

Comparative genomic analysis of the IDD genes in five Rosaceae species and expression analysis in Chinese White Pear (*Pyrus bretschneideri*) (#31623)

1

First submission

Editor guidance

Please submit by **9 Nov 2018** for the benefit of the authors (and your \$200 publishing discount).



Structure and Criteria

Please read the 'Structure and Criteria' page for general guidance.



Raw data check

Review the raw data. Download from the [materials page](#).



Image check

Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the [materials page](#).

13 Figure file(s)

9 Table file(s)

1 Other file(s)



Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. BASIC REPORTING
2. EXPERIMENTAL DESIGN
3. VALIDITY OF THE FINDINGS
4. General comments
5. Confidential notes to the editor

 You can also annotate this PDF and upload it as part of your review

When ready [submit online](#).

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your [guidance page](#).

BASIC REPORTING

-  Clear, unambiguous, professional English language used throughout.
-  Intro & background to show context. Literature well referenced & relevant.
-  Structure conforms to [PeerJ standards](#), discipline norm, or improved for clarity.
-  Figures are relevant, high quality, well labelled & described.
-  Raw data supplied (see [PeerJ policy](#)).

EXPERIMENTAL DESIGN

-  Original primary research within [Scope of the journal](#).
-  Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
-  Rigorous investigation performed to a high technical & ethical standard.
-  Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

-  Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
-  Data is robust, statistically sound, & controlled.
-  Speculation is welcome, but should be identified as such.
-  Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

3



The best reviewers use these techniques

Tip

Support criticisms with evidence from the text or from other sources

Example

Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.

Give specific suggestions on how to improve the manuscript

Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).

Comment on language and grammar issues

The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.

Organize by importance of the issues, and number your points

1. Your most important issue
2. The next most important item
3. ...
4. The least important points

Please provide constructive criticism, and avoid personal opinions

I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC

Comment on strengths (as well as weaknesses) of the manuscript

I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.

Comparative genomic analysis of the IDD genes in five Rosaceae species and expression analysis in Chinese White Pear (*Pyrus bretschneideri*)

Xueqiang Su¹, Yu Zhao¹, Guohui Li¹, Xi Cheng¹, Muhammad Abdullah¹, Xu Sun¹, Yongping Cai^{Corresp., 1}, Yi Lin^{Corresp., 1}

¹ School of Life Science, Anhui Agricultural University, Hefei, China

Corresponding Authors: Yongping Cai, Yi Lin

Email address: 810730810@qq.com, linyi992547404@163.com

The INDETERMINATE DOMAIN (IDD) gene family encodes hybrid transcription factors with distinct zinc finger motifs and appears to be found in all higher plant genomes. In the model plants *Arabidopsis* and rice, IDD genes have been identified throughout the genome, and the functions of many members have been studied. However, this gene family has been studied less in Rosaceae species, among which genome-wide identification has only been completed in apple, while there is still no comprehensive research in pear. This study focuses on comparative genomic analysis of the IDD gene family in five Rosaceae species (pear, strawberry, plum, raspberry and cherry). We identified a total of 68 IDD genes: 16 genes in Chinese pear (*Pyrus bretschneideri*), 14 genes in strawberry (*Fragaria vesca*), 13 genes in plum (*Prunus mume*), 14 genes in raspberry (*Rubus occidentalis*) and 11 genes in cherry (*Prunus avium*). The evolution of the IDD genes of these five Rosaceae species was revealed by constructing a phylogenetic tree, tracking gene duplication events, sliding window analysis, and conserved microsynteny analysis. The expression analysis of different tissues showed that most of the pear IDD genes had a very high transcription level in fruit, flower, and bud. Combining our results with previous research, we speculate that *PbIDD2* and *PbIDD8* may participate in plant flowering induction in pear. Temporal expression analysis showed that the expression patterns of *PbIDD3* and *PbIDD5* were completely opposite to the accumulation pattern of fruit lignin and stone cell contents. Combining the results of the composite phylogenetic tree and expression pattern analysis showed that *PbIDD3* and *PbIDD5* might be involved in the metabolism of lignin and secondary cell wall (SCW) formation. In summary, we provide basic information about the IDD genes of five Rosaceae species, providing a theoretical basis for deeper study of the functions of these genes.

Comparative genomic analysis of the IDD genes in five Rosaceae species and expression analysis in Chinese White Pear (*Pyrus bretschneideri*)

Xueqiang Su[#], Yu Zhao[#], Guohui Li, Xi Cheng, Muhammad Abdullah, Xu Sun, Yongping Cai*, Yi Lin*

School of Life Science, Anhui Agricultural University, No. 130, Changjiang West Road, Hefei 230036, China;

[#] This authors have contributed equally to this work.

*Corresponding author: ypcaiah@163.com (Yongping Cai);

Co-Corresponding author: linyi1957@126.com (Yi Lin);

Abstract

The INDETERMINATE DOMAIN (IDD) gene family encodes hybrid transcription factors with distinct zinc finger motifs and appears to be found in all higher plant genomes. In the model plants *Arabidopsis* and rice, IDD genes have been identified throughout the genome, and the functions of many members have been studied. However, this gene family has been studied less in Rosaceae species, among which genome-wide identification has only been completed in apple, while there is still no comprehensive research in pear. This study focuses on comparative genomic analysis of the IDD gene family in five Rosaceae species (pear, strawberry, plum, raspberry and cherry). We identified a total of 68 IDD genes: 16 genes in Chinese pear (*Pyrus bretschneideri*), 14 genes in strawberry (*Fragaria vesca*), 13 genes in plum (*Prunus mume*), 14 genes in raspberry (*Rubus occidentalis*) and 11 genes in cherry (*Prunus avium*). The evolution of the IDD genes of these five Rosaceae species was revealed by constructing a phylogenetic tree, tracking gene duplication events, sliding window analysis, and conserved microsynteny analysis. The expression analysis of different tissues showed that most of the pear IDD genes had a very high transcription level in fruit, flower, and bud. Combining our results with previous research, we speculate that *PbIDD2* and *PbIDD8* may participate in plant flowering induction in pear. Temporal expression analysis showed that the expression patterns of *PbIDD3* and *PbIDD5* were completely opposite to the accumulation pattern of fruit lignin and stone cell contents. Combining the results of the composite phylogenetic tree and expression pattern analysis showed that *PbIDD3* and *PbIDD5* might be involved in the metabolism of lignin and secondary cell wall (SCW) formation. In summary, we provide basic information about the IDD genes of five Rosaceae species, providing a theoretical basis for deeper study of the functions of these genes.

Keywords: *INDETERMINATE DOMAIN (IDD)* genes, pear, phylogenetic analysis, microsynteny, lignin synthesis, SCW formation

INTRODUCTION

Zinc finger proteins are transcription factors with a finger-like domain. They are widely distributed in animals and microorganisms, as well as the plant kingdom (Miller et al., 1985). Zinc finger proteins have been identified and functionally

analysed in several plant species, such as *Jatropha curcas* (Shi et al., 2018), *Oryza sativa* (Zhang et al., 2018), *Musa acuminata* (Chen et al., 2014), *Gossypium hirsutum* (Ren et al., 2018). One group of this large family of proteins, the INDETERMINATE DOMAIN (IDD) proteins, have a highly conserved ID domain (Colasanti et al., 1998), which has typical C2H2 and C2HC zinc finger motifs (Wu et al., 2008). C2H2 zinc finger transcription factors are one of the most thoroughly studied transcription factor families (Agarwal et al., 2007; Wei et al., 2016). The C2H2 zinc finger transcription factors contain tandem repeat segments of approximately 30 amino acids, all of which have a highly conserved amino acid sequence: (F/Y)-XC-X2-5-C-X3-(F/Y)-X5-psi-X2-H-X3-5-H (wherein C and H represent cysteine and histidine, respectively, X means any amino acid at this position, and psi means a hydrophobic residue) (Parraga et al., 1988). This particular sequence structure can bind to Zn^{2+} and fold to form a structure with two β -sheets and an α -helix. Zn^{2+} mainly plays a role in linking individual amino acid chains and is also the key to the function of zinc finger protein (Islam et al., 2009; Tian et al., 2010).

The INDETERMINATE DOMAIN (IDD) protein family is a special type of C2H2 zinc finger subgroup. These proteins have a highly conserved ID domain with a nuclear localization signal and four distinct zinc finger motifs (ZF1, ZF2, ZF3, ZF4) (Colasanti et al., 2006; Kozaki et al., 2004). A quantitative study has shown that members of the IDD gene family have diverse biological functions in plants, most of which involve plant growth and development. The functions of many IDD genes in Arabidopsis have been reported. *AtIDD1* (ENHYDROUS) promotes seed germination by regulating the signalling pathways of light and hormones (Feurtado et al., 2011). *AtIDD3* (MAGPIE) and *AtIDD10* (JACKDAW) are closely related to the development of roots, which can regulate tissue boundaries and asymmetric cell division (Welch et al., 2007). *AtIDD5* regulates the expression of the SS4 gene while positively regulating the biosynthesis of starch granules (Ingkasuwan et al., 2012). *AtIDD8* interacts with the promoter of the SUS4 gene to regulate plant flowering and photoperiod signalling (Seo et al., 2011). At the same time, *AtIDD8* can also be inhibited by the key gene ANIN10 in the sugar signalling pathway, causing changes in downstream gene expression, thereby regulating the plant flowering process (Jeong et al., 2015). *AtIDD14*, *15* and *16* cooperate to regulate lateral organ morphogenesis and gravitropism by promoting auxin biosynthesis and transport (Cui et al., 2013). In maize, *ZmID1* regulates a leaf-generated signal required for the transition to flowering. *ZmID1* produces a transmissible signal in the leaf, causing the transformation of the stem tip in reproductive development (Colasanti et al., 1998). Also in maize, two homologous genes were identified from the naked endosperm (nkd) mutants by map-based cloning, named *nkd1* and *nkd2*, which were subsequently shown to be maize *ZmIDDveg9* and *ZmIDD9*, respectively (Yi et al., 2015). In rice, *OsIDD10* mediates ammonium absorption and N metabolism in roots (Xuan et al., 2013). *OsID1* plays an important role in the development of flowers from the vegetative stage to the reproductive stage (Park et al., 2008).

Pear, belonging to the Rosaceae family, is widely cultivated around the world and is popular among people because of its unique flavour. The history of eating pears in China is very long, and the variety of pears grown is very diverse. 'Dangshan Su' pear (*Pyrus bretschneideri* cv. Dangshan Su) belongs to the white pear variety, a characteristic pear variety produced in Dangshan County, Anhui Province, China. Because of its good quality and high medicinal efficacy, it has a very high market value. The stone cells in the pear pulp are one of the important sources of this pear's unique taste, and they determine the quality of the pear fruit (Cai et al., 2010; Yan et al., 2014). The content and size of stone cells in pear fruit have a great influence on the taste. The main constituent of pear stone cells is lignin, which accounts for 18% of the content of stone cells (Lee et al., 2007). Because stone cells are part of a secondary cell wall thickening process accompanied by lignin accumulation, lignin synthesis and secondary cell wall (SCW) formation are closely related to the formation of stone cells (Liu et al., 2012; Zhao et al., 2013). Some zinc finger proteins are involved in plant SCW formation. The CCCH-tandem zinc finger protein IIP4 interacts with NAC29 and NAC31 (NAC29 and NAC31 can regulate the formation of SCW formation), resulting in inhibition of plant SCW formation and thus altered plant mechanical strength (Zhang et al., 2018). Arabidopsis C3H14 and C3H15 have similar functions, which can

regulate the SCW formation and are closely related to the formation of pollen (Chai et al., 2015). Although the functions of the zinc finger protein family have been analysed in depth, the members of the IDD gene family have rarely been reported in lignin biosynthesis and SCW formation; such a role has only been reported in rice. *OsIDD2* can inhibit the expression of cinnamyl alcohol dehydrogenase (CAD2, CAD3), a key enzyme gene for lignin biosynthesis, thereby negatively regulating SCW formation and lignin biosynthesis (Huang et al., 2018). Although much research has been done on the function of IDD gene family members, whether IDD genes regulate lignin biosynthesis and SCW formation in pears is still unknown. The functions of most pear IDD family genes are unknown, though their basic function can be illustrated by the evolution of the pear IDD family genes.

Currently, complete genome-wide identification of the IDD gene family has been done only in the model plants *Arabidopsis*, rice, and maize (Colasanti et al., 2006). Among the Rosaceae species, only apples have had their IDD genes identified genome-wide (Fan et al., 2017), and this gene family is still unstudied in pear. In this study, we performed genome-wide identification and analysis of the pear IDD gene family. At the same time, to help us better understand the characteristics of the IDD gene family of pears, we also selected four Rosaceae species (strawberry, plum, raspberry, cherry) for comparative genomic analysis with pear. We identified 68 putative IDD genes from pear (16), strawberry (14), plum (13), raspberry (14), and strawberry (11). We analysed the evolutionary relationships, gene structures, conserved motifs, chromosomal locations, and duplication events of these five Rosaceae species' IDD genes. We analysed the cis-acting elements of the pear IDD gene promoter. We studied the expression patterns of the 16 IDD genes in pear at different stages of pear fruit development, including 15 DAF, 39 DAF, 47 DAF, 55 DAF, 63 DAF, 79 DAF, and 145 DAF, and in different pear tissues, including fruit, flower, bud, stem, leaf (pear fruit development represented by 15 DAF). Our research provides an important theoretical basis for understanding the pear IDD gene family.

MATERIALS AND METHODS

Identification of IDD Genes in Five Rosaceae Plants

In this study, the pear genome database was obtained from the Pear Genome Project. (<http://peargenome.njau.edu.cn/>). The genome databases of the other four Rosaceae plants (Strawberry, plum, Raspberry and Cherry) were downloaded from GDR (<https://www.rosaceae.org/>). DNATOOLS software was used to create a local database containing the amino acid sequences of the five Rosaceae plants' IDD genes (Cao et al., 2016a). To identify the IDD genes of these five Rosaceae species, we used the following methods. First, the *Arabidopsis* (16) and apple (20) IDD gene sequences were collected for use as query sequences in TBlastN with the default E-value (Supplementary Table S1). We compared these sequences with the local database sequences of the above five Rosaceae species. Then, the IDD candidate gene sequences initially screened by BLAST were tested using SMART online software to test whether they contained a zinc finger domain (<http://smart.embl-heidelberg.de/>) (Letunic et al., 2012). Finally, each candidate sequence containing a zinc finger domain was manually checked for an IDD domain. We discarded the protein sequences lacking a full IDD domain and redundant sequences. Finally, we isolated the candidate IDD gene sequences. The ExPASy online site was used for investigating the molecular weights of IDD proteins (<http://web.expasy.org/protparam/>) (Artimo et al., 2012). Prediction of subcellular localization results for all IDD proteins was performed on the WoLFPSORT website (<http://www.genscript.com/wolf-psort.html>) (Horton et al., 2007).

Phylogenetic Analysis

Sequence alignment of all IDD proteins was done using the ClustalW tool in MEGA 7.0 software. The phylogenetic tree was constructed using MEGA 7.0 software using the maximum likelihood (ML) (bootstrap = 1000) (Kumar et al., 2016).

sequence of apple used in the phylogenetic tree can from the study of Fan (Fan et al., 2017). The sequences of *Arabidopsis thaliana*, rice, and maize were obtained from the article by Colasanti (Colasanti et al., 2006).

IDD gene Structures and Conserved Motif Prediction

IDD gene structures from pear, strawberry, plum, raspberry and cherry were compared using Gene Structure Server (<http://gsds.cbi.pku.edu.cn>) (Guo et al., 2007). The motifs of the IDD genes in these five Rosaceae species were analysed using MEME online analysis software (http://meme.sdsc.edu/meme4_3_0/intro.html) (Bailey et al., 2015). Parameters for the conserved motif prediction were motif width greater than 6 and less than 200. The number of identified motifs was 20.

Chromosomal Location, Gene Duplication and Ka/Ks Ratio Analysis

Pear, strawberry, plum, raspberry and cherry chromosome start positions and other relevant information about the IDD genes were obtained from the public genome database. The chromosomal physical locations of the IDD genes of all five Rosaceae species were mapped using MapInspect software (<http://mapinspect.software.informer.com>) (Niu et al., 2016). To define gene duplication events, we mainly relied on the following principles: 1, The similarity of the two gene-coding sequences (CDS) was more than 80%. 2, If the two genes were located on the same chromosome and separated by at least 200 kb, we considered these two genes tandem-duplicated genes. 3, If the two genes were located on different chromosomes, they were called segmentally duplicated genes (Long & Thornton, 2001). DnasP v5.0 software was used to calculate non-synonymous (Ka) and synonymous substitution (Ks) values and perform sliding window analysis. Parameters for sliding window analysis were window size 150 bp and step size 9 bp (Librado & Rozas, 2009).

Microsynteny Analysis and Pear IDD Gene Promoter *Cis*-acting Element Analysis

We obtained the promoter sequence of each IDD gene from the Chinese white pear genome database, including the DNA sequence of the initiation codon (ATG) located 1500 bp upstream of each IDD gene. The online software Plantcare was used to analyse the *cis*-acting elements of the promoter region (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (Rombauts et al., 1999). The microsynteny of 6 Rosaceae species (we added the IDD genes of apple) was identified using the Multiple Collinearity Scan toolkit (MCSscanX) (Abdullah et al., 2018). IDD genes of pear, strawberry, plum, raspberry, cherry and apple were ordered according to the phylogenetic tree.

RNA Extraction and qRT-PCR

The plant material used in this experiment was the 'Dangshan Su' pear, which grows in the Yuanyichang agricultural park (Suzhou, Anhui, China). 'Dangshan Su' pear belongs to the white pear variety. In April 2017, we began collecting materials of various tissues of pears. ~~After pollination, we began collecting pear fruit samples.~~ The fruits were picked at 15 DAF, 39 DAF, 47 DAF, 55 DAF, 63 DAF, 79 DAF, and 145 DAF. RNA was extracted from plant material (including fruits and various tissues) using the Tiangen (Beijing, China) plant RNA extraction kit. Reverse transcription was performed using a PrimeScript™ RT reagent kit with gDNA Eraser (Takara, Japan). Each reaction used 1 µg of RNA. The qRT-PCR primers for the pear IDD genes were designed using Beacon Designer 7 software (Supplementary Table S8). The qRT-PCR system consisted of 20 µL: 10 µL of SYBR® Premix Ex Taq™ II (2×) (Takara, Japan), 2 µL of cDNA and 0.8 µL of PbIDD-F and PbIDD-R. The reaction procedure followed the  action manual, and it was run in three repetitions. We used the 2^{-ΔΔCt} method for the calculation of relative expression (Livak & Schmittgen, 2001).

RESULTS

1, Identification and chemical characterization of IDD gene family in five Rosaceae species

To identify the IDD gene family of pears and 4 other Rosaceae species, we use apple IDD as a query sequence against the pear, strawberry, plum, raspberry and cherry genome databases with the BLASTP programme with the default E-value (Cao et al., 2017). Subsequently, we manually checked candidate IDD proteins and discarded proteins with incomplete IDD domains. A total of 68 IDD proteins were identified and used for further analysis (Table 1; Supplementary Table S2). In Chinese pear, we identified 16 IDD proteins (*PbIDD1-PbIDD16*). We identified a total of 52 IDD proteins in the other four species, including 14 in strawberry (*FvIDD1-FvIDD14*), 13 in plum (*PmIDD1-PmIDD13*), 14 in raspberry (*RoIDD1-RoIDD14*), and 11 in cherry (*PaIDD1-PaIDD11*). Though only 11 IDD proteins were identified in cherry, the numbers of IDD proteins in the other species were very similar, but less than the number of IDD proteins in apple (20).

2, Phylogenetic Analysis of the IDD genes in pear, strawberry, plum, raspberry and cherry

Phylogenetic and evolutionary analysis of the five Rosaceae species was performed. At the same time, we added 20 IDD genes from apple and the genes from Arabidopsis, maize, and rice whose functions are known. A phylogenetic tree of a total of 98 IDD proteins was built using the maximum likelihood (ML) (Figure 1). Based on the phylogenetic tree, we identified four phylogenetic groups (groups1-4). Group1 had the most members, containing all six Rosaceae species and 44 IDD genes: pear (8), strawberry (7), plum (6), raspberry (7), cherry (6) and apple (10). In group3 and group2 we identified 21 and 14 IDD genes, respectively. Group4 had the fewest members: 2 pear genes, 2 strawberry, 2 plum, 2 raspberry, 1 cherry, and no apple. *ZmID1* was individually grouped into a class. Notably, three Arabidopsis genes (*AtIDD3*, *AtIDD8*, *AtIDD10*) and *OsIDD2* were found in group1. *AtIDD3* (MAGPIE) and *AtIDD10* (JACKDAW) are closely related to the development of roots, *AtIDD8* regulates flowering and photoperiod signals in plants, and *OsIDD2* can inhibit key genes of lignin biosynthesis to negatively regulate plant SCW formation. These results suggest that Rosaceae IDD genes in group1 may be involved in plant root development, flowering regulation, and SCW formation. *PbIDD3*, *PbIDD5*, *FvIDD3*, *RoIDD4*, *PmIDD3*, *PaIDD4* were included in group1 and were closely related to *OsIDD2*, suggesting that they are also likely to participate in plant SCW formation. Members of group2 may be related to plant absorption and metabolism of material, because *OsIDD10* was included in this group, which is involved in ammonium uptake and nitrogen (N) metabolism. *AtIDD1* was in group3, indicating that the IDD genes in group3 may be involved in the germination of plant seeds. *AtIDD15* and *OsIDD14* were identified in group4. These two genes are involved in plant gravitational induction and regulating plant architecture. Therefore, the functions of the 9 IDD genes in group4 may also be closely related to these aspects.

3, Structural and conserved motif analysis of IDD proteins

To more comprehensively analyse the structural diversity of the five Rosaceae species' IDD genes, we created exon-intron pattern maps of all 68 IDD genes (Figure 2). Of the 34 IDD genes in group1, 25 genes had 3 exons, and 6 genes had two exons. Additionally, *FvIDD7* (4), *RoIDD6* (5), *PmIDD8* (6) had a large number of exons. In group2, *FvIDD11* had two exons, *PbIDD7* had four exons, and all other members had three exons. Group3 was a subfamily with a complex gene structure: eleven genes had four exons, and four genes had three exons. All members of Group4 had a gene structure of 3 exons. The lack of high similarity between the 68 IDD genes allowed us to better understand the conserved motifs of these IDD genes. We used MEME software to identify 20 conserved motifs (Figure 2; Supplementary Table S3). Motif 2, Motif 3, Motif 1, and Motif 4 were identified to encode the ID domain, while the remaining motifs did not have functional annotations. As shown in Figure 3, Motif 2, Motif 3,

Motif 1, Motif 4 represented zinc fingers ZF1, ZF2, ZF3 and ZF4, respectively. The four conserved motifs constituted a conserved ID domain, while these four conserved motifs represented two C2H2 and two C2HC structures (ZF1 and ZF2 belonged to C2H2, and ZF3 and ZF4 belonged to C2HC). Two C2H2 and two C2HC structures were identified by sequence alignment of the conserved ID domains of five species, and the results showed that the ID domain was highly conserved. Cysteine (C) and histidine (H) in each zinc finger domain showed conservation in each species, in line with previous studies (Supplementary Figure S1-5). All 68 IDD genes had Motifs 1-4. These four motifs were thus the most conserved motifs. At the same time, we found that some members of the same group and the more closely related members had highly similar motif compositions (e.g., *FvIDD7* and *RoIDD6* in group1, *PmIDD10* and *PaIDD9* in group2). These genes were closely related and had exactly the same motif compositions, which may suggest similarities in the functions of these genes. We also identified certain group-specific motifs, such as motif 12, motif 17 and motif 20 in members of group 4 and motif 19 in some members of group 1. These group-specific motifs may be key factors that distinguish the IDD genes of different groups and play important roles in the distinct functions of these group members. In addition to the conserved IDD domain, we found two conserved sequences near the C-terminus of some IDD genes (Figure 2; Supplementary Figure S1-5). The sequence MSATALLOKAA (motif 6) was found at the C-terminus of most IDD genes and encoded a conserved sequence. Another conserved sequence was TRDFLG, encoded by motif 7. All 9 members of group 4 lacked both motif 6 and motif 7. In group1, *PaIDD10* did not have motif 6, and *PbIDD14* and *PmIDD8* lacked motif 7. All members of group2 and group3 contained these two motifs. The lack of these two conserved sequences at the C-terminus of some IDD genes may be the reason that the members of group4 were clustered into a single category and may also underlie differences in the functions of these proteins.

4, Chromosomal Location and Duplication Events of IDD Family Genes in Rosaceae

According to the complete genome sequences of the five Rosaceae species, the exact chromosomal locations of all 68 IDD genes were identified (Figure 4). In pear, the 16 IDD genes were distributed on chromosomes 3, 4, 8, 9, 11, 12, 14, 15, 16 and 17. In strawberry, the IDD genes were distributed on chromosomes 1, 2, 3, 4, 5 and 6. In raspberry, the IDD genes were mainly distributed on chromosomes 2, 4 and 6. Several other IDD genes were distributed on chromosomes 1, 3 and 5. In plum, except *PaIDD5*, which could not be mapped to any chromosome, the IDD genes were distributed on chromosomes 1, 2, 4, 6, 7 and 8. However, in cherry, 4 IDD genes were distributed on chromosome 1 and the remaining 7 IDD genes were distributed on chromosomes 3, 5, 6, 7 and 8. ~~Gene duplication events are the main driving force for the expansion of gene family members, among which tandem replication and fragment replication are the major types. To elucidate how the IDD genes were expanded after the five Rosaceae species had diverged, we investigated the gene duplication events of the five Rosaceae species' IDD gene families. Only 4 pairs of duplicated genes were identified in pear (Figure 4). However, we did not identify duplicated gene pairs in strawberry, plum, raspberry and cherry. It is worth noting that the four pairs of duplicated genes identified in the Chinese white pear were segmental duplications, indicating that segmental duplication has been the main driving force for the expansion of the IDD gene family in pears.~~

To clarify the driving forces of gene duplication and explore the impact of these genes on evolutionary processes, we calculated the Ka, Ks and Ka/Ks ratios for the 4 duplicated gene pairs in Chinese Pear. Ka/Ks=1 is the cut-off value that indicates neutral selection, Ka/Ks<1 represents negative selection and Ka/Ks>1 represents positive selection (Bitocchi et al., 2017). For all four duplicated gene pairs identified in Chinese white pear, the Ka/Ks values were less than 0.2726 (Supplementary Table S4). These results indicate that in Chinese white pear, the IDD gene family was expanded due to gene duplication events and that these IDD genes have experienced intense purifying selection. In the evolutionary processes of genes, positive selection may be overshadowed by strong negative selection (Han et al., 2016). Therefore, to comprehensively explain the selection pressure of IDD genes, we performed sliding window analysis of the 4 pairs of IDD paralogues in pears (Supplementary Figure S6). As

shown in Figure 5, the Ka/Ks ratios of the duplicated gene pair *PbIDD9-PbIDD10* were all less than 1 in the coding region, while for the other 3 duplicated gene pairs, the Ka/Ks ratios of most coding regions were less than 1 (the exceptions were a few distinct peaks with Ka/Ks > 1). Compared with the regions with higher Ka/Ks ratios, the conserved IDD domains had very low Ka/Ks ratios.

5, Microsynteny Analysis of IDD family genes

The locations of homologous genes can be determined by microsynteny analysis. In this study, we performed microsynteny analysis in pear, strawberry, plum, raspberry, cherry and apple. A total 30 orthologous gene pairs were identified in the crosses of the pear and strawberry, pear and plum, pear and raspberry, pear and cherry, pear and apple (Figure 5; Supplementary Table S5). There were 1 collinear gene pair between pear and strawberry, 8 pairs between pear and plum, 3 pairs between pear and raspberry, 6 pairs between pear and cherry, and 12 pairs between pear and apple. Interestingly, *PbIDD1* had collinear relationships with some members of the IDD gene families of several other species. This phenomenon may be because *PbIDD1* appeared before the divergence of common ancestor of these Rosaceae species. We identified some collinear gene pairs among pear and plum, such as *PbIDD6/PmIDD6* and *PbIDD12/PmIDD10*, while we could not identify any between pear and raspberry or pear and cherry, suggesting that these orthologous genes developed after plum diverged from the common ancestor of raspberry and cherry. At the same time, we identified some collinear genes in pear and cherry, such as *PbIDD8/PaIDD7* and *PbIDD7/PaIDD3*, but did not identify any in pear and plum. It is worth noting that *PbIDD2* also had a collinear relationship with the members of the *MdIDD*, *PmIDD* and *PaIDD* families. This suggests that *PbIDD2* appeared before the divergence of the common ancestor of pear, apple, plum, and cherry. On the other hand, two IDD genes from pear matched with one pair of plum or raspberry genes. For example, *PbIDD9* and *PbIDD10* were orthologous to *PmIDD9*, while *PbIDD9* and *PbIDD10* were to orthologous to *RoIDD7*. These findings suggest that they may belong to paralogous gene pairs.

6, Analysis of Cis-acting elements in the promoters of the IDD genes in Chinese pear

Plant IDD genes may be involved in multiple biological processes, and hormones affect their expression. We identified the promoter sequences of the 16 IDD genes in Chinese white pear (promoter length=1500 bp) (Supplementary Table S6). We identified the type and number of cis-acting elements in all 16 IDD gene promoter sequences using Plantcare online software (Figure 6; Supplementary Table S7). We further analysed the possible roles of the Chinese white pear IDD genes in various processes in plants. The results showed that Chinese white pear IDD gene promoters contained a variety of cis-acting elements, involved in a variety of biological processes. Most of the IDD genes (except *PbIDD7*, 9 and 10) contained a G-Box, which is a light-responsive element, indicating that the expression of these genes may be regulated by light. Half of the IDD genes contained HSE components (such as *PbIDD8*, 9, 12 and 16), which enhance plant tolerance to high temperature by activating HSP gene expression (Zhao et al., 2018). These IDD genes may play an important role in the regulation of high-temperature stress in plants. Approximately 70% of the Chinese white pear IDD genes had an MBS, suggesting that IDD gene expression may be regulated by MYB transcription factors. TC-rich repeats, which are cis-acting elements involved in defence and stress responsiveness, were identified in *PbIDD1*, 6, 7, 9, 10, 12, 14 and 16. In addition, there were many cis-acting elements related to the responses to hormones, including abscisic acid (ABRE), MeJA (CGTCA-motif, TGACG-motif), gibberellin (P-box, GARE-motif), auxin (TGA-element), salicylic acid (TCA-element). The auxin-related TGA-element was identified in only five members, which was the least number of homeopathic components. There were 41 cis-acting elements associated with MeJA, which were the most abundant cis-acting elements, suggesting that the expression of the IDD genes may be mainly regulated by MeJA. Of the 16 pear IDD genes, 12 contained 19 CGTCA-motifs, and TCA-elements appeared in the promoter regions of 10 members, for

a total of 22. These two cis-acting elements were abundant, indicating that *PbIDDs* play an important role in the plant stress response, possibly helping resist stress by promoting the biosynthesis of plant lignin. In addition, CAT-box and CCGTCC-box were found in 11 members (*PbIDD1*, 2, 3, 4, 5, 6, 7, 8, 9, 12 and 14), which are cis-acting elements related to meristem expression.

7, Expression Characteristics of Pear IDD Genes

To deepen our understanding of the Chinese white pear IDD gene family, we investigated the expression patterns of *PbIDDs* in different tissues (Figure 7). Among the 16 members, *PbIDD4*, 9 and 11 were highly expressed in bud, but in other tissues their expression was extremely low. *PbIDD1*, 10 was mainly expressed in flower. *PbIDD2*, 6, 8 and 15 were mainly expressed in fruit, flower and bud. Among them, *PbIDD2*, 8 had the highest expression levels in bud, while *PbIDD6*, 15 had the highest expression levels in fruit. *PbIDD5*, 12, 16 were detected in multiple tissues. The main difference was the expression levels in the leaves: *PbIDD12* was highly expressed in leaves, while the expression level of *PbIDD5* in leaves was the lowest among its 5 expressed tissues. *PbIDD16* had almost no expression in the leaves. *PbIDD3*, 7, 13 were mainly expressed in the fruit of Chinese white pear, and *PbIDD3*, 7 had almost no expression detected in other tissues. We speculate that *PbIDD3*, 7 are specifically expressed in fruits. However, *PbIDD13* was detected to a certain degree in bud.

To clarify the expression patterns of IDD gene family members in 'Dangshan Su' pear fruits, we studied the expression patterns of all 16 IDD genes in pear at 7 different stages of fruit development, including 15 DAF, 39 DAF, 47 DAF, 55DAF, 63 DAF, 79 DAF and 145 DAF (Figure 7). According to our results, *PbIDD14* was a special member of the pear IDD gene family because no expression of *PbIDD14* was detected in the developmental stages of 'Dangshan Su' pear fruit in various tissues. qRT-PCR showed that *PbIDDs* gene family members were expressed during the development of pear fruit and reached a peak of expression at a certain developmental stage. The expression level of *PbIDD1* peaked at 39 DAF and was low during the other six fruit developmental stages. The peak of *PbIDD2*, 7, 8, 11, 13, 16 expression also appeared at 39 DAF, but high expression was also detected in the early developmental stages (15, 47 DAF). In the late developmental stages of fruit (79, 145 DAF), these genes were expressed at a medium level. *PbIDD6*, 15 had a similar expression pattern to the above genes, except that *PbIDD6*, 15 were detected in the early developmental stages of fruit but were expressed at extremely low levels in the late developmental stages of fruit. The expression pattern of *PbIDD9* was special: no expression was detected in the early developmental stages of fruit, and higher expression was detected in the late developmental stages. However, *PbIDD9* had a lower expression level at 79 DAF. In contrast, *PbIDD12* showed the highest amounts of transcriptional accumulation only at 79 DAF, while in the other 6 developmental stages *PbIDD12* maintained at a low expression level. *PbIDD10* had two expression peaks, at 39 and 55 DAF. *PbIDD3*, 5 had the same expression pattern: high expression early with a peak at 15 DAF, a gradual decline, and a trough of expression at 47, 55, 63 DAF. Finally, the expression of *PbIDD3*, 5 increased at maturity.

DISCUSSION

The INDETERMINATE DOMAIN (IDD) gene family encode hybrid transcription factors that have four zinc finger structures and one nuclear positioning signal. IDD proteins play important roles in regulating plant flowering, development, stress resistance, secondary metabolism, and other processes (Wong & Colasanti, 2007; Matsubara et al., 2008). The plant IDD gene can also regulate the expression of the key enzyme genes of lignin synthesis, which affects the biosynthesis of lignin and regulates the biosynthesis of SCW formation (Huang et al., 2018).

In this study, we identified 16 IDD genes in Chinese white pears, which is more than the number of IDD genes we identified

from strawberry, plum, raspberry and cherry. Among the Rosaceae species, pear had the second-most IDD gene family members to apple. This may be because in early evolution, a whole-genome duplication occurred in pear that amplified the chromosome number from 9 to 17 (Velasco et al., 2010; Cao et al., 2017). A total of 68 IDD proteins were identified in five Rosaceae species here, all with four distinct zinc finger motifs, consistent with Arabidopsis, rice and other species (Colasanti et al., 2006). To analyse Rosaceae species' IDD gene family evolutionary relationships, we constructed a phylogenetic tree. According to the phylogenetic tree, all IDD genes could be divided into four groups (1-4). Group4 included members of the pear, strawberry, plum, raspberry, cherry, Arabidopsis, and rice IDD families, but not of apple, even though the apple IDD gene family is the largest of the Rosaceae species (Fan et al., 2017). Its absence from group4 may have been due to the different differentiation patterns of some genes in the IDD gene family after the initial ancestors appeared, resulting in these genes having different evolutionary trajectories. A similar phenomenon has been found in another gene family in pear. In the pear PRX gene family, the 15th subfamily includes members of strawberry, peach, and plum, but no PRX gene of Chinese white pear appears (Cao et al., 2016b). The *ZmIDD1* gene was clustered into its own class in the phylogenetic tree. No homologous genes were found in the Arabidopsis and 6 Rosaceae species, but the *ZmIDD1* homologue *OsIDD1* was present (Colasanti et al., 2006), suggesting the *IDD1* gene may be unique to gramineous plants. Interestingly, five Rosaceae species (pear, strawberry, plum, raspberry, cherry) had at least one IDD gene present in each clade and had a strong bootstrap. These results show that the rapid duplication of IDD genes occurred before these Rosaceae species diverged.

The gene structure and conserved motif composition may be closely related to the diversity of gene functions (Swarbreck et al., 2008), so we analysed the IDD genes' exon-intron structures and conserved motif components in the five Rosaceae species of interest. Genes in the same subfamily tended to have very similar gene structures, reflecting that these genes may have similar functions. For example, *PbIDD12* and *PmIDD10* in group2 had the same gene structure (3 exons, 2 introns), and their exon lengths are also basically the same. Furthermore, based on MEME analysis, we found that members of the same subfamily had roughly the same conserved protein motifs, but there were differences in motif composition between different subfamilies. The 68 IDD genes we identified all had motifs 1-4, which encoded the conserved ID domain, consisting of four zinc fingers (Z1, Z2, Z3 and Z4). Through sequence alignment, we found that the five Rosaceae species' ID domains were highly conserved, particularly the cysteine (C) and histidine (H) in each zinc finger domain. However, we found that multiple of the 68 IDD genes lacked the nuclear localization sequence at the N-terminal border (composed of lysine (K) and arginine (R) residues), such as *PbIDD13*, *14*, *FvIDD10*, *11*, *PmIDD11*, *RoIDD8*, *9*, *PaIDD6*, *10*. We also found many IDD gene family members in apple lacked the nuclear localization sequence, but IDD genes in Arabidopsis, maize, and rice all had nuclear localization sequences. Perhaps this lack of a nuclear localization sequence is unique to the Rosaceae IDD genes. In addition to the highly conserved ID domain, we found two small domains in the C-terminal region based on previous studies. The conservation of these two small domains was much lower than that of the ID domain (Colasanti et al., 2006). The first was a TRDFLG domain, which existed only in the IDD genes of higher plants, but this motif did not exist in all IDD genes; for instance, group4 did not contain this motif. This condition existed in rice (*OsIDD9*), maize (*ZmIDD6*, *ZmIDD8*), and Arabidopsis (*AtIDD15*). The MSATALLQKAA motif was present in most IDD genes, but similar to the TRDFLG motif, we did not identify this motif in the members of group4. The lack of these two conserved sequences at the C-terminus may be the reason why group4 members clustered into a single subfamily, and it may also result in differences in the function of these proteins. These sequences show similarities and major differences in IDD genes beyond the ID domain, suggesting that these genes differentiated early in the evolution of Rosaceae plants and produced some independent and different functions.

Chromosome localization results showed that the chromosomal distribution of IDD genes of the five Rosaceae species was irregular. In Chinese white pear, IDD genes were distributed on 10 chromosomes and Chr16 had the most members (3), while

Chr8, 12, 14, 17 had only one member. A similar distribution has been found in apple. Strawberry, plum, raspberry and cherry IDD genes were also spread across the chromosomes. Functional divergence of genes often occurs after gene duplication, accompanied by the introduction of new functions or losses of original functions (Chao et al., 2017). At the same time, gene duplication events are the main driving force for gene family expansion, which is an important way plants adapt to changing climates and environments (Tang et al., 2016). Gene duplication events have been identified in multiple gene families, such as the PKS gene family in *Gossypium hirsutum* (Su et al., 2017), the HSF gene family in *Sesamum indicum* (Dossa et al., 2016), the WRKY gene family in *Musa acuminata* (Goel et al., 2016). A total of 4 PbIDD gene pairs were identified as duplications in this study and these duplicated gene pairs all were created by segmental duplication. Four pairs of segmentally duplicated genes were also identified in apple, verifying the reliability of our results. However, we did not identify gene duplication events in strawberry, plum, raspberry or cherry. This may be because strawberry, plum, raspberry and cherry have only undergone one (old) whole-genome duplication (WGD), but the pear has had a recent WGD in addition to an old WGD (Zhang et al., 2012; Wu et al., 2013). The four pairs of duplicated genes that appeared in pears had a very positive effect on the expansion of the IDD gene family and contributed to the diversification of the function of the IDD gene. We then performed a sliding window analysis of the four pairs of segmentally duplicated genes in Chinese white pear. The results showed that these genes have experienced intense purifying selection. IDD gene involved in hormone pathways simultaneously regulate multiple processes in plant growth and development. For example, *AtIDD3* can be combined with SRDX (a plant-specific repression domain) and interact with DELLA proteins to regulate GA signalling (Yoshida et al., 2014). *AtIDD14, 15, 16* regulate lateral organ morphogenesis and gravitropism by promoting the biosynthesis of auxin (Cui et al., 2013). The IDD family genes have a large number of responsive hormone-related cis-acting elements (Fan et al., 2017), so we analysed the cis-acting elements of all 16 IDD gene promoters in pear. Salicylic acid is an endogenous signalling molecule necessary for plant resistance to systemic stresses (Klessig & Malamy, 1994) and MeJA can be used as a signalling molecule to induce plant defence gene expression (Titarenko et al., 1997). Many MeJA response elements and SA response elements were found in pear, such as *PbIDD1, 5, 6* which all contained both kinds of cis-acting elements. From this we infer that members of the pear IDD gene have resistance mechanisms that rely on MeJA and SA. GA plays an important role in the process of plant flower induction (Porri et al., 2012; Zhang et al., 2016), as GA response elements were found in 12 members. This suggests that pear IDD genes may be closely involved in flower induction.

~~The expression pattern of genes are closely related to their functions. We analysed the expression patterns of the 16 IDD genes in Chinese white pears to analyse the potential functions of pear IDD genes. We studied their expression patterns in different plant parts, including fruit, flower, bud, stem and leaf. Previous research has shown that IDD genes are most highly expressed in the leaf and bud (Fan et al., 2017), but we found the highest pear IDD gene expression in fruit, flower and bud. Only *PbIDD12* had high expression in leaf. Interestingly, we found that *PbIDD2, 5, 6, 8, 9, 12, 16* showed higher transcription levels in both flower and bud, the expression in bud was higher than flowers. In *Arabidopsis, rice and maize*, IDD genes are involved in plant flower induction and are regulated by hormones and sugars (Matsubara et al., 2008; Wu et al., 2008; Jeong et al., 2015). GA can regulate plant flowering by inhibiting flowering genes. The promoter regions of *PbIDD2, 6, 8, 12, 16* have the GA-response cis-acting element, but *PbIDD6, 16* were mainly expressed in fruit and *PbIDD12* mainly in leaf. Based on our above findings, we speculate that *PbIDD2, 8* may be involved in the regulation of plant flowering induction in Chinese white pears. *PbIDD1, 10* were mainly expressed in the flowers of Chinese white pears, probably because these two genes are involved in the synthesis of bioactive substances in flowers.~~

The stone cells and lignin in 'Dangshan Su' pear fruit are mainly formed during the period of 15 to 63 DAF and reach their peak levels at 55 DAF (Zhang et al., 2017). The expression analysis of these 16 *PbIDDs* at seven developmental stages of fruit showed that the expression pattern of any of gene was consistent with the developmental law of 'Dangshan Su' pear fruit stone

cells and lignin. However, there were two more specific genes, *PbIDD3* and *PbIDD5*. The expression patterns of these two genes showed a completely opposite rule to the accumulation of 'Dangshan Su' pear stone cells and lignin. *PbIDD3* and *PbIDD5* showed extremely high expression levels at 15 DAF. At 15 to 47 DAF, a stage of massive accumulation of lignin and stone cells, the expression of these two genes showed a gradually decreasing trend. At the peak of lignin accumulation (47, 55, 63 DAF), *PbIDD3* and *PbIDD5* showed very low levels of transcription. In the late stage of fruit development (79, 145 DAF), the expression level increased slightly. According to the phylogenetic tree, *PbIDD3*, *PbIDD5* and *OsIDD2* were most closely related. In rice, *OsIDD2* regulates SCW formation while negatively regulating the expression of key enzyme genes in lignin synthesis to inhibit lignin biosynthesis, such as cinnamyl alcohol dehydrogenase (CAD) (Huang et al., 2018). In Chinese white pear, *CAD* is predicted to regulate the biosynthesis of fruit lignin. The expression pattern of *PbCAD* is exactly the same as the accumulation pattern of stone cells and lignin (Cheng et al., 2017). The expression patterns of *PbIDD3* and *PbIDD5* were opposite to those of *PbCAD*. *PbIDD3* and *PbIDD5* expression was low during the high expression period of *PbCAD* and higher during the low expression period of *PbCAD*. *PbIDD3* and *PbIDD5* exhibited extremely high levels of transcription in fruits. We speculate that *PbIDD3* and *PbIDD5* are mainly expressed in fruits and have similar functions to *OsIDD2*, regulating SCW formation in pear fruit cells and inhibiting lignin biosynthesis by inhibiting the expression of lignin synthesis key enzyme genes. *PbIDD6* and *OsIDD2* also had a close relationship, but their expression patterns in fruits was not related to the accumulation of lignin. *PbIDD6* may not be involved in the biosynthesis of lignin, but it might participate in the synthesis of SCW formation or regulate the biosynthesis of lignin in other tissues of pear.

CONCLUSION

We identified 68 IDD genes in five Rosaceae species (pear, strawberry, plum, raspberry, and cherry), which we systematically assessed by bioinformatic analysis. According to the phylogenetic tree, the 68 IDD genes were divided into 4 groups. In each class, we found that the structures of the genes and the compositions of the conserved motifs were very similar. Through chromosomal localization, conserved microsynteny, and gene duplication analysis, we revealed the evolution and family expansion of the IDD genes of these five Rosaceae species. According to qRT-PCR, pear IDD genes have high expression levels in fruit, flower and bud. Further, we believe that *PbIDD2*, 8 are involved in the regulation of plant flowering induction in pear and that *PbIDD3*, 5 are involved in the regulation of SCW formation and lignin biosynthesis in pear fruit. All in all, this study reveal the basic information of the five Rosaceae species' IDD genes and predicts the potential functions of some pear IDD proteins. These results will provide an important theoretical basis for improving the quality of 'Dangshan Su' pear.

References

- Abdullah, M., Cao, Y. P., Cheng, X., Shakoor, A., Su, X. Q., Gao, J. S., Cai, Y. P. (2018). Genome-Wide Analysis Characterization and Evolution of SBP Genes in *Fragaria vesca*, *Pyrus bretschneideri*, *Prunus persica* and *Prunus mume*. *FRONT. GENET.* 9, 64.
- Agarwal, P., Arora, R., Ray, S., Singh, A. K., Singh, V. P., Takatsuji, H., Kapoor, S., Tyagi, A. K. (2007). Genome-wide identification of C2H2 zinc-finger gene family in rice and their phylogeny and expression analysis. *PLANT. MOL. BIOL.* 65, 467-485.
- Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., Castro, E., Duvaud, E., Flegel, V., Fortier, A.,

411 Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K.,
 412 Redaschi, N., Rossier, G., Xenarios, I., Stockinger, H. (2012). ExPASy: SIB bioinformatics resource portal. *NUCLEIC.*
 413 *ACIDS. RES.* **40**, 597-603.

414 Bailey, T. L., Johnson, J., Grant, C. E., Noble, W. S. (2015). The MEME Suite. *NUCLEIC. ACIDS. RES.* **43**, 39-49.

415 Bitocchi, E., Rau, D., Benazzo, A., Bellucci, E., Goretti, D., Biagetti, E., Panziera, A., Laidò, G., Rodriguez, M., Gioia, T.,
 416 Attene, G., McClean, P., Lee, R. K., Jackson, S. A., Bertorelle, G., Papa, R. (2017). High Level of Nonsynonymous
 417 Changes in Common Bean Suggests That Selection under Domestication Increased Functional Diversity at Target Traits.
 418 *FRONT PLANT SCI.* **7**, 2005.

419 Cai, Y. P., Li, G. Q., Nie, J. Q., Lin, Y., Nie, F., Zhang, J. Y. and Xu, Y. L. (2010). Study of the structure and biosynthetic
 420 pathway of lignin in stone cells of pear. *Sci Hortic-Amsterdam.* **125**, 374-379.

421 Cao, Y. P., Han, Y. H., Li, D. H., Lin, Y., Cai, Y. P. (2016a). Systematic Analysis of the 4-Coumarate: Coenzyme A Ligase
 422 (4CL) Related Genes and Expression Profiling during Fruit Development in the Chinese Pear. *GENES.* **7**, 89-106.

423 Cao, Y. P., Han, Y. H., Meng, D. D., Li, D. H., Jin, Q., Lin, Y., Cai, Y. P. (2016b). Structural, Evolutionary, and Functional
 424 Analysis of the Class III Peroxidase Gene Family in Chinese Pear (*Pyrus bretschneideri*). *FRONT. PLANT. SCI.* **7**, 1874.

425 Cao, Y. P., Han, Y. H., Meng, D. D., Li, D. H., Jin, Q., Lin, Y., Cai, Y. P. (2017). Genome-wide analysis suggests high level
 426 of microsynteny and purifying selection affect the evolution of EIN3/EIL family in Rosaceae. *PEERJ.* **5**, e3400.

427 Chai, G. H., Kong, Y. Z., Zhu, M., Yu, L., Qi, G., Tang, X. F., Wang, Z. G., Cao, Y. P., Yu, C. J., Zhou, G. K. (2015).
 428 Arabidopsis C3H14 and C3H15 have overlapping roles in the regulation of secondary wall thickening and anther
 429 development. *J. EXP. BOT.* **66**, 2595-2609.

430 Cheng, X., Li, M. L., Li, D. H., Zhang, J. Y., Jin, Q., Sheng, L. L., Cai, Y. P., Lin, Y. (2017). Characterization and analysis of
 431 CCR and CAD gene families at the whole-genome level for lignin synthesis of stone cells in pear (*Pyrus bretschneideri*)
 432 fruit. *BIOL. OPEN.* **6**, 1602-1613.

433 Chao, N., Li, N., Qi, Q., Li, S., Lv, T., Jiang, X. N., Gai, Y. (2017). Characterization of the cinnamoyl-CoA reductase (CCR)
 434 gene family in *Populus tomentosa* reveals the enzymatic active sites and evolution of CCR. *PLANTA.* **245**, 61-75.

435 Chen, J., Wang, J. N., Kuang, J. F., Shan, W., Chen, J. Y., Lu, W. J. (2014). Molecular characterisation of a fruit ripening-
 436 related INDETERMINATE DOMAIN transcription factor in banana (*Musa acuminata*). *J. HORTIC. SCI. BIOTECH.* **89**,
 437 373-380.

438 Colasanti, J., Tremblay, R., Wong, A. Y. M., Coneva, V., Kozaki, A., Mable, B. K. (2006). The maize INDETERMINATEI
 439 flowering time regulator defines a highly conserved zinc finger protein family in higher plants. *BMC. GENOMICS.* **7**, 158-
 440 175.

441 Colasanti, J., Yuan, Z. and Sundaresan, V. (1998). The indeterminate Gene Encodes a Zinc Finger Protein and Regulates a
 442 Leaf-Generated Signal Required for the Transition to Flowering in Maize. *Cell.* **93**, 593-603.

443 Cui, D. Y., Zhao, J. B., Jing, Y. J., Fan, M. Z., Liu, J., Wang, Z. C., Xin, W., Hu, Y. X. (2013). The Arabidopsis IDD14,

444 IDD15, and IDD16 Cooperatively Regulate Lateral Organ Morphogenesis and Gravitropism by Promoting Auxin
445 Biosynthesis and Transport. *PLOS. GENET.* **9**, e1003759.

446 **Dossa, K., Diouf, D., Cisse, N.** (2016). Genome-Wide Investigation of Hsf Genes in Sesame Reveals Their Segmental
447 Duplication Expansion and Their Active Role in Drought Stress Response. *FRONT. PLANT. SCI.* **7**, 1522.

448 **Fan, S., Zhang, D., Xing, L. B., Qi, S. Y., Du, L. S., Wu, H. Q., Shao, H. X., Li, Y. M., Ma, J. J., Han, M. Y.** (2017).
449 Phylogenetic analysis of IDD gene family and characterization of its expression in response to flower induction in Malus.
450 *MOL. GENET. GENOMICS.* **292**, 755-771.

451 **Feurtado, J. A., Huang, D., Wicki-Stordeur, L., Hemstock, L. E., Potentier, M. S., Tsang, E. W., Cutler, A. J.** (2011). The
452 Arabidopsis C2H2 zinc finger INDETERMINATE DOMAIN1/ENHYDROUS promotes the transition to germination by
453 regulating light and hormonal signaling during seed maturation. *Plant. Cell.* **23**, 1772-1794.

454 **Goel, R., Pandey, A., Trivedi, P. K., Asif, M. H.** (2016). Genome-Wide Analysis of the Musa WRKY Gene Family: Evolution
455 and Differential Expression during Development and Stress. *FRONT. PLANT. SCI.* **7**, 299.

456 **Guo, A. Y., Zhu, Q. H., Chen, X., Luo, J. C.** (2007). [GSDS: a gene structure display server]. *Yi. Chuan.* **29**, 1023-1026.

457 **Han, Y. H., Ding, T., Su, B., Jiang, H. Y.** (2016). Genome-Wide Identification, Characterization and Expression Analysis of the
458 Chalcone Synthase Family in Maize. *INT. J. MOL. SCI.* **17**, 161.

459 **Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., Nakai, K.** (2007). WoLF PSORT:
460 protein localization predictor. *NUCLEIC. ACIDS. RES.* **35**, 585-587.

461 **Huang, P., Yoshida, H., Yano, K., Kinoshita, S., Kawai, K., Koketsu, E., Hattori, M., Takehara, S., Huang, J., Hirano, K.,
462 Ordonio, R. L., Matsuoka, M., Ueguchi-Tanaka, M.** (2018). OsIDD2, a zinc finger and INDETERMINATE DOMAIN
463 protein, regulates secondary cell wall formation. *J. INTEGR. PLANT. BIOL.* **60**, 130-143

464 **Ingkasuwan, P., Netrphan, S., Prasitwattanaseree, S., Tanticharoen, M., Bhumiratana, S., Meechai, A., Chaijaruwanich,
465 J., Takahashi, H., Cheevadhanarak, S.** (2012). Inferring transcriptional gene regulation network of starch metabolism in
466 Arabidopsis thaliana leaves using graphical Gaussian model. *BMC. SYST. BIOL.* **6**, 100-121.

467 **Islam, M. S., Hur, J. H., Wang, M. H.** (2009). The Influence of Abiotic Stresses on Expression of Zinc Finger Protein Gene in
468 Rice. *RUSS. J. PLANT. PHYSIOL.* **56**, 695-701.

469 **Jeong, E. Y., Seo, P. J., Woo, J. C., Park, C. M.** (2015). AKIN10 delays flowering by inactivating IDD8 transcription factor
470 through protein phosphorylation in Arabidopsis. *BMC. PLANT. BIOL.* **15**, 110.

471 **Klessig, D. F., Malamy, J.** (1994). The salicylic acid signal in plants. *PLANT. MOL. BIOL.* **26**, 439-458.

472 **Kozaki, A., Hake, S., Colasanti, J.** (2004). The maize ID1 flowering time regulator is a zinc finger protein with novel DNA
473 binding properties. *Nucleic. Acids. Res.* **32**, 1710-1720.

474 **Kumar S, Stecher G, Tamura K.** (2016). MEGA 7.0: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger
475 Datasets. *MOL. BIOL. EVOL.* **33**, 1870-1874.

476 Lee, S. H., Choi, J. H., Kim, W. S., Park, Y. S., Gemma, H. (2007). Effects of calcium chloride spray on peroxidase activity
477 and stone cell development in pear fruit (*Pyrus pyrifolia* 'Niitaka'). *J. JPN. SOC. HORTIC. SCI.* **76**, 191-196.

478 Letunic, I., Doerks, T., Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. *NUCLEIC.*
479 *ACIDS. RES.* **40**, 302-305.

480 Librado, P., Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.
481 *BIOINFORMATICS.* **25**, 1451-1452.

482 Liu, C. J. (2012). Deciphering the Enigma of Lignification: Precursor Transport, Oxidation, and the Topochemistry of Lignin
483 Assembly. *MOL. PLANT.* **5**, 304-317.

484 Livak, K. J., Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-
485 Delta Delta C(T)) Method. *Methods.* **25**, 402-408.

486 Long M, Thornton K. (2001). Gene duplication and evolution. *Science.* **293**, 1551.

487 Matsubara, K., Yamanouchi, U., Wang, Z. X., Minobe, Y., Izawa, T., Yano, M. (2008). Ehd2, a Rice Ortholog of the Maize
488 INDETERMINATE1 Gene, Promotes Flowering by Up-Regulating Ehd1. *PLANT. PHYSIOL.* **148**, 1425-1435.

489 Miller, J., McLachlan, A. D., Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from
490 *Xenopus* oocytes. *EMBO. J.* **4**, 1609-1614.

491 Niu, E., Cai, C., Zheng, Y., Shang, X., Fang, L., Guo, W. (2016). Genome-wide analysis of *CrRLKIL*, gene family in
492 *Gossypium*, and identification of candidate *CrRLKIL*, genes related to fiber development. *MOL. GENET. GENOMICS.* **291**,
493 1137-1154.

494 Park, S. J., Kim, S. L., Lee, S., Je, B. I., Piao, H. L., Park, S. H., Kim, C. M., Ryu, C. H., Park, S. H., Xuan, Y. H.,
495 Colasanti, J., An, G., Han, C. D. (2008). Rice Indeterminate 1 (*OsId1*) is necessary for the expression of Ehd1 (Early
496 heading date 1) regardless of photoperiod. *PLANT. J.* **56**, 1018-1029.

497 Parraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T., Klevit, R. E. (1988). Zinc-dependent structure
498 of a single-finger domain of yeast ADR1. *Science.* **241**, 1489-1492.

499 Porri, A., Torti, S., Romera-Branchat, M., Coupland, G. (2012). Spatially distinct regulatory roles for gibberellins in the
500 promotion of flowering of Arabidopsis under long photoperiods. *DEVELOPMENT.* **139**, 2198-2209.

501 Rombauts, S., Déhais P, Van Montagu, M, Rouzé, P. (1999). PlantCARE, a plant cis-acting regulatory element database.
502 *NUCLEIC. ACIDS. RES.* **27**, 295-296.

503 Seo, P. J., Ryu, J., Kang, S. K., Park, C. M. (2011). Modulation of sugar metabolism by an INDETERMINATE DOMAIN
504 transcription factor contributes to photoperiodic flowering in Arabidopsis. *PLANT. J.* **65**, 418-429.

505 Shi, X. D., Gu, Y. X., Dai, T. W., Wu, Y., Wu, P., Xu, Y., Chen, F. (2018). Regulation of trichome development in tobacco by
506 *JcZFP8*, a C2H2 zinc finger protein gene from *Jatropha curcas* L. *GENE.* **658**, 47-53.

507 Su, X. Q., Sun, X., Cheng, X., Wang, Y. N., Abdullah, M., Li, M. L., Li, D. H., Gao, J. S., Cai, Y. P., Lin, Y. (2017).

- Comparative genomic analysis of the PKS genes in five species and expression analysis in upland cotton. *PEERJ*. **5**, e3974.
- Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R., Ploetz, L., Radenbaugh, A., Singh, S., Swing, V., Tissier, C., Zhang, P., Huala, E.** (2008). The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *NUCLEIC. ACIDS. RES.* **36**, 1009-1014.
- Ren, G. P., Li, L. F., Huang, Y. H., Wang, Y. Q., Zhang, W. B., Zheng, R. Y., Zhong, C. M., Wang, X. J.** (2018). *GhWIP2*, a WIP zinc finger protein, suppresses cell expansion in *Gerbera hybrida* by mediating crosstalk between gibberellin, abscisic acid, and auxin. *NEW. PHYTOL.* **219**, 728-742.
- Tang, R. M., Zhu, W. J., Song, X. Y., Lin, X. Z., Cai, J. H., Wang, M., Yang, Q.** (2016). Genome-Wide Identification and Function Analyses of Heat Shock Transcription Factors in Potato. *FRONT. PLANT. SCI.* **7**, 490.
- Tian, Z. D., Zhang, Y., Liu, J., Xie, C. H.,** (2010). Novel potato C2H2-type zinc finger protein gene, *StZFP1*, which responds to biotic and abiotic stress, plays a role in salt tolerance. *PLANT. BIOLOGY.* **12**, 689-697.
- Titarenko, E., Rojo, E., Leon, J., Sanchez-Serrano, J. J.** (1997). Jasmonic acid-dependent and -independent signaling pathways control wound-induced gene activation in *Arabidopsis thaliana*. *PLANT. PHYSIOL.* **115**, 817-826.
- Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., et al.** (2010). The genome of the domesticated apple (*Malus x domestica* Borkh.). *NAT. GENET.* **42**, 833-839.
- Wei, K., Pan, S., Li, Y.** (2016). Functional Characterization of Maize C2H2, Zinc-Finger Gene Family. *PLANT. MOL. BIOL. REP.* **34**, 761-776.
- Welch, D., Hassan, H., Blilou, I., Immink, R., Heidstra, R., Scheres, B.** (2007). Arabidopsis JACKDAW and MAGPIE zinc finger proteins delimit asymmetric cell division and stabilize tissue boundaries by restricting SHORT-ROOT action. *GENE. DEV.* **21**, 2196-2204.
- Wong, A. Y., Colasanti, J.** (2007). Maize floral regulator protein INDETERMINATE1 is localized to developing leaves and is not altered by light or the sink/source transition. *J. EXP. BOT.* **58**, 403-414.
- Wu, J., Wang, Z., Shi, Z., Zhang, S., Ming, R., Zhu, S., et al.** (2013). The genome of the pear (*Pyrus bretschneideri* Rehd). *Genome. Res.* **23**, 396-408.
- Wu, C. Y., You, C. J., Li, C. S., Long, T., Chen, G. X., Byrne, M. E., Zhang, Q. F.** (2008). *RID1*, encoding a Cys2/His2-type zinc finger transcription factor, acts as a master switch from vegetative to floral development in rice. *P. NATL. ACAD. SCI. USA.* **105**, 12915-12920.
- Xuan, Y. H., Priatama, R. A., Huang, J., Je, B. I., Liu, J. M., Park, S. J., Piao, H. L., Son, D.Y., Lee, J. J., Park, S. H., Jung, K. H., Kim, T. H., Han, C. D.** (2013). Indeterminate domain 10 regulates ammonium-mediated gene expression in rice roots. *NEW. PHYTOL.* **197**, 791-804.
- Yan, C. C., Yin, M., Zhang, N., Jin, Q., Fang, Z., Lin, Y., Cai, Y.P.** (2014). Stone cell distribution and lignin structure in various pear varieties. *SCI. HORTIC-AMSTERDAM.* **174**, 142-150.

- Yi, G., Neelakandan, A. K., Gontarek, B. C., Vollbrecht, E., Becraft, P. W. (2015). The naked endosperm Genes Encode Duplicate INDETERMINATE Domain Transcription Factors Required for Maize Endosperm Cell Patterning and Differentiation. *PLANT. PHYSIOL.* **167**, 443-456.
 - Yoshida, H., Hirano, K., Sato, T., Mitsuda, N., Nomoto, M., Maeo, K., Koketsu, E., Mitani, R., Kawamura, M., Ishiguro, S., Tada, Y., Ohme-Takagi, M., Matsuoka, M., Ueguchi-Tanaka, M. (2014). DELLA protein functions as a transcriptional activator through the DNA binding of the INDETERMINATE DOMAIN family proteins. *P. NATL. ACAD. SCI. USA.* **111**, 7861-7866.
 - Zhang, D. M., Xu, Z. P., Cao, S. X., Chen, K. L., Li, S., Liu, X. L., Gao, C. X., Zhang, B. C., Zhou, Y. H. (2018). An Uncanonical CCCH-Tandem Zinc-Finger Protein Represses Secondary Wall Synthesis and Controls Mechanical Strength in Rice. *MOL. PLANT.* **11**, 163-174.
 - Zhang, J. Y., Cheng, X., Jin, Q., Su, X. Q., Li, M. L., Yan, C. C., Jiao, X. Y., Li, D. H., Lin, Y., Cai, Y. P. (2017). Comparison of the transcriptomic analysis between two Chinese white pear (*Pyrus bretschneideri* Rehd.) genotypes of different stone cells contents. *PLOS. ONE.* **12**, e0187114.
 - Zhang, Q., Chen, W., Sun, L., Zhao, F., Huang, B., Yang, W., et al. (2012). The genome of *Prunus mume*. *Nat. Commun.* **3**, 1318.
 - Zhang, S. W., Zhang, D., Fan, S., Du, L. S., Shen, Y. W., Xing, L. B., Li, Y. M., Ma, J. J., Han, M. Y. (2016). Effect of exogenous GA(3) and its inhibitor paclobutrazol on floral formation, endogenous hormones, and flowering-associated genes in 'Fuji' apple (*Malus domestica* Borkh.). *PLANT. PHYSIOL. BIOCH.* **107**, 178-186.
 - Zhao, P., Wang, D. D., Wang, R. Q., Kong, N. N., Zhang, C., Yang, C. H., Wu, W. T., Ma, H. L., Chen, Q. (2018). Genome-wide analysis of the potato Hsp20 gene family: identification, genomic organization and expression profiles in response to heat stress. *BMC. GENOMICS.* **19**, 61.
 - Zhao, S. G., Zhang, J.G., Zhao, Y. P., Zhang, Y. X. (2013). New discoveries of stone cell differentiation in fruitlets of 'Yali' pears (*Pyrus bretschneideri* Rehd.). *J. FOOD. AGRIC. ENVIRON.* **11**, 937-942.
- Table 1.** The IDD genes of pear identified in this study are listed.
- Figure 1.** Phylogenetic relationships and subfamily designations in IDD proteins from pear, strawberry, plum, raspberry, cherry, apple, Arabidopsis, rice and maize. Groups1-4 are represented by shades of red, green, blue and cyan, respectively.
- Figure 2.** Predicted pear, strawberry, plum, raspberry and cherry IDD protein conserved motifs and exon-intron structures. (A) Gene structures of the IDD genes. Black wedge indicates exon, black line indicates intron and red wedge indicates UTR. (B) Distribution of MEME motifs in IDD genes. (C) The colour and corresponding number of each motif box.
- Figure 3.** Conserved ID domain composition. All 68 IDD genes had a characteristic ID domain. The ID domain consists of four zinc fingers (Z1, Z2, Z3 and Z4). Alignment analysis of the 68 IDD gene sequences using the ClustalW tool in MEGA 7.0 software. These domain diagrams were plotted using the online WebLogo tool.
- Figure 4.** Chromosomal locations of IDD genes in pear (A), strawberry (B), raspberry (C), plum (D), and cherry (E). Duplicated

gene pairs are connected with coloured lines.

Figure 5. Microsynteny of regions among pear, strawberry, plum, raspberry, cherry and apple. The chromosome numbers are indicated by differently coloured boxes and are labelled by Pb, Fv, Pm, Ro, Pa, and Md. The differently coloured boxes also represent the sequence lengths of chromosomes in megabases. The black line indicates the syntenic relationship among the IDD regions.

Figure 6. Potential cis-elements in the 5' regulatory sequences of the 16 PbIDD genes.

Figure 7. Expression patterns of IDD genes of Chinese white pear in different tissues and in fruit at different developmental stages. (a-o) Expression patterns of IDD genes in Chinese white pear in different tissues. (A-O) Expression patterns of IDD genes in Chinese white pear at different developmental stages.

Supplemental Figure S1. Pear IDD protein sequence alignment. Black underline indicates zinc finger domain (Z1, Z2, Z3 and Z4). Red triangle indicates a conserved C residue, and blue triangle indicates a conserved H residue. The yellow underline indicates the NLS sequence in the N-terminal region of the IDD gene. Green box means the MSATALLQKAA domain, and purple box indicates the TRDFLG domain.

Supplemental Figure S2. Strawberry IDD protein sequence alignment. Black underline indicates zinc finger domain (Z1, Z2, Z3 and Z4). Red triangle indicates a conserved C residue, and blue triangle indicates a conserved H residue. The yellow underline indicates the NLS sequence in the N-terminal region of the IDD gene. Green box means the MSATALLQKAA domain, and purple box indicates the TRDFLG domain.

Supplemental Figure S3. Plum IDD protein sequence alignment. Black underline indicates zinc finger domain (Z1, Z2, Z3 and Z4). Red triangle indicates a conserved C residue, and blue triangle indicates a conserved H residue. The yellow underline indicates the NLS sequence in the N-terminal region of the IDD gene. Green box means the MSATALLQKAA domain, and purple box indicates the TRDFLG domain.

Supplemental Figure S4. Raspberry IDD protein sequence alignment. Black underline indicates zinc finger domain (Z1, Z2, Z3 and Z4). Red triangle indicates a conserved C residue, and blue triangle indicates a conserved H residue. The yellow underline indicates the NLS sequence in the N-terminal region of the IDD gene. Green box means the MSATALLQKAA domain, and purple box indicates the TRDFLG domain.

Supplemental Figure S5. Cherry IDD protein sequence alignment. Black underline indicates zinc finger domain (Z1, Z2, Z3 and Z4). Red triangle indicates a conserved C residue, and blue triangle indicates a conserved H residue. The yellow underline indicates the NLS sequence in the N-terminal region of the IDD gene. Green box means the MSATALLQKAA domain, and purple box indicates the TRDFLG domain.

Supplemental Figure S6. Sliding window plots of duplicated IDD genes in Chinese white pear. The grey shaded portion indicates conserved ID domain. The X-axis indicates the synonymous distance within each gene.

Supplementary Table S1. Gene sequence involved in the article.

605 **Supplemental Table S2.** The IDD genes of strawberry, plum, raspberry and cherry identified in this study are listed.

606 **Supplementary Table S3.** Detailed information of the 20 motifs in the 68 IDD proteins.

607 **Supplementary Table S4.** Ka/Ks analysis of the duplicated IDD paralogues from Chinese white pear.

608 **Supplemental Table S5.** Synteny data in pear, strawberry, plum, raspberry, cherry, apple.

609 **Supplementary Table S6.** Promoter sequence of 16 IDD genes in pear.

610 **Supplementary Table S7.** Complete information on Cis-acting elements of the 16 PbIDD genes.

611 **Supplementary Table S8.** Primers used in qRT-PCR.

Table 1(on next page)

The IDD genes of pear identified in this study are listed.

The IDD genes of pear identified in this study are listed.

1 **Table 1.** The IDD genes of pear identified in this study are listed.

Gene name	Gene ID	AA	KD	pI	GRAVY	Predicted subcellular localization
<i>PbIDD1</i>	Pbr029706.1	489	51.5	8.17	-0.460	nucl
<i>PbIDD2</i>	Pbr006167.1	535	55.7	8.98	-0.427	nucl
<i>PbIDD3</i>	Pbr032492.1	533	58.0	8.98	-0.659	nucl
<i>PbIDD4</i>	Pbr021137.1	553	60.4	9.31	-0.634	nucl
<i>PbIDD5</i>	Pbr028264.1	526	57.4	8.89	-0.693	nucl
<i>PbIDD6</i>	Pbr019853.1	537	59.1	8.67	-0.807	nucl
<i>PbIDD7</i>	Pbr012193.2	567	61.7	8.92	-0.772	nucl
<i>PbIDD8</i>	Pbr020403.1	524	57.3	9.07	-0.740	nucl
<i>PbIDD9</i>	Pbr009954.1	465	50.9	9.19	-0.664	nucl
<i>PbIDD10</i>	Pbr012907.1	465	51.1	8.99	-0.629	nucl
<i>PbIDD11</i>	Pbr008330.1	464	51.0	9.35	-0.727	nucl
<i>PbIDD12</i>	Pbr012192.1	607	64.7	9.13	-0.723	nucl
<i>PbIDD13</i>	Pbr025606.1	426	47.2	9.25	-0.646	nucl
<i>PbIDD14</i>	Pbr029012.1	381	42.7	9.36	-0.655	nucl
<i>PbIDD15</i>	Pbr038802.1	502	56.4	9.02	-0.914	nucl
<i>PbIDD16</i>	Pbr012170.1	492	54.4	8.93	-0.805	nucl

2

Figure 1

Phylogenetic relationships in IDD proteins from pear, strawberry, plum, raspberry, cherry, apple, Arabidopsis, rice and maize.

Phylogenetic relationships and subfamily designations in IDD proteins from pear, strawberry, plum, raspberry, cherry, apple, Arabidopsis, rice and maize. Groups1-4 are represented by shades of red, green, blue and cyan, respectively.

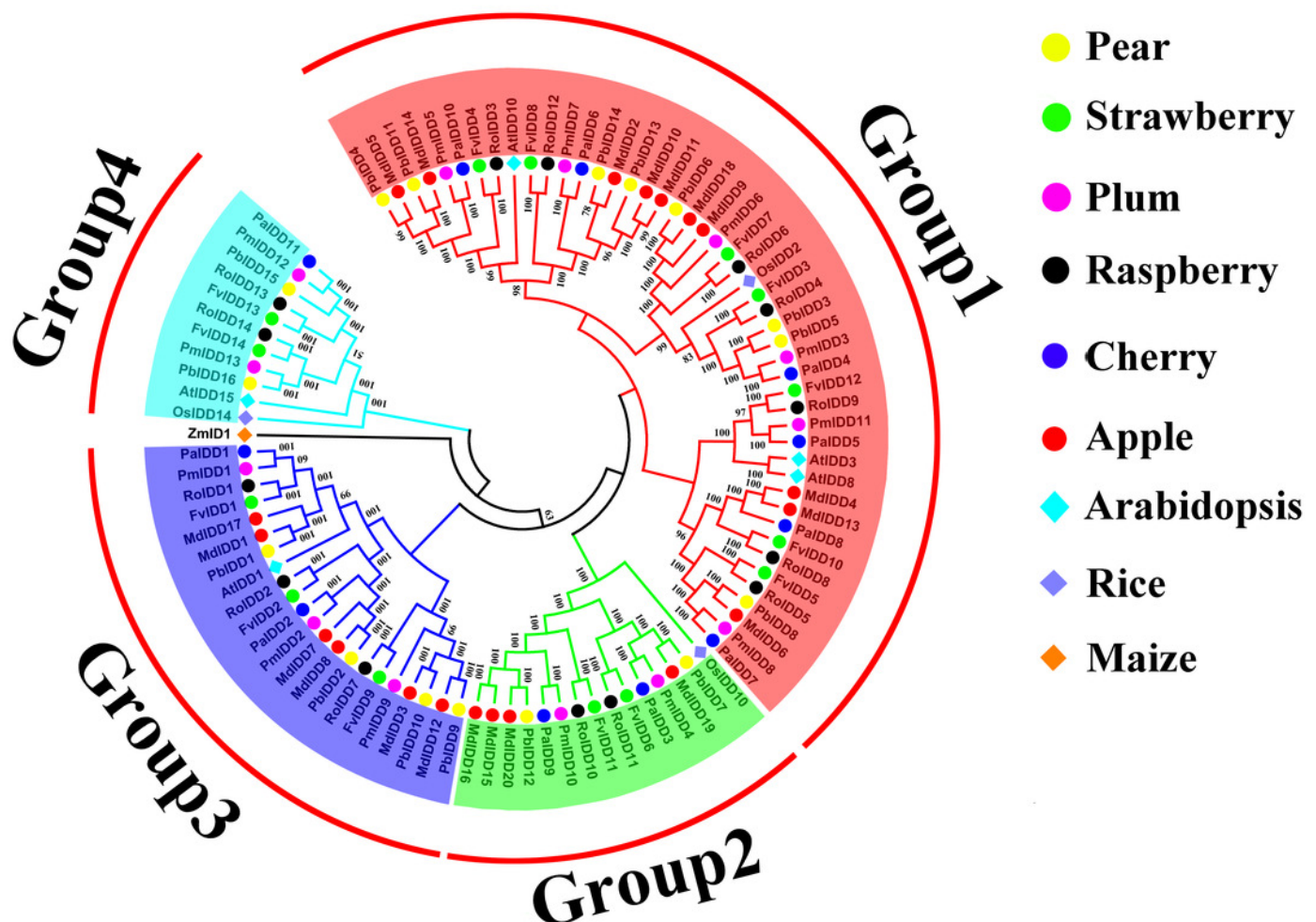


Figure 2

Predicted pear, strawberry, plum, raspberry and cherry IDD protein conserved motifs and exon-intron structures.

Predicted pear, strawberry, plum, raspberry and cherry IDD protein conserved motifs and exon-intron structures. (A) Gene structures of the IDD genes. Black wedge indicates exon, black line indicates intron and red wedge indicates UTR. (B) Distribution of MEME motifs in IDD genes. (C) The colour and corresponding number of each motif box.

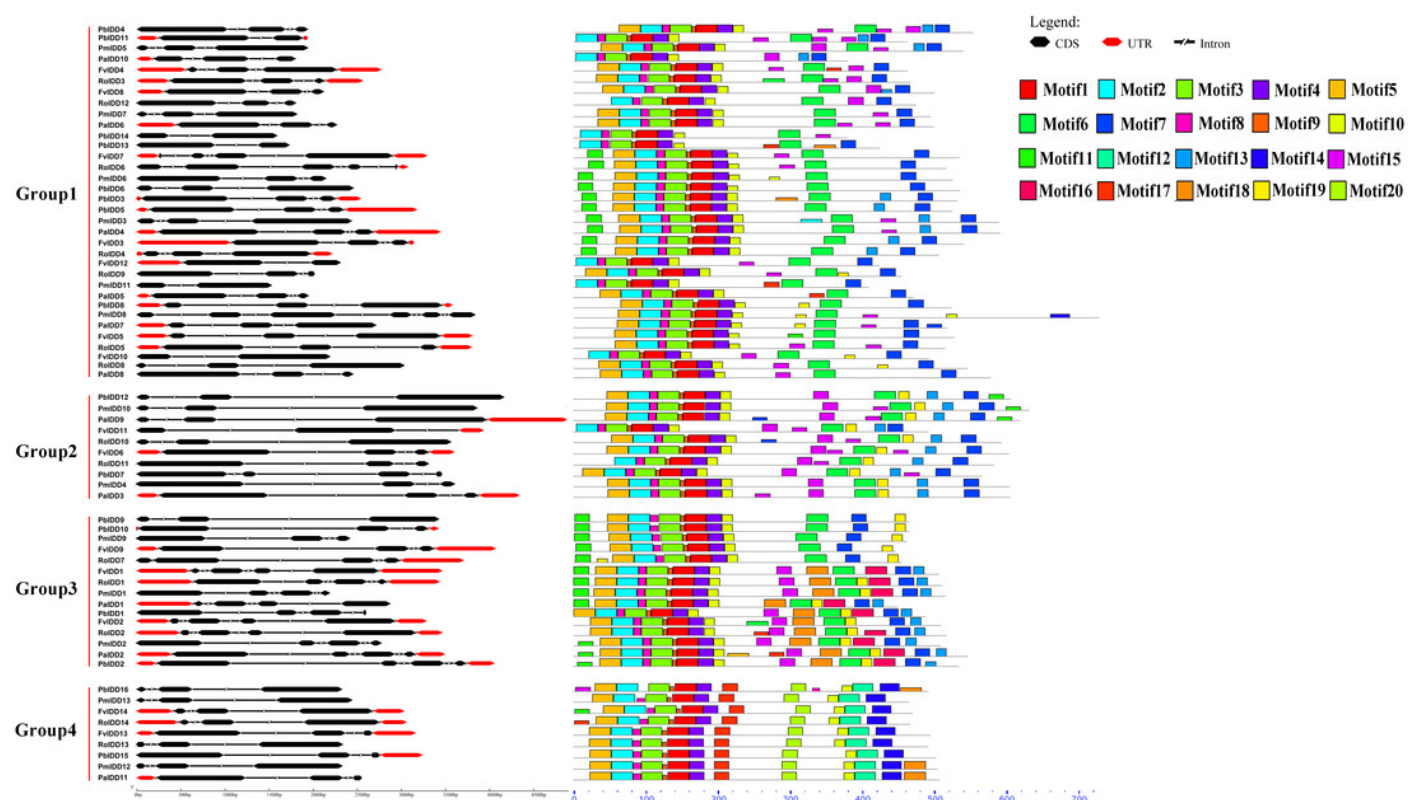


Figure 3

Conserved ID domain composition.

Conserved ID domain composition. All 68 IDD genes had a characteristic ID domain. The ID domain consists of four zinc fingers (Z1, Z2, Z3 and Z4). Alignment analysis of the 68 IDD gene sequences using the ClustalW tool in MEGA 7.0 software. These domain diagrams were plotted using the online WebLogo tool.

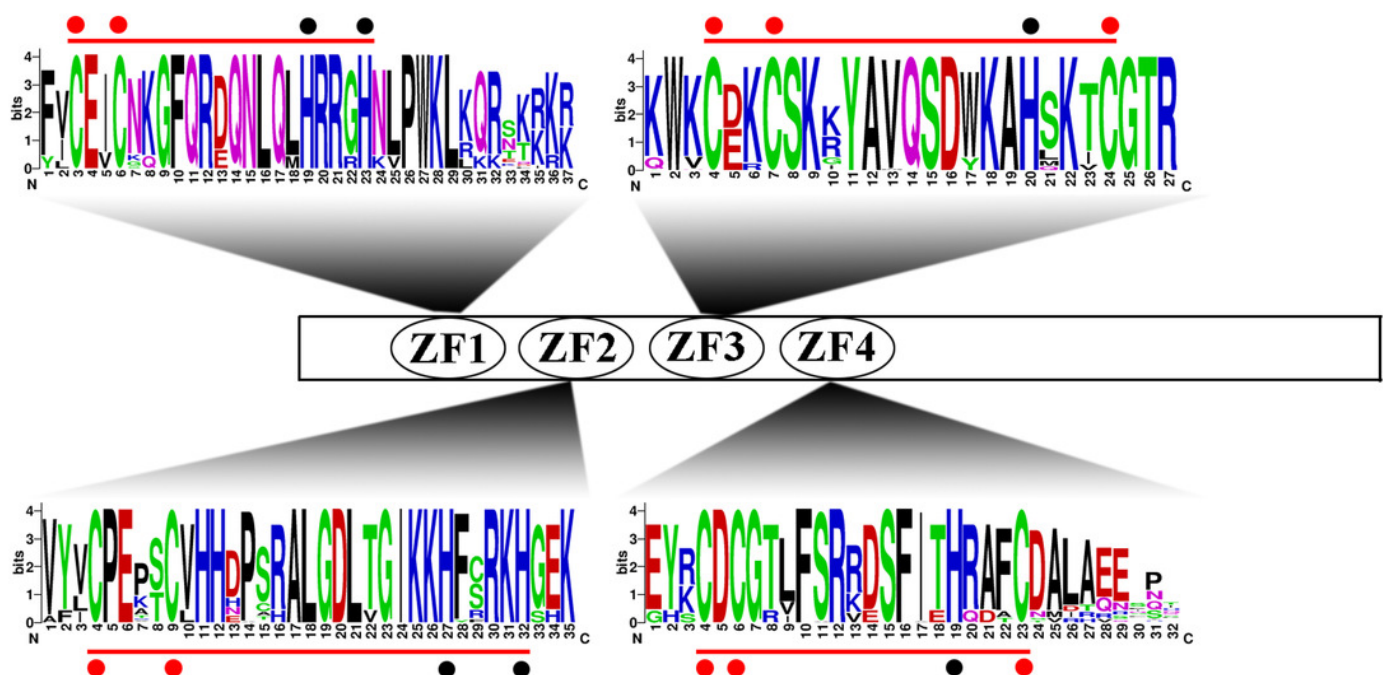


Figure 4

Chromosomal locations of IDD genes in pear, strawberry, raspberry, plum, and cherry.

Chromosomal locations of IDD genes in pear (A), strawberry (B), raspberry (C), plum (D), and cherry (E). Duplicated gene pairs are connected with coloured lines.

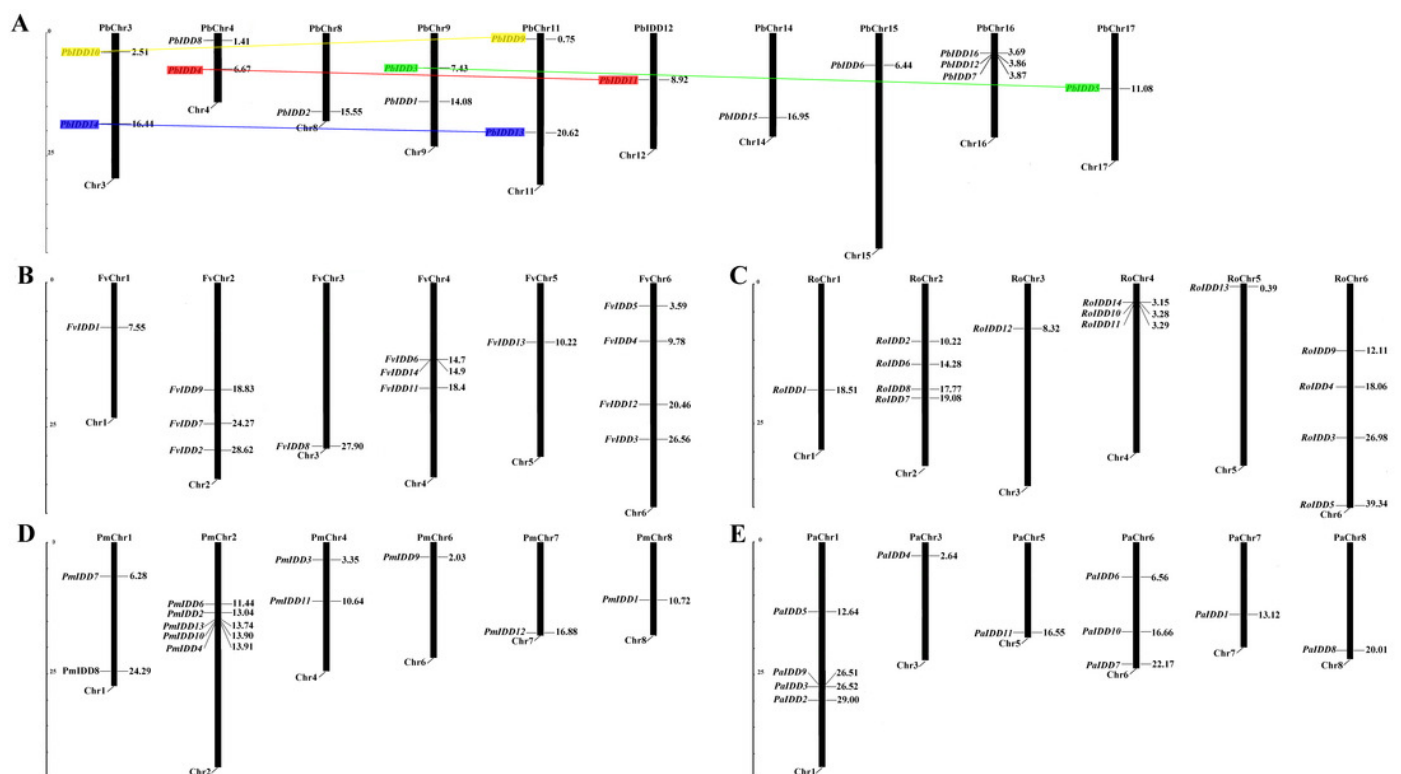


Figure 5

Microsynteny of regions among pear, strawberry, plum, raspberry, cherry and apple.

Microsynteny of regions among pear, strawberry, plum, raspberry, cherry and apple. The chromosome numbers are indicated by differently coloured boxes and are labelled by Pb, Fv, Pm, Ro, Pa, and Md. The differently coloured boxes also represent the sequence lengths of chromosomes in megabases. The black line indicates the syntenic relationship among the IDD regions.

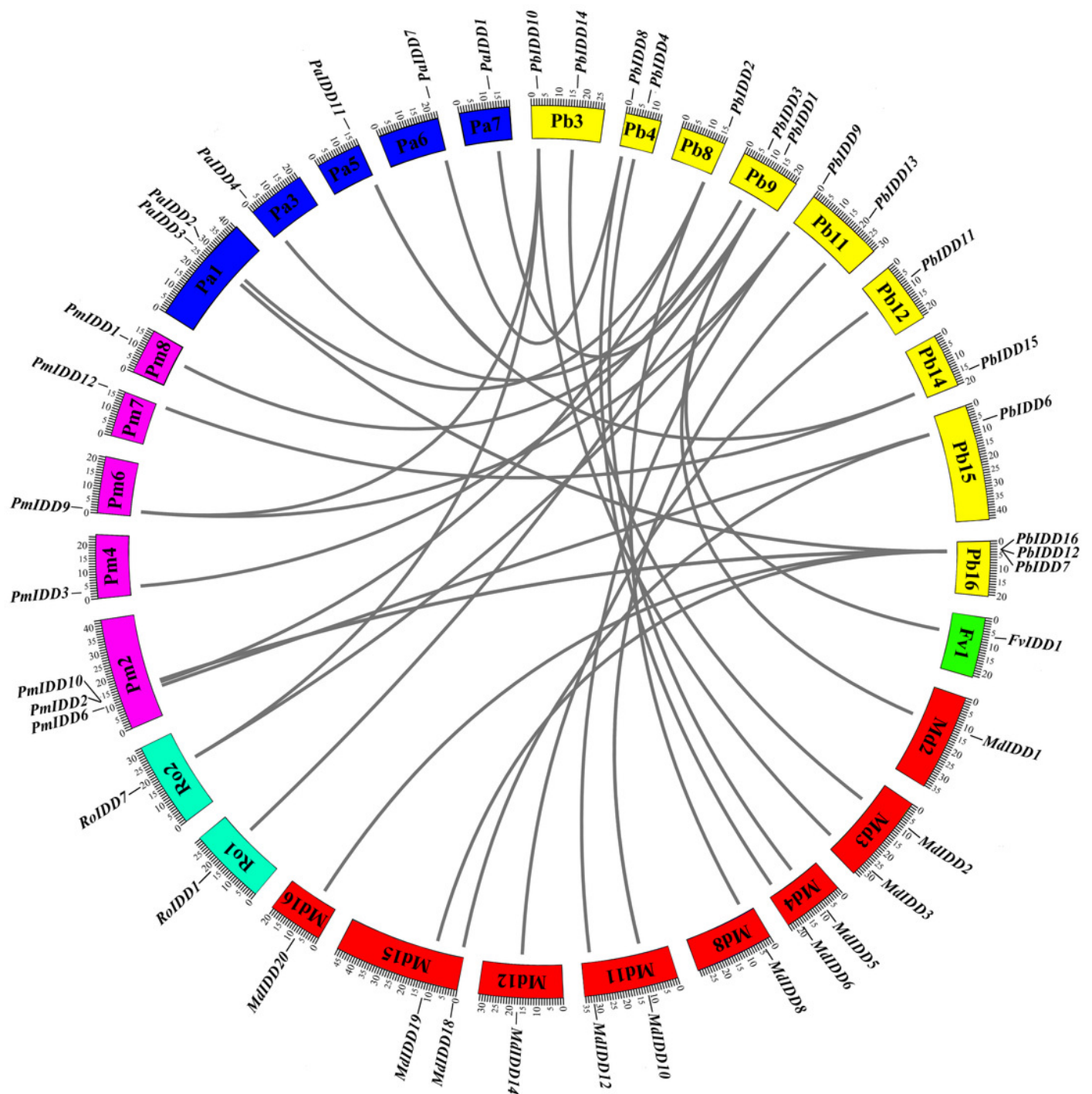


Figure 6

Potential cis-elements in the 5' regulatory sequences of the 16 PbIDD genes.

Potential cis-elements in the 5' regulatory sequences of the 16 PbIDD genes.

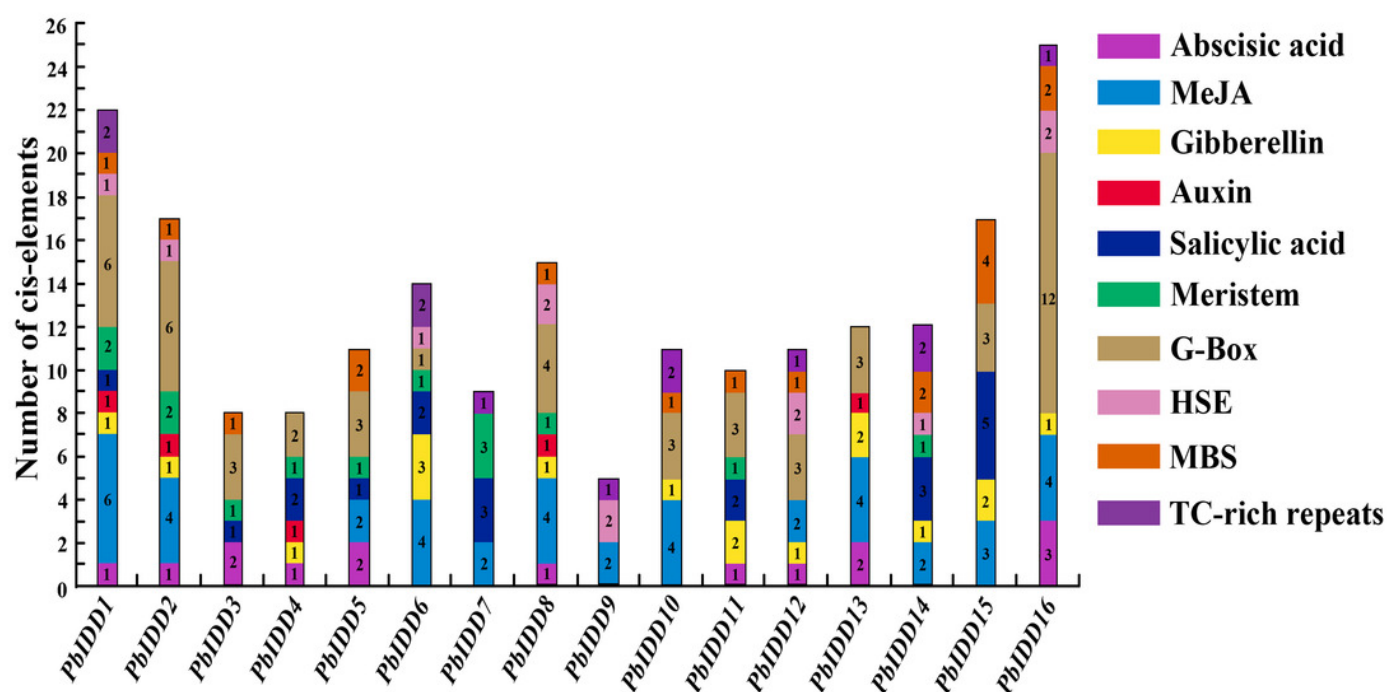
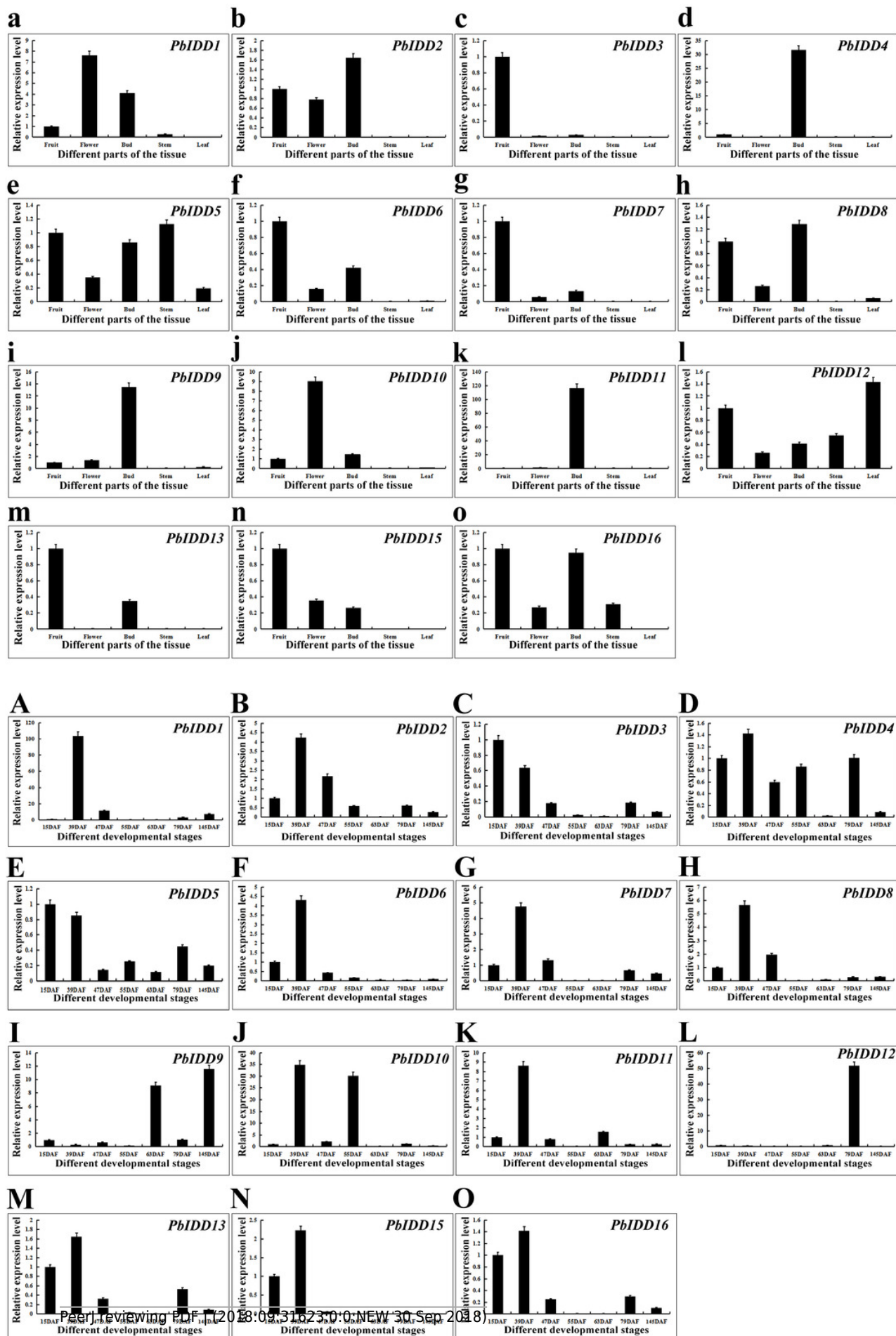


Figure 7

Expression patterns of IDD genes of Chinese white pear in different tissues and in fruit at different developmental stages.

Expression patterns of IDD genes of Chinese white pear in different tissues and in fruit at different developmental stages. (a-o) Expression patterns of IDD genes in Chinese white pear in different tissues. (A-O) Expression patterns of IDD genes in Chinese white pear at different developmental stages.



Review Tips

The English should be improved. As a Chinese writer, the Chinglish in this manuscript is obvious. The language statement is not refined. There are some repeated sentences can be found both in Result and Discussion. Most of these kinds of language problems I found have been labeled with wavy underline or linear underline in the PDF.

Line 44-58: The research progress of IDD gene functions in plants were described in too much detail and look like a simple list. I suggest you improve this by providing a conclusive evaluation of IDD gene functions study progress with few sentences.

Line 59-77: In this paragraph, there is a long text about research background of the molecular mechanism of SCW in pear which seems to be too narrow compared with the title of this manuscript. I suggest you review abundant researches on roles of IDD in fruit development and phylogenetic and evolutionary studies of IDD in other plant taxa.

Line 80-89: There are detail descriptions of study results which should be described in the "Results" section. I suggest you shorten this paragraph and briefly explain the purpose and significance of the current study based on the previous research background.

Line 148-152: The number of the IDD identified in different species should be described in a simple and clear way. I suggest you integrate this paragraph with the following paragraph of phylogenetic analysis. Besides, this paragraph lacks basic information about the IDD amino acid sequence and protein. You can refer to the literature (Fan et al., 2017) in your references list.

Line 160-170: The speculation of different IDD genes based on phylogenetic clustering is inadequate. You should combine all of the analyses results including the gene structure, motif (Line 194 and 197) , location to provide a more reliable speculation, and **these should be described in the "Discussion" section**. Here, what you can do is just describe the evolutionary relationship of these IDD genes in terms of species and the number of IDD genes based on what you can see according to phylogenetic tree.

Line 224: I suggest integrate his paragraph with "chromosomal location" and " phylogenetic analysis " to better analyze the findings.

Line 230-239: I suggest you include the Arabidopsi, rice and maize in the Microsyneny Analysis (you might do an additional analysis for pear and these three species) as you do in the phylogenetic analysis, which can lead you to draw a **complete and clear** evolutionary events of **all of IDDS genes** (paralogs, ortholog) identified in this study before and after divergence of Rosaceae species, rather than few examples in the current manuscript. I would like to see a conclusion of "hypothetical evolutionary mode" of IDDs genes based on the combined results in the end of this article in the form of a picture.

Line 247-259: Improve the description of the results with clear (e.g. the exact of number of genes)

and concise comparisons among or between different IDD genes. There are a lot of speculations which should be discussed in the “Discussion” section.

Discussion:

Line 296: The number of IDD genes identified in apple and pear (both belong to Maloideae) are different, and the author needs to do some explanation.

Line 303-30: The explanation is weak, the possible evolutionary processes of gene should be proposed based on combined results of this study and other related studies.

Line 321-324: The results of nuclear localization prediction are not found in the “Results”, and the related figures of tables are not shown here. The explanation of “Perhaps this lack of a nuclear localization sequence is unique to the Rosaceae IDD genes” is weak and contradictory to “These proteins have a highly conserved ID domain with a nuclear localization signal” (Line 41).

Line 330-331: The same sentence has been described in the Results. And the “explanation” or speculation of why group 4 lacks these two conserved sequences is actually not provided.

Line 363-365: You should compare the gene expression pattern of a pair of the most homologous IDDS genes in apple and pear to get some congruence or contradiction, and provide corresponding explanation.

Line 366-367: Since the IDD gene from Arabidopsis, rice and maize are included in this study, you should clearly state which IDD in the three species may have similar function to those (PbIDD2, 6, 8, 12, 16) found in pear, and provide stronger support from the other evidences besides the cis element.

Line 374-376: “The expression analysis of these 16 *PbIDDs* at seven developmental stages of fruit showed that the expression pattern of any one gene was consistent with the developmental law of ‘Dangshan Su’ pear fruit stone cells and lignin.” I can’t see and understand the “consistence” from the Figure 7.

Line 390-392: PbIDD6 is closely related to MdIDD9 and MdIDD18, so you should compare expression pattern of PbIDD6 with MdIDD9 and MdIDD18 (Fan et al., 2017). And, you will find contradiction. Since you did not carry out related laboratory work related to lignin in any tissue, it is too weak to make such a speculation based on a phylogenetic relationship only.

For the Discussion section, you should combine all of the results including expression pattern, phylogenetic, chromosomal location, microsynteny, gene structure and motif to draw a conclusive picture of speculated mode of evolution and function diversification of all of the IDD genes identified in this study. To do this, my suggestion is divide your Discussion into several organized topics rather than a similar arrangement to the Results. Since you did not carry out any laboratory work on either lignin or other interesting compounds, it is strange that you just pick two

genes that are possibly related to lignin based on expression pattern during fruit development to make some discussions (also in the introduction) . According to the title (topic) of the study, you should not ignore the other IDD genes that might participate in important pathways of fruit development. If you are most interested in lignin, you might carry out related work to provide a reliable evidence.