Pak. J. Bot., 48(1): 155-165, 2016.

GENOME-WIDE IDENTIFICATION AND ANALYSIS OF GENES ENCODING PHD-FINGER PROTEIN IN TOMATO

XUEJIN CHEN[1,2], SIKANDAR HAYAT[1] AND ZHIHUI CHENG[1*]

[1]*College of Horticulture, Northwest A&F University, Yangling, 712100, Shaanxi, PR China*
[2]*School of Horticulture Landscape Architecture, Henan Institute of Science and Technology, Xinxiang, 453003, Henan, PR China*
[*]*Corresponding author e-mail: chengzh@nwsuaf.edu.cn*

Abstract

The PHD-finger proteins are conserved in eukaryotic organisms and are involved in a variety of important functions in different biological processes in plants. However, the function of PHD fingers are poorly known in tomato (*Solanum lycopersicum* L.). In current study, we identified 45 putative genes coding PHD-finger protein in tomato distributed on 11 chromosomes except for chromosome 8. Some of the genes encode other conserved key domains besides PHD-finger. Phylogenetic analysis of these 45 proteins resulted in seven clusters. Most PHD finger proteins were predicted to PML body location. These PHD-finger genes displayed differential expression either in various organs, at different development stages and under stresses in tomato. Our study provides the first systematic analysis of PHD-finger genes and proteins in tomato. This preliminary study provides a very useful reference information for PHD-finger proteins in tomato. They will be helpful for cloning and functional study of tomato PHD-finger genes.

Key words: Tomato, PHD-finger, Identification, Analysis.

Introduction

Tomato (*Solanum lycopersicum* L.), as an important vegetable (Siddiq *et al.*, 2009), is now grown and consumed worldwide in diverse ways including raw in salads, ketchup and vegetable. In addition, tomato is an excellent model for other berry crops due to its well-known genetic characteristics. It has a relatively moderate size of diploid genome (~900 Mb, n=12) (Consortium, 2012), has hundreds of molecular markers, abundant collections of germplasm and mutants, and it is transformable (Hasan *et al.*, 2008). Genetic and functionally genomic study in tomato will support breeding references to tomato, other Solanaceae family members and even fleshy berry plants in order to satisfy requirements for both consumers and farmers. Though more and more important genes in tomato have been cloned and identified (Martin *et al.*, 1993; Powell *et al.*, 2012), the function of many genes still remains largely unknown, especially the ones determining key agronomic characters. To study gene function, the first and basal strategy is to analyze gene family from genomic database using bioinformatics and choose potential candidates as research targets.

The PHD (Plant homeodomain) finger was first discovered by Schindler (Schindler *et al.*, 1993), who noted a Cys4-His-Cys3 motif of HAT3 protein in *Arabidopsis thaliana* in 1993. This PHD finger domain, composed of approximately 50 to 80 amino acids, can bind two zinc ions as metal binding RING domain through so-called 'cross-brace' motif. In general, the PHD finger adopts a globular fold, consisting of a two-stranded beta-sheet and an alpha-helix. The region consisting of these secondary structures and the residues involved in coordinating the zinc-ions are highly conserved among eukaryotic species. The loop regions 1 and 2 formed by PHD finger are variable and could contribute functional specificities to the different PHD fingers. It occurs as a single or in clusters of two or three, and it also occurs together with other domains, such as the chromodomain and the bromodomain (Ragvin *et al.*, 2004). As conserved domain in eukaryotic organisms, PHD-finger family tends to be found in nuclear proteins that have a role in epigenetics (Sanchez & Zhou, 2011), histone modification and chromatin-mediated transcriptional regulation (Aasland *et al.*, 1995). Some views have emphasized the ideas that PHD finger domains, such as related LIM and RING finger domains, are probably involved in protein-protein, protein-DNA and protein-RNA interactions (Linder *et al.*, 2000; Lyngso *et al.*, 2000). Some evidences suggested that PHD finger showed E3 ubiquitin ligase activity and it was described primarily as E3 ubiquitin ligases (Aravind *et al.*, 2003; Coscoy & Ganem, 2003). In plant kingdom, many PHD finger proteins were identified as important components or regulators in plant development, fertility, photoperiod pathways and vernalization (Mussig *et al.*, 2000; Sung & Amasino, 2004; Sung *et al.*, 2006). For example, ALFIN-LIKE 6, a PHD finger protein, is involved in root hair elongation during phosphate deficiency in *Arabidopsis* (Chandrika *et al.*, 2013). Another two PHD finger proteins (OBE1 and OBE2) in *Arabidopsis* were identified as crucial components for apical meristem maintenance and embryonic meristem initiation (Saiga *et al.*, 2008; Saiga *et al.*, 2012). In rice, *PTC1* encodes a PHD-finger protein that is required for tapetal cell death and pollen development (Li *et al.*, 2011). Two PHD-finger proteins in soybean (GmZF-HD1 and GmZF-HD2) function as transcription factors to activate *GmCaM4* gene expression in response to pathogen (Park *et al.*, 2007). These findings indicate the key roles of PHD finger in plants.

Whether PHD fingers have specific biological function in tomato is unclear. There was no literature reporting PHD finger function in tomato and no systematic analyses of PHD finger family in tomato were carried out previously. Here, we reported that 45 genes coding PHD finger proteins in tomato were identified through bioinformatics strategy. Multiple sequences alignment, phylogenetic tree, prediction of sub-nuclear location and gene expression profile in different tissues

and response to stresses provided a better understanding of PHD finger family in tomato. Our preliminary results can be utilized as a reference for cloning PHD finger genes and further exploring the function of PHD finger protein family in tomato.

Materials and Methods

Database searches for genes encoding PHD-finger proteins: To identify genes encoding PHD-finger proteins in tomato, multiple strategies were employed to search candidates from public tomato genomic database ITAG2.3 (International Tomato Annotation Group) (http://solgenomics.net/organism/Solanum_lycopersicum/genome) and non-redundant protein database from NCBI (National Center for Biotechnology Information) (http://www.ncbi.nlm.nih.gov/). The sequence of published PHD-finger genes form *Arabidopsis thaliana* and *Oryza sativa* were used to run BLASTN in ITAG2.3 for comprehensive collection. Sequences with *E* values above 1e-6 and maximum identity less than 30% were excluded from our dataset. Self-BLAST of the sequences was carried out manually to remove the redundancy. Sequence of candidates was reconfirmed using Pfam (http://pfam.sanger.ac.uk/search) and conserved domain database (CDD) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

Gene annotations and protein domain composition: The information of tomato PHD-finger genes, including gene locus, chromosomal location, length, number of exons and protein description were retrieved from ITAG2.3 database. The conserved domains were determined by Pfam and CDD. Illustrator CS5 software was used to draw protein model and gene distribution along chromosomes.

Multiple sequence alignment of PHD-finger domain: Complete amino acid sequences of PHD-finger genes were downloaded from ITAG2.3 database and transformed to FASTA format. Multiple sequence alignments were performed using Clustal W. Subsequently, alignments of PHD-finger domain were adjusted using Bioedit 7.0 software. Typical 'cross-brace' model of the PHD-finger domain was designed by Illustrator CS5 with a reference to previous literature (Kosarev *et al.*, 2002).

Phylogenetic analysis of the PHD-finger proteins: Full length protein sequence was employed to generate phylogenetic tree by the neighbor-joining method using MEGA 5.0 software (Tamura *et al.*, 2011). The bootstrap test was carried out with 1000 replicates to assess the reliability of the interior nodes.

Prediction of protein sub-nuclear location: PHD-finger proteins were predicted to analyze sub-nuclear locations using Sub-nuclear Compartments Prediction System v2.0 (http://array.bioengr.uic.edu/subnuclear.htm) (Lei & Dai, 2005; Lei & Dai, 2006).

Plant materials, treatments and gene expression analysis: Seeds of tomato (cv. Micro-Tom) were germinated and grown in grown mix (peat : vermiculite = 1:1) in climate-controlled chamber under long day conditions (16 h light at 25°C /8 h dark at 20°C cycle).

Three-week-old seedlings were transplanted to pots and various tissues including root, stem, leaf, flower and green fruit were collected at fruit formation stage. Abiotic stress treatments including heat (38°C, 4 h), cold (4°C, 4 h), salt (150 mM NaCl, 4 h) and drought (Dry air blow, 3 h) were carried out when the seedlings were 28-day-old. Leaf tissues were harvested and frozen immediately in liquid nitrogen after treatments. 5 representative PHD-finger genes (*Solyc05g032760*, *Solyc06g062850*, *Solyc07g041190*, *Solyc07g049210*, *Solyc10g078520*) were chosen to carry out real time RT-PCR for gene expression analysis. The primer sequences are listed as follows:

Solyc05g032760-F: 5'-AAGAGGAGGAAGAGCATGG-3'
Solyc05g032760-R: 5'-CTCTTGTTGCTGCATGATGG-3'
Solyc06g062850-F: 5'-AGTTAGCAAAGTTGCAGGC-3'
Solyc06g062850-R: 5'-AATTTTCTCCACAGGCACC-3'
Solyc07g041190-F:5'-AGATGAACATGGCGACACTC-3'
Solyc07g041190-R: 5'-CTTTATATGCTCGGCCTTGG-3'
Solyc07g049210-F: 5'-GATGAGACCATCTGATTCTG-3'
Solyc07g049210-R 5'-CTGCACATCAAAGTGATCTG-3'
Solyc10g078520-F: 5'-AGCTTCCTGCAGATGAACC-3'
Solyc10g078520-R: 5'-AGTCTCTGTCACGTTCTCG-3'

About 100 mg tissues were employed to extract total RNA using Column Plant RNA out reagent kit according to manufacturer's protocol (TIANDZ, Beijing, China). RNA concentration was measured by NanoDrop2000 (Thermo SCIENTIFIC, USA) to further normalize RNA template among different samples. The first strand cDNAs were synthesized (~0.4 μg RNA as template) using Super Quick RT cDNA Synthesis kit (CWBIO, Beijing, China). Real-time quantification RT-PCR reactions were performed in iQ^{TM}5 machine (Bio-Rad, USA) using the SYBR *Premix Ex Taq* II (TaKaRa, Dalian, China) according to the manufacturer's instructions. Each PCR reaction (25 μl) contained 15 μl Mix, 0.8 μl of each primer, and appropriately diluted cDNA. The PCR program was 95°C for 30 s followed by 40 cycles of 95°C for 20 s, 58°C for 30 s, and 68°C for 45 s. The *Actin 2/7* (de Jong *et al.*, 2009) was used as internal reference for all the RT-PCR analysis. Each treatment was repeated three times independently.

Microarray expression profiles of 10 PHD-finger genes (*Solyc01g102750*, *Solyc01g102760*, *Solyc01g110020*, *Solyc03g121930*, *Solyc05g032760*, *Solyc06g051420*, *Solyc06g062850*, *Solyc07g041190*, *Solyc07g049210*, *Solyc10g078520*) were obtained from online server Genevestigator (Hruz *et al.*, 2008). Detailed experimental information, plant materials and methods can be acquired at the website (http://www.genevestigator.com/gv/).

Results

Identification of genes encoding PHD-finger protein in tomato: Using integrated database mining, 45 putative genes encoding PHD-finger protein in tomato were identified from ITAG2.3 database. As shown in Table 1, the gene length varied from 1482 to 27813 bp and their encoding proteins varied from 97 to 2002 (aa) in length. The number of exon differed from 2 to 23 and the number of intron differed widely (Data not shown) indicating gene diversity in component. Besides PHD-finger domain, some genes encoded other conserved domains, such as BAH (Bromo-Adjacent Homology) domain, ING (Inhibitor of Growth) domain, RING finger domain and so on (Table 1). According to protein annotation from the database, some PHD-finger proteins played roles in methylation, demethylation and DNA banding activity (Table 1) indicating their various biochemical property. According to chromosomal location from ITAG2.3, these PHD-finger genes were distributed from chromosome 1 to 12 with an exception of chromosome 8 (Fig. 1). Chromosome 1, 9 and 10 were rich in the PHD-finger genes relatively. There were two gene aggregates on the bottom of chromosome 1 and 10 respectively (Fig. 1). Interestingly, the gene *Solyc01g102750.2.1* and *Solyc01g102760.2.1* were quite close (~ 1.4 kb in between) to each other on the bottom of chromosome 1 suggesting that they might belong to the same gene cluster (Fig. 1).

Table 1. List of 45 PHD-finger genes identified in tomato.

Gene locus	Length (bp)	Exons	Chr.	Protein length (aa)	Protein description
Solyc01g095890.2.1	12179	23	1	1048	Histone-lysine N-methyltransferase
Solyc01g100130.2.1	3864	3	1	925	PHD-finger family
Solyc01g102750.2.1	5458	5	1	249	PHD-finger family
Solyc01g102760.2.1	6606	6	1	248	PHD-finger family
Solyc01g107410.2.1	7422	9	1	796	PHD-finger family
Solyc01g110020.2.1	6389	5	1	216	Contain BAH-PHD domain
Solyc02g069980.2.1	9696	8	2	494	Contain RING finger domain
Solyc02g081130.1.1	3549	6	2	211	Contain BAH-PHD domain
Solyc03g005170.2.1	11108	9	3	1104	PHD-finger family
Solyc03g071550.1.1	6959	10	3	963	Chromodomain helicase DNA binding
Solyc03g083410.2.1	10406	24	3	981	Histone-lysine N-methyltransferase
Solyc03g116440.1.1	2438	3	3	709	PHD-finger family
Solyc03g121930.2.1	4828	5	3	276	PHD-finger family
Solyc04g007270.2.1	3852	6	4	703	PHD-finger family
Solyc04g008420.1.1	3085	3	4	678	PHD-finger family
Solyc04g058190.2.1	14502	6	4	603	Lysine-specific demethylase
Solyc05g005640.2.1	11531	9	5	2002	Contain methyl-CpG binding domain
Solyc05g032760.2.1	6043	3	5	97	PHD-finger family
Solyc05g042030.2.1	7733	5	5	212	Contain BAH-PHD domain
Solyc06g010220.1.1	7015	8	6	637	Chromodomain helicase DNA binding
Solyc06g051420.2.1	12723	5	6	245	PHD-finger family
Solyc06g062850.2.1	5428	5	6	257	PHD-finger family
Solyc06g069360.2.1	6633	3	6	650	PHD-finger family
Solyc06g074640.1.1	7050	9	6	1569	Chromodomain helicase DNA binding
Solyc07g008730.1.1	8703	10	7	707	Pathogenesis-related homeodomain
Solyc07g041190.2.1	4016	5	7	246	PHD-finger family
Solyc07g049210.2.1	4719	5	7	216	Contain BAH-PHD domain
Solyc07g062100.2.1	1482	2	7	302	Bromodomain and PHD finger
Solyc07g062600.2.1	7260	9	7	1141	PHD-finger family
Solyc09g005360.2.1	8419	5	9	241	PHD-finger family
Solyc09g008520.2.1	8670	21	9	930	Chromodomain helicase DNA binding
Solyc09g031580.2.1	8932	11	9	1213	Chromodomain helicase DNA binding
Solyc09g065340.2.1	10651	8	9	1360	PHD-finger family
Solyc09g076010.2.1	11175	8	9	1364	PHD-finger family
Solyc09g091190.2.1	4822	7	9	236	Inhibitor of growth protein
Solyc10g005490.2.1	27813	32	10	1258	Protein strawberry notch homolog
Solyc10g076690.1.1	5960	5	10	240	PHD-finger family
Solyc10g078520.1.1	3377	7	10	258	Inhibitor of growth protein
Solyc10g083800.1.1	10624	21	10	906	Chromodomain helicase DNA binding
Solyc10g084650.1.1	12643	19	10	1349	PHD-finger family
Solyc10g085180.1.1	4964	5	10	240	PHD-finger family
Solyc11g065140.1.1	3101	3	11	717	Chromodomain helicase DNA binding
Solyc11g066480.1.1	8880	17	11	1064	PHD-finger family
Solyc12g014540.1.1	10471	6	12	817	Lysine-specific demethylase
Solyc12g096040.1.1	7690	5	12	254	PHD finger family

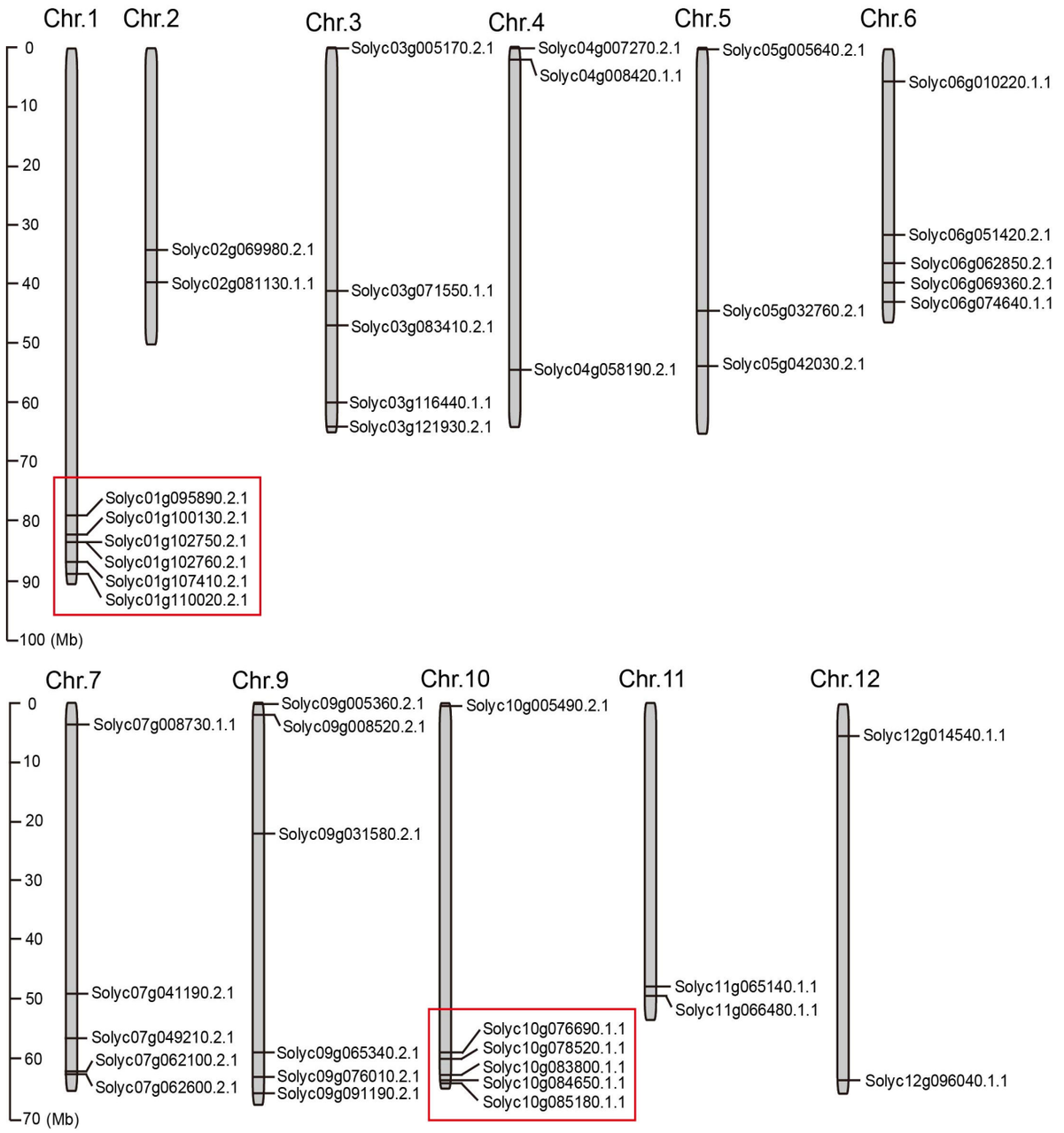

Fig. 1. Chromosomal distribution of PHD-finger genes in tomato.
Tomato chromosomes are represented by grey bars. The approximate location of each candidate PHD-finger genes was deduced from ITAG2.3 database. The two red boxes indicate gene aggregates.

Multiple sequence alignment of PHD-finger domain: PHD-finger domain, containing typical Cys4-His-Cys3 zinc banding motif, is highly conserved in more than 400 eukaryotic organisms (Kaadige & Ayer, 2006). To investigate conservative property of Cys4-His-Cys3 motif of these PHD-finger proteins, full-length amino acid sequence was used to generate multiple sequence alignment. All of the PHD-finger domains, consisting of approximately 60 residues, had a highly consensus Cys4-His-Cys3 motif (Fig. 2A). The amino acid sequences between cysteines and histidine were also relatively conserved, which was considered to be a typical property of PHD-finger domain (Bienz, 2006). The PHD-finger domain can bind two zinc ions as metal binding RING domain through the so-called 'cross-brace' model (Fig. 2B). The Loop 1 (between C2 and C3) and Loop 2 (between C5 and C6) of 45 PHD-fingers varied in length and constituted two different surfaces of the domain according to structure model (Fig. 2). Loop 2 (12-19 residues) was longer compared with loop 1 (4-18 residues) and contained related functional residues, such as Leu and Thr, suggesting that loop2 constituted a prime interaction surface of these PHD fingers (Bienz, 2006).

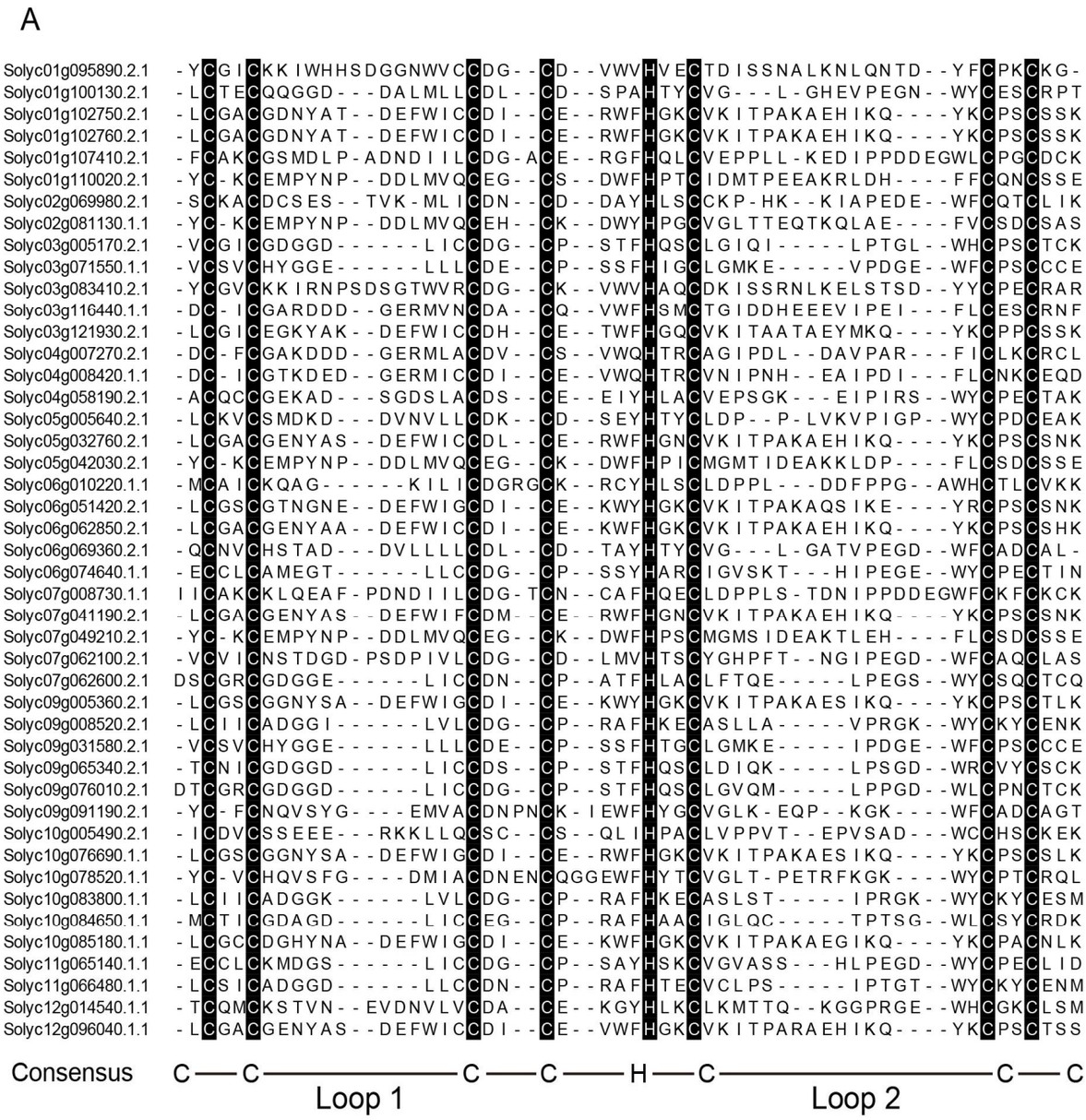

Fig. 2. Multiple sequence alignment (A) and structure of PHD-finger domain (B).
C4HC3 zinc-finger-like motif is shaded black (A). Schematic presentation ('cross-brace' model) of the structure of a PHD-finger domain (B).

Analysis of domain composition: Protein function is determined by its domain composition. So it is necessary to analyze domain composition as references for protein function study. Protein model of our identified PHD-fingers in tomato is shown in Fig. 3. Besides PHD-finger domain, other conserved domains were observed through Pfam and CDD of NCBI. Many of them contained only one PHD-finger domains without any other conserved domains. Differently, the gene *Solyc04g058190.2.1* encoded two PHD-finger domains in cluster. Ten PHD-finger proteins contained DUF3594 (domain of unknown function 3594) domain which was functionally uncharacterized (Fig. 3). Conserved DUF3594 domain family was found in eukaryotes and is approximately 140 amino acids in length. Interestingly, DUF3594 domain was found association with PHD finger, implying strong relationship between them. There were five PHD-finger proteins connecting with BAH domain which was found in proteins such as eukaryotic DNA methyltransferases and transcriptional regulation. The BAH domain might therefore play an important role in linking DNA methylation, replication and transcriptional regulation (Callebaut *et al.*, 1999). ING (Inhibitor of Growth) domain and DDT (DNA binding homeobox and different transcription factors) domain serving as transcript regulation and chromosome remodeling (Doerks *et al.*, 2001; Chruscicki *et al.*, 2010) were also observed in 45 PHD-finger proteins. PHD-finger like domain, such as PHD-2, zf-HC5HC2H-2 and zf-RING-2 were tandem to PHD-finger and might have the similar zinc ion banding activity. Other conserved domains, such as WHIM1, PWWP, SET and etc. joint together with PHD-finger domain. Combined with many other conserved domains, these 45 PHD-fingers might play multiple roles in tomato.

Phylogenetic analysis of the PHD-finger proteins in tomato: Phylogenetic tree was generated from full-length amino acid sequence to examine the evolutionary relationships among 45 PHD-finger family members in tomato. These PHD-finger members can be grouped into 7 clusters based on the ≥ 50% bootstrap values (Fig. 4). Two main clusters were composed with 11 and 10 members respectively. The third largest group contained 4 members and the forth group contained 3 members. There were 3 small groups composed with 2 members. The grouped PHD-fingers in tomato may have strong evolutionary relationships among subfamily members. There were 11 unclassified clades indicating relative far evolutionary relationships between each other (Fig. 4). In all, the 45 PHD-finger gene family may have undergone multiple duplications during the evolution history.

Prediction of protein sub-nuclear location: Identified PHD-finger proteins were known to be nuclear in model organism (Conti *et al.*, 1998; Huh *et al.*, 2003). So these 45 PHD-finger proteins were predicted to analyze sub-nuclear locations using Sub-nuclear Compartments Prediction System (Version 2.0) which can predict locations in cell nucleus (Lei & Dai, 2005; Lei & Dai, 2006). According to the prediction results, most of PHD-finger proteins (64%) were thought to be PML body which was dynamic nuclear protein aggregates interspersed between chromatin (Fig. 5). PML bodies show discrete nuclear foci, 0.2-1.0 micrometer wide, were present in most mammalian cell nuclei and typically number 1 to 30 bodies per nucleus, depending on the cell type, cell-cycle phase and differentiation stage. Eight members of them (18%) were predicted to localize at chromatin which was the combination of DNA and proteins that make up the contents of the nucleus of a cell (Fig. 5). There were seven PHD-finger proteins harboring at nucleoplasm which was a highly viscous liquid that surrounds the chromosomes and nucleoli. Protein encoded by *Solyc12g096040.1.1* was considered to localize at nuclear lamina which function in regulating important cellular events such as DNA replication and cell division. Additionally, nuclear lamina participated in chromatin organization and anchored the nuclear pore complexes embedded in the nuclear envelope. Different sub-nuclear location results indicated that PHD-finger proteins might have diverse biochemical function in cell nucleus.

Gene expression analysis in various tissues and at different developmental stages: To observe expression profiles of the PHD-finger genes in tomato development, the expression of these genes was analyzed under normal growth conditions by Genevestigator (Hruz *et al.*, 2008). Ten representative genes were chosen to create standard heat map in various tissues including seedling, inflorescence, racemose cyme, flower, fruit, shoot, stem, leaf, hypocotyl, cotyledon, roots, unspecified root and root tip. The gene *Solyc01g102760.2.1* had higher expression in racemose cyme and flower. The expression level of the gene *Solyc10g078520.1.1* was higher in seedling (Fig. 6A). In addition, we also identified the expression profiles of PHD-finger family genes at different developmental stages through analysis of public microarray database. The whole growth stages were divided into six periods and similar heat map was generated. Obviously, all the 10 genes investigated showed relatively high expression level at final stage (Fruit ripening complete) suggesting these genes may play a role in fruit maturity or plant senescence (Fig. 6B).

Gene expression analysis under stresses: To further understand the gene function, the change of gene expression induced by stresses including salt, heat, drought and three pathogens of tomato was analyzed. Fold change of 10 genes was investigated through Genevestigator (Hruz *et al.*, 2008). Many of them showed down-regulation trend after stress treatment suggesting their negative role in stress responses (Fig. 7). The expression level of *Solyc07g041190.2.1* was significantly down regulated by drought and *Colletotrichum coccodes* infection, indicating that the gene may play a negative role in response to the two stresses.

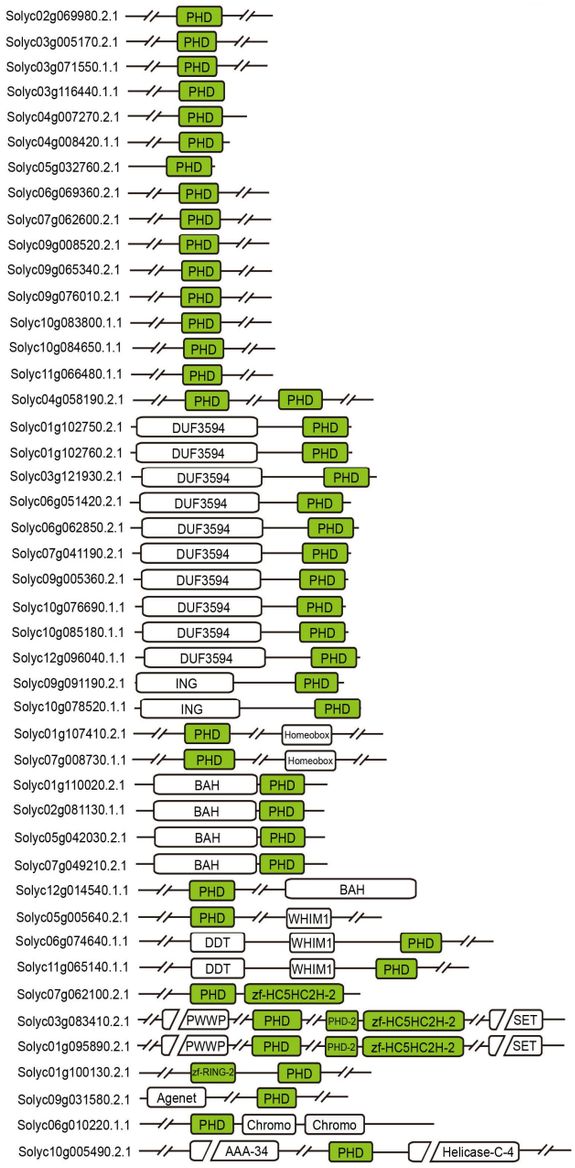

Fig. 3. Domain composition of 45 PHD-finger proteins.
Proteins with similar domain composition are listed together. The length of the lines and boxes is determined by protein length (aa). The green box is used to represent PHD-finger and RING domain.

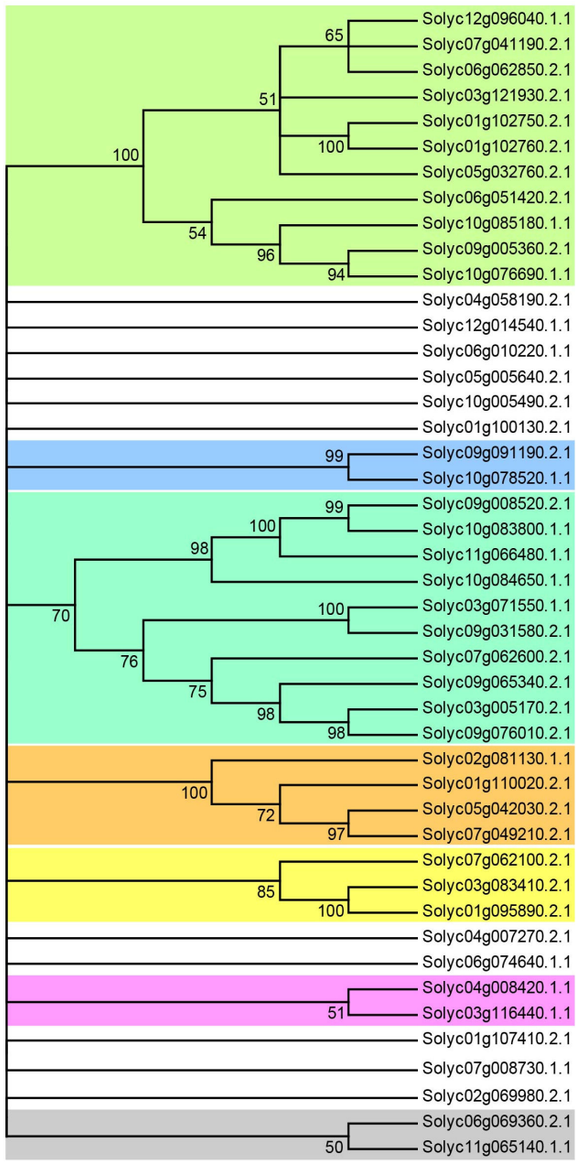

Fig. 4. Phylogenetic tree of 45 PHD-finger proteins.
Neighbor-joining tree was created using MEGA 5.0 program with 1000 bootstrap. Bootstrapping values are indicated as percentages (when more than or equal to 50%) along the branches. Different clusters are shown with various colors.

Gene expression analysis in model tomato cv. Micro-Tom: In tomato germplasm, a miniature tomato cultivar of *Solanum lycopersicum* cv. Micro-Tom is regarded as an excellent model system (Kobayashi *et al.*, 2013). Micro-Tom has characteristics that make it suitable for experimental study, such as small size, short generation time and ease of transformation. Currently, many Micro-Tom background mutants have been developed through various mutagenized strategies (Matsukura *et al.*, 2008). Thus, analysis of gene expression in the model cultivar would provide references to study gene function in tomato. On the other hand, gene expression pattern may differ in different cultivar as a result of diverse genetic background. To better understand the expression of PHD-finger genes in Micro-Tom, real time RT-PCR was carried out in various plant tissues and under stresses. Expression level of four genes investigated in leaf tissue was relatively higher compared with other tissues except for *Solyc07g049210.2.1* (Fig. 8A). The gene *Solyc07g049210.2.1* was slightly up-regulated induced by cold and drought stress (Fig. 8B). Other four genes were all down regulated in response to cold, heat, salt and drought stresses (Fig. 8B). Different expression pattern compared with microarray data might result from differed cultivar or experimental conditions. Taken together, PHD-finger genes may play multiple roles in tomato because of various expression patterns.

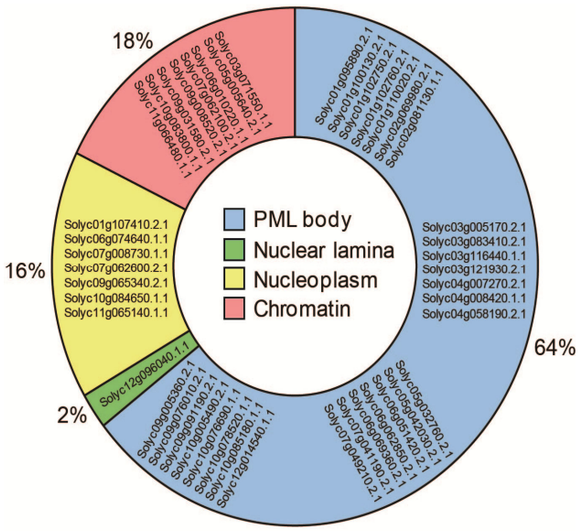

Fig. 5. Prediction of protein sub-nuclear location.
Sub-nuclear Compartments Prediction System (Version 2.0) was used to predict locations in cell nucleus. Prediction results are presented by colors in the form of pie chart. Percentage of location in 45 proteins is calculated.

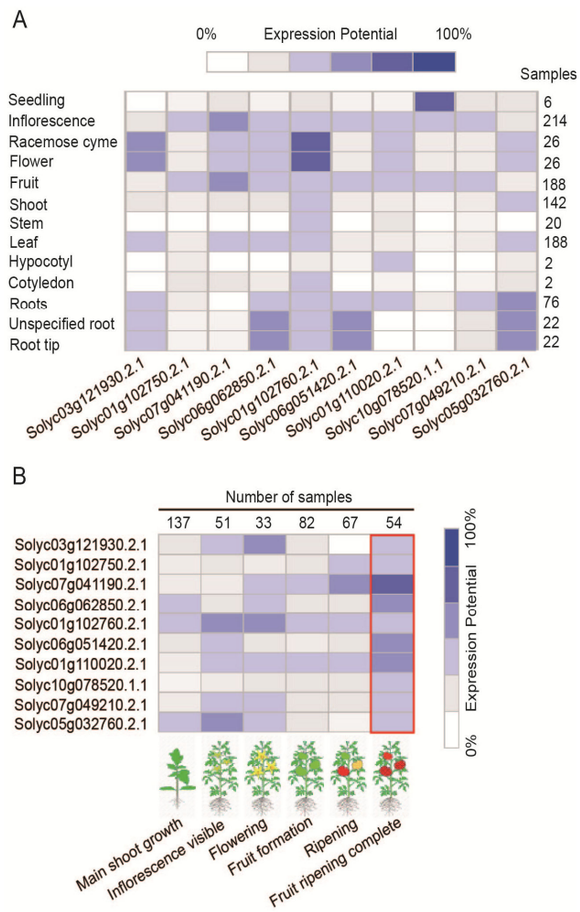

Fig. 6. Expression potential of PHD-finger genes in various tissues (A) and at different developmental stages (B).
Degree of blue color stands for various expressions potential. The complete fruit ripening stage is highlighted by red box.

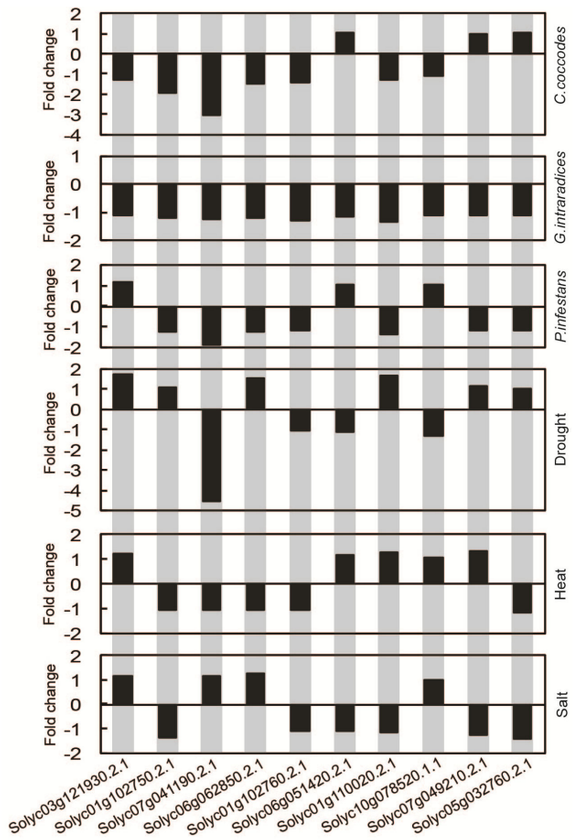

Fig. 7. Variation of PHD-finger gene expression induced by stresses.
Fold change is presented by black bars. The same gene is shaded gray for easy observation. Values above or below zero mean up or down regulation of gene expression respectively.

Discussion

As the progress of several whole-genome sequencing projects, including human, mouse, fruit fly, *Arabidopsis* and so on, more and more genomic DNA sequences have become available. When it comes to *Solanum* species, new version with high-quality genome sequence in tomato was released publicly in 2012 (Consortium, 2012), which provide a powerful support to tomato researchers. These genomic sequencing projects make it possible to analyze gene families in one species systematically. One of the well-known strategies for gene family analysis is to detect all the gene models first in one genome through bioinformatics tools (Andolfo *et al.*, 2012; Cai *et al.*, 2013; Chen *et al.*, 2013). Here, key words searching and BLAST program were undertaken to screen genes coding PHD-finger domain in tomato genome. In total, 45 putative PHD-fingers were identified and further analyzed. The gene and their coded protein varied from each other in length. Protein annotation suggested their roles in methylation, demethylation and DNA binding activity. The distribution of genes, in general, is nonrandom in plant genomes. Our analyses showed that the PHD finger genes were differentially represented on each of the tomato chromosomes except number 8. It seems that there were two gene aggregates on the bottom of chromosome 1 and 10 respectively indicating a nonrandom distribution to some extent.

The PHD finger is a common structural motif found in all eukaryotic organisms (Bienz, 2006). So PHD finger domain from 45 proteins in tomato was generated to multiple sequence alignment to investigate typical conserved Cys4-His-Cys3 motif. According to the results, all of them showed identical Cys4-His-Cys3 motif forming Loop1 and Loop2. The PHD finger binds two zinc ions through the so-called 'cross-brace' motif and is thus structurally related to the RING finger (Bienz, 2006). Besides PHD-finger domain, these genes encoded other conserved domains, such as BAH domain, ING domain, RING finger domain and so on, indicating multiple biochemical function. The PHD finger was thought to be involved in epigenetics, E3 binding activity, protein interaction and chromatin-mediated transcriptional regulation (Sanchez & Zhou, 2011). Phylogenetic tree was generated to examine the evolutionary relationships among different PHD-finger family members in tomato. These PHD-finger members can be grouped into 7 clusters (Fig. 4), suggesting the gene family may have undergone multiple duplications during the evolution history.

At present, more and more online servers can be used to predict protein sub-cellular location providing useful clues before experimental investigation (Emanuelsson *et al.*, 2000). The PHD finger proteins are nuclear in some model species including budding yeast and fruit fly (Bienz, 2006). So it is speculated that our newly identified PHD finger proteins in tomato might have the potential of nuclear location. The predicted results provide a better understand the functional mechanism of PHD fingers. For instance, PHD finger can have nucleosome-binding activity such as p300 and ACF1 (Eberharter *et al.*, 2004; Ragvin *et al.*, 2004). Based on our prediction, 8 members from PHD fingers in tomato localized at chromatin (Fig. 5) indicating their potential of nucleosome-binding activity. According to gene expression profile from Genevestigator and real time RT-PCR results in Micro-Tom, some PHD finer genes might regulate fruit maturation or plant senescence and respond to stresses.

In plant kingdom, some literatures have proved that PHD fingers play roles in plant development, fertility, photoperiod pathways and vernalization (Mussig *et al.*, 2000; Sung & Amasino, 2004; Sung *et al.*, 2006). For instance, Rice gene *PTC1* encodes a PHD-finger protein that is required for tapetal cell death and pollen development (Chandrika *et al.*, 2013). However, the relationship between PHD finger and their biological roles remain largely unknown in plant world. As worldwide vegetable, tomato is crucial to human diet and health because of its source of vitamin and active chemical lycopene functioning in antitumor (Ishfaq *et al.*, 2012). Genetic and functionally genomic study will support tomato breeding profoundly. However, the function of a great number genes in tomato still remains largely unknown, especially those ones determining key agronomic characters. Our preliminary results provided a supportive information of gene cloning and a better understanding of PHD finger gene family in tomato. However, further study is needed to determine the functions of the PHD fingers by additional biological experiments.

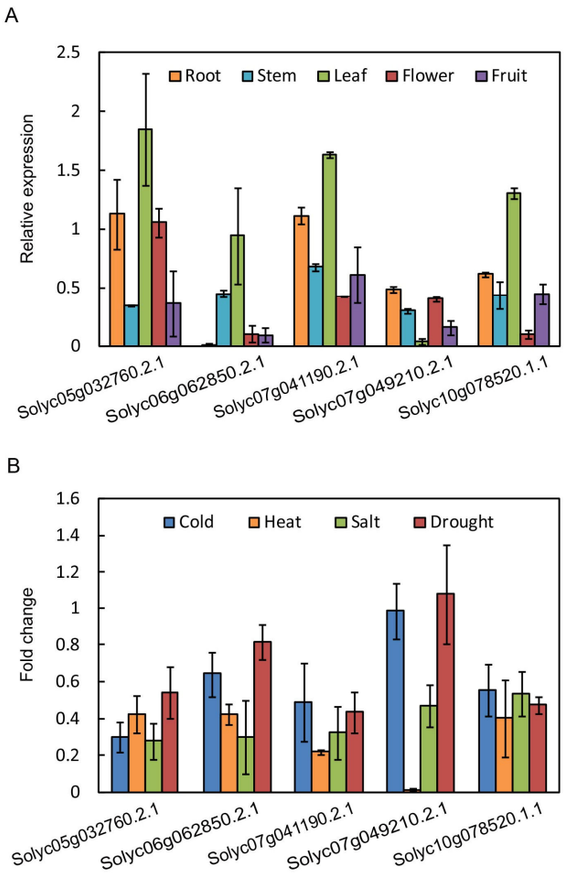

Fig. 8. Gene expression analysis in model tomato cv. Micro-Tom. Relative expression of 5 representative genes in various tissues were determined by real-time qRT-PCR. Gene expression after cold, heat, salt and drought treatments was determined through the same strategy. Expression levels were normalized by *Actin 2/7* (de Jong *et al.*, 2009). Error bars represent means of three replicates ± SD. Fold change of gene expression was compared with control (No treatment). Up and down regulation were determined by values above and below 1 respectively. One of three independent experimental replicates with similar results is shown.

Conclusions

Our study first provides the systematic analysis of the *Solanum lycopersicum* PHD-finger genes and proteins they encoded. From tomato genomic database, 45 putative PHD-finger genes were identified through multiple bioinformatics strategies. The PHD-finger gene family distributed among 11 chromosomes and might have undergone multiple duplications during the evolution history. The Cys4-His-Cys3 motif in these PHD-finger proteins was highly identical. In addition to PHD-finger domain, other conserved domains were observed among these proteins. Based on phylogenetic analysis, these PHD-finger proteins were grouped into 7 clusters. As nuclear protein, most of PHD-finger proteins were predicted to be involved in PML body. The PHD-finger genes might play multiple roles in tomato because of various expression patterns. Our results presented here may provide a starting point and supportive information for the functional dissection of PHD-finger family in tomato.

Acknowledgments

We thank Mr. Ren Lei (University of Saskatchewan, Saskatoon, Canada) for critical reading of the manuscript. We are grateful for financial support to Zhihui Cheng from China National Key Technology Support Program [2006BAD07B02].

References

Aasland, R., T.J. Gibson and A.F. Stewart. 1995. The phd finger: Implications for chromatin-mediated transcriptional regulation. *Trends Biochem. Sci.*, 20(2): 56-59.

Andolfo, G., W. Sanseverino, S. Rombauts, Y. Van de Peer, J.M. Bradeen, D. Carputo, L. Frusciante and M.R. Ercolano. 2012. Overview of tomato (*solanum lycopersicum*) candidate pathogen recognition genes reveals important solanum r locus dynamics. *New Phytol.*, 197(1): 223-237.

Aravind, L., L.M. Iyer and E.V. Koonin. 2003. Scores of rings but no phds in ubiquitin signaling. *Cell Cycle*, 2(2): 123-126.

Bienz, M. 2006. The phd finger, a nuclear protein-interaction domain. *Trends Biochem. Sci.*, 31(1): 35-40.

Cai, X., Y. Zhang, C. Zhang, T. Zhang, T. Hu, J. Ye, J. Zhang, T. Wang, H. Li and Z. Ye. 2013. Genome-wide analysis of plant-specific dof transcription factor family in tomato. *J. Integr. Plant Biol.*, 55(6): 552-566.

Callebaut, I., J.C. Courvalin and J.P. Mornon. 1999. The bah (bromo-adjacent homology) domain: A link between DNA methylation, replication and transcriptional regulation. *FEBS Lett.*, 446(1): 189-193.

Chandrika, N.N., K. Sundaravelpandian, S.M. Yu and W. Schmidt. 2013. Alfin-like 6 is involved in root hair elongation during phosphate deficiency in arabidopsis. *New Phytol.*, 198(3): 709-720.

Chen, X., B. Ghazanfar, A.R. Khan, S. Hayat and Z. Cheng. 2013. Comparative analysis of putative orthologues of mitochondrial import motor subunit: Pam18 and pam16 in plants. *PLoS One*, 8(10): e78400.

Chruscicki, A., V.E. Macdonald, B.P. Young, C.J. Loewen and L.J. Howe. 2010. Critical determinants for chromatin binding by saccharomyces cerevisiae yng1 exist outside of the plant homeodomain finger. *Genetics*, 185(2): 469-477.

Consortium, T.T.G. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485(7400): 635-641.

Conti, E., M. Uy, L. Leighton, G. Blobel and J. Kuriyan. 1998. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell*, 94(2): 193-204.

Coscoy, L. and D. Ganem. 2003. Phd domains and e3 ubiquitin ligases: Viruses make the connection. *Trends Cell Biol.*, 13(1): 7-12.

de Jong, M., M. Wolters-Arts, R. Feron, C. Mariani and W.H. Vriezen. 2009. The *solanum lycopersicum* auxin response factor 7 (slarf7) regulates auxin signaling during tomato fruit set and development. *Plant J.*, 57(1): 160-170.

Doerks, T., R. Copley and P. Bork. 2001. Ddt -- a novel domain in different transcription and chromosome remodeling factors. *Trends Biochem. Sci.*, 26(3): 145-146.

Eberharter, A., I. Vetter, R. Ferreira and P.B. Becker. 2004. Acf1 improves the effectiveness of nucleosome mobilization by iswi through phd-histone contacts. *EMBO J.*, 23(20): 4029-4039.

Emanuelsson, O., H. Nielsen, S. Brunak and G. von Heijne. 2000. Predicting subcellular localization of proteins based on their n-terminal amino acid sequence. *J. Mol. Biol.*, 300(4): 1005-1016.

Hasan, M., A.J. Khan, S. Khan, A.H. Shan, A.R. Khan and B. Mirza. 2008. Transformation of tomato (*Lycopersicon esculentum* Mill.) with *Arabidopsis* early flowering gene *APETALAI* (*API*) Through *Agrobacterium* infiltration of ripened fruits. *Pak. J. Bot.*,40(1): 161-173.

Hruz, T., O. Laule, G. Szabo, F. Wessendorp, S. Bleuler, L. Oertle, P. Widmayer, W. Gruissem and P. Zimmermann. 2008. Genevestigator v3: A reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinformatics*, 2008: 420747.

Huh, W.K., J.V. Falvo, L.C. Gerke, A.S. Carroll, R.W. Howson, J.S. Weissman and E.K. O'Shea. 2003. Global analysis of protein localization in budding yeast. *Nature*, 425(6959): 686-691.

Ishfaq, M., I.A. Nasir, N. Mahmood and M. Saleem. 2012. *In vitro* induction of mutation in tomato (*Lycopersicon esculentum* L.) cv. Roma by using chemical mutagens. *Pak. J. Bot.*, 44(1): 311-314.

Kaadige, M.R. and D.E. Ayer. 2006. The polybasic region that follows the plant homeodomain zinc finger 1 of pf1 is necessary and sufficient for specific phosphoinositide binding. *J. Biol. Chem.*, 281(39): 28831-28836.

Kobayashi, M., H. Nagasaki, V. Garcia, D. Just, C. Bres, J.P. Mauxion, M.C. Le Paslier, D. Brunel, K. Suda, Y. Minakuchi, A. Toyoda, A. Fujiyama, H. Toyoshima, T. Suzuki, K. Igarashi, C. Rothan, E. Kaminuma, Y. Nakamura, K. Yano and K. Aoki. 2013. Genome-wide analysis of intraspecific DNA polymorphism in 'micro-tom', a model cultivar of tomato (*solanum lycopersicum*). *Plant Cell Physiol.*, 55(2): 445-454.

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.

Lei, Z. and Y. Dai. 2005. An svm-based system for predicting protein subnuclear localizations. *BMC Bioinformatics*, 6: 291.

Lei, Z. and Y. Dai. 2006. Assessing protein similarity with gene ontology and its use in subnuclear localization prediction. *BMC Bioinformatics*, 7: 491.

Li, H., Z. Yuan, G. Vizcay-Barrena, C. Yang, W. Liang, J. Zong, Z.A. Wilson and D. Zhang. 2011. Persistent tapetal cell1 encodes a phd-finger protein that is required for tapetal cell death and pollen development in rice. *Plant Physiol.*, 156(2): 615-630.

Linder, B., R. Newman, L.K. Jones, S. Debernardi, B.D. Young, P. Freemont, C.P. Verrijzer and V. Saha. 2000. Biochemical analyses of the af10 protein: The extended lap/phd-finger mediates oligomerisation. *J. Mol. Biol.*, 299(2): 369-378.

Lyngso, C., G. Bouteiller, C.K. Damgaard, D. Ryom, S. Sanchez-Munoz, P.L. Norby, B.J. Bonven and P. Jorgensen. 2000. Interaction between the transcription factor spbp and the positive cofactor rnf4. An interplay between protein binding zinc fingers. *J.Biol. Chem.*, 275(34): 26144-26149.

Martin, G.B., S.H. Brommonschenkel, J. Chunwongse, A. Frary, M.W. Ganal, R. Spivey, T. Wu, E.D. Earle and S.D. Tanksley. 1993. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science*, 262(5138): 1432-1436.

Matsukura, C., K. Aoki, N. Fukuda, T. Mizoguchi, E. Asamizu, T. Saito, D. Shibata and H. Ezura, 2008. Comprehensive resources for tomato functional genomics based on the miniature model tomato micro-tom. *Curr. Genomics*, 9(7): 436-443.

Mussig, C., A. Kauschmann, S.D. Clouse and T. Altmann. 2000. The arabidopsis phd-finger protein shl is required for proper development and fertility. *Mol. Gen. Genet.*, 264(4): 363-370.

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(11): 3612-3623.

Powell, A.L., C.V. Nguyen, T. Hill, K.L. Cheng, R. Figueroa-Balderas, H. Aktas, H. Ashrafi, C. Pons, R. Fernandez-Munoz, A. Vicente, J. Lopez-Baltazar, C.S. Barry, Y. Liu, R. Chetelat, A. Granell, A. Van Deynze, J.J. Giovannoni and A.B. Bennett. 2012. Uniform ripening encodes a golden 2-like transcription factor regulating tomato fruit chloroplast development. *Science*, 336(6089): 1711-1715.

Ragvin, A., H. Valvatne, S. Erdal, V. Arskog, K.R. Tufteland, K. Breen, O.Y. AM, A. Eberharter, T.J. Gibson, P.B. Becker and R. Aasland. 2004. Nucleosome binding by the bromodomain and phd finger of the transcriptional cofactor p300. *J. Mol. Biol.*, 337(4): 773-788.

Saiga, S., C. Furumizu, R. Yokoyama, T. Kurata, S. Sato, T. Kato, S. Tabata, M. Suzuki and Y. Komeda. 2008. The arabidopsis oberon1 and oberon2 genes encode plant homeodomain finger proteins and are required for apical meristem maintenance. *Development*, 135(10): 1751-1759.

Saiga, S., B. Moller, A. Watanabe-Taneda, M. Abe, D. Weijers and Y. Komeda. 2012. Control of embryonic meristem initiation in arabidopsis by phd-finger protein complexes. *Development*, 139(8): 1391-1398.

Sanchez, R. and M.M. Zhou. 2011. The phd finger: A versatile epigenome reader. *Trends Biochem. Sci.*, 36(7): 364-372.

Schindler, U., H. Beckmann and A.R. Cashmore. 1993. Hat3.1, a novel arabidopsis homeodomain protein containing a conserved cysteine-rich region. *Plant J.*, 4(1): 137-150.

Siddiq, S., M. Yaseen, S.A.R. Mehdi, A. Khalid, S.U.R. and Kashif. 2009. Growth and yield response of tomato (*Lycopersicon esculentum* Mill.) to soil applied calcium carbide and L-Methionine. *Pak. J. Bot.*, 41(5):2455-2464.

Sung, S. and R.M. Amasino. 2004. Vernalization in arabidopsis thaliana is mediated by the phd finger protein vin3. *Nature*, 427(6970): 159-164.

Sung, S., R.J. Schmitz and R.M. Amasino. 2006. A phd finger protein involved in both the vernalization and photoperiod pathways in *Arabidopsis*. *Genes. Dev.*, 20(23): 3244-3248.

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(10): 2731-2739.

(Received for publication 3 February 2015)