



Systematic analysis and comparison of the *PHD-Finger* gene family in Chinese pear (*Pyrus bretschneideri*) and its role in fruit development

Yunpeng Cao¹ · Yahui Han² · Dandan Meng¹ · Muhammad Abdullah¹ · Dahui Li¹ · Qing Jin¹ · Yi Lin¹ · Yongping Cai¹

Received: 8 July 2017 / Revised: 2 April 2018 / Accepted: 9 April 2018 / Published online: 20 April 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

PHD-finger proteins, which belongs to the type of zinc finger family, and that play an important role in the regulation of both transcription and the chromatin state in eukaryotes. Currently, PHD-finger proteins have been well studied in animals, while few studies have been carried out on their function in plants. In the present study, 129 non-redundant PHD-finger genes were identified from 5 Rosaceae species (pear, apple, strawberry, mei, and peach); among them, 31 genes were identified in pear. Subsequently, we carried out a bioinformatics analysis of the PHD-finger genes. Thirty-one *PbPHD* genes were divided into 7 subfamilies based on the phylogenetic analysis, which are consistent with the intron-exon and conserved motif analyses. In addition, we identified five segmental duplication events, implying that the segmental duplications might be a crucial role in the expansion of the PHD-finger gene family in pear. The microsynteny analysis of five Rosaceae species showed that there were independent duplication events in addition to the genome-wide duplication of the pear genome. Subsequently, ten expressed PHD-finger genes of pear fruit were identified using qRT-PCR, and one of these genes, *PbPHD10*, was identified as an important candidate gene for the regulation of lignin synthesis. Our research provides useful information for the further analysis of the function of PHD-finger gene family in pear.

Keywords Gene family · Lignin synthesis · PHD-finger · Pear · Systematic analysis

Introduction

Zinc finger proteins contain large family members and are widely distributed in eukaryotic organisms. Although rich in histidines or cysteines, the “finger” structural domains have

been divided into different types, such as really interesting new genes (RING); Lin11, Isl-1, and Mec-3 (LIM); and plant homeodomain (PHD) according to their amino acids alignment (Takatsuji 1998). PHD-finger proteins are among the common zinc finger proteins with one or several PHD-finger

Yunpeng Cao and Yahui Han contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10142-018-0609-9>) contains supplementary material, which is available to authorized users.

✉ Yongping Cai
swkx12@ahau.edu.cn

Yunpeng Cao
xfcpeng@126.com

Yahui Han
hyahui@163.com

Dandan Meng
2925743098@qq.com

Dahui Li
ldh@ahau.edu.cn

Qing Jin
qingjin@ahau.edu.cn

Yi Lin
linyi320722@163.com

¹ School of Life Sciences, Anhui Agricultural University, Hefei 230036, China

² State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei 230036, China

domains, which contain approximately 60 amino acids and are characterized by the Cys4-His-cys3 zinc-binding motif (Kaadige and Ayer 2006). Among these PHD domains, the numbers of amino acids between the cysteine residues or between cysteine and histidine are usually conserved, and the second amino acid residue before the last pair of cysteines is usually aromatic amino acids, such as tryptophan (Bienz 2006). A cysteine-rich domain was first observed in the PHD-finger proteins of *Arabidopsis thaliana* HAT3.1 and maize *Zmhox1* (Schindler et al. 1993). Subsequently, PHD-finger proteins were studied in humans, drosophila, nematodes, and yeasts (Gibbons et al. 1997; Kehle et al. 1998; Martin et al. 2006; Papoulas et al. 1998).

Many studies reported that the PHD-finger proteins regulated the chromatin by binding their PHD-finger domains to nucleosomes, such as ATP-dependent chromatin assembly factor 1 (ACF1), bromodomain PHD finger transcription factor (BPTF), and BHC80 (also known as PHF21A) in humans (Lan et al. 2007). The human BPTF showed binding activity with H3K4 me_{2/3} (Wysocka et al. 2006), while the BHC80 protein specifically bound to non-methylated H3K4 (H3K4 me₀) (Lan et al. 2007). In yeast, Yng1, Yng2, and PHO23 of the YNG family exhibited a strong H3K4 me_{2/3} binding activity in the PHD domain (Shi et al. 2007). Some PHD-finger proteins could bind to protein subunits that lack histone in the PHD-finger domain. For example, the PHD-finger domain of the trans-activator protein Pygopus in flies bound to Legless/BCL9 (Kramps et al. 2002). In addition, the PHD-finger domain of the human transcriptional repressor KAP-1 bound to an isomer of the Mi-2 alpha subunit (Schultz et al. 2001). PHD-finger transcription factors are also involved in the regulation of plant growth and development. For example, MS1 (male sterility 1) is related to the development of pollen and the viscoelastic layer (Yang et al. 2007); VIN3 (Vernalization insensitive 3), VRN5 (Vernalization 5) (Greb et al. 2007), SHL (short life), and MMD1 (male meiocyte death 1) are involved in the regulation of genes during the developmental process and meiosis (Yang et al. 2003).

The PHD-finger gene family has been identified in some model plant species, such as *Arabidopsis thaliana*, maize (*Zea mays*), and poplar (*Populus trichocarpa*). However, the PHD-finger genes in Rosaceae have not been studied as compared to the extensive surveys that have been conducted in other plants. Recently, the whole genomes of apple (*Malus × domestica*) (Velasco et al. 2010), pear (*Pyrus bretschneideri*) (Wu et al. 2013), peach (*Prunus persica*) (Verde et al. 2013), mei (*Prunus mume*) (Zhang et al. 2012), and strawberry (*Fragaria vesca*) (Shulaev et al. 2011) have been fully sequenced. This resource provided an excellent opportunity for further understanding of gene structure and function of the PHD-finger gene family in five Rosaceae species. In the present study, the gene structures, chromosomal localizations, and collinearity analysis of the PHD-finger genes in pear and other

Rosaceae species were analyzed. Furthermore, their expression profiles during the development of pear fruit were assessed. These results suggested that the expression trend of *PbPHD10* was consistent with the changing trend of stone cell content in pear. Some previous studies have suggested that the content of the stone cell is an important factor that have influences on the quality of pear fruit (Jin et al. 2013). Current strategies focus to reduce lignin content for improving the quality of pear fruit (Cai et al. 2010; Jin et al. 2013). Our study will contribute by revealing the effect of the PHD-finger genes on lignin synthesis during the development of pear fruit and providing a reference for improving the quality of other fruits that contain lignin.

Materials and methods

Identification of PHD-finger containing genes

To identify the PHD-finger genes in five Rosaceae species, we obtained the Hidden Markov Model (HMM) profile of PHD-finger proteins from the Pfam database (<http://pfam.sanger.ac.uk/>). These sequences in PHD-finger HMM profiles were used as a query to search against the fully sequenced genome databases of the pear, peach, apple, strawberry, and mei using the BlastP program (p value = 0.001). Subsequently, the SMART database (<http://smart.embl-heidelberg.de/>) (Letunic et al. 2012) and Pfam database (<http://xfam.org/>) (Punta et al. 2011) were used to determine whether each candidate PHD-finger protein sequence was a member of the PHD-finger family. Subsequently, we removed all potentially redundant PHD-finger sequences according to the results of the sequence alignments. According to the starting positions of the PHD-finger genes on the pear chromosomes that were extracted from the genome annotation document ‘*Pyrus bretschneideri*_Chr_gene.gff_V121010,’ which can be obtained from the website (<http://gigadb.org/dataset/100083>), the chromosome location images of these genes in the pear were generated using Circos (<http://circos.ca/>) (Krzywinski et al. 2009). At the same time, the exon-intron structures of the pear PHD-finger genes were produced using the online GSDS website (<http://gsds.cbi.pku.edu.cn/>) (Guo et al. 2007) according to the genome annotations. The conserved motifs that were encoded by each PHD-finger gene were also identified in pear using the online tool MEME (<http://meme-suite.org/tools/meme>) (Bailey et al. 2015) with the optimum motif width ranging from 15 to 200 and the maximum number of motifs set to 20.

Phylogenetic analysis

ClustalX was used with the default parameters to perform a coding sequence alignment. A neighbor-joining (N-J) tree was constructed using MEGA5 (Tamura et al. 2011), and the parameters were set as follows: the bootstrap was set to 1000,

and the pairwise deletion option and the Poisson correction were set to the distance model. The tree included five types of Rosaceae species and was constructed by MEGA5 according to the data that were downloaded from the National Center for Biotechnology Information (NCBI) Taxonomy Common Tree (<http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi>).

Collinearity analysis

In this research study, the pear genome sequences were downloaded to our local server to identify the collinearity. The microsynteny relationships between each pair of chromosomes were ascertained using MCscanX (Wang et al. 2012). Consequently, the microsynteny chains were evaluated with an E value of $< e^{-10}$ (Jain and Das 2016).

Ka/Ks analysis

The synonymous (Ks) and nonsynonymous (Ka) substitution rates were estimated using DnaSP 5.0 (Librado and Rozas 2009). Subsequently, the selection modes of the *PbPHD* paralogs were confirmed by analyzing the Ka/Ks ratios.

Functional annotation analysis

The functional annotations of the PbPHDs were performed using Blast2GO software (Conesa et al. 2005) with the default parameters. Based on the cellular localizations, molecular functions, and biological processes, 31 pear PHD-finger proteins were classified using the WEGO web site (Ye et al. 2006). The KEGG pathways of these PbPHDs were carried out using KEGG database (<http://www.kegg.jp>) (Kanehisa and Goto 1999).

Orthologous analysis

The position of each *PbPHD* was marked on the chromosomes using a Perl script. The orthologous *PbPHDs* among the pear (*Pyrus bretschneideri*), apple (*Malus x domestica*), strawberry (*Fragaria vesca*), mei (*Prunus mume*), and peach (*Prunus persica*) were identified using OrthoMCL (Li et al. 2003) with E values less than $1e^{-5}$. The orthologous pairs among the five species were drawn using Circos (Krzywinski et al. 2009).

Expression analysis based on the RNA-sequencing data

The RNA-seq raw data of five developmental periods of pear fruits (25 DAF, 55 DAF, 85 DAF, 115 DAF, and 145 DAF) were retrieved from the National Center for Biotechnology Information (NCBI: Accession number SRP008545) (Xie

et al. 2013). An analysis of these data was performed according to a previous method (Muthamilarasan et al. 2014). The NGS toolkit was used to filter the RNA-seq data for delete low quality reads (Patel and Jain 2012). The RNA-seq data mapped onto the reference genome of pear using the Bowtie2 software. The expression levels of the *PbPHD* genes in the exon model were estimated as reads per kilobase per million mapped reads (RPKM). In the present study, we considered a *PbPHD* gene was expressed if the RPKM value was $>$ or $= 1$ in an expression atlas based on the previously published papers (Lin et al. 2012). The hierarchical clustering was performed with R software (<http://www.bioconductor.org/>) to analyze the expression profile of the *PHD* genes in pear.

Quantitative real-time RT-PCR

To investigate the expression patterns of the *PbPHD* genes, pear root, stem, leaf, and fruit samples (15, 39, 47, 55, 63, 79, 102, and 145 days after flowering (DAF)) were collected in 2013. Three biological replicates of the samples were harvested from 40-year-old pear trees grown in a horticultural field in Dangshan, Anhui, China during each developmental stage. All fruit samples were taken to the laboratory, immediately frozen in liquid nitrogen, and stored at -80 °C for RNA isolation. The total RNA of all the samples was extracted using the RNAPrep pure Plant Kit (Tiangen, Beijing) according to the manufacturer's instructions with DNase I (Tiangen, Beijing) to remove any genomic DNA contamination. Then, the RNA (1 μ g) was used for the reverse transcription using the SuperScript III reverse transcription kit (Invitrogen). Gene-specific primers (Table S1) for the qRT-PCR were designed using the Beacon Designer 7 software, and the tubulin gene (Wu et al. 2012) was used as an internal control. These primers were synthesized by General Biosystems, Inc. (Chuzhou, China). The qRT-PCR was conducted on ABI7500 (Applied Biosystems) using the FastFire qPCR PreMix (Tiangen, Beijing) based on the manufacturer's instructions. The $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001) was used to assess the relative expression levels of these *PbPHD* genes.

Results

Sequence collection and identification

To identify the PHD-finger family genes, the Hidden Markov Model (HMM) and BLASTP searches were performed against the reference genomes of pear (Wu et al. 2013), apple (Kalyanaraman et al. 2010), peach (Verde et al. 2013), mei (Zhang et al. 2012), and strawberry (Shulaev et al. 2011) using DNATools software. Totally, 156 candidate PHD-finger genes were identified from five Rosaceae species. Subsequently, all

candidate PHD-finger proteins were surveyed to further determine whether they contained the PHD-finger domain using the Pfam database. Due to the absence of the PHD-finger binding domain, 27 candidate genes were discarded. Finally, 129 non-redundant *PHD* genes were identified, including 31 apple *PHDs*, 19 mei *PHDs*, 25 peach *PHDs*, 23 strawberry *PHDs*, and 31 pear *PHDs* (Fig. 1, Tables S2 and S3). Remarkably, 7 genes encode proteins that contain 3 PHD-finger domains, and 12 genes encode proteins that contain 2 PHD-finger domains in pear (Table S2). A phylogenetic tree of these PHD-finger genes from the five Rosaceae species was constructed to deduce the genome-wide duplication between the progenitors of pear and apple (Fig. 1). Additionally, due to several genes of this gene family occurred alternative splicing, the longest splicing isoforms were selected for following analysis except specific description.

Phylogenetic and structures analysis of the *PbPHD* genes

To investigate the evolutionary relationships of the PHD-finger family in plant species, the PHD-finger genes from the monocots *Oryza sativa* and *Zea mays* or the dicots *Arabidopsis thaliana*, *Vitis vinifera*, *Populus trichocarpa*, *Malus × domestica*, *Pyrus bretschneideri*, *Prunus persica*, *Prunus mume*, and *Fragaria vesca* were used to construct an N-J phylogenetic tree with the MEGA 5.0 software (Tamura et al. 2011). Bootstrapping tests were performed on this tree. Based on the bootstrap values and the topology of the tree, all the PHD-finger proteins were divided into the following five main clades: I, II, III, IV, and V (Figs. 2 and S1). Furthermore, clade I was categorized into the following four subfamilies: I (a)–I (d); clade II classified into five subfamilies: II (a)–II (e); and clade IV divided into the following five subfamilies: IV (a)–IV (e), according to the phylogenetic tree (Figs. 2 and S1). The I (c) subfamily was not present in the monocots *Oryza sativa* and *Zea mays*; however, the I (c) subfamily was identified in the dicots species, including pear species. In pear

species, the II (a), II (c), II (d), II (e), IV (c), and IV (d) subfamilies were absent (Figs. 2 and S1).

Previous studies have shown that the gene structure diversity provides the primary resource for the evolution of multi-gene families (Cao et al. 2016a, b). To obtain further insights into the structural diversity of the PHD-finger genes in pear, an exon-intron analysis was performed. In the present study, the exon-intron structures (Fig. 3) showed that the intron numbers in *PbPHDs* varied from 1 to 23, and *PbPHD23* contained the most introns (23). Moreover, most members clustered in the same group shared highly similar exon-intron structures, including the numbers and lengths of the introns. For instance, *PbPHD14* and *PbPHD24*, which are located in subfamily III, all contained 22 introns, while all genes in subfamily I (a) included seven introns. In contrast, the *PbPHDs* in subfamily I (c) showed a high diversity in their intron lengths and numbers (Fig. 3). To sum up, the exon-intron structures of the *PHD* genes in pear substantially conformed to their phylogenetic relationships.

As a common mechanism, alternative splicing could produce different subtypes of proteins, which could likely affect the diversities of proteomic and transcriptomics, and finally affecting the regulation of protein function and gene expression (Black 2003; Pan et al. 2008). In the present study, we also revealed the occurrence of alternative splicing events (such as *PbPHD22* and *PbPHD23*) in the *PHD* family during the evolutionary process. The mRNAs of *PbPHD22* and *PbPHD23*, which are produced by variable splicing, are different in the 3'-end (Fig. S2). Their first 20 exons and introns are all identical; however, an intron is inserted into the twenty-first exon, resulting in *PbPHD23* being longer than *PbPHD22* by 12 nucleotides.

The domain sequence and conserved motifs analysis of the *PbPHD* proteins

In general, the protein domain is associated with protein function; therefore, the domain sequence characteristics of 31 *PbPHDs* were analyzed. The sequence alignment suggested

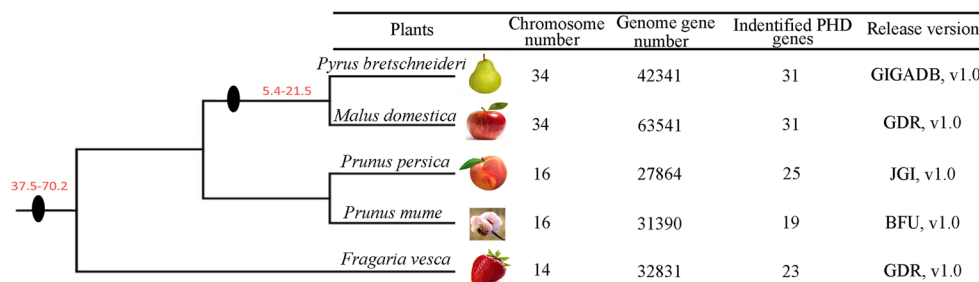
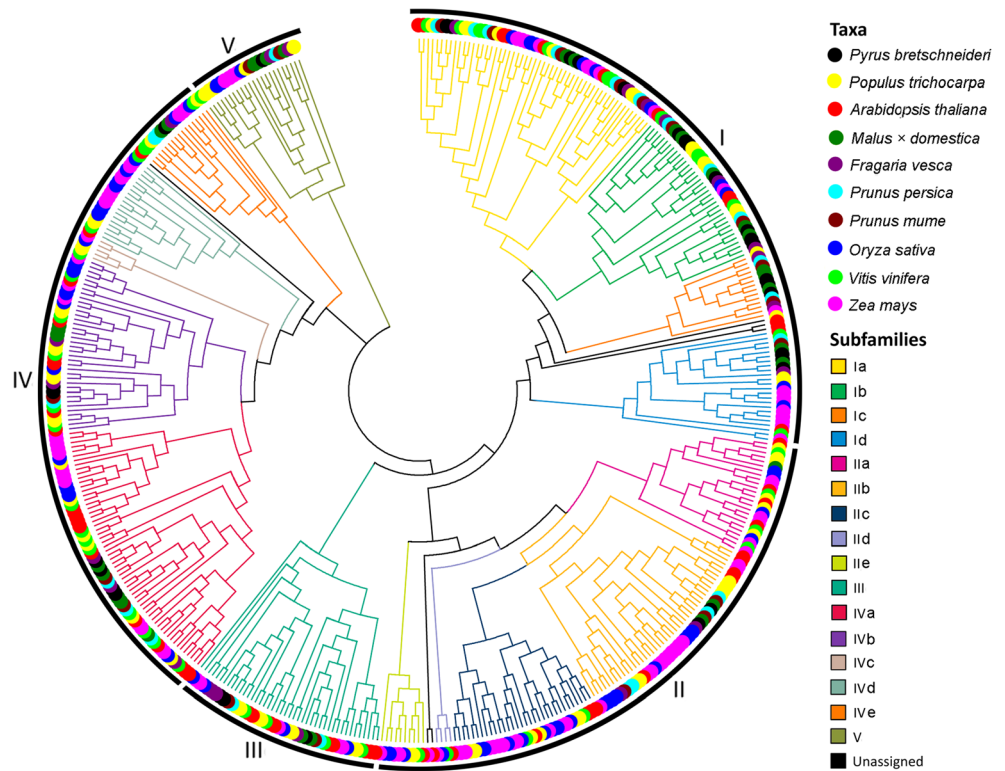


Fig. 1 Comparison of the PHD-finger genes from five Rosaceae species (*Pyrus bretschneideri*, *Malus × domestica*, *Fragaria vesca*, *Prunus mume*, and *Prunus persica*) in phylogenetic relationships. In the present study, the genomes of five Rosaceae species were investigated. GIGADB:

<http://gigadb.org/site/index>; GDR: <http://www.rosaceae.org/>; BFU: <http://prunusmumegenome.bjfu.edu.cn/index.jsp>; JGI: (<http://www.jgi.doe.gov/>)

Fig. 2 Phylogenetic tree of the PHD-finger proteins in the monocots *Oryza sativa* and *Zea mays* and the dicots *Arabidopsis thaliana*, *Populus trichocarpa*, *Vitis vinifera*, *Malus × domestica*, *Pyrus bretschneideri*, *Prunus persica*, *Prunus mume*, and *Fragaria vesca*. For the ten-species PHD-finger gene tree, the PHD-finger proteins of *Oryza sativa* and *Vitis vinifera* were obtained from the PLAZA 3.0 database (Proost et al. 2015). The neighbor-joining tree was constructed using MEGA5 software



that the length of the PHD-finger domains ranged from 40 to 60 amino acids and contained 7 cysteine and 1 histidine residues for the binding of Zn ions and maintaining the stability of

the domain. Additionally, the conserved residues on both sides are thought to be related to the different histone methylation patterns. In contrast, there are great differences between the

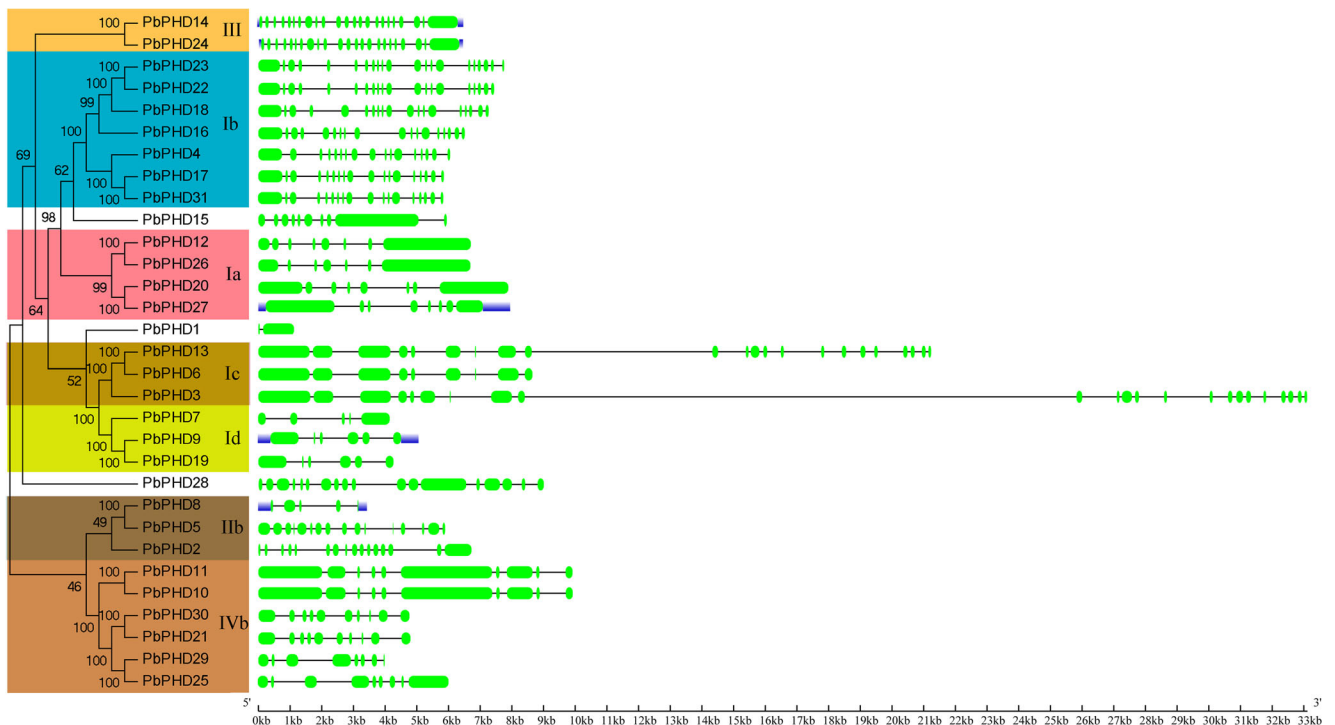


Fig. 3 Phylogenetic relationships and gene structures of pear PHD-finger genes. The numbers next to the branches represent the bootstrap values based on 1000 replications. Blue boxes indicate UTRs, green boxes

represent exons, and the introns are indicated by gray lines. The scale on the bottom is used to estimate the sizes of the exons and introns

two loop areas, including their lengths and the constitutions of amino acids. Moreover, the PbPHD proteins that contained common sequences were C-X (1–2)–C-X (8–14)–C-X (2–3)–C-X 4–H-x 2–C x (12–18) C-X 2–C (Figs. S3 and 4), which formed the C4HC3 type zinc finger structure. This common sequence was basically consistent with the conserved residues reported in a previous study (Capili et al. 2001). Remarkably, it was also found that there was a great difference in the nucleotide composition among the PHD domain genes, except for a few conserved T or G nucleotides (Fig. S4).

Regarding the fact that the PHD-finger proteins function via their conserved domains, a study on the PHD-finger protein domains in pear was carried out. In total, 31 pear PHD-finger proteins were analyzed by MEME. As a result, 20 motifs were identified. The detailed information regarding each motif is shown in Table S4. The results showed that members with highly conserved motifs were included in the same subgroup, and most members of the same subgroup had at least one homologous protein with completely coincident motifs (Fig. 5), such as PbPHD13/PbPHD31 and PbPHD22/PbPHD23. Because of their sequence similarity, these paralogous proteins possibly exhibited a functional redundancy, resulting in a difficulty in analyzing their functions using single gene mutants. In the case of most PHD-finger proteins within diverse subgroups, the proteins were different not only in their motifs but also in their locations (Fig. 5). The known domains, except for the PHD-finger domain, were identified and annotated using SMART and Pfam (Fig. 5 and Table S4). Among these domains, motif 5, which encodes the TPL domain (De Geyter et al. 2012; Pauwels et al. 2010), was identified in all members of groups I (a), I (b), and I (c), suggesting that these subgroup members might be involved in regulating the jasmine acid signaling pathway. Motif 9 encoded the DDT domain (Doerks et al. 2001), which was predicted to be a DNA binding domain in different transcription and chromosome remodeling factors and was only distributed in group I (c). Both PbPHD10 and PbPHD11 in group I (d) contained motifs 18 and 19, which encode an SNF domain (Eisen et al. 1995; Linder et al. 2004), suggesting that they might be related to the transcriptional regulation of genes (Eisen et al. 1995; Linder et al. 2004). Additionally, a number of other interesting

structural domains were also identified, such as SANT and PWWP. The identification of these motifs and domains will be helpful for further studies on the biological functions of the PHD-finger family in pear.

Functional annotations analysis of the PbPHD proteins

In our study, to better understand the function of the different PbPHDs, a GO annotation analysis was performed to predict the functions of the putative PbPHD proteins. In total, 24 of the 31 PbPHD proteins were classified into 19 functional categories within 3 main ontologies, including molecular function, biological process, and cellular component (Fig. 6, Table S5). The GO annotation of the PbPHD proteins showed that the binding, metabolic process, cell, catalytic, and cell part annotations were predominant among the functional categories. Several proteins were assigned to the macromolecular complex, organelle part, anatomical structure formation, and cellular component biogenesis. Furthermore, the molecular function annotations suggested that all PbPHDs functioned in the binding and catalytic categories, while the cellular component annotations suggested that these proteins were predominantly localized in the nucleus.

To further evaluate the active pathways of *PbPHDs*, these genes were mapped to the canonical pathway in KEGG database. Subsequently, we found that the most represented pathways by these *PbPHDs* were cellular processes, cell cycle, environmental information processing, and spliceosome (Table S5). The pathway annotations of PbPHDs not only support the GO enrichment analysis but also provide a valuable resource for gaining insights into the functions, specific processes, and pathways in pear growth and development.

Chromosomal locations and duplications of the *PbPHD* genes

Pear is the third important fruit tree after grape and apple, which has very important economic value (Wu et al. 2013). At least one whole genome duplication (WGD) event was discovered in the pear genome during the evolutionary process (Cao et al. 2016c; Wu et al. 2013). In order to further

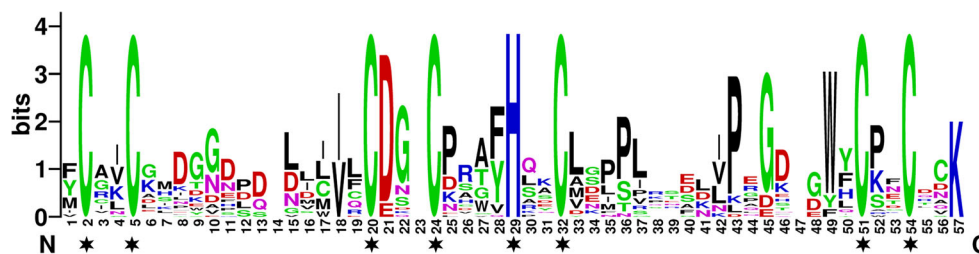
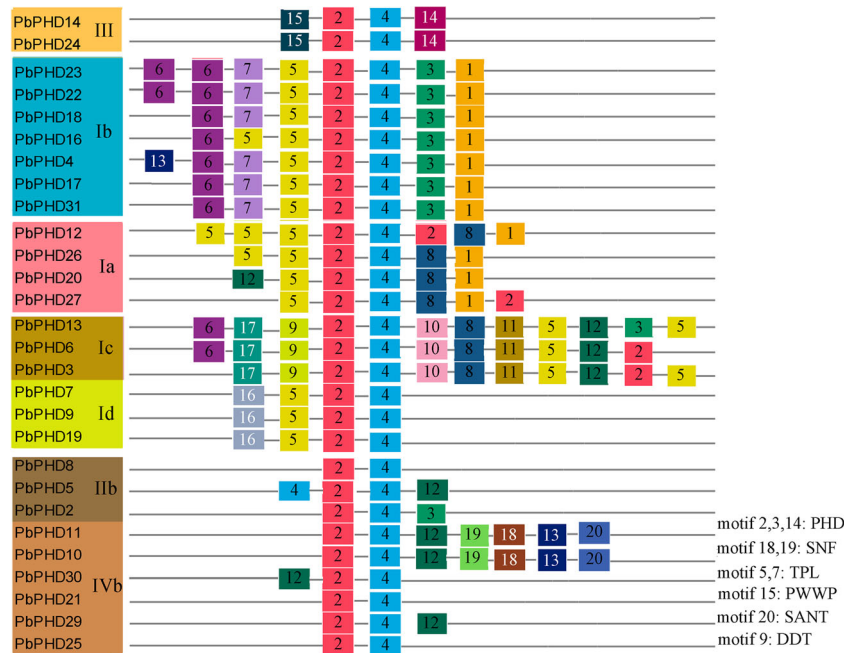


Fig. 4 The sequence logos of the PHD-finger domains. These logos were based on a multiple alignment analysis of 31 PbPHDs (Fig. S2). The bit score indicates the information content of each position in the sequence.

The asterisks indicate that the typically conserved PHD finger domains in pear were C-X (1–2)–C-X (8–14)–C-X (2–3)–C-X 4–H-x 2–C x (12–18) C-X 2–C

Fig. 5 Distribution of the motifs in the *PbPHD* family members. The domains in the A, B, C, D, and E subgroups were determined by both SMART and Pfam searches. Different motifs are represented by different color boxes. The length of each box did not relate to the actual size of each motif. The detailed information regarding each motif is annotated on the right



understand the effect of gene differentiation and gene duplication on the *PbPHD* gene family, we investigated the localization of *PbPHD* in the pear genome and the types of its gene duplication events. According to the initial positions of the genes in the chromosomes, 21 of the 31 *PbPHD* genes were unevenly distributed on 11 chromosomes in pear, while *PbPHD22*, -23, -24, -25, -26, -27, -28, -29, -30, and -31 were mapped to currently unattributed scaffolds. The number of *PbPHD* genes that were mapped on each chromosome varied from 1 to 5 (Fig. 7). Chromosome 11 contained five *PbPHD* genes, while chromosomes 4, 5, 6, 8, 9, 12, and 15 included only one *PbPHD* gene (Fig. 7). These uneven distribution patterns were also observed in the genomes of maize, poplar,

and *Arabidopsis thaliana* (Fan et al. 2009; Wang et al. 2015; Wu et al. 2016b).

In plants, segmental duplication or tandem duplication are the main ways to expand the number of family members. To understand the expanded mechanism of the PHD-finger gene family in pear during the evolution process, we investigated their duplication event potential. Based on the collinearity analysis of the *PbPHD* genes, eight gene pairs (*PbPHD4/17*, *PbPHD5/8*, *PbPHD7/19*, *PbPHD9/19*, *PbPHD10/11*, *PbPHD12/26*, *PbPHD14/24*, and *PbPHD17/31*) were identified as involved in the segmental duplication events (Fig. 7), among them *PbPHD4/17* and *PbPHD7/19* were involved in WGD events (Table 1). However, we did not find any tandem

Fig. 6 Gene ontology classification of the PHD-finger proteins of pear. The Y-axis (left) shows the percentages of genes, and the Y-axis (right) indicates the actual gene number. In total, 31 pear *PHD-finger* genes were annotated in three main categories (cellular component, molecular function, and biological process (X-axis)) using the Blast2GO software

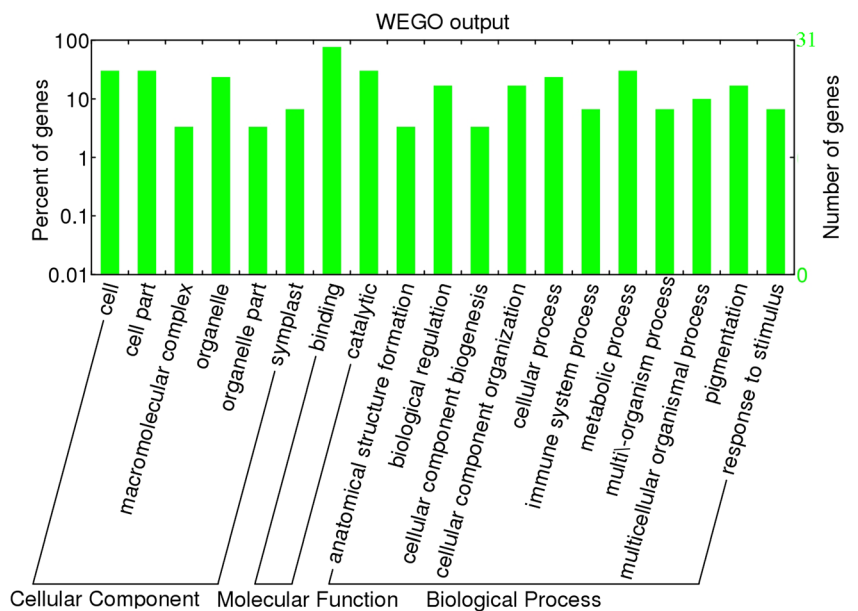
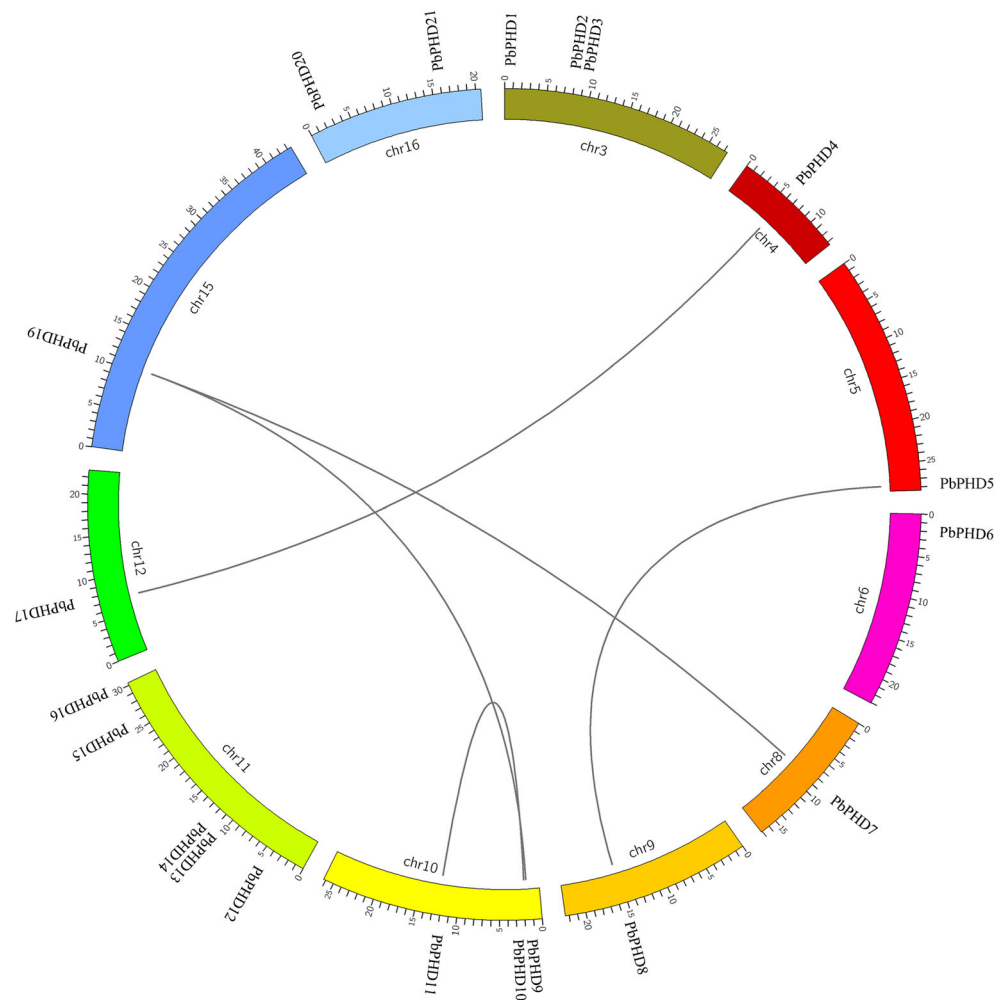


Fig. 7 Distribution and gene duplication events of *PHDs* in pear chromosomes. Gray lines connect the corresponding duplicated genes. The numbers of chromosomes are displayed



duplication events. *PbPHD12/26*, *PbPHD14/24*, and *PbPHD17/31* were not shown in the figure because *PbPHD26*, *PbPHD24*, and *PbPHD31* are mapped on the unattributed scaffolds.

To evaluate the driving force underlying the PHD-finger gene evolution, the synonymous (K_s) substitution rate ratio versus the nonsynonymous (K_a) were calculated for the duplicated *PbPHD* genes. Generally, $K_a/K_s > 1$ indicated a

positive selection with an accelerated evolution; $K_a/K_s < 1$ indicated functional constraints with a negative selection; and $K_a/K_s = 1$ signified a neutral selection. The results showed that the K_a/K_s ratios of eight duplicated *PbPHD* gene pairs ranged from 0.090 to 0.696 with an average of 0.343 (Table 1). Additionally, the majority of the ratios were less than 0.4, suggesting that the duplicated *PbPHD* genes were under a strong negative selection during evolution.

Table 1 K_a/K_s analysis and estimated divergence time for the duplicated *PbPHD* paralogs

Paralogous pairs	K_s	K_a	K_a/K_s	Purifying selection	Duplication date (Mya)	Duplicate type
<i>PbPHD4-PbPHD17</i>	0.1738	0.0456	0.2624	No	57.9333	(WGD) Segmental
<i>PbPHD5-PbPHD8</i>	1.1535	0.104	0.0902	No	384.5	Segmental
<i>PbPHD7-PbPHD19</i>	0.1768	0.0795	0.4497	No	58.9333	(WGD) Segmental
<i>PbPHD9-PbPHD19</i>	0.0153	0.0056	0.366	No	5.1	Segmental
<i>PbPHD10-PbPHD11</i>	0.0087	0.0019	0.2184	No	2.9	Segmental
<i>PbPHD12-PbPHD16</i>	0.7825	0.5445	0.6958	No	260.8333	Segmental
<i>PbPHD14-PbPHD24</i>	0.0141	0.0044	0.3121	No	4.7	Segmental
<i>PbPHD17-PbPHD31</i>	0.0176	0.0061	0.3466	No	5.8667	Segmental

Microsynteny analysis among pear, apple, peach, mei, and strawberry

To further insight into the evolutionary history of the *PbPHD* genes, we carried out the microsynteny analysis among five Rosaceae species. Finally, orthologous of all 90 pear genome *PHDs* were found in the genomes of the other Rosaceae species (i.e., apple, peach, mei, and strawberry) (Table S6). Among these orthologous, 25 orthologous gene pairs were identified in pear, apple, and peach, while 20 gene pairs were found in pear, mei, and strawberry (Fig. S4 and Table S6). In addition, to better observe the microsynteny of the *PHDs* in the different species, *PbPHD8* was used as examples to display the microsynteny relationship in different Rosaceae species (Fig. 8).

Remarkably, we observed that certain collinear gene pairs in pear and apple were missing between pear and peach, mei, and strawberry, such as *PbPHD6*/MDP0000120221, *PbPHD18*/MDP0000232962, *PbPHD22*/MDP0000125254, *PbPHD23*/MDP0000125254, *PbPHD25*/MDP0000246432, *PbPHD28*/MDP0000272113, and *PbPHD30*/MDP0000311077. Similarly, certain uniquely duplicated genes also appeared between pear, peach, mei, and strawberry (Fig. S5 and Table S6). Moreover, the result indicated that two or more PHD-finger genes in apple, peach, mei, and strawberry matched one pear PHD-finger gene, i.e., *mrna00537* and *mrna19503*, *Pm009703* and *Pm015135*, and *ppa011288m* and *ppa011295m* are orthologous to *PbPHD8*, and *MDP0000193753* and *MDP0000120221* are orthologous to *PbPHD6* (Table S6).

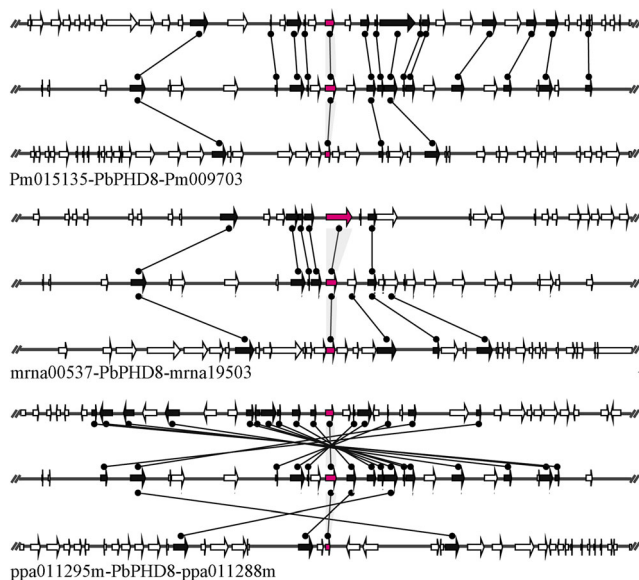


Fig. 8 Microsynteny relationships of the *PHD*-finger genes in pear. The relative positions of all flanking protein-coding genes were defined by the anchored *PHD*-finger genes and are highlighted in red. The broad line with the arrowhead represents the gene transcriptional orientation. The conserved gene pairs among the segments were connected by a gray line

Expression analysis of the *PbPHD* genes

To determine the role of the PHD-finger gene family during the development of the pear fruit, the expression of the *PbPHD* genes were analyzed using Illumina RNA-Seq data. Figure S5 shows that the transcripts of the most *PbPHD* genes (67.74%) were not detected in each developmental stage of pear fruit, proposing that these genes might be express in other organs, such as the leaf, root, and flower or these genes are pseudogenes. Ten *PbPHDs* (32.26%) were expressed in one or more developmental stages. *PbPHD10* and *PbPHD11* were highly expressed in all stages (Fig. S6). Six of the ten PHD-finger genes that expressed during pear fruit development belonged to subgroup E (Fig. S5), indicating that this subfamily might play an important role in the developmental period of pear fruit.

In order to study the dynamic of *PbPHD* genes, the expression levels of ten *PbPHDs* in the root, stem, and leaf, and during the pear fruit development were investigated by qRT-PCR. The results showed that *PbPHD12*, 15, and 29 were expressed predominantly in the leaf, while the expression of *PbPHD11* was not detected in the leaf (Fig. 9). The expression of *PbPHD8* and *PbPHD25* was higher at the 102 DAF (days after flowering) stage as compared to other like root, stem, and leaf in other developmental stages (15 DAF, 39 DAF, 47 DAF, 55 DAF, 63 DAF, 79 DAF, and 145 DAF) of pear fruit. Figure 9 demonstrated the expression of *PbPHD5*, 10, 11, and 25 was higher during the different developmental stages of pear fruit than that in the root, stem, or leaf. Remarkably, the expression of *PbPHD10* was basically consistent with the lignin content of the pear fruit during the different developmental stages, which increased rapidly during the early stages of pear fruit development, highest peak in the middle stage of the development (55 DAF), and decrease gradually during the mature period (Cai et al. 2010; Jin et al. 2013).

Discussion

Due to the completion of pear genome sequencing, the analysis of a gene family in terms of its biological information has laid a steady foundation for exploring the origin, evolution, and gene function of pear. Currently, numerous transcription factor families have been identified in the pear genome, such as the *WRKY*, *MYB*, *MADS* families, etc. However, comprehensive analysis of the PHD-finger transcription factor gene family has not been reported in pear. The PHD-finger transcription factor is a zinc finger protein with a characteristic Cys4-His-Cys3 structure, which widely exists in transcriptional regulatory proteins in eukaryotes. The PHD-finger proteins, a class of transcription factors, have been thought to play an important role in plant development process.

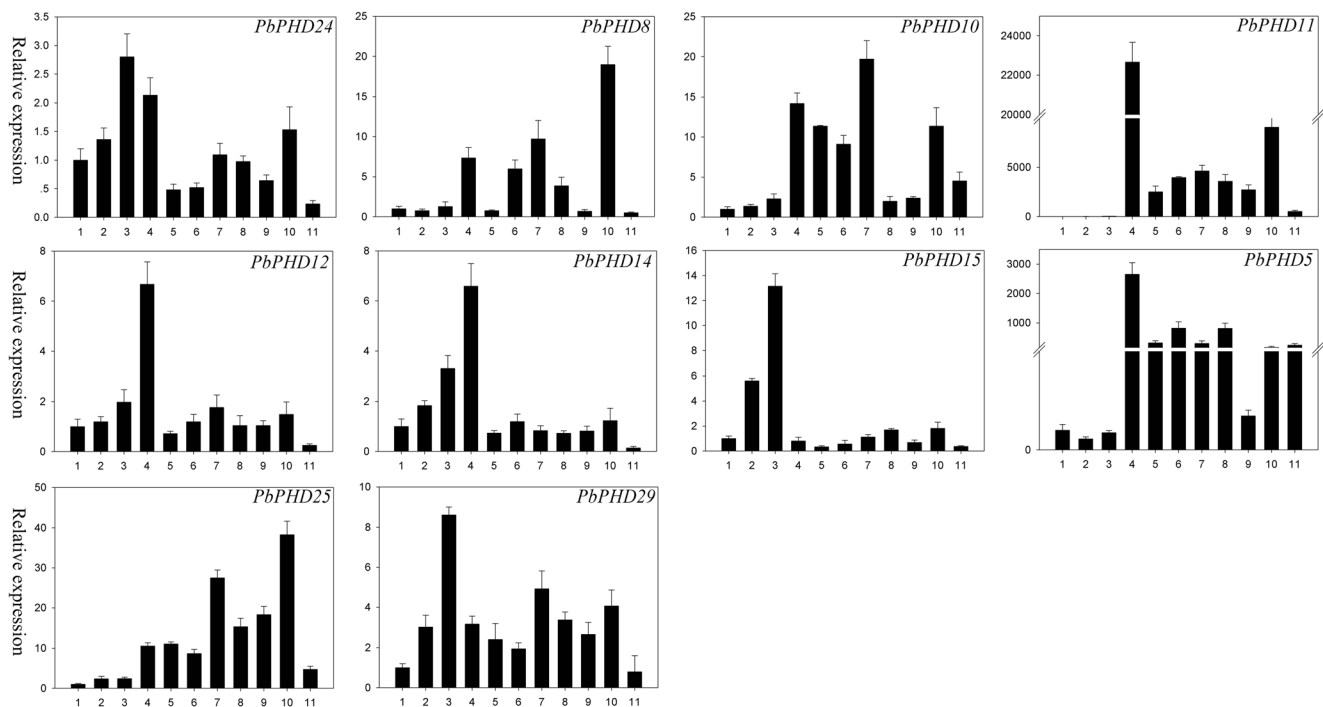


Fig. 9 Expression analysis of *PbPHD* genes in different plant tissues. The sample identities are as follows: 1, root; 2, stem; 3, leaf; 4, 15 days after flowering (DAF); 5, 39 DAF; 6, 47 DAF; 7, 55 DAF; 8, 63 DAF; 9,

79 DAF; 10, 102 DAF; and 11, 145 DAF. Error bars show the standard error between three replicates

In our study, 129 PHD-finger genes were screened from 5 Rosaceae species, among them various numbers of PHD-finger genes were identified in pear (31), apple (31), mei (20), peach (25), and strawberry (23). The number of PHD-finger genes in pear and apple were almost two-fold duplications as compare to mei, peach, and strawberry. This gene family has undergone an expansion at different extensions, which corresponds to the variation in chromosome number (Shulaev et al. 2011; Velasco et al. 2010; Verde et al. 2013; Wu et al. 2013; Zhang et al. 2012). For example, the chromosome number in apple and pear is 34, but there are only 16 chromosomes in peach and mei and 14 chromosomes in the strawberry. A recent genome-wide duplication event was shared by the apple and pear 30–45 MYA (million years ago) (Fig. 1) (Wu et al. 2013), the PHD-finger gene family expansion in the apple and pear might occur at that time. Remarkably, we also observed that there was no positive association between genome sizes and the number of PHD-finger gene family members in these Rosaceae species. For example, although there was no significant difference in the genome size of pear (271.9 Mb) and strawberry (240 Mb), while the number of PHD-finger genes obviously changed. In contrast, the number of PHD-finger genes in peach (224.6 Mb) and strawberry (240 Mb) have a positive correlation with their genome size.

To better understand the evolutionary relationships, a comparative phylogenetic analysis was performed. Then this tree was divided into 5 clades with 16 subfamilies. We found that

most of the subfamilies were dominated by monocots and dicots in addition to subfamily I (c). These results suggested that the evolution of the PHD-finger gene families in plants (monocots and dicots) was relatively conservative. However, the PHD-finger genes from dicots, such as the poplar and pear, were classified in subfamily I (c), indicating that the PHD-finger genes in these subfamilies might have species-specific roles. These genes were lost in the monocots maize and rice or obtained in dicot lineages after the divergence from their last common ancestor. In addition, the phylogenetic tree was also showed that pear have strong association/relationship with apple as compared to other Rosaceae species. This result was supported by the interspecific microsynteny analysis. For example, 25 orthologous genes were found between pear and apple, but we found only 20 orthologous genes between pear and mei and 20 orthologous genes between pear and strawberry.

Gene duplication is the main way to produce new genes, which provide the raw material for the evolution of species (Faris et al. 2008; Moore and Purugganan 2003). Gene duplications, including tandem duplication, segmental duplication, and transposition events, were the primary driving force of the gene family expansion during the evolution process, such as the *GRF*, *MYB*, and *WRKY* family in pear (Cao et al. 2016c, d; Huang et al. 2015), the *CHS* family in maize (Han et al. 2016), the *HD-Zip* family in soybean (Belamkar et al. 2014), the 14-3-3 gene family in cotton (Sun et al. 2011), and γ -gliadin gene family in wheat (Sun et al. 2011). In the present study, the

PHD-finger genes were assigned to segmental/WGD duplication blocks, and no PHD-finger genes were assigned to tandem duplication blocks, which indicated that segmental/WGD duplication was the main expansion mechanism of this PHD-finger gene family. Certain PHD-finger subfamilies have increased rapidly during the evolution process. In addition, the Ka/Ks ratios of eight *PbPHD* duplications were < 1 , indicating that the *PbPHD* gene family has gone through purifying selection and highly conserved evolution. Remarkably, previous studies have indicated that the genome of the pear has undergone two genome-wide duplication events. The first genome-wide duplication event occurred at ~ 140 MYA (Ks ~ 1.5 – 1.8) (Fawcett et al. 2009), while the latest duplication event occurred in the period of 30–45 MYA (Ks ~ 0.15 – 0.3) (Wu et al. 2013). As shown in Table 1, these PHD-finger gene duplications may have been derived from the same recent genome-wide duplication events and the same ancient genome-wide duplication events, indicating that the PHD-finger family is an ancient gene family that expanded with the recent genome-wide duplication events.

Previous studies showed that the consensus sequences of the PHD-finger domain were C-X(1–2)-C-X(8–19)-C-X(2–4)-C-X(4–6)-H-X2-C-X(11–26)-C-X(2–3)-C in *Arabidopsis* (Fan et al. 2009), C-X (1–2)-C-X(8–28)-C-X(2–4)-C-X(4–6)-H-X 2-C-X(12–34)-C-X 2-C in poplar (Wu et al. 2016a), C-X(1–2)-C-X (8–19)-C-X(2–4)-C-X(4–6)-H-X2-C-X(11–26)-C-X(2–3)-C in maize (Wang et al. 2015), and C-X (1–2) -C-X (8–14) -C-X (2–8) -C-X (4–6) -H-X 2 -C-X (11–34) -C-X 2 -C in carrot (Wu et al. 2016b). In the present study, we found that the consensus sequence of the PHD-finger domain was C-X (1–2) -C-X (8–14) -C-X (2–3) -C-X 4- H-X 2 -C- X (12–18) -C-X 2 -C in pear. The difference among the five plants proposed that there have been significant changes in the PHD-finger proteins during the evolution process. The *PbPHDs* were divided into seven subfamilies, and the majority of these genes were similar in their exon-intron structures and included conserved motifs. The characteristics of each PHD sequence may be highly diverse due to the different numbers of PHD-finger motifs and other motifs, which is consist of many domains, such as the SNF domain (Eisen et al. 1995; Linder et al. 2004), DDT domain (Doerks et al. 2001), and PWWP domain (Qiu et al. 2002). In general, most *PbPHDs* in the same subgroup contain similar motifs, which was consistent with the phylogenetic analysis. Interestingly, several motifs only existed in specific-subgroups, which indicated that these motifs might be involved in the functional differentiation of PHD-fingers in pear.

The gene expression patterns could provide important clues for exploring gene function (Budak and Zhang 2017; Neilson et al. 2017; Thomas et al. 2018). Previous studies have confirmed that the expression of the PHD-finger genes was affected by salt, PEG, and ABA in maize or drought stress in poplar (Wang et al. 2015; Wu et al. 2016a). However, still nobody

reported the PHD-finger genes roles during the fruit development. The stone cell is an important factor that affects the quality of pear, and lignin is the main component of stone cells (Cai et al. 2010; Jin et al. 2013). In the present study, ten genes were identified to be expressed in five developmental stages of pear fruit based on Illumina RNA-Seq data, suggesting their important roles in the development of pear fruit. Furthermore, the qRT-PCR showed that the expression level of *PbPHD10* was consistent with the changing trend of lignin contents during different stages of pear fruit development (Cai et al. 2010; Jin et al. 2013). Interestingly, the expression of *PbPHD10* was obviously increased at 55 DAF, and it showed an expression pattern that was similar to that of key genes that are involved in the regulation of the lignin synthesis pathway (Xie et al. 2013). These results strongly implied that *PbPHD10* might regulate lignin synthesis in pear fruit. Remarkably, we found that a segmental duplication gene pair (*PbPHD14* and *PbPHD24*) exhibited similar expression levels not only in the root, stem, and leaf but also during pear fruit development, indicating that these genes did not diverge much along with the evolution of each gene after the duplication.

Conclusions

In the present study, 129 non-redundant PHD-finger members were identified in 5 Rosaceae species. Among these members, 31 PHD-finger proteins were from the pear. Using integrated methods, including a phylogenetic, gene structure, sequence feature, duplication event, and orthologous analyses, it was found that although with a certain degree of divergence during the evolution process, certain well-conserved motifs were sustained in members of the PHD-finger family. The expression analysis suggested that the *PbPHD* genes exhibited different expression patterns during pear fruit development. In addition, *PbPHD10* was found to be possibly involved in regulating lignin synthesis in pear fruit. *PbPHD10* may be selected as a target for a functional investigation to improve the quality of pear fruit in the future.

Acknowledgments This study was supported by the National Natural Science Foundation of China (grant 31640068) and the 2017 Graduate Innovation Fund of the Anhui Agriculture University (grant 2017yis-31).

Author contributions Yunpeng Cao and Yahui Han conceived and designed the experiments; Yunpeng Cao and Yahui Han performed the experiments; Yunpeng Cao, Qing Jin, Yahui Han, and Dandan Meng analyzed the data; Yunpeng Cao, Yahui Han, Dahui Li, Yi Lin, and Yongping Cai contributed reagents/materials/analysis tools; Yunpeng Cao and Yahui Han wrote the paper.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

- Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME suite. *Nucleic Acids Res* 43:W39–W46
- Belamkar V, Weeks NT, Bharti AK, Farmer AD, Graham MA, Cannon SB (2014) Comprehensive characterization and RNA-Seq profiling of the HD-Zip transcription factor family in soybean (*Glycine max*) during dehydration and salt stress. *BMC Genomics* 15:1
- Bienz M (2006) The PHD finger, a nuclear protein-interaction domain. *Trends Biochem Sci* 31:35–40
- Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 72:291–336
- Budak H, Zhang B (2017) MicroRNAs in model and complex organisms. *Funct Integr Genomics* 17:1–4
- Cai Y, Li G, Nie J, Lin Y, Nie F, Zhang J, Xu Y (2010) Study of the structure and biosynthetic pathway of lignin in stone cells of pear. *Sci Hortic* 125:374–379
- Cao Y, Han Y, Li D, Lin Y, Cai Y (2016a) Systematic analysis of the 4-Coumarate:coenzyme a ligase (4CL) related genes and expression profiling during fruit development in the Chinese pear. *Genes* 7:89
- Cao Y, Han Y, Meng D, Li D, Jin Q, Lin Y, Cai Y (2016b) Structural, evolutionary, and functional analysis of the class III peroxidase gene family in Chinese pear (*Pyrus bretschneideri*). *Front Plant Sci* 7:1874
- Cao YP, Han Y, Jin Q, Lin Y, Cai Y (2016c) Comparative genomic analysis of the GRF genes in Chinese pear (*Pyrus bretschneideri* Rehd), poplar (*populus*), grape (*Vitis vinifera*), *Arabidopsis* and rice (*Oryza sativa*). *Front Plant Sci* 7:1750. <https://doi.org/10.3389/fpls.2016.01750>
- Cao YP, Han Y, Li D, Lin Y, Cai Y (2016d) MYB transcription factors in Chinese pear (*Pyrus bretschneideri* Rehd.): genome-wide identification, classification and expression profiling during fruit development. *Front Plant Sci* 7:577. <https://doi.org/10.3389/fpls.2016.00577>
- Capili AD, Schultz DC, Rauscher FJ, Borden KL (2001) Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. *EMBO J* 20:165–177
- Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676
- De Geyter N, Gholami A, Goormachtig S, Goossens A (2012) Transcriptional machineries in jasmonate-elicited plant secondary metabolism. *Trends Plant Sci* 17:349–359
- Doerks T, Copley R, Bork P (2001) DDT – a novel domain in different transcription and chromosome remodeling factors. *Trends Biochem Sci* 26:145–146
- Eisen JA, Sweder KS, Hanawalt PC (1995) Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res* 23:2715–2723
- Fan Z, ZhongNan Y, Sen Z (2009) Genome-wide analysis of PHD-finger protein family in *Arabidopsis thaliana*. *Acta Bot Yunnanica* 31: 227–238
- Faris JD, Zhang ZC, Fellers JP, Gill BS (2008) Micro-colinearity between rice, *Brachypodium*, and *Triticum monococcum* at the wheat domestication locus Q. *Funct Integr Genomics* 8:149–164
- Fawcett JA, Maere S, Van de Peer Y (2009) Plants with double genomes might have had a better chance to survive the cretaceous–tertiary extinction event. *Proc Natl Acad Sci* 106:5737–5742
- Gibbons RJ, Bachoo S, Picketts DJ, Afimos S, Asenbauer B, Bergoffen JA, Berry SA, Dahl N, Fryer A, Keppler K, Kurosawa K, Levin ML, Masuno M, Neri G, Pierpont ME, Slaney SF, Higgs DR (1997) Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain. *Nat Genet* 17:146–148
- Greb T, Mylne JS, Crevillen P, Geraldo N, An H, Gendall AR, Dean C (2007) The PHD finger protein VRN5 functions in the epigenetic silencing of *Arabidopsis* FLC. *Curr Biol* 17:73–78
- Guo A, Zhu Q, Chen X, Luo J (2007) GSDB: a gene structure display server. *Yi Chuan* 29:1023–1026
- Han Y, Ding T, Su B, Jiang H (2016) Genome-wide identification, characterization and expression analysis of the Chalcone synthase family in maize. *Int J Mol Sci* 17:161
- Huang X, Li K, Xu X, Yao Z, Jin C, Zhang S (2015) Genome-wide analysis of WRKY transcription factors in white pear (*Pyrus bretschneideri*) reveals evolution and patterns under drought stress. *BMC Genomics* 16:1
- Jain A, Das S (2016) Synteny and comparative analysis of miRNA retention, conservation, and structure across Brassicaceae reveals lineage- and sub-genome-specific changes. *Funct Integr Genomics* 16:253–268
- Jin Q, Yan C, Qiu J, Zhang N, Lin Y, Cai Y (2013) Structural characterization and deposition of stone cell lignin in Dangshan Su pear. *Sci Hortic* 155:123–130
- Kaadige MR, Ayer DE (2006) The polybasic region that follows the plant homeodomain zinc finger 1 of Pfl is necessary and sufficient for specific phosphoinositide binding. *J Biol Chem* 281:28831–28836
- Kalyanaraman A et al (2010) The genome of the domesticated apple (*Malus domestica* Borkh.). *Nat Genet* 42:833–839
- Kanehisa M, Goto S (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30
- Kehle J, Beuchle D, Treuheit S, Christen B, Kennison JA, Bienz M, Müller J (1998) dMi-2, a hunchback-interacting protein that functions in polycomb repression. *Science* 282:1897–1900
- Kramps T, Peter O, Brunner E, Nellen D, Froesch B, Chatterjee S, Murone M, Züllig S, Basler K (2002) Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear β -catenin-TCF complex. *Cell* 109:47–60
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* 19:1639–1645
- Lan F, Collins RE, de Cegli R, Alpatov R, Horton JR, Shi X, Gozani O, Cheng X, Shi Y (2007) Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* 448:718–722
- Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 40:D302–D305
- Li L, Stoeckert CJ, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13:2178–2189
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- Lin F, Xing K, Zhang J, He X (2012) Expression reduction in mammalian X chromosome evolution refutes Ohno's hypothesis of dosage compensation. *Proc Natl Acad Sci U S A* 109:11752–11757
- Linder B, Cabot RA, Schwickert T, Rupp RAW (2004) The SNF2 domain protein family in higher vertebrates displays dynamic expression patterns in *Xenopus laevis* embryos. *Gene* 326:59–66
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25:402–408
- Martin DG et al (2006) The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. *Mol Cell Biol* 26:7871–7879
- Moore RC, Purugganan MD (2003) The early stages of duplicate gene evolution. *Proc Natl Acad Sci* 100:15682–15687
- Muthamilarasan M, Khandelwal R, Yadav CB, Bonthala VS, Khan Y, Prasad M (2014) Identification and molecular characterization of MYB transcription factor superfamily in C 4 model plant foxtail millet (*Setaria italica* L.)
- Neilson J et al (2017) Gene expression profiles predictive of cold-induced sweetening in potato. *Funct Integr Genomics* 17:1–18

- Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* 40:1413–1415
- Papoulas O, Beek SJ, Moseley SL, McCallum CM, Sarte M, Shearn A, Tamkun JW (1998) The *Drosophila* trithorax group proteins BRM, ASH1 and ASH2 are subunits of distinct protein complexes. *Development* 125:3955–3966
- Patel RK, Jain M (2012) NGS QC toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7:544–548
- Pauwels L, Barbero GF, Geerinck J, Tillemans S, Grunewald W, Pérez AC, Chico JM, Bossche RV, Sewell J, Gil E, García-Casado G, Witters E, Inzé D, Long JA, de Jaeger G, Solano R, Goossens A (2010) NINJA connects the co-repressor TOPLESS to jasmonate signalling. *Nature* 464:788–791
- Proost S, Van BM, Vanechoutte D, Van dPY, Inzé D, Mueller-Roeber B, Vandepoel K (2015) PLAZA 30: an access point for plant comparative genomics. *Nucleic Acids Res* 43:D974
- Punta M et al. (2011) The Pfam protein families database. *Nucleic Acids Res* gkr1065
- Qiu C, Sawada K, Zhang X, Cheng X (2002) The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat Struct Biol* 9:217–224
- Schindler U, Beckmann H, Cashmore AR (1993) HAT3. 1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *Plant J* 4:137–150
- Schultz DC, Friedman JR, Rauscher FJ (2001) Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the mi-2 α subunit of NuRD. *Genes Dev* 15:428–443
- Shi X, Kachirskaia I, Walter KL, Kuo JHA, Lake A, Davrazou F, Chan SM, Martin DGE, Fingerman IM, Briggs SD, Howe LA, Utz PJ, Kutateladze TG, Lugovskoy AA, Bedford MT, Gozani O (2007) Proteome-wide analysis in *Saccharomyces cerevisiae* identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J Biol Chem* 282:2450–2455
- Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton JM, Rees DJG, Williams KP, Holt SH, Rojas JJR, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troglio M, Viola R, Ashman TL, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Girona EL, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arus P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). *Nat Genet* 43:109–116
- Sun G, Xie F, Zhang B (2011) Transcriptome-wide identification and stress properties of the 14-3-3 gene family in cotton (*Gossypium hirsutum* L.). *Funct Integr Genomics* 11:627–636
- Takatsuji H (1998) Zinc-finger transcription factors in plants. *Cell Mol Life Sci CMLS* 54:582–596
- 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
- Thomas J, Bowman MJ, Vega A, Kim HR, Mukherjee A (2018) Comparative transcriptome analysis provides key insights into gene expression pattern during the formation of nodule-like structures in *Brachypodium*. *Funct Integr Genomics*:1–12
- Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troglio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niaz F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mráz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchiatti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagné D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouzé P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel CE, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus [times] domestica* Borkh.). *Nat Genet* 42:833–839
- Verde I et al (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet* 45:487–494
- Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH (2012) MScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res* 40:e49–e49
- Wang Q, Liu J, Wang Y, Zhao Y, Jiang H, Cheng B (2015) Systematic analysis of the maize PHD-finger gene family reveals a subfamily involved in abiotic stress response. *Int J Mol Sci* 16:23517–23544
- Wu T, Zhang R, Gu C, Wu J, Wan H, Zhang S, Zhang S (2012) Evaluation of candidate reference genes for real time quantitative PCR normalization in pear fruit. *Afr J Agric Res* 7:3701–3704
- Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, Khan MA, Tao S, Korban SS, Wang H, Chen NJ, Nishio T, Xu X, Cong L, Qi K, Huang X, Wang Y, Zhao X, Wu J, Deng C, Gou C, Zhou W, Yin H, Qin G, Sha Y, Tao Y, Chen H, Yang Y, Song Y, Zhan D, Wang J, Li L, Dai M, Gu C, Wang Y, Shi D, Wang X, Zhang H, Zeng L, Zheng D, Wang C, Chen M, Wang G, Xie L, Sovero V, Sha S, Huang W, Zhang S, Zhang M, Sun J, Xu L, Li Y, Liu X, Li Q, Shen J, Wang J, Paull RE, Bennetzen JL, Wang J, Zhang S (2013) The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome Res* 23:396–408
- Wu S, Wu M, Dong Q, Jiang H, Cai R, Xiang Y (2016a) Genome-wide identification, classification and expression analysis of the PHD-finger protein family in *Populus trichocarpa*. *Gene* 575:75–89
- Wu XJ, Li MY, Que F, Wang F, Xu ZS, Xiong AS (2016b) Genome-wide analysis of PHD family transcription factors in carrot (*Daucus carota* L.) reveals evolution and response to abiotic stress. *Acta Physiol Plant* 38:1–15
- Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C, Allis CD (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442:86–90
- Xie M, Huang Y, Zhang Y, Wang X, Yang H, Yu O, Dai W, Fang C (2013) Transcriptome profiling of fruit development and maturation in Chinese white pear (*Pyrus bretschneideri* Rehd.). *BMC Genomics* 14:823
- Yang X, Makaroff CA, Ma H (2003) The Arabidopsis MALE MEIOCYTE DEATH1 gene encodes a PHD-finger protein that is required for male meiosis. *Plant Cell* 15:1281–1295
- Yang C, Vizcay-Barrena G, Conner K, Wilson ZA (2007) Male sterility 1 is required for tapetal development and pollen wall biosynthesis. *Plant Cell* 19:3530–3548
- Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J (2006) WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res* 34:W293–W297
- Zhang Q et al (2012) The genome of *Prunus mume*. *Nat Commun* 3:187–190