Genome-wide Identification and Expression Pattern Analysis of Zinc-finger Homeodomain Transcription Factors in Tomato under Abiotic Stress

Jingkang Hu, Yingmei Gao, Tingting Zhao, Jingfu Li, Meini Yao, and Xiangyang Xu¹

Horticultural Sciences, Northeast Agricultural University, Harbin 150030, Heilongjiang, China

ADDITIONAL INDEX WORDS. expression characteristics, gene structures, *Solanum lycopersicum*, stress resistance, subfamily distribution, ZF-HD

ABSTRACT. Members of the zinc-finger homeodomain (ZF-HD) family play a key role in the control of plant growth and development, which are involved in plant responses to stress. Although many functional studies of this gene family have been performed in different plants, the features of this family in tomato (*Solanum lycopersicum*) remain unknown. In this study, we identified 22 ZF-HD genes in the tomato genome and classified them into seven groups located on six chromosomes. Expression of 15 ZF-HD genes in tomato was studied in different tissues to identify their putative functions in many aspects of plant growth and development. Based on previous phylogenetic analyses in arabidopsis (*Arabidopsis thaliana*), our results showed that some tomato SL-ZH (*S. lycopersicum* zinc-finger homeodomain) genes cluster into the same neighbor-joining (NJ) branch as arabidopsis, indicating that these genes may share similar structures and functions in these plants. Gene expression analysis demonstrated that the tomato ZF-HD gene may be involved in abiotic stress responses, the SL-ZH13 gene in cold stress and the SL-ZH15 gene in drought stress; almost all tomato ZF-HD genes were responsive to salt stress, except for SL-ZH7, -ZH8, and -ZH22. However, the structures and functions of unknown groups require further research. In conclusion, this study identified tomato ZF-HD genes and analyzed their gene structures, subfamily distribution, and expression characteristics. These experiments combined with previous research findings reveal significant information and insight for future studies on the agronomic features and stress resistance in tomato.

Plant growth is systematically regulated by a system of specialized genes that encode various proteins, with certain genes being capable of influencing vegetative growth and stress tolerance, including cold, drought, and salinity. Transcription factors (TFs), which bind to specific nucleotide sequences, have been shown to play important roles in the regulatory networks of numerous developmental processes in plants. The first TF was discovered in maize (*Zea mays*), and a large number of TFs in vascular plants were later identified (Huang et al., 2015). For example, the GRAS family regulates development in plants (Song et al., 2017), and the MYB family can regulate fruit color (Czemmel et al., 2012; Espley et al., 2007). TFs contain motifs for activation or repression of target genes through direct binding to gene regulatory elements or motifs. Therefore, different TF families have evolved their unique DNA-binding domains, which confer binding specificity.

The homeodomain (HD) is a conserved DNA-binding domain of ≈60 amino acids (Burglin, 1994), which folds into a characteristic three-helix structure that acts as a recognition helix and can pack against the DNA major groove to promote specific contact (Gehring et al., 1994; Wolberger, 1996). The HD domain binds DNA through an unordered N-terminal arm on helix I and interactions made by helix III (Akin and Nazarali, 2005; Hunter and Rhodes, 2005). Several HD-containing proteins/TFs are known to play crucial roles in plant or animal development (Ito et al., 2002; Williams, 1998); however, plants

have evolved the specific HD-Zip TF family (Ariel et al., 2007), the members of which bear a unique leucine zipper domain at the C-terminus. Zinc fingers, which consist of two pairs of conserved cysteine and/or histidine residues coordinating a single zinc ion to form a finger-shaped loop (Klug and Schwabe, 1995) are necessary motifs that are found widely in regulatory proteins (Krishna et al., 2003; Takatsuji, 1999). There are many categories of zinc fingers, such as those containing one (C₂H₂, C₂C₂, and C₃H) or two (ring finger in plants and the Lin/Is1/Mec domain in animals) zinc ions (Englbrecht et al., 2004; Halbach et al., 2000; Kosarev et al., 2002; Li et al., 2001; Yanagisawa, 2004). One protein can have one or many zinc fingers and domains. There are six subfamilies of HD-containing proteins, as determined via advanced studies of model plants, such as rice (*Oryza sativa*) and arabidopsis. These subfamilies include PHD (connected to a finger domain), Bell (based on the distinctive Bell domain), WOX (associated with Wuchel), KNOX (including knotted), HD-Zip (containing a leucine zipper), and ZF-HD (containing a zinc finger) (Ariel et al., 2007).

The ZF-HD TFs were first identified by Windhövel et al. (2001) in the C4 plant *Flaveria bidentis*. These proteins influence transcription of the C4 phosphoenolpyruvate carboxylase gene. ZF-HD proteins have a special structure, including a conserved domain at the N-terminal that consists of cysteine and histidine residues and coordinates zinc and canonical homeodomains associated with the C-terminal domain (Windhövel et al., 2001). A number of studies in plants have focused on ZF-HD TFs and model organisms including soybean [*Glycine max* (Park et al., 2007)], arabidopsis (Hong et al., 2010), and rice (Figueiredo et al., 2012). Arabidopsis contains 14 ZF-HD family members, some of which can regulate floral development (Windhövel et al., 2001). In addition, three mini zinc

Received for publication 22 Aug. 2017. Accepted for publication 25 Oct. 2017. This work was supported by National Key R&D Program of China (2017YFD0101900), the China Agriculture Research System (CARS - 25 - A - 15) and the Breeding of New Staple Vegetable Varieties of Heilongjiang Province (GA15B103).

¹Corresponding author. E-mail: xxy709@126.com.

finger (MIF) genes have been found in arabidopsis; even lacking the HD domain, these genes present sequence similarity with the ZF domain of ZF-HD proteins (Hu and Ma, 2006; Hu et al., 2008). Nonetheless, the connection between ZHD and MIF genes and the evolutionary pathway remain unclear to date. ZF-HD genes were renamed ZHD according to arabidopsis gene nomenclature guidelines and review of the history of the family. Arabidopsis ZF-HD proteins can form homo- and heterodimers through binding to the core DNA sequence consensus ATTA (Tan and Irish, 2006). It has been reported that an arabidopsis ZHD protein named AtZHD1 can bind specifically to the promoter of Early Response To Dehydration Stress1 (ERD1), which is related to salt, dehydration, and abscisic acid stress, demonstrating that AtZHD1 can be influenced by adverse conditions (Tran et al., 2007). Recently, research in different plants, such as soybean and rice, has expanded to ZHD functions. On pathogen stimulation, expression of two genes in soybean (GmZF-HD1 and GmZF-HD2) is upregulated (Park et al., 2007), and four rice ZHD genes have also been implicated in gene regulation (Figueiredo et al., 2012). This research has been done on the role of ZF-HD under different abiotic stresses in tomato. Tomato is one of the most popular crops worldwide; it presents good flavor characteristics and high nutritional value. With the publication of the entire genome sequence of tomato, a whole gene family can now be conveniently studied (Consortium, 2012). In our present study, we identified 22 tomato ZHD genes by downloading their sequences and analyzing their gene structures, chromosomal locations, gene duplicates, and evolutionary mechanisms. We also discuss the kinetics of the corresponding transcription patterns in response to different stress treatments (cold, drought, and salt), as assessed by quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR). We report that some ZF-HD genes are connected to tissue growth and

some gene expression was controlled by different abiotic stress reactions. What is interesting is an unknown (UN) subfamily in tomato. These findings contribute to a more comprehensive understanding of the expansion and evolution of tomato.

Materials and Methods

IDENTIFICATION OF ZF-HD GENES IN TOMATO. To identify ZF-HD genes in the tomato genome, ZF-HD sequences were downloaded from the Sol genomics network (SOL, 2015), and 22 genes were found when searched against the Pfam ZF-HD Hidden Markov Model [PF04770 (Artimo et al., 2012)]. Several tools [e.g., HMMER (European Bioinformatics Institute, 2015), NCBI (National Center for Biotechnology Information, 1988)] were used to identify putative ZF-HD proteins in tomato (Finn et al., 2011).

ANALYSIS OF PHYLOGENY, CONSERVED MOTIFS, AND GENE STRUCTURE. ClustalX and MEGA5 software were used to construct NJ distance trees by comparing tomato ZF-HD and arabidopsis protein domain sequences (Tamura et al., 2011). The details of arabidopsis ZF-HD sequences were searched in the Arabidopsis Information Resource (Artimo et al., 2012; Phoenix Bioinformatics Corporation, 2013). The online tool Multiple Expectation-maximizations for Motif Elicitation (MEME) (National Institutes of Health, 2015) was used to identify conserved ZF-HD motifs in tomato (Bailey et al., 2009) using the default parameters, except that the maximum number of motifs was set to 12. The structure of tomato ZF-HD genes was analyzed by the online program GSDS (Center for Bioinformatics, 2015), which displays coding sequence (CDS) and DNA sequences (Guo et al., 2007).

CHROMOSOMAL LOCATION AND DUPLICATION OF ZF-HD GENES. To determine the physical locations of ZF-HD genes on tomato chromosomes, we used Phytozome (Energy's Joint Genome Institute, 1997) to identify the initial site of each ZF-HD gene. MapInspect software was used to generate images of ZF-HD gene locations (Song et al., 2014). We aligned 22 ZF-HD genes in tomato with two requirements to identify potential gene duplications: an alignable nucleotide sequence covering over 70% of the longer sequence and identity between amino acid sequences over 70% (Zhou et al., 2004).

ANALYSIS OF ZF-HD GENES IN DIFFERENT TISSUES AND ISOLATION OF RNA AND QUANTITATIVE REAL-TIME PCR ANALYSIS. The tomato cultivar Money-maker cultivated at the Northeast Agriculture University Horticulture Experimental Station (Harbin, China) was used for the experiments. Seeds were sown in vermiculite mixture soil (3:1) with 60% relative humidity (measured by soil hygrometer) at the greenhouse of Northeast Agricultural University. We chose 69 healthy seedlings for subsequent analysis (15 seedlings for tissue analysis and 54 seedlings for cold, salt, and drought treatments). Five tissues of the tomato plants, the root, stem, leaf,

Table 1. Characteristics of zinc-finger homeodomain (ZF-HD) genes in the tomato genome.^z

Gene locus	Gene name	Length (aa)	MW (Da)	PI	Gravy	AI	DL
Solyc01g014970	SL-ZH1	229	25,789.04	8.12	-0.713	71.57	21-69
Solyc01g102980	SL-ZH2	290	32,600.98	7.76	-0.759	62.52	64-119
Solyc01g103810	SL-ZH3	116	13,144.06	9.46	-0.364	58.02	14-67
Solyc01g103820	SL-ZH4	94	10,361.66	8.27	-0.463	50.85	13-67
Solyc01g103830	SL-ZH5	210	24,492.28	9.68	-0.702	72.43	11-64
Solyc01g103840	SL-ZH6	84	9,529.72	6.27	-0.405	63.81	11-64
Solyc02g067310	SL-ZH7	292	32,086.75	8.20	-0.871	53.46	62-119
Solyc02g067320	SL-ZH8	331	36,516.54	7.75	-0.858	54.53	53-110
Solyc02g067330	SL-ZH9	102	11,208.36	6.18	-0.467	53.53	30-85
Solyc02g085160	SL-ZH10	293	33,312.95	8.14	-1.103	44.61	90-145
Solyc02g087970	SL-ZH11	90	10,104.24	9.07	-0.797	42.22	22-76
Solyc03g061620	SL-ZH12	88	9,648.53	5.76	-0.855	42.16	23-79
Solyc03g098060	SL-ZH13	179	19,951.89	8.47	-0.936	44.75	8-62
Solyc03g116070	SL-ZH14	83	9,037.14	8.75	-0.436	59.88	21-74
Solyc04g014260	SL-ZH15	247	27,722.19	7.16	-0.832	62.83	54-106
Solyc04g074990	SL-ZH16	143	16,096.89	5.97	-0.916	51.19	44-100
Solyc04g080490	SL-ZH17	290	31,547.28	8.72	-0.767	56.17	58-112
Solyc05g007580	SL-ZH18	297	32,814.23	7.26	-0.926	47.31	51-109
Solyc05g018740	SL-ZH19	119	13,674.54	9.30	-0.555	61.43	14-67
Solyc05g020000	SL-ZH20	119	13,702.57	9.44	-0.510	58.15	14-67
Solyc05g051420	SL-ZH21	167	19,300.40	9.75	-0.751	65.93	3-51
Solyc09g089550	SL-ZH22	805	90,496.49	6.97	-0.587	68.29	554-608

^zaa = amino acid; MW = molecular weight; Da = Dalton; PI = isoelectric point; Gravy = grand average of hydropathy; AI = aliphatic index; DL = domain location.

and a final step at 72 °C for 2 min. Relative expression was calculated using the $2^{-\Delta\Delta}$ superscript method, and the results were illustrated using GraphPad Prism 6 (Kenan et al., 2017) software and OmiShare Tools (Gene Denovo, 2015).

Results

IDENTIFICATION AND CLASSIFICATION OF ZF-HD GENES IN TOMATO. In this study, 22 tomato ZF-HD genes were identified using the programs mentioned earlier. The genes were named SL-ZH1 to SL-ZH22, and the amino acid sequence used to identify ZF-HD gene characteristics in the tomato genome is shown in Table 1. The length of the inferred ZF-HD proteins ranged from 83 (SL-ZH14) to 331 (SL-ZH8) amino acids, except for SL-ZH22, which is 805 amino acids in length and contains two DUF4283 domains that are unrelated to ZF-HD. The molecular weights vary from 9037.14 (SL-ZH14) to 90496.49 (SL-ZH22), and PI values range from 5.76 (SL-ZH12) to 9.75 (SL-ZH21). Thus, different ZF-HD proteins carried out different functions in different microenvironments.

The evolutionary relationship between tomato and arabidopsis was assessed with respect to ZF-HD TFs and visualized with a NJ phylogenetic tree as shown in Fig. 1. The subfamilies assigned are based on earlier studies (Wang et al., 2016). Finger-homeodomain proteins (ZHD) II and ZHD V only have three members; the largest subfamily is MIF, with 12 members. As shown in Fig. 1, 9 ZF-HD genes that are homologous to arabidopsis genes were identified. Thus, there are some tomato homogenous ZF-HD proteins that were regarded as species-specific in previous publications. Three genes (SL-ZH1, SL-ZH5, and SL-ZH6) were clustered in an unknown subfamily, the function of which requires further investigation. These results suggest that a number of these ZF-HD genes from tomato may have the same function as the arabidopsis gene, whereas other genes may have species-specific functions in tomato, such as genes in the UN subfamily.

CONSERVED DOMAINS AND MOTIFS OF THE ZF-HD FAMILY AND STRUCTURE IN TOMATO. The DNAMAN tool was used for multiple sequence alignment of ZF-HD proteins to identify domain structures. Sequence conservation was verified by the WebLogo online tool. ZF-HD proteins possess two conserved domains: the ZF and HD domains as shown in Fig. 2. ZHD genes harbor both the ZF and HD domains, but MIF genes only have the ZF domain. Based on the results, all ZHD homeodomains have a conserved W as the 25th residue and an N as the 55th, similar to a previous report (Hu et al., 2008).

In this study, we used MEME to identify conserved motifs among tomato ZF-HD proteins as shown in Fig. 3. We

Fig. 1. The neighbor-joining phylogenetic tree of tomato and arabidopsis zinc-finger homeodomain (ZF-HD) genes. Seven subfamilies were identified in tomato: ZHD I, ZHD II, ZHD III, ZHD IV, ZHD V, mini zinc finger (MIF), and an unknown subfamily (UN).

flower, and fruit, were analyzed, and three tomato tissue samples were obtained from three different plants. As cold treatment, seedlings at the four-leaf stage were moved to growth chambers set at 4 °C under the same light treatment and daylength. For abiotic stresses, cold, salt, and drought treatments, seedlings in 100-mL conical flasks were grown in nutrient solution for 24 h under normal growth conditions [the experiment was carried out in a greenhouse under a 16/8 h (day/night) photoperiod, a photon flux density of 120 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, day and night temperature of 20 °C, and 65% relative humidity] and then irrigated with 250 mM NaCl and 15% polyethylene glycol, respectively, for 24 h. Each treatment had three seedlings, and all the treatments had three biological replicates. At 1.5, 3, 6, 12, and 24 h after treatment, new tomato leaf samples were gathered and stored at -80 °C for analysis. Total RNA was isolated from 100 mg of frozen tomato leaves using the TRIzol method (Rio et al., 2010), and each RNA sample was reverse transcribed into cDNA using a reverse transcription system (TransGen Biotech, Beijing, China). Primers for tomato ZF-HD genes were designed using Primer 5.0 (Wang et al., 2016) based on tomato transcriptome sequencing data. The qRT-PCR experiments were executed with three technical and biological replicates using an IQ5 (Shanghai Zhiyan Scientific Instrument Co., Shanghai, China) system in a 20- μL reaction system. The fluorescent dye 2 $\times$ Power SYBR Green PCR Master Mix (Vazyme, Nanjing, China) was used, and the reaction protocol was as follows: 95 °C for 6 min; followed by 40 cycles of 95 °C for 10 s, 58 °C for 20 s, and 72 °C for 15 s;

Fig. 2. The zinc finger (ZF) domain of the ZHD subfamily, homeodomain (HD) of the ZHD subfamily, and ZF domain of the mini zinc finger (MIF) subfamily. All conserved domain can be matched with the WebLogo picture, the size of amino acid show the conservatism of domain, and the bigger, the more conservative.

identified 12 motifs, referred to as motif 1 through motif 12. Most of the closely related members in the NJ tree have similar motif compositions. For example, the MIF group harbors motif 1 and motif 2 but lacks motif 3 and motif 4. This result is consistent with that of previous research showing that the MIF subfamily has a ZF domain but lack HD. Almost all ZF-HD members possess motif 2, demonstrating that this motif may comprise the ZF domain. The different motifs distributed among different subfamilies of ZF-HD genes provide evidence of their functional divergence.

The number of introns varies from one to three, as shown in Fig. 3, and this finding is consistent with ZF-HD genes in arabidopsis. SL-ZH22 has the longest CDS and intron region of all tomato ZF-HD genes, and it is also the only gene in the MIF subfamily with introns. In fact, we found that almost all ZF-HD genes lack an intron, which is consistent with previous research (Hu et al., 2008).

CHROMOSOMAL LOCATIONS AND GENE DUPLICATION ANALYSIS OF ZF-HD GENES. Gene duplication plays a key role in genomic expansion and consists of two types: tandem and segmental. Two or more duplicated genes appearing on the same chromosome are defined as a tandem duplication, whereas other types are defined as segmental duplications (Song et al., 2014). We detected four tandem duplications (SL-ZH3 and SL-ZH4, SL-ZH5 and SL-ZH6, SL-ZH7 and SL-ZH8, SL-ZH19 and SL-ZH20) on three

chromosomes as shown in Fig. 4. A gene cluster is a chromosomal region with two or more genes within 200 kb of the sequence (Holub, 2001), and we found seven ZF-HD genes that form two gene clusters: SL-ZH3–6 and SL-ZH7–9 located on Ch1 and Ch2, respectively. No clusters were found on Ch3, Ch4, Ch5, or Ch9. In addition, tomato ZF-HD genes display an uneven distribution across six chromosomes (Ch1, Ch2, Ch3, Ch4, Ch5, and Ch9). Ch1 carries the largest number of ZF-HD genes ($n = 6$), whereas 5, 3, 3, 4, and 1 ZF-HD genes were found on Ch2, Ch3, Ch4, Ch5, and Ch9, respectively. Our chromosomal location and gene duplication analysis indicated disproportionality with regard to the number of genes on chromosomes and the manner of gene duplication.

EXPRESSION ANALYSIS OF ZF-HD GENES IN DIFFERENT TISSUES. Expression patterns of 15 ZF-HD genes in six tissues of tomato ‘MoneyMaker’ plants were investigated by qRT-PCR. The expression levels of these tomato genes were clustered, and the results are presented as a heatmap as shown in Fig. 5. For 22 tomato ZF-HD genes, transcripts of 15 genes were obtained from at least one tissue. As opposed to arabidopsis (Windhövel et al., 2001), not all tomato ZF-HD genes

were found to be expressed in floral tissues as shown in Fig. 5. Moreover, expression of all 15 genes was low in the stem and fruit, with striking differences in different tissues. Several genes such as SL-ZH14 and SL-ZH16 were only expressed in the root and leaf, respectively, and can be considered tissue-specific. SL-ZH21 was highly expressed only in the petal, showing that it may be involved in development of the petal. Several genes were highly expressed in the root (8), stem (4), leaf (7), and petal (5), which may indicate that these genes not only participate in the formation of these tissues but are also involved in development and growth.

EXPRESSION ANALYSIS OF ZF-HD GENES UNDER DIFFERENT ABIOTIC STRESSES. The mode of gene expression is often connected with gene function. To determine whether ZF-HD genes in leaves respond to abiotic stress, qRT-PCR was used to evaluate transcript levels under cold, drought, and salt stress conditions.

We chose 15 of the 22 ZF-HD genes and examined expression. Under cold conditions, almost all genes were downregulated, with the largest change observed for SL-ZH2 and SL-ZH16; three genes (SL-ZH6, SL-ZH8, and SL-ZH15) were found to be hyperresponsive to cold as shown in Figs. 6 and 7. Following drought stress, three genes (SL-ZH1, SL-ZH2, and SL-ZH19) were downregulated, three genes (SL-ZH8, SL-ZH15, and SL-ZH16) were upregulated,

Fig. 3. Phylogenetic relationships, conserved motif compositions, and gene structure of tomato and arabidopsis zinc-finger homeodomain (ZF-HD) proteins. The neighbor-joining tree of tomato and arabidopsis ZF-HD genes and their motif locations. Motifs 1–12 are indicated by different colored boxes. The order of the motif corresponds to their positions in individual protein sequences. Introns, exons (CDS), and upstream/downstream are displayed as black, yellow, and blue lines, respectively.

Fig. 4. Positions of zinc-finger homeodomain (ZF-HD) gene members along tomato chromosomes (ch). Genes with tandem duplications are shown in different colors, and gene clusters were marked by black lines.

four genes (SL-ZH10, SL-ZH17, SL-ZH20, and SL-ZH21) showed an expression trend of rising and then falling, and three genes (SL-ZH7, SL-ZH14, and SL-ZH22) exhibited no change as shown in Figs. 6 and 7.

ZF-HD genes, except for SL-ZH7, SLZH-8, and SL-ZH22, were considerably downregulated using salt treatment at 6 and 12 h, and low expression was observed at other time points. Regarding salt treatment, limited changes were observed in ZF-HD genes compared with the levels observed with cold and drought treatments as shown in Figs. 6 and 7. These results demonstrate that ZF-HD genes in tomato are closely associated with salt stress, which indicates that these tomato TFs are extremely salt-sensitive families.

Discussion

Tomato is an important plant cultivated worldwide. However, multiple stresses have negative effects on the production and quality of tomatoes, and the ZF-HD family is important for plant growth and resistance to stress conditions. ZF-HD genes have been discovered and classified in a number of plants, such as arabidopsis, chinese cabbage

Fig. 5. Heatmap representation of tomato zinc-finger homeodomain (ZF-HD) genes in the root, stem, leaf, petal, and fruit. The bar on the right of heat map represents the relative expression value.

(*Brassica pekinensis*), and grape (*Vitis vinifera*). However, the evolution and function of tomato ZF-HD genes are poorly understood, especially in response to abiotic stresses. Accordingly, we undertook this experiment to identify ZF-HD genes of tomato and describe the features of their structure, expression, and evolution as well as their relationship with abiotic stresses. In previous studies, ZF-HD genes have been identified as plant-specific genes encoding a group of TFs, including land plants covering a large number of seed plants as well as *Selaginella* and *Physcomitrella*, with vascular and nonvascular properties, respectively. Conversely, these genes are not observed in other organisms, such as yeast, prokaryotes, and the single-celled alga *Chlamydomonas*. These findings demonstrate that ZF-HD genes came from a common ancestor who originated after the differentiation of land plants and evolved through land plant lineages (Bhattacharya and Medlin, 1998; Karol et al., 2001). In this analysis, more ZF-HD proteins were identified in tomato than in arabidopsis.

In this study, we identified all 22 ZF-HD genes in the tomato genome, including 9 ZHD genes and 10 MIF genes. Although the tomato genes present a similar classification as arabidopsis ZF-HD genes, the number of genes in each tomato group was different from that in arabidopsis, thus demonstrating that different plants bear variable duplications. ZHD and MIF genes have a common ZF domain, although MIF genes do not harbor

an HD domain, indicating that MIF genes originated from ZHD genes via elimination of the HD domain or that ZHD genes were derived from MIF genes via addition of an HD domain. Moreover, our exon and intron analysis showed that 13 of the ZF-HD genes in tomato are intronless as shown in Fig. 3; by contrast, all 17 genes in arabidopsis are intronless (Hu et al., 2008). This result indicates that ZF-HD TF genes have a close evolutionary connection with ZF-HD proteins. ZF-HD is not the only gene family with intronless features. For example, introns are not observed in F-box TFs (Jain et al., 2007) or in the small Auxin-up RNA family (Jain et al., 2006) and DEAD-box RNA family (Aubourg et al., 1999). A lack of introns is a classic feature of prokaryotic genomes, and there are three explanations for this phenomenon: retrocession of an intronless gene, derivation of a gene from ancient prokaryotes (horizontal transfer), and duplication of an existing intronless gene (Zou et al., 2011). In a study by Zhang, horizontal gene transfer was used to explain the beginning of plant GRAS genes from the prokaryotic genomes of bacteria (Zhang et al., 2012); therefore, ZF-HD genes may be similar to GRAS genes.

Tandem, segmental, and whole genome duplications have played an important role in the evolution of many plants. The gene duplication analysis performed in this study identified four groups of tandem duplication events (SL-ZH3 and SL-ZH4, SL-ZH5 and SL-ZH6, SL-ZH7 and SL-ZH8, SL-ZH19 and SL-ZH20) and 2 ZF-HD gene clusters (SL-ZH3–6 and SL-ZH7–9). The results showed that tandem duplications and gene clusters primarily contribute to the conservation and expansion of the ZF-HD TF family in tomato, thereby resulting in their quantitative and functional diversification. To contribute genomic knowledge about plants, Lyons performed a comparison between genes with well-characterized features and genes of an infrequently studied plant, *Chaenomeles cathayensis* (Lyons et al., 2008). This process represents a rapid and effective method of genomic comparison, and it was used between the model plant arabidopsis and tomato to predict the functions of ZF-HD genes in tomato. Certain gene groups, such as SL-ZH7 and SL-ZH8 (highly expressed in leaf) and SL-ZH14, SL-ZH19, and SL-ZH20 (highly expressed in root), and their expression patterns were found to be similar to those observed in previous studies, such as Vv-ZHD gene expression. For example, the AtZHD10 and SL-ZH7 proteins are similar, as indicated by their presence within the same phylogenetic group (ZHD IV). AtZHD10 is expressed preferentially in inflorescences and seed tissues (Hu et al., 2008). All 15 genes had low expression in stem and fruit;

Fig. 6. Expression of tomato zinc-finger homeodomain (ZF-HD) genes visualized as heatmaps exposed to different stresses: (A) cold stress, (B) drought stress, (C) salt stress. The bar on the right of heatmap represents the relative expression value; C = cold, D = draught, S = salt, CK = contrast.

this proved that ZF-HD gene have no influence in the development of stem and fruit, or a little. Compared with roots and fruit, the closely related SL-ZH7 showed a higher expression level in leaves and flowers, which is similar to SL-ZH8 expression pattern. SL-ZH14 and SL-ZH16 can be considered as root and leaf specific. SL-ZH21 is a key gene in development of the petal according to the results. These genes not only showed interesting features in different tissues, but also improved that there are some differences between arabidopsis and tomato because the arabidopsis ZF-HD family members are all expressed in floral tissues (Windhövel et al., 2001), tomato does not. This indicated that ZF-HD genes in tomato have its own principle to regulate the development and growth of tissues, may be more diversified.

In this study, phylogenetic analysis of the tomato and arabidopsis genomes showed that five pairs of ZF-HD genes (AtZHD1, AtZHD2, and SL-ZH16; AtZHD3, AtZHD4, and SL-ZH2; AtZHD8–12, and SL-ZH7, SL-ZH 8, and SL-ZH 18; AtZHD13, AtZHD14, and SL-ZH13; ATMIF1–3 and SL-ZH14) belong to a syntenic genomic region as shown in Fig. 1; other genes may correspond to other AtZHD genes. The lack of clustering of SL-ZH1, SL-ZH5, and SL-ZH6 into any syntenic group does not indicate that they are stand-alone genes. A number of experiments have been performed to identify the functions of arabidopsis genes, such as AtZHD1 (Tran et al., 2007), AtZHD5 (Hong et al., 2010), and AtZHD12 (Tan and Irish, 2006), which may help in clarifying the possible

Fig. 7. Expression analysis of 15 zinc-finger homeodomain (ZF-HD) genes in response to abiotic stress, CK = contrast, 1.5, 3, 6, 12, 24 = treat time (unit = hour). The x-axis and y-axis represent the name of the ZF-HD gene and various abiotic treatments, respectively.

functions of tomato ZF-HD. For example, a comparison with AtZHD1 indicates that SL-ZH16 may have similar functions in controlling floral structure because these genes are in the same syntenic group.

TFs are involved in the regulation of development and in the response to abiotic stresses in plants, and ZF-HD TFs exhibit similar functions. Nakashima reported that expression of dehydration in transgenic *Arabidopsis* can be promoted by overexpression of both ZF-HD and NAC proteins (Nakashima and Yamaguchi-Shinozaki, 2006). We found that SL-ZH13 is sensitive to cold stress and that SL-ZH15 is highly expressed during drought, results that may demonstrate that SL-ZH13 is involved in cold responses and that SL-ZH15 is involved in drought responses. Moreover, two of the SL-ZH genes (16 and 21) show relatively high expression levels in response to salt or drought, and all SL-ZH genes were upregulated by salt treatment. This result is consistent with that of a previous study reporting that drought resistance can be enhanced by overexpression of NAC and AtZHD1 and that salt stress can induce AtZHD1 expression (Tran et al., 2007). We found that SL-ZH2 was highly expressed under salt treatment (12 h), which showed

that the SL-ZH2 gene might be involved in salt-stress regulation because it is homologous to the AtZHD4 gene. Moreover, SL-ZH1 and SL-ZH6 also exhibited high levels of expression in response to salt stress but were hyperresponsive to cold and drought, indicating that the UN subfamily in tomato not only possesses the basic features of ZF-HD genes but also traits that are unknown and remain to be discovered.

To date, many functions of ZF-HD TFs have been characterized in plants such as *Arabidopsis*, grape, and soybean, yet more ZF-HD gene family members in other plants remain to be further investigated. In this study, 22 ZF-HD TFs were analyzed in tomato. Further studies of ZF-HD protein structure will reveal information regarding the functions of these genes, and studies of ZF-HD gene phylogeny and expression can contribute to the synthetic functional characterization of this family to identify their putative roles in abiotic stress. Our study is a comprehensive analysis of ZF-HD TFs in tomato, the results presented here provide insight for future investigations of the agronomic features of and stress resistance in tomato.

Literature Cited

- Akin, Z.N. and A.J. Nazarali. 2005. Hox genes and their candidate downstream targets in the developing central nervous system. *Cell. Mol. Neurobiol.* 25:697–741.
- Ariel, F.D., P.A. Manavella, C.A. Dezar, and R.L. Chan. 2007. The true story of the HD-Zip family. *Trends Plant Sci.* 12:419–426.
- Artimo, P., M. Jonnalagedda, K. Arnold, D. Baratin, G. Csardi, C.E. De, S. Duvaud, V. Flegel, A. Fortier, and E. Gasteiger. 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* 40:597–603.
- Aubourg, S., M. Kreis, and A. Lecharny. 1999. The DEAD box RNA helicase family in *Arabidopsis thaliana*. *Nucleic Acids Res.* 27:628–636.
- Bailey, T.L., M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li, and W.S. Noble. 2009. MEME suite: Tools for motif discovery and searching. *Nucleic Acids Res.* 37:202–208.
- Bhattacharya, D. and L. Medlin. 1998. Algal phylogeny and the origin of land plants. *Plant Physiol.* 116:9–15.
- Burglin, T.R. 1994. A comprehensive classification of homeobox genes, p. 27–71. In: D. Dubou-le (ed.). *Guidebook to the homeobox genes*. Oxford Univ. Press, New York, NY.
- Center for Bioinformatics. 2015. Gene structure display server. 17 May 2017. <<http://gsds.cbi.pku.edu.cn/>>.
- Consortium, T.T.G. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485:635–641.
- Czemmel, S., S.C. Heppel, and J. Bogs. 2012. R2R3 MYB transcription factors: Key regulators of the flavonoid biosynthetic pathway in grapevine. *Protoplasma* 249:109–118.
- Energy's Joint Genome Institute. 1997. Phytozome. 21 June 2017. <<https://phytozome.jgi.doe.gov/pz/portal.html>>.
- Englbrecht, C.C., H. Schoof, and S. Böhm. 2004. Conservation, diversification and expansion of C2H2 zinc finger proteins in the *Arabidopsis thaliana* genome. *BMC Genomics* 5(1):39.
- Espley, R., R. Hellens, J. Putterill, D.A. Stevenson, and A. Allan. 2007. Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant J.* 49:414–427.
- European Bioinformatics Institute. 2015. Biosequence analysis using profile hidden Markov models. 5 Jan. 2017. <<http://www.ebi.ac.uk/Tools/hmmer/search/hmmscan>>.
- Figueiredo, D.D., P.M. Barros, A.M. Cordeiro, T.S. Serra, T. Lourenço, S. Chander, M.M. Oliveira, and N.J. Saibo. 2012. Seven zinc-finger transcription factors are novel regulators of the stress responsive gene OsDREB1B. *J. Expt. Bot.* 63:3643–3656.
- Finn, R.D., J. Clements, and S.R. Eddy. 2011. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* 39:29–37.

- Gehring, W.J., M. Affolter, and T. Brclin. 1994. Homeodomain proteins. *Annu. Rev. Biochem.* 63:487–526.
- Gene Denovo. 2015. Omicshare. 8 May 2017. <<http://www.omicshare.com/tools/Home/Soft/heatmap>>.
- Guo, A.Y., Q.H. Zhu, X. Chen, and J.C. Luo. 2007. GSDS: A gene structure display server. *Hereditas* 29:1023–1026.
- Halbach, T., N. Scheer, and W. Werr. 2000. Transcriptional activation by the PH-D finger is inhibited through an adjacent leucine zipper that binds 14-3-3 proteins. *Nucleic Acids Res.* 28:3542–3550.
- Holub, E.B. 2001. The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat. Rev. Genet.* 2:516–227.
- Hong, S.Y., O.K. Kim, S.G. Kim, M.S. Yang, and C.M. Park. 2010. Nuclear import and DNA binding of the ZHD5 transcription factor is modulated by a competitive peptide inhibitor in *Arabidopsis*. *J. Biol. Chem.* 286:1659–1668.
- Hu, W., C.W. Depamphilis, and H. Ma. 2008. Phylogenetic analysis of the plant-specific zinc finger-homeobox and mini zinc finger gene families. *J. Integr. Plant Biol.* 50:1031–1045.
- Hu, W. and H. Ma. 2006. Characterization of a novel putative zinc finger gene MIF1: Involvement in multiple hormonal regulation of *Arabidopsis* development. *Plant J. Cell. Mol. Biol.* 45:399–422.
- Huang, W., Z. Xian, K. Xia, N. Tang, and Z. Li. 2015. Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato. *BMC Plant Biol.* 15(1):209.
- Hunter, C.S. and S.J. Rhodes. 2005. LIM-homeodomain genes in mammalian development and human disease. *Mol. Biol. Rep.* 32(2): 67–77.
- International Solanaceae Initiative. 2015. The SOL genomics network (SGN). 2 Jan. 2017. <<https://sgn.cornell.edu/search/locus>>.
- Ito, M., Y. Sato, and M. Matsuoka. 2002. Involvement of homeobox genes in early body plan of monocot. *Intl. Rev. Cytol.* 218:1–35.
- Jain, M., A. Nijhawan, R. Arora, P. Agarwal, S. Ray, P. Sharma, S. Kapoor, A.K. Tyagi, and J.P. Khurana. 2007. F-Box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. *Plant Physiol.* 143:1467–1483.
- Jain, M., A.K. Tyagi, and J.P. Khurana. 2006. Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive SAUR gene family in rice (*Oryza sativa*). *Genomics* 88:360–371.
- Karol, K.G., R.M. Mccourt, M.T. Cimino, and C.F. Delwiche. 2001. The closest living relatives of land plants. *Science* 294:2351–2353.
- Klug, A. and J.W. Schwabe. 1995. Protein motifs 5. Zinc fingers. *Federation Amer. Soc. Expt. Biol.* 9:597–604.
- Kenan, K., T. dem, B. Fatih, T. Suat, S. Suleyman, and O.G. Ahmet. 2017. Synthesis, structural and thermal characterizations, dielectric properties and in vitro cytotoxic activities of new 2,2,4,4-tetra(4'-oxy-substituted-chalcone)-6,6-diphenylcyclotriphosphazene derivatives. *Med. Chem. Res.* 26:962–974.
- Kosarev, P., K.F. Mayer, and C.S. Hardtke. 2002. Evaluation and classification of RING-finger domains encoded by the *Arabidopsis* genome. *Genome Biol.* 3(4):1–12.
- Krishna, S.S., I. Majumdar, and N.V. Grishin. 2003. Structural classification of zinc fingers: Survey and summary. *Nucleic Acids Res.* 31:532–550.
- Li, J., D. Jia, and X. Chen. 2001. HUA1, a regulator of stamen and carpel identities in *Arabidopsis*, codes for a nuclear RNA binding protein. *Plant Cell* 13:2269–2281.
- Lyons, E., B. Pedersen, J. Kane, M. Alam, R. Ming, H. Tang, X. Wang, J. Bowers, A. Paterson, and D. Lisch. 2008. Finding and comparing syntenic regions among *Arabidopsis* and the outgroups papaya, poplar, and grape: CoGe with rosids. *Plant Physiol.* 148:1772–1781.
- Nakashima, K. and K. Yamaguchi-Shinozaki. 2006. Regulons involved in osmotic stress-responsive and cold stress-responsive gene expression in plants. *Physiol. Plant.* 126:62–71.
- National Center for Biotechnology Information. 1988. Search for conserved domains within a protein or coding nucleotide sequence. 2 Mar. 2017. <<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>>.
- National Institutes of Health. 2015. Mupile Em for motif elicitation. 15 May 2017. <<http://meme-suite.org/tools/meme>>.
- Park, H.C., M.L. Kim, S.M. Lee, J.D. Bahk, D.J. Yun, C.O. Lim, J.C. Hong, S.Y. Lee, M.J. Cho, and W.S. Chung. 2007. Pathogen-induced binding of the soybean zinc finger homeodomain proteins GmZF-HD1 and GmZF-HD2 to two repeats of ATTA homeodomain binding site in the calmodulin isoform 4 (GmCaM4) promoter. *Nucleic Acids Res.* 35:3612–3623.
- Phoenix Bioinformatics Corporation. 2013. The arabidopsis information resource. 20 Apr. 2017. <http://www.arabidopsis.org/servlets/Search?action=new_search&type=gene>.
- Rio, D.C., M. Ares, G.J. Hannon, and T.W. Nilsen. 2010. Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harb. Protoc.* 2010 (6):pdb.prot5439.
- Song, H., P. Wang, Z. Nan, and X. Wang. 2014. The WRKY transcription factor genes in *Lotus japonicus*. *Intl. J. Genomics* 2014:420128.
- Song, L., L. Tao, H. Cui, L. Ling, and C. Guo. 2017. Genome-wide identification and expression analysis of the GRAS family proteins in *Medicago truncatula*. *Acta Physiol. Plant.* 39:93.
- Takatsui, H. 1999. Zinc-finger proteins: The classical zinc finger emerges in contemporary plant science. *Plant Mol. Biol.* 39:1073–1078.
- Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28:2731–2739.
- Tan, Q.K. and V.F. Irish. 2006. The *Arabidopsis* zinc finger-homeodomain genes encode proteins with unique biochemical properties that are coordinately expressed during floral development. *Plant Physiol.* 140:1095–1108.
- Tran, L.S.P., K. Nakashima, Y. Sakuma, Y. Osakabe, F. Qin, S.D. Simpson, K. Maruyama, Y. Fujita, K. Shinozaki, and K. Yamaguchi-Shinozaki. 2007. Co-expression of the stress-inducible zinc finger homeodomain ZFHD1 and NAC transcription factors enhances expression of the ERD1 gene in *Arabidopsis*. *Plant J. Cell. Mol. Biol.* 49:46–63.
- Wang, W., P. Wu, Y. Li, and X. Hou. 2016. Genome-wide analysis and expression patterns of ZF-HD transcription factors under different developmental tissues and abiotic stresses in Chinese cabbage. *Mol. Genet. Genomics* 291:1465.
- Williams, R.W. 1998. Plant homeobox genes: Many functions stem from a common motif. *Bioessays News Rev. Mol. Cell. Dev. Biol.* 20:280–282.
- Windhvel, A., I. Hein, R. Dabrowa, and J. Stockhaus. 2001. Characterization of a novel class of plant homeodomain proteins that bind to the C4 phosphoenolpyruvate carboxylase gene of *Flaveria trinervia*. *Plant Mol. Biol.* 45:201–214.
- Wolberger, C. 1996. Homeodomain interactions. *Curr. Opin. Struct. Biol.* 6:62–68.
- Yanagisawa, S. 2004. Dof domain proteins: Plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant Cell Physiol.* 45:386–391.
- Zhang, D., L.M. Iyer, and L. Aravind. 2012. Bacterial GRAS domain proteins throw new light on gibberellic acid response mechanisms. *Bioinformatics* 28:2407–2411.
- Zhou, T., Y. Wang, J.Q. Chen, H. Araki, Z. Jing, K. Jiang, J. Shen, and D. Tian. 2004. Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Mol. Genet. Genomics* 271:402–415.
- Zou, M., B. Guo, and S. He. 2011. The roles and evolutionary patterns of intronless genes in deuterostomes. *Comp. Funct. Genomics* 2011:680673.