also demonstrated to be localized to the plasma membrane (Ishimaru et al., 2005). *Medicago MtZIP1*, *MtZIP3*, *MtZIP4*, and *MtZIP5* genes were upregulated in roots/leaves under Zn deficiency and encoded a protein of 350–372 amino acids with eight TM domains (Lopez-Millan et al., 2004). Nine *ZIP* member genes were demonstrated to function in Zn/Fe transport in maize, which encode 359–490 amino acids with 6–9 putative TM domains (Li et al., 2013). Moreover, many other studies have also reported the similar general features of Zn transporters in the ZIP family (Tiong et al., 2015; Milner et al., 2013). In the present study, we have also found similar topological features and physicochemical properties in identified Zn transporter sequences, thereby showing consistency with the current

Table. List of zinc transporter genes in 18 plant species.

Gene symbol[a] / Homolog[b]	Product	E-value[c] (blastp)	Gene ID (Phytozome)	NCBI accession	Exon number	Protein length	MW (kDa)	*pI*	TM domain
AtZIP1/AtZIP12	Zinc transporter 1	0.0	AT3G12750	O81123.1	2	355	37.8	6.22	9
AtZIP2/AtZIP11	Zinc transporter 2	0.0	AT5G59520	Q9LTH9.1	2	353	38.3	5.98	9
AtZIP3/AtZIP5	Zinc transporter 3	0.0	AT2G32270	Q9SLG3.1	3	339	36.0	6.78	9
AtZIP4/AtIRT3	Zinc transporter 4, chloroplastic	0.0	AT1G10970	O04089.1	4	408	43.1	6.11	6
AtZIP5/AtZIP3	Zinc transporter 5	0.0	AT1G05300	O23039.1	3	360	38.1	6.26	9
AtZIP6/OsZIP6	Zinc transporter 6, chloroplastic	0.0	AT2G30080	O64738.1	2	341	36.0	5.71	8
AtZIP7/AtIRT1	Zinc transporter 7	0.0	AT2G04032	NP_178488.1	3	365	39.3	5.90	9
AtZIP9/AtIRT3	Fe(ii) transporter isolog family protein	0.0	AT4G33020	NP_195028.1	2	344	36.1	6.18	6
AtZIP10/AtIRT1	Probable zinc transporter 10	0.0	AT1G31260	Q8W245.2	3	364	39.4	8.40	9
AtZIP11/AtZIP2	Putative zinc transporter zip2 - like protein	0.0	AT1G55910	NP_564703.1	2	326	35.4	5.67	9
AtZIP12/AtZIP1	Probable zinc transporter 12	0.0	AT5G62160	Q9FIS2.1	2	355	37.5	6.79	9
AtIRT3/AtZIP4	Fe(2+) transport protein 3, chloroplastic	0.0	AT1G60960	Q8LE59.3	4	425	45.0	6.19	6
LOC100822027/OsZIP1	Zinc transporter 1-like	0.0	Bradi2g04020	XP_003569091.1	2	378	39.9	6.11	8
LOC100833512/OsZIP2	Zinc transporter 2	0.0	Bradi1g60110	XP_003557730.1	4	358	36.9	5.26	9
LOC100831108/OsZIP3	Zinc transporter 3-like isoform x1	0.0	Bradi5g21580	XP_003581657.2	3	360	37.5	6.89	6
LOC100827894/OsZIP4	Zinc transporter 4-like	0.0	Bradi3g17900	XP_003571530.1	3	416	42.5	8.34	7
LOC100822110/OsZIP5	Zinc transporter 5-like	0.0	Bradi2g22520	XP_003566152.1	3	362	37.3	6.75	7
LOC100827984/OsZIP6	Zinc transporter 6	0.0	Bradi2g34560	XP_003568854.1	2	399	42.1	5.95	8
LOC100826969/OsZIP7	Zinc transporter 7-like	0.0	Bradi2g33110	XP_003568766.1	4	388	40.6	7.79	6
LOC100822361/OsZIP8	Zinc transporter 8-like	0.0	Bradi1g53680	XP_003557360.1	3	366	37.8	5.92	7
LOC100826853/OsZIP9	Zinc transporter 9-like	0.0	Bradi2g22530	XP_003568257.1	3	369	38.1	6.16	9
LOC100830270/OsZIP10	Zinc transporter 10-like isoform x2	0.0	Bradi1g37667	XP_003560649.1	5	408	42.6	5.99	6
LOC100830895/OsIRT1	Fe(2+) transport protein 1-like	0.0	Bradi1g12860	XP_010229515.1	2	367	38.5	8.39	8
LOC103870229/AtZIP1	Zinc transporter 1	0.0	Brara.E02829	XP_009146590.1	2	353	37.6	5.94	8
LOC103851574/AtZIP2	Zinc transporter 2	0.0	Brara.B01094	XP_009126693.1	2	356	38.6	6.75	9
LOC103865184/AtZIP3	Zinc transporter 3-like	0.0	Brara.D01950	XP_009141213.1	3	342	36.4	6.25	9
LOC103843273/AtZIP4	Zinc transporter 4, chloroplastic	0.0	Brara.I05200	XP_009118225.1	3	413	43.4	6.08	6
LOC103844142/AtZIP5	Zinc transporter 5	0.0	Brara.J00353	XP_009119168.1	3	357	37.8	6.56	8
LOC103867987/AtZIP6	Zinc transporter 6, chloroplastic	0.0	Brara.E01313	XP_009144334.1	2	337	35.6	5.34	8
LOC103853609/AtZIP8	Probable zinc transporter 8 isoform x1	0.0	Brara.B02814	XP_009128761.1	3	347	37.0	5.95	8
LOC103850142/AtZIP9	Zinc transporter 9	0.0	Brara.A00489	XP_009125109.1	3	385	40.6	5.56	6
LOC103833819/AtZIP10	Probable zinc transporter 10	0.0	Brara.H00723	XP_009108125.1	3	357	38.4	8.08	9
LOC103832550/AtZIP11	Zinc transporter 11	0.0	Brara.H00021	XP_009106819.1	2	322	34.7	5.32	9
LOC103854983/AtZIP12	Probable zinc transporter 12	0.0	Brara.F02137	XP_009130197.1	3	351	36.8	6.79	8
LOC103857703/AtIRT1	Fe(2+) transport protein 1-like	0.0	Brara.A01065	XP_009133175.1	3	339	36.0	6.14	8
LOC103861180/AtIRT2	Fe(2+) transport protein 2-like	0.0	Brara.K01776	XP_009137130.1	3	347	37.1	7.05	8
LOC103830215/AtIRT3	Fe(2+) transport protein 3, chloroplastic	0.0	Brara.A02506	XP_009104210.1	4	421	44.5	5.58	6
ZRT3	Zinc-nutrition responsive transporter	0.0	Cre13.g573950	XP_001693505.1	5	408	41.8	5.78	8
ZIP6	Zip family transporter	0.0	Cre06.g299600	XP_001691291.1	8	413	41.8	6.73	8
LOC101216720/AtZIP1	Zinc transporter 1-like	0.0	Cucsa.239910	XP_004139841.1	3	354	37.2	8.53	8
LOC101214729/AtZIP6	Zinc transporter 6, chloroplastic	0.0	Cucsa.139350	XP_004166326.1	2	334	35.6	5.73	8
LOC101204028/AtZIP7	Zinc transporter 7-like	0.0	Cucsa.046330	XP_004134951.1	1	348	38.0	5.58	7
LOC101208105/AtZIP11	Zinc transporter 11-like	0.0	Cucsa.390890	XP_004150296.1	3	337	36.2	6.02	8
LOC101218224/AtIRT3	Fe(2+) transport protein 3, chloroplastic	0.0	Cucsa.395110	XP_004142927.1	3	417	44.5	6.09	6
LOC104451775/AtZIP2	Zinc transporter 2-like	0.0	Eucgr.F02060	XP_010064670.1	2	349	37.5	5.96	9
LOC104436300/AtZIP4	Zinc transporter 4, chloroplastic	0.0	Eucgr.C00648	XP_010047335.1	4	422	45.1	6.03	6
LOC104433710/OsZIP5	Zinc transporter 5-like	0.0	Eucgr.A00916	XP_010044860.1	3	364	38.3	6.91	7
LOC104432631/AtZIP6	Zinc transporter 6, chloroplastic	0.0	Eucgr.B02234	XP_010043412.1	2	332	35.1	6.12	8
LOC104425170/OsZIP8	Zinc transporter 8-like	0.0	Eucgr.K01344	XP_010036079.1	4	374	39.6	7.25	6
LOC104451774//AtZIP11	Zinc transporter 11-like	0.0	Eucgr.F02059	XP_010064669.1	2	334	35.8	8.45	7
LOC104446499/AtIRT1	Fe(2+) transport protein 1-like	0.0	Eucgr.E01915	XP_010058645.1	3	357	37.9	7.15	8
LOC100805124/AtZIP1	Zinc transporter 1-like	0.0	Glyma.15G262800	XP_003546826.1	3	359	38.2	6.14	8
LOC100809940/OsZIP2	Zinc transporter 2-like	0.0	Glyma.08G328000	XP_003532170.1	3	349	37.5	6.06	9
LOC100809426/AtZIP4	Zinc transporter 4, chloroplastic-like	0.0	Glyma.17G228600	XP_006601231.1	3	393	43.1	5.73	7
LOC100797310/AtZIP6	Zinc transporter 6, chloroplastic-like	0.0	Glyma.14G196200	XP_003544893.1	2	324	33.8	6.35	8
IRT/AtZIP10	Probable zinc transporter 10-like precursor	0.0	Glyma.20G022500	NP_001274385.1	3	358	38.7	6.86	9
LOC100784872/OsIRT1	Fe(2+) transport protein 1-like	0.0	Glyma.02G126000	XP_003520144.1	3	360	38.4	7.62	6

Table. (Continued).

Gene symbol[a] / Homolog[b]	Product	E-value[c] (blastp)	Gene ID (Phytozome)	NCBI accession	Exon number	Protein length	MW (kDa)	*pI*	TM domain
LOC100812212/AtIRT3	Fe(2+) transport protein 3, chloroplastic-like	0.0	Glyma.06G052000	XP_003526000.2	4	478	51.4	6.57	7
-------------------/*AtZIP1*	Zinc transporter 1	1e-154	Gorai.002G110700	O81123.1	3	356	37.9	6.94	7
-------------------/*AtZIP4*	Zinc transporter 4, chloroplastic	9e-172	Gorai.009G044800	O04089.1	2	388	41.2	5.95	6
------------------/*OsZIP5*	Zinc transporter 5	4e-113	Gorai.001G093500	Q6L8G0.1	3	366	39.2	5.98	8
------------------/*AtZIP6*	Zinc transporter 6, chloroplastic	8e-160	Gorai.004G232700	O64738.1	2	325	34.6	7.16	8
------------------/*AtZIP7*	Zinc transporter 7	3e-142	Gorai.005G148500	Q8W246.1	3	355	38.3	7.56	7
-----------------/*AtZIP10*	Probable zinc transporter 10	2e-157	Gorai.011G049900	Q8W245.2	3	357	38.2	8.76	8
-----------------/*AtZIP11*	Zinc transporter 11	1e-145	Gorai.013G153700	Q94EG9.1	2	330	35.7	5.65	9
ZIP1/AtZIP1	Metal transport protein	0.0	Medtr2g064310	AAR08412.1	3	358	38.5	6.49	8
MTR_3g082050/OsZIP5	Zinc transporter	0.0	Medtr3g082050	XP_003601469.1	4	372	40.3	5.89	6
MTR_4g083570/AtZIP10	Zinc transporter	0.0	Medtr4g083570	XP_003607852.1	3	350	37.7	6.45	9
ZIP7/OsZIP2	Metal transport protein	0.0	Medtr3g058630	AAR08417.1	3	350	37.5	5.44	9
MTR_8g105030/OsIRT1	Zip zinc/iron transport family protein	0.0	Medtr8g105030	AET05397.2	5	396	43.0	5.90	9
OsZIP1/OsZIP2	Zinc transporter 1	0.0	LOC_Os01g74110	Q94DG6.1	2	352	37.4	8.90	9
OsZIP2/AtZIP11	Zinc transporter 2	0.0	LOC_Os03g29850	Q852F6.1	4	358	36.6	5.70	9
OsZIP5/OsZIP9	Zinc transporter 5	0.0	LOC_Os05g39560	Q6L8G0.1	3	353	36.7	6.35	7
OsZIP6/AtZIP6	Zinc transporter 6	0.0	LOC_Os05g07210	Q6L8F9.1	2	395	41.3	6.34	8
OsZIP7/AtZIP4	Zinc transporter 7	0.0	LOC_Os05g10940	Q6L8F7.1	4	384	39.7	6.56	6
OsZIP8/OsZIP5	Zinc transporter 8	0.0	LOC_Os07g12890	A3BI11.1	3	390	40.2	6.30	7
OsZIP9/OsZIP5	Zinc transporter 9	0.0	LOC_Os05g39540	Q0DHE3.3	3	362	37.8	5.97	8
OsZIP10/AtZIP4	Zinc transporter 10	0.0	LOC_Os06g37010	Q5Z653.2	4	404	41.5	6.53	6
PHAVU_006G055800g/OsZIP2	Hypothetical protein	0.0	Phvul.006G055800	XP_007146623.1	3	352	37.7	5.58	9
---------------------------/*OsZIP5*	Zinc transporter 5	1e-117	Phvul.008G290500	Q6L8G0.1	3	356	38.0	6.34	7
PHAVU_008G259200g/AtZIP6	Hypothetical protein	0.0	Phvul.008G259200	XP_007142183.1	2	325	33.9	6.11	8
PHAVU_003G262400g/OsIRT1	Hypothetical protein	0.0	Phvul.003G262400	XP_007156150.1	3	352	37.4	6.70	8
PHYPADRAFT_190331/OsZIP2	Hypothetical protein	0.0	Phpat.016G005000	XP_001771951.1	1	336	36.6	8.45	9
PHYPADRAFT_104780/AtZIP4	Zip family transporter	0.0	Phpat.006G078700	XP_001754592.1	5	367	39.5	6.51	8
POPTR_0009s07810g/AtZIP6	Zinc transporter 6 family protein	0.0	Potri.009G074100	XP_002313244.2	2	335	35.7	6.03	8
POPTR_0001s37480g/AtZIP11	Zinc transporter 11 precursor family protein	0.0	Potri.001G366100	XP_002300374.2	3	347	37.3	5.75	9
POPTR_0015s15730g/AtIRT1	Root iron transporter family protein	0.0	Potri.015G117900	XP_002322355.2	4	337	36.0	8.45	7
PRUPE_ppa006989mg /AtZIP4	Hypothetical protein	0.0	ppa006989m	XP_007223082.1	2	387	40.8	5.99	6
PRUPE_ppa027114mg/AtZIP10	Hypothetical protein	0.0	ppa027114m	XP_007214341.1	3	358	38.3	8.57	8
PRUPE_ppa019717mg/AtZIP11	Hypothetical protein	0.0	ppa019717m	XP_007214183.1	2	306	33.1	6.69	9
PRUPE_ppa020099mg /OsIRT1	Hypothetical protein	0.0	ppa020099m	XP_007225019.1	3	361	38.3	7.63	8
PRUPE_ppa007995mg /AtZIP6	Hypothetical protein	0.0	ppa007995m	XP_007205467.1	2	349	36.9	6.18	7
LOC101255999/AtZIP2	Zinc transporter 2	0.0	Solyc06g005620.2	XP_004240368.1	3	337	36.7	5.80	9
LOC101260003/OsZIP3	Zinc transporter 3-like	0.0	Solyc02g081600.2	XP_004232649.1	3	352	38.5	8.40	7
LOC101259773/AtZIP4	Zinc transporter 4, chloroplastic	0.0	Solyc08g065190.2	XP_004245100.1	4	407	43.3	6.30	6
LOC101257981/OsZIP5	Zinc transporter 5-like	0.0	Solyc07g043230.2	XP_004243649.1	3	342	36.4	6.70	9
LOC101252338/AtZIP1	Zinc transporter 8-like	0.0	Solyc02g032100.2	XP_004231600.1	3	347	37.1	5.85	7
LOC101250568/AtIRT1	Fe(2+) transport protein 1-like	0.0	Solyc02g069210.2	XP_004233216.1	3	355	38.3	9.13	8
LOC543598/AtZIP10	Iron-regulated transporter 2 precursor	0.0	Solyc02g069190.2	NP_001234252.1	3	352	37.6	8.41	8
SORBIDRAFT_03g047340/OsZIP1	Hypothetical protein	0.0	Sobic.003G443700	XP_002459178.1	1	361	38.8	8.74	9
--------------------------------/*OsZIP2*	Zinc transporter 2	2e-177	Sobic.001G339600	Q852F6.1	4	358	37.4	5.62	9
LOC100248132/OsZIP2	Zinc transporter 11-like	0.0	GSVIVG01014656001	XP_010644553.1	4	322	34.4	5.31	8
LOC100244130/AtZIP4	Zinc transporter 4, chloroplastic isoform x1	0.0	GSVIVG01035402001	XP_002279424.1	5	353	37.6	7.07	6
LOC100250102/AtZIP6	Zinc transporter 6, chloroplastic	0.0	GSVIVG01024696001	XP_002276231.1	2	335	35.6	6.85	8
LOC100240875/AtZIP10	Probable zinc transporter 10	0.0	GSVIVG01026252001	XP_002273179.1	3	348	37.1	8.92	8
LOC100242590/OsIRT1	Fe(2+) transport protein 1	0.0	GSVIVG01026250001	XP_002273397.1	3	349	37.3	7.10	9
LOC100241788/OsIRT2	Fe(2+) transport protein 2	0.0	GSVIVG01014337001	XP_002282425.1	3	354	37.3	6.75	7
LOC103652993/OsZIP4	Zinc transporter 4-like	0.0	GRMZM2G111300	XP_008678193.1	3	386	38.6	7.22	7
LOC103630465/OsZIP5	Zinc transporter 5-like	0.0	GRMZM2G047762	XP_008649739.1	3	341	34.7	5.30	7
LOC100283249/OsZIP6	Zip zinc/iron transport family protein	0.0	GRMZM2G034551	NP_001149623.1	2	396	41.5	6.15	8
LOC100281849/OsZIP8	Zip zinc/iron transport family protein precur.	0.0	GRMZM2G093276	NP_001148241.1	3	397	40.6	6.33	7
LOC103644356/OsIRT2	Fe(2+) transport protein 2-like	0.0	GRMZM2G115190	XP_008665793.1	2	361	37.6	8.65	6

[a] Gene symbol shows the gene entry for corresponding Phytozome ID. ------------ means that there is no entry for that species in the NCBI database.

[b] Homolog shows the most homologous gene. Only if a homolog gene is available, then product name, e-value, and NCBI accession number specify the homolog.

[c] blastp was performed against UniProtKB/Swiss-Prot (swissprot) and nonredundant protein sequence (nr) databases of the NCBI.

literature. Overall, it may be reported that general gene and protein features of Zn transporter proteins of the ZIP family could be used in identification of Zn transporter homologs in different plant genomes. However, wet-lab studies are needed to confirm the Zn transport activities of these identified proteins. Furthermore, we have searched the experimentally characterized homologs of identified Zn transporters (Table). Based on a homology search, the retrieved sequences were identified as corresponding homologs of either *A. thaliana* or *O. sativa*.

Exon/intron analysis showed that most Zn transporter genes have 2–4 exons, with the exception of *C. sativa* (Cucsa.046330), *P. patents* (Phpat.016G005000), and *S. bicolor* (Sobic.003G443700) genes with only one exon; *B. distachyon* (Bradi1g37667), *C. reinhardtii* (Cre13.g573950), *M. truncatula* (Medtr8g105030), *P. patents* (Phpat.006G078700), and *V. vinifera* (GSVIVG01035402001) genes with 5 exons; and the *C. reinhardtii* (Cre06.g299600) gene with 8 exons. It could be suggested that these various exon/intron organizations may show divergences of Zn transporter genes in plants.

3.2. Conserved motif and sequence analysis

Motif analysis was performed for the most conserved five motif types (Figure 1). Motifs 1–3 were related to the ZIP protein family (PF02535), while motifs 4 and 5 did not relate to any protein families. In addition, motifs 1 and 3 were present in all sequences, and motif 2 was in 90 out of 112 sequences while motifs 4 and 5 were in 111 out of 112 sequences, proposing that motif structures of Zn transporter are well conserved in plants.

All identified 112 Zn transporter sequences in the 18 plant species were aligned by ClustalW, and identical and similar residues were shaded as black and gray, respectively (Figure 2). Approximate locations of TM domains have been indicated by lines and numbers above the aligned sequences. TM domains locations were determined based

Figure 1. Sequence logo of the five most conserved motifs in Zn transporter proteins in 18 plant species.

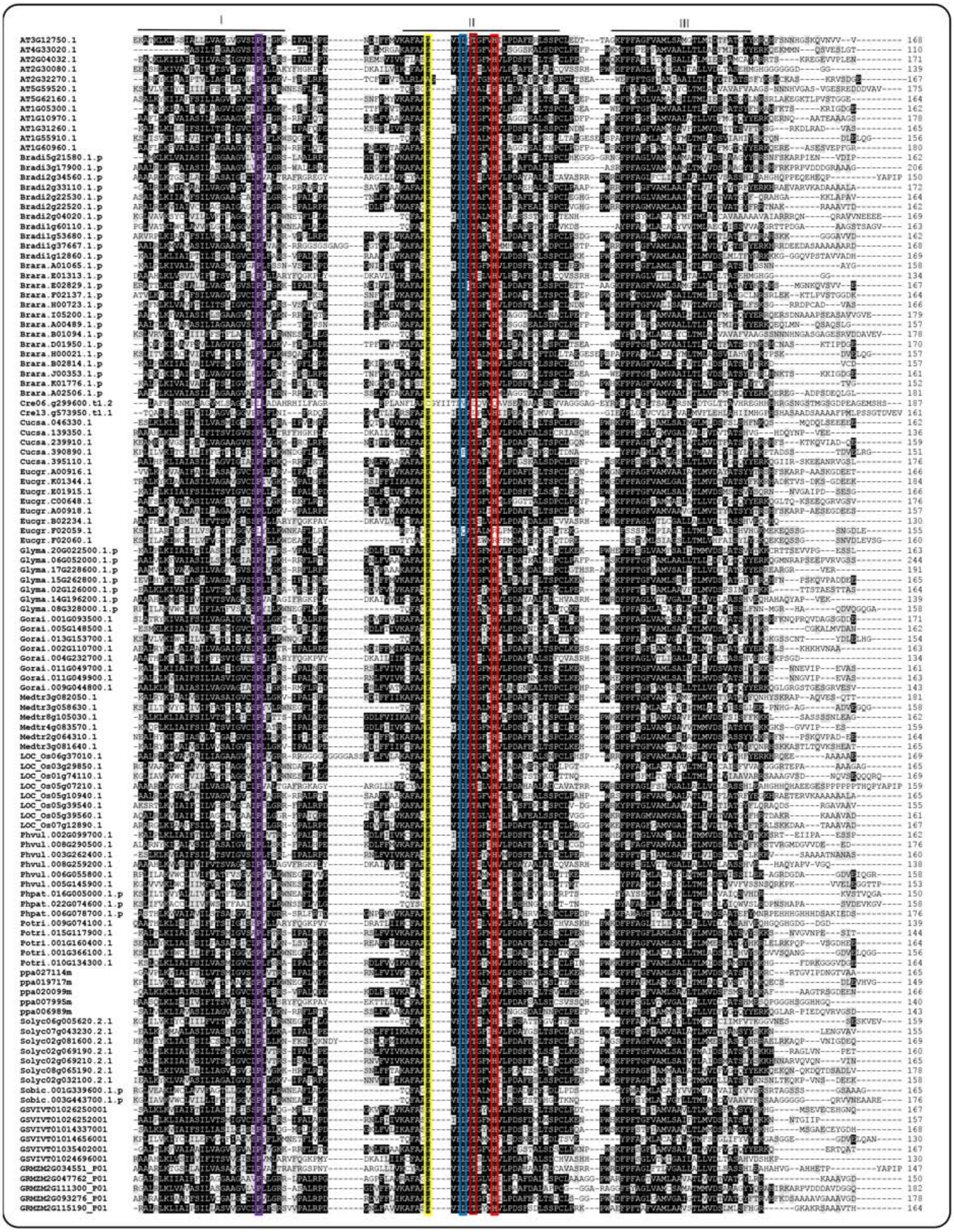

Figure 2. Sequence alignment of Zn transporter proteins in 18 plant species. Sequences were aligned by ClustalW, and identical and similar residues were shaded as black and gray, respectively. TM domains were indicated by lines and numbers above the sequences. Highly conserved residues were specified with rectangles in different colors: histidine with red, glycine with yellow, leucine with blue, phenylalanine with orange, proline with purple, threonine with dark red, and tyrosine with light green. The region between TM-3 and TM-4 shows the cytoplasmic variable region.

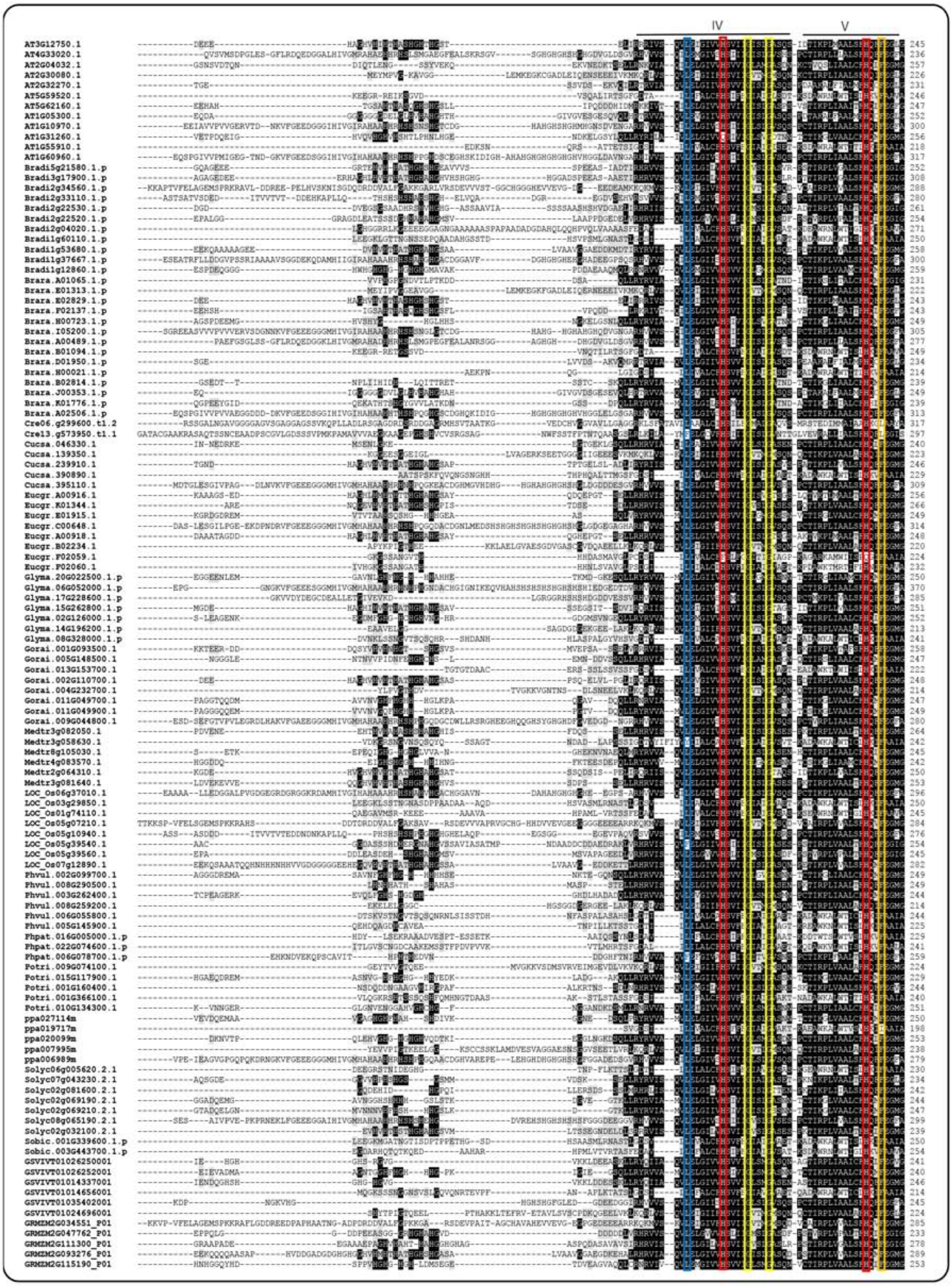

Figure 2. (Continued).

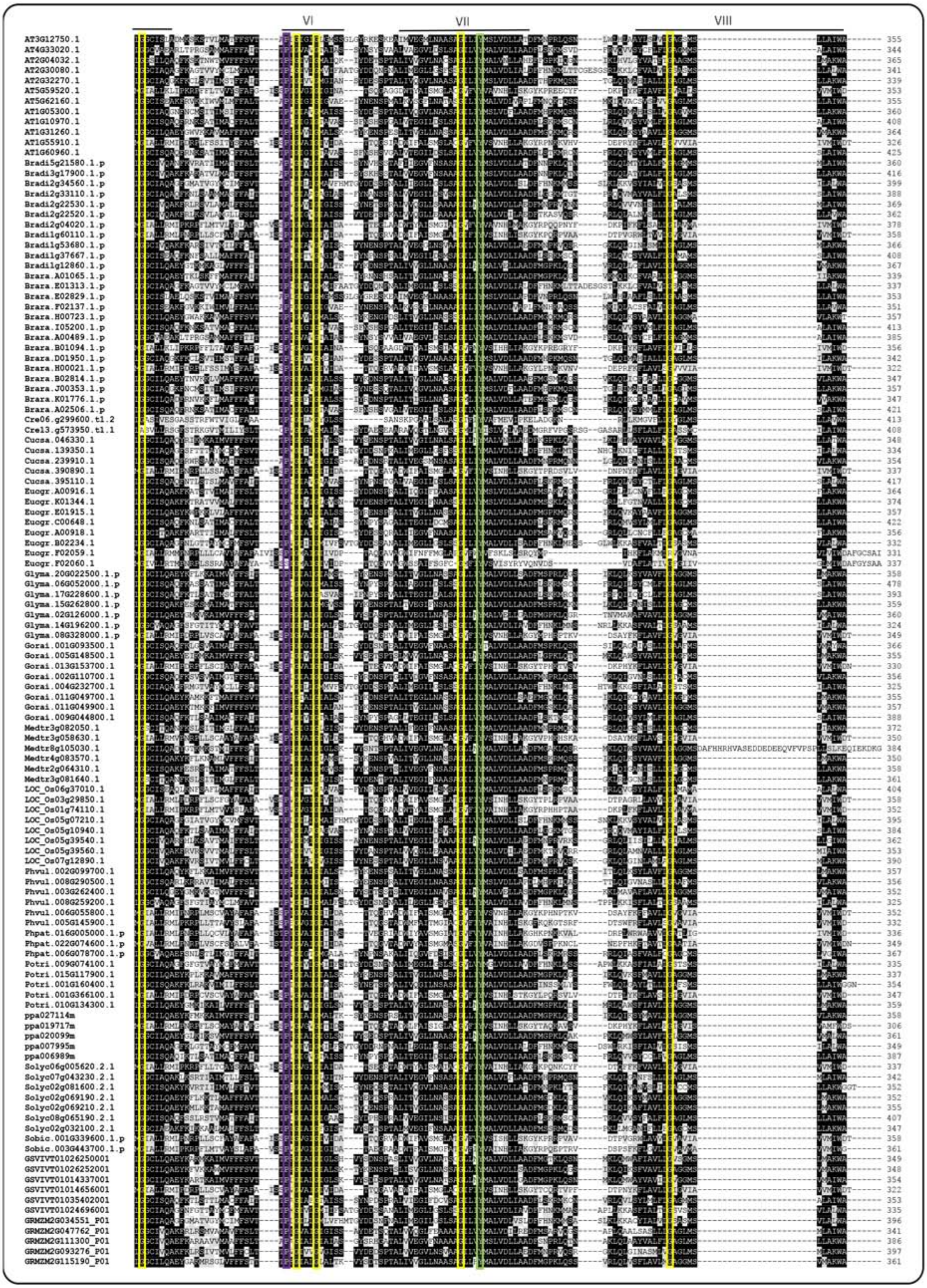

Figure 2. (Continued).

on general topological features of Zn transporter proteins in previous studies. Moreover, we specified the highly conserved residues on alignment with different colored rectangles: histidine with red, glycine with yellow, leucine with blue, phenylalanine with orange, proline with purple, threonine with dark red, and tyrosine with light green (Figure 2).

Studies have reported that ZIP proteins have similar topology with eight potential TM domains and they contain a variable cytoplasmic region between TM-3 and TM-4, including a potential His-rich metal-binding domain (Guerinot, 2000; Lopez-Millan et al., 2004; Li et al., 2013). In addition, TM-2, -4, and -5 were reported to contain a fully conserved histidine residue in all family members (Grotz et al., 1998; Guerinot, 2000). Particularly, sites of highly conserved histidine residues have been reported to have potential metal-binding activities (Guerinot, 2000). In aligned sequences, the cytoplasmic variable region between TM-3 and TM-4 has been clearly present. In addition, we have also identified the single row of fully conserved His residues in TM-2, -4, and -5. This implicitly indicates that these residues may have been universally conserved throughout all plant species. Moreover, these sites could also stand as potential metal-binding sites because of highly conserved His residues. Furthermore, we have manually searched the location of the three most conserved motifs in aligned sequences to find out the relationship between potential metal-binding sites and conserved motifs. Motif 1 was located within TM-4 and -5, motif 2 was located within TM-7, and motif 3 was located within TM-1 and -2. The presence of motifs 1 and 3 within the TM-2, -4, and -5 domains, which are predicted potential metal-binding sites, may indicate the well-conserved structure of metal-binding sites.

3.3. Phylogenetic analysis

Retrieved protein sequences were renamed along with their corresponding homologous gene symbols in *A. thaliana* or *O. sativa* because these two species have been experimentally well characterized. The homologous gene was used as a benchmark to analyze the clustering of protein sequences. We have constructed four separate phylogenetic trees (Figures 3A-3C and Figure 4). Three phylogenetic trees were constructed for *Arabidopsis* (Figure 3A), *Oryza* (Figure 3B), and *Arabidopsis*-*Oryza* (Figure 3C) to find out similarities of Zn transporter sequences between/among themselves. The main phylogenetic tree (Figure 4) was constructed from 112 potential Zn transporter sequences identified from 18 plant species.

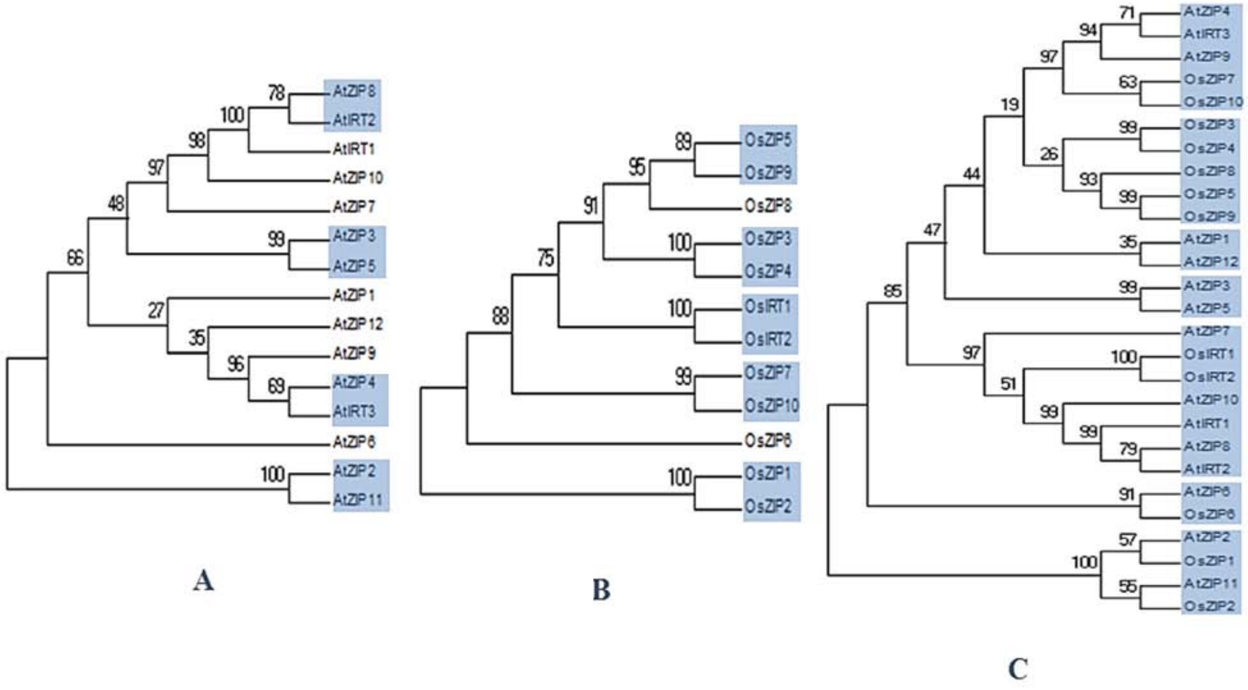

Figure 3. Phylogenetic trees of Zn transporter proteins in *A. thaliana* (A), *O. sativa* (B), and *Arabidopsis*-*Oryza* (C). Trees were constructed by MEGA5 software with the ML method for 1000 bootstrap values. Trees were used as benchmarks to analyze the clustering of Zn transporter sequences in 18 plant species. The *Arabidopsis* tree (A) was constructed with sequences of O81123.1, Q9LTH9.1, Q9SLG3.1, O04089.1, O23039.1, O64738.1, Q8W246.1, Q8S3W4.1, O82643.1, Q8W245.2, Q94EG9.1, Q9FIS2.1, Q38856.2, O81850.1, and QLE59.3 NCBI accession numbers. The *Oryza* tree (B) was constructed with sequences of Q94DG6.1, Q852F6.1, Q7XLD4.2, Q6ZJ91.1, Q6L8G0.1, Q6L8F9.1, Q6L8F7.1, A3BI11.1, Q0DHE3.3, Q5Z653.2, Q75HB1.1, and Q6L8G1.1 NCBI accession numbers.

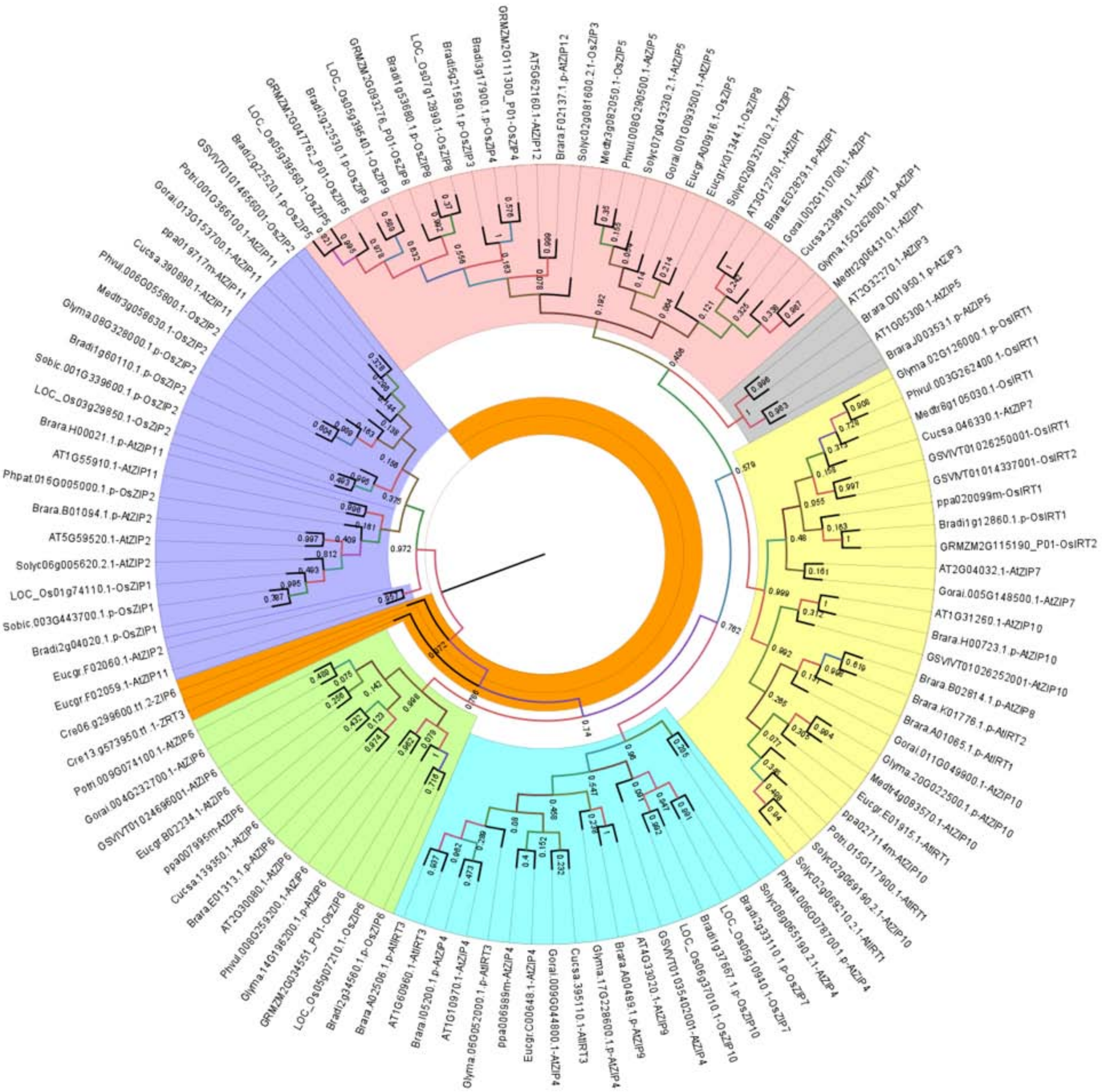

Figure 4. Phylogenetic tree of Zn transporter proteins in 18 plant species. The circular phylogenetic tree was constructed by MEGA5 software with the ML method for 1000 replicates bootstrap and visualized by FigTree. Blue, orange, green, cyan, yellow, gray, and pink colors represent groups A, B, C, D, E, F, and G, respectively.

The main tree was divided into seven major groups, namely A, B, C, D, E, F, and G (Figure 4). Group A (blue in the tree) included sequences that were homologous to AtZIP2 and -11 and OsZIP1 and -2. A monocot/dicot separation was not observed in this group. However, *E. grandis* sequences diverged from others with 0.972 bootstrap value. Trees of *Arabidopsis* and *Oryza* (Figures 3A–3C) showed that for AtZIP2 and -11 and OsZIP1 and -2 sequences are closely related between and among themselves, respectively. Thus, homologs of these sequences were clustered together in this group. Group B (orange in the tree) only contained *C. reinhardtii* sequences and this group did not demonstrate any particular similarity to *Arabidopsis* or *Oryza* sequences. This may indicate that *Chlamydomonas* Zn transporter genes may have independently evolved from other plant taxa during the evolutionary history. Group C (green in the tree) included AtZIP6/OsZIP6 homologs without any monocot/dicot separation. ZIP6 sequences were also observed to be clearly separated from other Zn transporter members in *Arabidopsis* and *Oryza* trees (Figures 3A–3C). This indicates that ZIP6 sequences could be very similar to each other and have a well-conserved structure in monocots and dicots. Group D (cyan in the tree) comprised AtZIP4, AtZIP9, AtIRT3, OsZIP7, and OsZIP10 homologs without any monocot/dicot separation. *P. patens* (moss) and *S. lycopersicum* sequences were found to diverge from the others with 0.96 bootstrap value. In *Arabidopsis* and *Oryza* trees (Figures 3A–3C), AtZIP4, AtZIP9, AtIRT3, OsZIP7, and OsZIP10 sequences were closely related between/among themselves. Moreover, AtIRT3 was also reported to be more similar to AtZIP4 (Grotz et al., 1998). We may thus report that AtZIP4 and AtIRT3 homologs are more similar to each other and closely related to AtZIP9, while sequences of OsZIP7 homologs are more similar to that of OsZIP10. Group E (yellow in the tree) was the second largest group with 14 species. AtZIP7, AtZIP8, AtZIP10, AtIRT1, AtIRT2, OsIRT1, and OsIRT2 homologs were clustered in this group without monocot/dicot separation. AtZIP7, AtZIP8, AtZIP10, AtIRT1, and AtIRT2 sequences in the *Arabidopsis* tree (Figure 3A) and OsIRT1 and OsIRT2 sequences in the *Oryza* tree (Figure 3B) have been observed to be similar to each other. Moreover, all these sequences were found to be closely related in the *Arabidopsis*-*Oryza* combined tree (Figure 3C). This explains why homologs of these sequences were closely clustered together in this group. Group F (gray in the tree) included two *A. thaliana* and *B. rapa* sequences that are homologous to AtZIP3 and AtZIP5. To mention a monocot/dicot separation for this group, it is necessary to construct more comprehensive phylogenies with more species. *Arabidopsis* and *Oryza* trees (Figures 3A–3C) also demonstrated that AtZIP3 and AtZIP5 are similar to each other, explaining the clustering of this group. Group G (pink in the tree) was the largest group in terms of species number. AtZIP1 and -12 and OsZIP3, -4, -5, -8, and -9 homologs were clustered in this group without monocot/dicot separation. AtZIP1 and AtZIP12 sequences in the *Arabidopsis* tree (Figure 3A) and OsZIP3 and OsZIP4 sequences in the *Oryza* tree (Figure 3B) were observed to be closely related. Additionally, in the *Arabidopsis-Oryza* combined tree (Figure 3C), AtZIP1 and -12 and OsZIP3, -4, -5, -8, and -9 sequences were clustered in the same clade, explaining the clustering of these sequence homologs in this group.

Overall, phylogenetic analysis implied that *Arabidopsis* and *Oryza* Zn transporter sequences could be used as references/benchmarks to identify the corresponding homologs of Zn transporters in various plant species.

In conclusion, we aimed to identify and comparatively analyze the Zn transporter genes and proteins in 18 different plant species by using bioinformatics tools in this study. Bioinformatics analyses showed that Zn transporters genes were well conserved and had similar physicochemical properties. Conserved TM and motif structures were also detected. The findings of this study will be valuable theoretical knowledge for future studies in terms of understanding the gene and protein features of Zn transporters in various plant species.

References

Assunção AG, Herrero E, Lin YF, Huettel B, Talukdar S, Smaczniak C, Immink RH, Eldik Mv, Fiers M, Schat H et al. (2010). *Arabidopsis thaliana* transcription factors bZIP19 and bZIP23 regulate the adaptation to zinc deficiency. P Natl Acad Sci USA 107: 10296-10301.

Connolly EL, Fett JP, Guerinot ML (2002). Expression of the IRT1 metal transporter is controlled by metals at the levels of transcript and protein accumulation. Plant Cell 14: 1347-1357.

Eng BH, Guerinot ML, Eide D, Saier MH Jr (1998). Sequence analyses and phylogenetic characterization of the ZIP family of metal ion transport proteins. J Membrane Biol 166: 1-7.

Fox TC, Guerinot ML (1998). Molecular biology of cation transport in plants. Annu Rev Plant Biol 49: 669-696.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005). Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The Proteomics Protocols Handbook. New York, NY, USA: Humana Press, pp. 571-607.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40: 1178-1186.

Grotz N, Fox T, Connolly E, Park W, Guerinot ML, Eide D (1998). Identification of a family of zinc transporter genes from *Arabidopsis* that respond to zinc deficiency. P Natl Acad Sci USA 95: 7220-7224.

Guerinot ML (2000). The ZIP family of metal transporters. BBA-Biomembranes 1465: 190-198.

Guo AY, Zhu QH, Chen X, Luo JC (2007). GSDS: a gene structure display server. Yi Chuan 29: 1023-1026 (in Chinese with abstract in English).

Ishimaru Y, Suzuki M, Kobayashi T, Takahashi M, Nakanishi H, Mori S, Nishizawa NK (2005). OsZIP4, a novel zinc-regulated zinc transporter in rice. J Exp Bot 56: 3207-3214.

Krogh A, Larsson B, von Heijne G, Sonnhammer ELL (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567-580.

Li S, Zhou X, Huang Y, Zhu L, Zhang S, Zhao Y, Chen R (2013). Identification and characterization of the zinc-regulated transporters, iron-regulated transporter-like protein (*ZIP*) gene family in maize. BMC Plant Biol 13: 114.

Lin YF, Liang HM, Yang SY, Boch A, Clemens S, Chen CC, Yeh KC (2009). *Arabidopsis IRT3* is a zinc-regulated and plasma membrane localized zinc/iron transporter. New Phytol 182: 392-404.

Lopez-Millan AF, Ellis DR, Grusak MA (2004). Identification and characterization of several new members of the zip family of metal ion transporters in *Medicago truncatula*. Plant Mol Biol 54: 583-596.

Marschner H (1995). Mineral Nutrition of Higher Plants. 2nd edition Boston, MA, USA: Academic Press.

Maser P, Thomine S, Schroeder JI, Ward JM, Hirschi K, Sze H, Guerinot ML (2001). Phylogenetic relationships within cation transporter families of *Arabidopsis*. Plant Physiol 126: 1646-1667.

Milner MJ, Seamon J, Craft E, Kochian LV (2013). Transport properties of members of the ZIP family in plants and their role in Zn and Mn homeostasis. J Exp Bot 64: 369-381.

Moreau S, Thomson RM, Kaiser BN, Trevaskis B, Guerinot ML, Udvardi MK, Puppo A, Day DA (2002). Gm*ZIP1* encodes a symbiosis-specific zinc transporter in soybean. J Biol Chem 15: 4738-4746.

Ramesh SA, Shin R, Eide DJ, Schachtman DP (2003). Differential metal selectivity and gene expression of two zinc transporters from rice. Plant Physiol 133: 126-134.

Romiti M (2010). Entrez nucleotide and entrez protein FAQs. Gene 1: 270.

Sonnhammer EL, Eddy SR, Durbin R (1997). Pfam: a comprehensive database of protein domain families based on seed alignments. Proteins 28: 405-420.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

Timothy L, Mikael BB, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37: 202-208.

Tiong J, McDonald G, Genc Y, Shirley N, Langridge P, Huang CY (2015). Increased expression of six ZIP family genes by zinc (Zn) deficiency is associated with enhanced uptake and root-to-shoot translocation of Zn in barley (*Hordeum vulgare*). New Phytol 207: 1097-1109.

Vert G, Briat JF, Curie C (2001). *Arabidopsis IRT2* gene encodes a root-periphery iron transporter. Plant J 26: 181-189.

Vert G, Grotz N, Dedaldechamp F, Gaymard F, Guerinot ML, Briat JF, Curie C (2002). IRT1, an Arabidopsis transporter essential for iron uptake from the soil and for plant growth. Plant Cell 14: 1223-1233.

Yu CS, Chen YC, Lu CH, Hwang JK (2006). Prediction of protein subcellular localization. Proteins 64: 643-651.