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

Xueqiang Su [1] , **Tiankai Meng** [2] , **Yu Zhao** [1] , **Guohui Li** [1] , **Xi Cheng** [1] , **Muhammad Abdullah** [1] , **Xu Sun** [1] , **Yongping Cai** [Corresp., 1] , **Yi Lin** [Corresp. 1]

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

[2] School of Life Sciences and Technology, TongJi University, Shanghai, China

Corresponding Authors: Yongping Cai, Yi Lin
Email address: ypcaiah@163.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. IDD genes have been identified throughout the genomes of the model plants *Arabidopsis thaliana* and *Oryza sativa*, and the functions of many members of this gene family have been studied. However, few studies have investigated the IDD gene family in Rosaceae species (among these species, a genome-wide identification of the IDD gene family has only been completed in *Malus domestica*). This study focuses on a comparative genomic analysis of the IDD gene family in five Rosaceae species (*Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis* and *Prunus avium*). We identified a total of 68 IDD genes: 16 genes in Chinese white pear, 14 genes in *Fragaria vesca*, 13 genes in *Prunus mume*, 14 genes in *Rubus occidentalis* and 11 genes in *Prunus avium*. The evolution of the IDD genes in these five Rosaceae species was revealed by constructing a phylogenetic tree, tracking gene duplication events, and performing a sliding window analysis and a conserved microsynteny analysis. The expression analysis of different organs showed that most of the pear IDD genes are found at a very high transcription level in fruits, flowers, and buds. Based on our results with those obtained in previous research, we speculated that *PbIDD2* and *PbIDD8* might participate in flowering induction in pear. A temporal expression analysis showed that the expression patterns of *PbIDD3* and *PbIDD5* were completely opposite to the accumulation pattern of fruit lignin and the stone cell content. The results of the composite phylogenetic tree and expression pattern analysis indicated 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 in five Rosaceae species and thereby provide a theoretical basis for studying the function of these IDD genes.

1 **Comparative genomic analysis of the IDD genes in five Rosaceae**

2 **species and expression analysis in Chinese white pear (*Pyrus***

3 ***bretschneideri*)**

4 Xueqiang Su#, Tiankai Meng#, Yu Zhao, Guohui Li, Xi Cheng, Muhammad Abdullah, Xu Sun, Yongping Cai*, Yi Lin*

5 1 School of Life Science, Anhui Agricultural University, Hefei, Anhui, China

6 2 School of Life Sciences and Technology, TongJi University, Shanghai, China

7   # This authors have contributed equally to this work.

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

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

Abstract

11   The INDETERMINATE DOMAIN (IDD) gene family encodes hybrid transcription factors with distinct zinc finger motifs
12 and appears to be found in all higher plant genomes. IDD genes have been identified throughout the genomes of the model plants
13 *Arabidopsis thaliana* and *Oryza sativa*, and the functions of many members of this gene family have been studied. However, few
14 studies have investigated the IDD gene family in Rosaceae species (among these species, a genome-wide identification of the
15 IDD gene family has only been completed in *Malus domestica*). This study focuses on a comparative genomic analysis of the
16 IDD gene family in five Rosaceae species (*Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis* and *Prunus
17 avium*). We identified a total of 68 IDD genes: 16 genes in Chinese white pear, 14 genes in *Fragaria vesca*, 13 genes in *Prunus
18 mume*, 14 genes in *Rubus occidentalis* and 11 genes in *Prunus avium*. The evolution of the IDD genes in these five Rosaceae
19 species was revealed by constructing a phylogenetic tree, tracking gene duplication events, and performing a sliding window
20 analysis and a conserved microsynteny analysis. The expression analysis of different organs showed that most of the pear IDD
21 genes are found at a very high transcription level in fruits, flowers, and buds. Based on our results with those obtained in previous
22 research, we speculated that *PbIDD2* and *PbIDD8* might participate in flowering induction in pear. A temporal expression
23 analysis showed that the expression patterns of *PbIDD3* and *PbIDD5* were completely opposite to the accumulation pattern of
24 fruit lignin and the stone cell content. The results of the composite phylogenetic tree and expression pattern analysis indicated
25 that *PbIDD3* and *PbIDD5* might be involved in the metabolism of lignin and secondary cell wall (SCW) formation. In summary,
26 we provide basic information about the IDD genes in five Rosaceae species and thereby provide a theoretical basis for studying
27 the function of these IDD genes.

INTRODUCTION

29   Zinc finger proteins are transcription factors with a finger-like domain that are widely distributed in animals,

30   microorganisms, and the plant kingdom (Miller et al., 1985). In fact, zinc finger proteins are found in all plant eukaryotic lineages,

31   which indicates that these proteins might have been derived from a common eukaryotic ancestor. Zinc finger proteins have been

32   identified and functionally analysed in several plant species, such as *Jatropha curcas* (Shi et al., 2018), *Oryza sativa* (Zhang et al.,

33   2018), *Musa acuminata* (Chen et al., 2014). One group of this large family of proteins, the INDETERMINATE DOMAIN (IDD)

34   proteins, has a highly conserved ID domain (Colasanti et al., 1998), which contains typical C2H2 and C2HC zinc finger motifs

35   (Wu et al., 2008). C2H2 zinc finger transcription factors, which are one of the most thoroughly studied transcription factor

36   families (Agarwal et al., 2007; Wei et al., 2016), contain tandem repeat segments of approximately 30 amino acids, all of which

37   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

38   cysteine and histidine, respectively, X represents any amino acid, and psi represents a hydrophobic residue) (Parraga et al., 1988).

39   The structure obtained from this particular sequence can bind to Zn^{2+} and fold to form a structure with two β-sheets and an α-

40   helix. Zn^{2+} mainly plays a role in the linking of individual amino acid chains and is also crucial to the function of zinc finger

41   proteins (Islam et al., 2009; Tian et al., 2010).

42   The IDD protein family is a special type of C2H2 zinc finger subgroup. These proteins have a highly conserved ID domain

43   with an N-terminal nuclear localization signal (NLS) and four distinct zinc finger motifs (ZF1, ZF2, ZF3, and ZF4) (Colasanti et

44   al., 2006; Kozaki et al., 2004). Two canonical zinc fingers are found in the IDD genes, and the encoded proteins also contain two

45   unusual CCHC fingers, one of which is also associated with the zinc finger domain in the *Saccharomyces cerevisiae* SW15

46   protein (Kozaki et al., 2004). Among the species for which a complete genome-wide identification of the IDD gene family has

47   been performed, *Malus domestica* has a higher number of IDD genes than *A. thaliana*, *O. sativa*, and *Zea mays* (Colasanti et al.,

48   2006; Fan et al., 2017). The functions of some IDD genes, particularly those from *A. thaliana* (IDD genes with similar functions

49   have also been reported in *O. sativa* and *Z. mays*), have been identified. The function of the IDD genes involves seed germination

50   (*AtIDD1*, *2*, *ZmIDD9*, *ZmIDDveg9*) (Feurtado et al., 2011; Yi et al., 2015), leaf and flower development (*AtIDD4*, *14*, *15*, *16*)

51   (Cui et al., 2013), flowering regulation (*ZmID1*, *AtIDD8*, *OsID1*) (Colasanti et al., 1998; Seo et al., 2011; Park et al., 2008),

52   starch metabolism (*AtIDD5*) (Ingkasuwan et al., 2012), the regulation of plant gravitropism (*AtIDD14*, *15*, *16*, *LPA1*, *OsIDD16*,

53   *18*), root development and nitrogen metabolism (*AtIDD3*, *8*, *9*, *10*) (Jeong et al., 2015; Coelho et al., 2018).

54   Pear, which belongs to the Rosaceae family, is widely cultivated throughout the world and is popular due to its unique

55   flavour. China has a long history of eating pears, and the pears cultivated in China are mainly *Pyrus bretschneideri*, *Pyrus*

56   *pyrifolia*, *Pyrus sinkiangensis* and *Pyrus ussuriensis* (Vavilov., 1951; Xue et al., 2018). 'Dangshan Su' pear (*Pyrus bretschneideri*

57   cv. Dangshan Su) is a diploid pear variety that originated in Dangshan County, Anhui Province, China, and due to its good

58   quality and high medicinal efficacy (Hamauzu et al., 2007; Huang et al., 2010; Zhang et al., 2017), the fruit has a very high

59   market value. The stone cells in pear are closely related to the flesh texture, and the content of stone cells is a key factor in

60   determining the quality of pear fruit (Cai et al., 2010; Yan et al., 2014). However, few studies have investigated the IDD genes

61   during fruit development, and the only results for *M. acuminata* have been reported (Chen et al., 2014). In *M. acuminata*, the

62   expression of *MaIDD* is closely related to fruit development, whereas, in *M. domestica*, the IDD gene also appears to play a role

63   in regulating flower induction. In addition, some zinc finger proteins, such as the CCCH-tandem zinc finger protein IIP4 (Zhang

64   et al., 2018), *A. thaliana* C3H14, C3H15 (Chai et al., 2015) and *OsIDD2* in *O. sativa* (Huang et al., 2018), are involved in plant

65   SCW formation. Although the function of several IDD genes has been studied and their potential functions can be explained by

66   bioinformatics, it remains unknown whether IDD genes participate in the regulation of flower induction, fruit development,

67   lignin biosynthesis, and SCW formation in *P. bretschneideri*.

68   Among Rosaceae species, genome-wide identification of IDD genes has only been performed in *M. domestica* (Fan et al.,

69   2017). Currently, there remain many questions about the IDD gene family in Chinese white pear. In this study, we identified the

70   IDD genes in five Rosaceae plants and performed a comparative genomic analysis. Woodland strawberry (*F. vesca*) is a perennial
71   herb with edible fruits that grows well in most parts of the Northern Hemisphere, and its genome-wide sequencing was first
72   reported in 2011 (Shulaev et al., 2011). Japanese apricot (*P. mume*) has great ornamental value and in some Asian countries, the
73   fruits of this tree are used for cooking and flavouring. Sweet cherry (*P. avium*), which originated in Europe, Anatolia, Maghreb
74   and western Asia, is an important economic tree species (Tavaud et al., 2004), and a new report on the de-genomic sequence of
75   this species was published in 2017 (Shirasawa et al., 2017). In addition, black raspberry (*R. occidentalis*), a specialty fruit crop
76   that is grown in the Pacific Northwest of the United States, has a unique flavour and rich nutritional value, and the whole-genome
77   sequencing of this species was published in 2017 (VanBuren et al., 2017). These four species and Chinese white pear belong to
78   the Rosaceae family and might be connected through evolutionary relationships. We compared the IDD gene family of Chinese
79   white pear with those of the four above-mentioned species to help us obtain a more comprehensive understanding of the IDD
80   gene family and to further predict and analyse the functions of the IDD gene family members in Chinese white pear.

81   # MATERIALS AND METHODS

82   **Identification of IDD Genes in Five Rosaceae Plants**

83   In this study, the Chinese white pear genome database was obtained from the Pear Genome Project
84   (http://peargenome.njau.edu.cn/). The genome databases of three Rosaceae plants (*F. vesca*, *R. occidentalis*, *P. avium*) were
85   downloaded from GDR (https://www.rosaceae.org/). The genome sequence data of *P. mume* were downloaded from the *P. mume*
86   genome project (http://prunusmumegenome.bjfu.edu.cn/). DNATOOLS software was used to create a local database containing
87   the amino acid sequences of the IDD genes in the five Rosaceae plants (Cao et al., 2016). To identify the IDD genes of these five
88   Rosaceae species, we used the following methods. First, the *A. thaliana* (16) and *M. domestica* (20) IDD gene sequences were
89   collected for use as query sequences in a TBlastN search with the default E-value (Supplementary Table S1). We compared these
90   sequences with the local database sequences of the above-mentioned five Rosaceae species. Using the SMART online software
91   program, the IDD candidate gene sequences initially screened by BLAST were then tested to determine whether they contained a
92   zinc finger domain (http://smart.embl-heidelberg.de/) (Letunic et al., 2012). Each candidate sequence containing a zinc finger
93   domain was subsequently manually checked for an IDD domain, and the protein sequences lacking a full IDD domain and
94   redundant sequences were discarded. We then isolated the candidate IDD gene sequences and used the ExPASy online site to
95   investigate the molecular weights of the IDD proteins (http://web.expasy.org/protparam/) (Artimo et al., 2012). The subcellular
96   localization of all IDD proteins was predicted using the WoLF PSORT website (http://www.g enscript.com/wolf-psort.html)
97   (Horton et al., 2007). Phtre2 website (http://www.sbg.bio.ic.ac.uk/phyre2) for protein three-dimensional structures prediction
98   (Kelley et al., 2007).

99   **Phylogenetic Analysis**

100   Sequence alignment of all IDD proteins was done using the ClustalW tool in MEGA 7.0 software. The phylogenetic tree
101   was constructed using MEGA 7.0 software using the maximum likelihood (ML) (bootstrap = 1000) (Kumar et al., 2016). The *M.*
102   *domestica* IDD gene sequence used in phylogenetic tree was derived from the results of Fan et al (Fan et al., 2017). The
103   sequences of *A. thaliana*, *O. sativa* and *Z. mays* were obtained from the article by Colasanti (Colasanti et al., 2006).

104   **IDD gene Structures and Conserved Motif Prediction**

105   IDD gene structures were compared using Gene Structure Server (http://gsds.cbi.pku.edu.cn) (Guo et al., 2007). The motifs

106   of the IDD genes in these five Rosaceae species were analysed using MEME online analysis software

107   (http://meme.sdsc.edu/meme4_3_0/intro.html) (Bailey et al., 2015). Parameters for the conserved motif prediction were motif

108   width greater than 6 and less than 200. The number of identified motifs was 20.

109   **Chromosomal Location, Gene Duplication and Ka/Ks Ratio Analysis**

110   Five Rosaceae species chromosome start positions and other relevant information about the IDD genes were obtained from

111   the public genome database. The chromosomal physical locations of the IDD genes of all five Rosaceae species were mapped

112   using MapInspect software (http://mapinspect.software.informer.com) (Hu et al., 2011; Lin et al., 2011; Zhu et al., 2015). To

113   define gene duplication events, we mainly relied on the following principles: 1, The similarity of the two gene-coding sequences

114   (CDS) was more than 80%. 2, If the two genes were located on the same chromosome and separated by at least 200 kb, we

115   considered these two genes tandem-duplicated genes. 3, If the two genes were located on different chromosomes, they were

116   called segmentally duplicated genes (Long & Thornton, 2001). DnasP v5.0 software was used to calculate non-synonymous (Ka)

117   and synonymous substitution (Ks) values and perform sliding window analysis. Parameters for sliding window analysis were

118   window size 150 bp and step size 9 bp (Librado & Rozas, 2009).

119   **Microsynteny analysis and Chinese white pear IDD gene promoter *Cis*-acting element analysis**

120   We obtained the promoter sequence of each IDD gene from the Chinese white pear genome database, including the DNA

121   sequence of the initiation codon (ATG) located 1500 bp upstream of each IDD gene. The online software Plantcare was used to

122   analyse the cis-acting elements of the promoter region (http://bioinformatics.psb.ugent.be/webtools/plantcarere/html/) (Rombauts

123   et al., 1999). The microsynteny of 6 Rosaceae species (we added the IDD genes of *Malus domestica*) was identified using the

124   Multiple Collinearity Scan toolkit (MCSscanX) (Abdullah et al., 2018).

125   **RNA extraction and qRT-PCR**

126   The plant material used in this experiment was the 'Dangshan Su' pear, which grows in the Yuanyichang agricultural park

127   (Dangshan County, Anhui Province, China). Samples of current-year flowers, flower buds, stems, mature leaves and fruits were

128   obtained from 40-year-old *Pyrus bretschneideri* cv. Dangshan Su trees. The fruits were picked 15 days after pollination (DAP),

129   39 DAP, 47 DAP, 55 DAP, 63 DAP, 79 DAP and 145 DAP, and the fruits obtained 39 DAP were used for the expression

130   analysis in different organs. 'Dangshan Su' pear buds were treated with gibberellin (GA) and sucrose. Specifically, on a clear

131   morning, terminal buds on current-year spurs were sprayed (<5 cm) with GA (700 mg•L^{-1}) and sucrose (20,000 mg•L^{-1}) using a

132   low-pressure hand wand sprayer. Some plant materials were collected immediately after spraying with GA and sucrose and

133   stored at -80℃ for use as controls. Materials collected 2 hours (H), 4 H, 6 H, 8 H, 12 H, and 24 H After GA and sucrose

134   treatment were stored at -80°C for gene expression analyses and samples collected 0 hours post-treatment (HPT), 2 HPT, 4 HPT,

135   6 HPT, 8 HPT, 12 HPT, and 24 HPT were also collected. RNA was extracted from the plant materials (including fruits and

136   various organs) using a plant RNA extraction kit from Tiangen (Beijing, China). Reverse transcription was performed using a

137   PrimeScript^{TM} RT reagent kit with gDNA Eraser (Takara, Japan), and each reaction consisted of 1 μg of RNA. The qRT-PCR

138   primers for the pear IDD genes were designed using Beacon Designer 7 software (Supplementary Table S8). Each 20-μL qRT-

139   PCR system consisted of 10 μL of SYBR® Premix Ex Taq™ II (2×) (Takara, Japan), 2 μL of cDNA, 6.4 μL of water and 0.8 μL

140   of PbIDD-F and PbIDD-R. The pear *Tubulin* gene (No. AB239680.1) (Wu et al., 2010) was used as an internal reference. The

141   reaction procedure was performed following the instruction manual, and three repetitions were run for each sample. The relative

142   expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001).

RESULTS

1, Identification, Characterization and Phylogenetic Analysis of IDD Genes

A total of 68 IDD proteins were identified and used for further analysis (Table 1; Supplementary Table S1), which included 16 IDD proteins (*PbIDD1-PbIDD16*) in Chinese white pear. We also identified a total of 52 IDD proteins in the other four species, including 14 in *F. vesca* (*FvIDD1-FvIDD14*), 13 in *P. mume* (*PmIDD1-PmIDD13*), 14 in *R. occidentalis* (*RoIDD1-RoIDD14*), and 11 in *P. avium* (*PaIDD1-PaIDD11*). Although only 11 IDD proteins were identified in *P. avium*, the numbers of IDD proteins in the other species were very similar but fewer than the number of IDD proteins in *M. domestica* (20). We used ExPASy software to calculate the physicochemical parameters of the IDD genes (Supplementary Table S2). We identified 68 IDD genes with isoelectric points (pIs) greater than 7. In these Rosaceae species, the lowest pI value was 7.76 (*PaIDD8*), whereas the highest pI value was 9.45 (*FvIDD4*). The molecular weights of these IDD genes were quite different, ranging from 42.0 kDa for *PaIDD10* to 79.6 kDa for *PmIDD8*. The numbers of amino acids in the IDD proteins were also quite different. *PbIDD14* contained the fewest (only 381) amino acids, whereas *PmIDD8* had the largest number (728). The grand average hydropathicity (GRAVY) of all 68 IDD genes was less than 1, and the smallest GRAVY value (-0.914) was obtained for *PbIDD15*.

A phylogenetic tree of 98 IDD proteins was constructed using the maximum likelihood (ML) method (Figure 1). Based on the phylogenetic tree, we identified four phylogenetic groups (groups 1-4). Group 1 had the most members and contained all six Rosaceae species and 44 IDD genes: *P. bretschneideri* (8), *F. vesca* (7), *P. mume* (6), *R. occidentalis* (7), *P. avium* (6), *M. domestica* (10). Groups 3, 2 included 21 and 14 IDD genes, respectively. Group 4 had the fewest members: two IDD genes in *P. bretschneideri*, two IDD genes in *F. vesca*, two IDD genes in *P. mume*, two IDD genes in *R. occidentalis*, one IDD gene found in *P. avium* and no IDD gene in *M. domestica*. *ZmID1* was grouped into a class by itself, and most *PbIDD* genes were more tightly grouped with *MdIDDs*. In addition, the IDD genes of some Rosaceae species were closely related to those of *A. thaliana* and *O. sativa*: *PbIDD3*, *PbIDD5*, *FvIDD3*, *RoIDD4*, *PmIDD3*, and *PaIDD4* were included in group 1 and were closely related to *OsIDD2*; *FvIDD1*2, *RoIDD9*, *PmIDD11*, *PaIDD5* and *AtIDD3* and *8* are closely related, and *AtIDD1* was included in group 3 and clustered with seven Rosaceae IDD genes.

2, Structural and Conserved Motif Analysis of IDD Proteins

To more comprehensively analyse the structural diversity of the IDD genes in the five Rosaceae species, we created exon-intron pattern maps of all 68 IDD genes (Figure 2). Twenty-five of the 34 IDD genes in group 1 had three exons and the other six genes had two exons. Additionally, *FvIDD7* (4), *RoIDD6* (5) and *PmIDD8* (6) had a large number of exons. The analysis of group 2 revealed that *FvIDD11* had two exons, *PbIDD7* had four exons and all the other members had three exons. Group 3 comprised a subfamily with a complex gene structure: 11 genes had four exons and four genes had three exons. All members of group 4 had a gene structure comprising three exons. The lack of high similarity among the 68 IDD genes allowed a better understanding of the conserved motifs in these IDD genes. We used MEME software to identify 20 conserved motifs (Figure 2; Supplementary Table S3). Motifs 2, 3, 1 and 4 encoded the ID domain, whereas the remaining motifs did not have functional annotations. As shown in Figure 3, motifs 2, 3, 1, 4 represented the zinc fingers ZF1, ZF2, ZF3, and ZF4, respectively. The four conserved motifs constituted a conserved ID domain, whereas 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 through a sequence alignment of the conserved ID domains of five species and the results showed that the ID domain was highly conserved. The cysteine (C) and histidine (H) residues in each zinc finger domain were conserved in each species,

181   consistent with the results of previous studies (Supplementary Figure S1-5). All 68 IDD genes had motifs 1-4, which were the

182   most conserved. In addition, some members of the same group and the more closely related members had highly similar motif

183   compositions (e.g., *FvIDD7* and *RoIDD6* in group 1; *PmIDD10* and *PaIDD9* in group 2). These genes were closely related and

184   had exactly the same motif compositions. We also identified certain group-specific motifs, such as motifs 12, 17, 20 in members

185   of group 4 and motif 19 in some members of group 1.

186   We also identified a NLS in the N-terminal border of these IDD genes (Supplementary Figure 6). According to previous

187   studies, this NLS is usually KKKR or KRKR (Colasanti et al., 2006). Not all members in the five Rosaceae species, *PbIDD13,*

188   *14*, *FvIDD10*, *11*, *12* and *PmIDD11*, have a NLS. In addition to the conserved IDD domain, we also found two C-terminal

189   domains (MSATALLQKAA and TRDFLG), which are encoded by motifs 6, 7, respectively, in some members (Figure 2;

190   Supplementary Figure S1-5). All nine members of group 4 lacked both motifs 6 and 7. Among the members of group 1,

191   *PaIDD10* did not contain motif 6, and *PbIDD14* and *PmIDD8* lacked motif 7. All members of groups 2, 3 contained these two

192   motifs. The lack of these two conserved sequences at the C terminus of some IDD genes might explain why the members of

193   group 4 were clustered into a single category and might also underlie the differences among the functions of these proteins

194   (Figure 1).

195   **3, Chromosomal Location and Duplication Events of IDD Family Genes in Rosaceae**

196   Based on the complete genome sequences of the five Rosaceae species, the exact chromosomal locations of all 68 IDD

197   genes were identified (Figure 4). In *P. bretschneideri*, the 16 IDD genes were distributed on chromosomes 3, 4, 8, 9, 11, 12, 14,

198   15, 16, 17. In *F. vesca*, the IDD genes were distributed on chromosomes 1, 2, 3, 4, 5, 6. In *R. occidentalis*, the IDD genes were

199   mainly distributed on chromosomes 2, 4 and 6. Several other IDD genes were distributed on chromosomes 1, 3 and 5. In *P. mume*,

200   except *PaIDD5*, which could not be mapped to any chromosome, the IDD genes were distributed on chromosomes 1, 2, 4, 6, 7, 8.

201   However, in *P. avium*, 4 IDD genes were distributed on chromosome 1 and the remaining 7 IDD genes were distributed on

202   chromosomes 3, 5, 6, 7, 8. Only 4 pairs of duplicated genes were identified in pear (Figure 4). All duplicated genes identified in

203   *P. bretschneideri* were segmental duplications.

204   To clarify the driving forces of gene duplication and explore the impact of these genes on evolutionary processes, we

205   calculated the Ka, Ks and Ka/Ks ratio for the 4 duplicated gene pairs in *P. bretschneideri*. Ka/Ks=1 is the cut-off value that

206   indicates neutral selection, Ka/Ks<1 represents negative selection and Ka/Ks>1 represents positive selection (Bitocchi et al.,

207   2017). For all four duplicated gene pairs identified in *P. bretschneideri*, the Ka/Ks values were less than 0.2726 (Supplementary

208   Table S4). In the evolutionary processes of genes, positive selection may be overshadowed by strong negative selection (Han et

209   al., 2016). Therefore, to comprehensively explain the selection pressure of IDD genes, we performed a sliding window analysis

210   of the 4 pairs of IDD paralogues in pears (Supplementary Figure S7). Most coding site Ka/Ks ratios were less than one, with

211   exceptions for one or several distinct peaks (Ka/Ks > 1). The Ka/Ks ratios of the conserved IDD domains were less than 1.

212   **4, Microsynteny and Evolutionary Analysis of the IDD Gene Family**

213   For the identification of interspecific microsynteny, we visualized the location of homologous and orthologous genes. A

214   total of 30 orthologous gene pairs were identified in five Rosaceae species (Figure 5; Supplementary Table S5) and these

215   included one, eight, three, six, and 12 collinear gene pairs between *P. bretschneideri* and *F. vesca*, *P. mume*, *R. occidentalis*, *P.*

216   *avium*, *M. domestica*, respectively. Notably, *PbIDD2* also showed a collinear relationship with the members of the *MdIDD*,

217   *PmIDD* and *PaIDD* families. In contrast, two IDD genes from pear matched one pair of *P. mume* or *R. occidentalis* genes. For

218   example, *PbIDD9* and *PbIDD10* were orthologous to *PmIDD9*, whereas *PbIDD9* and *PbIDD10* were orthologous to *RoIDD7*.

219 However, we did not identify any orthologous gene pairs between *P. bretschneideri* and *O. sativa*, *Z. mays*, and *A. thaliana*.
220 We also compared the other four Rosaceae species with these three species, and the results did not reveal any orthologous gene
221 pairs. To further study the evolution of IDD genes, we collected the IDD genes from four monocotyledons (*M. acuminata*,
222 *Sorghum bicolor*, *Z. mays*, and *O. sativa*) and two dicotyledons (*Vitis vinifera* and *A. thaliana*). A phylogenetic tree of the
223 genomes of these 12 species was obtained to perform an evolutionary analysis of the IDD gene family (Supplementary Figure
224 S8A). Several Rosaceae species exhibit a closer relationship and are more distant from monocots. This result is basically
225 consistent with the results of the phylogenetic tree constructed from all IDD genes of 12 species (Supplementary Figure S8B).
226 The results identified 12 classes (classes I-XII). The *P. bretschneideri* and *M. domestica*, *P. avium* and *P. mume* IDD genes often
227 cluster into one class, and the IDD genes of monocotyledons are closely related (i.e., classes X and XI are only composed of IDD
228 genes of monocotyledons). *ZmID1*, *OsID1*, and *SbID1* are closely related, but we did not identify an orthologous gene for the ID1
229 gene in several other species. Class III contained all species except *M. domestica*, and the IDD genes in this subfamily did not
230 have complete MSATALLQKAA and TRDFLG domains.

231 **5, Analysis of *Cis*-acting Elements in the Promoters of the IDD Genes in Chinese White Pear**

232 Plant IDD genes may be involved in multiple biological processes, and hormones affect their expression. We predicted the
233 cis-acting elements of 16 Chinese white pear IDD gene promoters (Figure 6; Supplementary Table S7). Thirteen IDD genes
234 (except *PbIDD7*, *9*, *10*) contained at least one G-Box, which is a light-responsive element, indicating that the expression of these
235 genes may be regulated by light. Eight *P. bretschneideri* IDD genes had HSE, and 11 members contained MBS components. TC-
236 rich repeats, which are cis-acting elements involved in defence and stress responsiveness, were identified in PbIDD1, 6, 7, 9, 10,
237 12, 14, and 16. In addition, there were many cis-acting elements related to the responses to hormones, including responses to
238 abscisic acid (ABRE), methyl jasmonate (MeJA) (CGTCA-motif, TGACG-motif), gibberellin (P-box, GARE-motif), auxin
239 (TGA-element), and salicylic acid (TCA-element). The auxin-related TGA-element was identified in only five members
240 (*PbIDD1*, *2*, *4*, *8*, 13), which was the least number of elements components found. A total of 41 *C*is-acting elements associated
241 with MeJA were found in 9 members, and these were the most abundant *C*is-acting elements. *C*is-acting elements responding to
242 gibberellin were the most widespread, and twelve of the *P. bretschneideri* 16 IDD genes had this element. The related *C*is-acting
243 elements of abscisic acid and salicylic acid were identified in 10 and 9 members, respectively. We found that *PbIDD1* has *C*is-
244 acting elements related to the above five hormones, whereas *PbIDD9* only has cis-acting elements that are responsive to MeJA.
245 Of the 16 pear IDD genes, 12 contained 19 CGTCA-motifs, and TCA-elements appeared in the promoter regions of 10 members,
246 for a total of 22 genes. In addition, the CAT-box and the CCGTCC-box were found in 11 members (*PbIDD1*, *2*, *3*, *4*, *5*, *6*, *7*, *8*, *9*,
247 *12*, 14); these are *C*is-acting elements related to meristem expression.

248 **6, Expression Characteristics of Chinese white pear IDD Genes**

249 To obtain a more in-depth understanding of the Chinese white pear IDD gene family, we investigated the expression
250 patterns of *PbIDDs* in different organs (Figure 7). No *PbIDD14* expression was detected in any of the organs of Chinese white
251 pear. Among the remaining 15 members, *PbIDD4*, *9*, *11* were highly expressed in buds but showed extremely low expression in
252 other organs. *PbIDD1* and *10* were mainly expressed in flowers. *PbIDD2*, *6*, *8*, and *15* were mainly expressed in fruits, flowers,
253 and buds. Among these four genes, *PbIDD2* and *8* showed the highest expression levels in buds, whereas *PbIDD6* and *15*
254 exhibited the highest expression levels in fruit. *PbIDD5*, *12*, and *16* were detected in multiple organs, but the main difference in
255 the expression of these genes was found in the leaves: (1) *PbIDD12* was highly expressed in leaves; (2) although *PbIDD5* was
256 expressed in all five organs, its lowest expression was detected in the leaves; and (3) *PbIDD16* showed almost no expression in
257 leaves. *PbIDD3*, *7*, and *13* were mainly expressed in the fruit of Chinese white pear, and *PbIDD3* and *7* showed almost no

258   expression in other organs.

259   We investigated the expression patterns of 16 IDD genes in Chinese white pear fruit at seven developmental stages (Figure
260   7). The expression level of *PbIDD1* peaked at 39 DAP and was low during the other six fruit developmental stages. The
261   expression peaks of *PbIDD2*, *7*, *8*, *11*, *13*, *16* also appeared at 39 DAP, but high expression of these genes was also detected at
262   the early developmental stages (15 and 47 DAP). At the late developmental stages of fruit (79 and 145 DAP), these genes were
263   expressed at a moderate level. *PbIDD6, 15* exhibited a similar expression pattern as the above-mentioned genes, with the
264   exception that *PbIDD6, 15* were detected at the early developmental stages of fruit but were expressed at extremely low levels at
265   the late developmental stages of fruit. The expression level of *PbIDD4* was higher only at the early stage of fruit development
266   (15-55 DAP) and at 79 DAP. *PbIDD9* had a high transcription level at 63 and 145 DAP, and *PbIDD12* was mainly expressed at
267   79 DAP. *PbIDD10* had two expression peaks, one at 39 DAP and another at 55 DAP. *PbIDD3* and *5* had the same expression
268   pattern: these two genes were mainly expressed at the early stage of fruit development, and their expression level then gradually
269   decreased.

270   **7, Gibberellin and Sugar Response Pattern Analysis of *PbIDDs***

271   Chinese white pear *PbIDD2, 5, 6, 8, 9, 12, 16* are mainly expressed in bud and flower. We treated the bud of Chinese white
272   pear with exogenous GA and sucrose to investigate the expression pattern of *PbIDD2*, *5*, *6*, *8*, *9*, *12*, *16* under GA and sucrose
273   treatment (Supplementary figure S9). Under exogenous GA treatment, *PbIDD2*, *5*, *6*, *8*, *9*, *12*, *16* showed three different
274   expression patterns (activation, inhibition and invariance). There was no significant change in the transcriptional level of *PbIDD5*,
275   *9* under exogenous GA treatment. The expression levels of *PbIDD2*, *8* continued to decrease after exogenous GA treatment,
276   indicating that the expression of *PbIDD2*, *8* was inhibited by GA. *PbIDD6*, *12* showed a significant increase in expression after
277   GA treatment (reaching a peak at 3HPT); thereafter, the expression level gradually decreased to untreated levels. *PbIDD16*
278   reached the maximal expression levels at 2, 4HPT. Overall, the expression of *PbIDD6*, *12*, *16* showed activation by GA. The
279   transcription levels of *PbIDD2*, *8* increased gradually after sucrose treatment in Chinese white pear bud. The expressions of
280   *PbIDD5*, *12*, *16* were not always induced by sucrose. The lowest expression levels of *PbIDD5*, *12* were found at 4HTP, while the
281   lowest expression level of *PbIDD16* was observed at 1HTP. The expressions of *PbIDD6*, *9* were inhibited by sucrose.

DISCUSSION

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

288   In this study, we identified 68 IDD genes in five Rosaceae species (Table 1; Supplementary Table S2). The pI values of all
289   IDD protein were greater than 7, indicating that all these proteins are alkaline proteins. The proteins were all hydrophilic, as
290   demonstrated by the finding that their GRAVY values were less than 1. Among the Rosaceae species, *P. bretschneideri* had the
291   second greatest number of IDD gene family members after *M. domestica* (Fan et al., 2017), and this difference might be traced
292   back to the origin of the IDD gene family, during which time the numbers of IDD genes in *P. bretschneideri* and *M. domestica*
293   were different. Previous studies have shown that plant WGD and chromosome rearrangement alter the sequences of genes, and
294   this alteration is accompanied by widespread gene loss (Adams & Wendel., 2005; Tang et al., 2008). The common ancestor of

295   Rosaceae species has nine chromosomes (Shulaev et al., 2011; Wu et al., 2013). *P. bretschneideri* and *M. domestica* both
296   experienced two WGDs 130 million years ago (Mya) and 30-45 Mya, but the number of chromosomes was only 17 (Jaillon et al.,
297   2007; Guo et al., 2013; Wu et al., 2013). This finding indicates that the nine chromosomes of Rosaceae species ancestors
298   experienced doubling, breaking and fusion during the long process and ultimately formed the 17 chromosomes of *P.*
299   *bretschneideri* and *M. domestica*. During this process, the genome of a species might become unstable and prone to gene
300   replacement, leading to chromosome rearrangement and gene loss (Paterson et al., 2004). It is possible that the IDD genes in *P.*
301   *bretschneideri* were lost during this process, which explains why the number of IDD genes in *P. bretschneideri* was less than that
302   in *M. domestica*. The phylogenetic tree divided 68 IDD genes into four groups (Figure 1). *ZmID1* comprised a separate class in
303   the phylogenetic tree, and we did not identify a closely related gene (Figure 1). At present, *ZmID1* homologous genes have been
304   identified only in *O. sativa* and *S. bicolor*, which belong to the same family of Gramineae (Colasanti et al., 2006). This finding
305   might indicate that the *ID1* gene is unique to Gramineae plants. Interestingly, five Rosaceae species had at least one IDD gene
306   present in each clade, and this finding had strong bootstrap support. These results showed that the rapid duplication of IDD genes
307   occurred before the divergence of these Rosaceae species.

308   The gene structure and conserved motif composition might be closely related to the diversity of gene functions (Swarbreck
309   et al., 2008). Genes in the same subfamily tend to have very similar gene structures, which indicates that these genes might have
310   similar functions (Figure 2). For example, *PbIDD12* and *PmIDD10* in group 2 had the same gene structure (three exons and two
311   introns), and their exon lengths were almost the same. Furthermore, based on the MEME analysis (Figure 2), we found that
312   members of the same subfamily had roughly the same conserved protein motifs, but some differences in motif composition were
313   found between different subfamilies. An N-terminal NLS is absent in many Rosaceae IDD genes (Supplementary Figure 6). All
314   IDD genes in *A. thaliana* have a NLS. *OsIDD10*, which belongs to the *O. sativa* IDD gene family, also lacks a NLS (Colasanti et
315   al., 2006). the results indicate that the lack of NLS sequence in some members of the IDD gene family might be due not to the
316   mutation or deletion of this entire characteristic sequence but rather to an extensive deletion of the N-terminal sequence. The
317   ancestors of the special IDD genes in these Rosaceae species might have appeared after the differentiation of these Rosaceae
318   species from the common ancestors of *A. thaliana*, and some of the N-terminal sequences, which includes the NLS, were lost
319   after the evolutionary process. In addition to the highly conserved ID domain, we found two small domains (Supplementary
320   Figure 1-5) in the C-terminal region (except in members of group 4). The MSATALLQKAA and TRDFLG motifs are conserved
321   regions outside the ID domain but are not characteristically conserved domains of IDD genes. Obviously, the members of group
322   4 lost these two small domains during evolution, which resulted in the generation of functional differences. We found that group
323   4 contained genes in *A. thaliana* (*AtIDD15*) and *O. sativa* (*OsIDD14*). *AtIDD14, 15, 16* synergistically regulate spatial auxin
324   biosynthesis by direct targeting YUCCA5 (YUC5), TRYPTOPHAN AMINOTRANSFERASE of ARABIDOPSIS1 (TAA1) and
325   PINFORMED1 (PIN1) and thereby controlling aerial organ morphogenesis and gravitropic responses in plants (Cui et al., 2013).
326   *OsIDD14* is the functional orthologue gene of *AtIDD15* and plays a role in the regulation of shoot gravitropism in *O. sativa*.
327   Members of this group might have evolved to have functions in regulating the geotropism of plant organs. However, this group
328   does not contain *M. domestica* IDD genes, which indicates that MSATALLQKAA and TRDFLG domains might have been
329   preserved during the evolutionary process or that IDD genes do not play a role in the regulation of organ geotropism in
330   *M.domestica*.

331   The functional divergence of genes often occurs after gene duplication and is accompanied by the introduction of new
332   functions or the loss of original functions (Chao et al., 2017). In addition, gene duplication events are the main driving force for
333   gene family expansion and thus constitute an important method used by plants to adapt to changing climates and environments
334   (Tang et al., 2016). Gene duplication events have been identified in multiple gene families, such as the PKS gene family in

335 *Gossypium hirsutum* (Su et al., 2017), the HSF gene family in *Sesamum indicum* (Dossa et al., 2016), and the WRKY gene

336 family in *Musa acuminate* (Goel et al., 2016). A total of four *PbIDD* gene pairs were identified as duplications in this study, and

337 these duplicated gene pairs were all created by segmental duplication (Supplementary Table 4). Four pairs of segmentally

338 duplicated genes were also identified in *M. domestica*, verifying the reliability of our results. However, we did not identify gene

339 duplication events in *F. vesca*, *P. mume*, *R. occidentalis* or *P. avium*. This finding might be due to the fact that *F. vesca*, *P. mume*,

340 *R. occidentalis* and *P. avium* underwent only one WGD, but *P. bretschneideri* had a recent WGD in addition to an old WGD

341 (Zhang et al., 2012; Wu et al., 2013). The four pairs of duplicated genes that appeared in pears exerted a very positive effect on

342 the expansion of the IDD gene family and contributed to the functional diversification of IDD genes. We then performed a

343 sliding window analysis of the four pairs of segmentally duplicated genes in Chinese white pear, and the results showed that

344 these genes experienced intense purifying selection to maintain their functional stability.

345   We did not identify the orthologous gene pair of IDD genes between several dicotyledons and monocotyledons. Therefore,

346 we preliminarily hypothesized that IDD genes might have appeared after the differentiation of the common ancestors of these

347 species differentiated. After the differentiation of monocotyledons and dicotyledons, two canonical zinc fingers and two unusual

348 CCHC fingers formed a special zinc finger domain (ID domain), which defines IDD genes in all plants. Based on the results, we

349 drew a hypothetical IDD genes evolutionary model map of Rosaceae species (Supplementary figure S10). The ancestors of

350 dicotyledons evolved through complex evolutionary processes to produce Rosaceae species. After an old WGDs, *F. vesca*, *R.*

351 *occidentalis*, *P. avium* and *P. mume* formed the current IDD gene family. *M. domestica* and *P. bretschneideri* formed the IDD

352 gene family after two WGDs. Subsequently, *P. bretschneideri* and another Rosaceae species differentiated into two IDD gene

353 types (with or without the MSATALLQKAA and TRDFLG domains). The IDD gene lacking these two domains may have

354 derived the function of regulation of plant organ geotropism. *M. domestica* did not detect such genes, perhaps because *M.*

355 *domestic*a does not rely on IDD genes to participate in the regulation of organ geotropism.

356   Previous research has shown that IDD genes are most highly expressed in leaves and buds (Fan et al., 2017), but the highest

357 pear IDD gene expression was detected in fruit, flower and bud organs. *PbIDD7* and *MdIDD19* represent a pair of homologous

358 genes that are closely related. *MdIDD19* was mainly expressed in the leaf, fruit, and bud, whereas *PbIDD7* was mainly expressed

359 in fruit and bud organs and showed almost no expression in leaves (Figure 7). The difference between the two homologous gene

360 expression patterns was mainly concentrated in the leaf expression pattern. Many IDD genes are associated with leaf

361 development and photosynthesis. However, we did not identify cis-acting elements related to the light response in the promoter

362 region of *PbIDD7* (Figure 6), whereas abundant light-responding elements have been found in the promoter region of *MdIDD19*

363 (Fan et al., 2017). This finding might be the main reason for the large difference in the expression patterns between *PbIDD7* and

364 *MdIDD19* in leaves. We found that *PbIDD4* showed a high transcription level only in buds but exhibited a very low expression

365 level in fruits and other organs (Figure 7). This gene might not be the main gene involved in various physiological metabolism in

366 fruit but plays an important role in the development of buds. Interestingly, we found that *PbIDD2*, *5*, *6*, *8*, *9*, *12*, *16* showed

367 higher transcription levels in flowers and buds, although the expression in buds was higher than that in flowers. In *A. thaliana*, *O.*

368 *sativa, Z. mays*, IDD genes, such as *AtIDD8*, *ZmID1*, and *OsID1*, are involved in plant flower induction and are regulated by

369 hormones and sugars (Matsubara et al., 2008; Wu et al., 2008; Jeong et al., 2015). GA can regulate plant flowering by inhibiting

370 flowering genes (Zhang et al. 2016; Galvao et al. 2012). The promoter regions of *PbIDD2*, *6*, *8*, *12*, *16* contain the GA-response

371 cis-acting element. Chinese white pear buds were treated with exogenous GA, and the resulting expression patterns of *PbIDD2*, *5*,

372 *6*, *8*, *9*, *12*, *16* were detected, which revealed that these genes were mainly expressed in the buds and flowers of Chinese white

373 pears (Supplementary Figure S9). *PbIDD5, 9* were not induced by GA, probably because these genes do not contain GA-related

374 cis-acting elements. After exogenous GA treatment, the expression level of *PbIDD2*, *8* decreased continuously, indicating that

375 these genes were inhibited by GA. However, the expression levels of these two genes increased gradually after sucrose treatment,
376 indicating that *PbIDD2*, *8* were induced by sucrose. We also identified a large number of light-responsive cis-acting elements in
377 the promoter regions of *PbIDD2, 8*, which suggested that these two genes might respond to photoperiodic signals. Similar results
378 have been found in *M. domestica*: *MdIDD7* is inhibited by GA in response to sucrose (Fan et al., 2017). In summary, the results
379 suggested that *PbIDD2* and *8* might participate in the GA flowering pathway during physiological flower bud differentiation and
380 regulate the plant flowering process in response to sugar and photoperiod signals. *PbIDD1* and *10* were mainly expressed in the
381 flowers of Chinese white pear, probably because these two genes are involved in the synthesis of bioactive substances in flowers.

382   The stone cells and lignin in 'Dangshan Su' pear fruit are mainly formed during the period from 15 to 63 DAP and reach
383 their peak levels at 55 DAP (Zhang et al., 2017). The expression analysis of these 16 *PbIDDs* at seven developmental stages of
384 fruit showed that the expression pattern of any one gene was not consistent with the trends obtained for the 'Dangshan Su' pear
385 fruit stone cell and lignin contents (Figure 7). However, there were two additional specific genes (*PbIDD3* and *PbIDD5*) and the
386 expression patterns of these two genes showed completely opposite trends compared with the accumulation of 'Dangshan Su' pear
387 stone cells and lignin. *PbIDD3* and *5* showed extremely high expression levels at 15 DAP. From 15 to 47 DAP, which is a stage
388 characterized by the massive accumulation of lignin and stone cells, the expression of these two genes gradually decreased. At
389 the peak of lignin accumulation (47, 55 and 63 DAP), *PbIDD3* and *5* showed very low transcription levels, whereas at the late
390 stage of fruit development (79 and 145 DAP), the expression levels of these genes increased slightly. According to the
391 phylogenetic tree (Figure 1), *PbIDD3*, *PbIDD5* and *OsIDD2* were the most closely related genes. In *O. sativa*, *OsIDD2* regulates
392 SCW formation while negatively regulating the expression of genes encoding key lignin synthesis enzymes, such as cinnamyl
393 alcohol dehydrogenase (CAD), to inhibit lignin biosynthesis (Huang et al., 2018). In Chinese white pear, CAD is predicted to
394 regulate the biosynthesis of fruit lignin. The expression pattern of *PbCAD* is exactly the same as the accumulation patterns of
395 stone cells and lignin (Cheng et al., 2017). The expression patterns of *PbIDD3* and *PbIDD5* were opposite to those found for
396 *PbCAD*. Moreover, *PbIDD3* and *PbIDD5* expression was low during the high expression period of *PbCAD* and was higher
397 during the low expression period of *PbCAD*. In addition, *PbIDD3* and *PbIDD5* exhibited extremely high transcription levels in
398 fruit. Subsequently, we predicted the three-dimensional structures of *PbIDD3, 5* and *OsIDD2* (Supplementary figure S11). As
399 shown in Supplementary figure S11, the three-dimensional structures of *PbIDD3, 5* exhibit high similarity with that of *OsIDD2*,
400 indicating that these proteins might have similar functions and molecular mechanisms in regulating downstream gene expression.
401 We speculate that *PbIDD3*, *5* are mainly expressed in fruit and have similar functions to *OsIDD2* in regulating SCW formation in
402 pear fruit cells and inhibiting lignin biosynthesis by inhibiting the expression of the genes encoding key lignin synthesis enzymes.
403 In future studies, the overexpression and gene silencing of *PbIDD3, 5* in 'Dangshan Su' pear will be performed to determine their
404 functions, and the interactions between *PbIDD3, 5* and *PbCAD* will be explored through a yeast one-hybrid assay.

405 # CONCLUSION

406   We identified 68 IDD genes in five Rosaceae species (*P. bretschneideri*, *F. vesca*, *P. mume*, *R. occidentalis* and *P. avium*),
407 which we systematically assessed by bioinformatic analysis. According to the phylogenetic tree, the 68 IDD genes were divided
408 into 4 groups. In each class, we found that the structures of the genes and the compositions of the conserved motifs were very
409 similar. Through a series of bioinformatics analyses, we explained the possible evolutionary patterns of IDD genes in these
410 Rosaceae species. According to qRT-PCR, Chinese white pear IDD genes have high expression levels in fruit, flower and bud.
411 Further, *PbIDD2*, *8* are mainly expressed in the buds of Chinese white pear and are responsive to the induction of GA and
412 sucrose. That *PbIDD3*, *5* are may be involved in the regulation of SCW formation and lignin biosynthesis in Chinese white pear

413 fruit. All in all, this study reveal the basic information of the five Rosaceae species' IDD genes and predicts the potential
414 functions of some pear IDD proteins. These results will provide an important theoretical basis for improving the quality of
415 'Dangshan Su' pear.

ACKNOWLEDGEMENTS

417 The authors are deeply grateful to Prof. Yongping Cai and Prof. Yi Lin, who provided the sample used in the study and very
418 effective direction. The authors also thank Dr. Xu Sun, Dr. Muhammad Abdullah, Dr. Xi Cheng, Dr. Guohui Li, Dr. Yu Zhao for
419 providing valuable suggestions and comments. Finally, author would like to thank Dr. Meng for his help in revising the paper.

References

421 **Abdullah, M., Cao, Y. P., Cheng, X., Shakoor, A., Su, X. Q., Gao, J. S., Cai, Y. P.** (2018). Genome-Wide Analysis
422 Characterization and Evolution of SBP Genes in *Fragaria vesca*, *Pyrus bretschneideri*, *Prunus persica* and *Prunus mume*.
423 *FRONT. GENET.* **9**, 64.

424 **Adams, K. L., Wendel, J. F.** (2005). Polyploidy and genome evolution in plants. *Curr. Opin. Plant. Biol.* **8**, 135-141.

425 **Agarwal, P., Arora, R., Ray, S., Singh, A. K., Singh, V. P., Takatsuji, H., Kapoor, S., Tyagi, A. K.** (2007). Genome-wide
426 identification of C2H2 zinc-finger gene family in rice and their phylogeny and expression analysis. *PLANT. MOL. BIOL.* **65**,
427 467-485.

428 **Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., Castro, E., Duvaud, E., Flegel, V., Fortier, A.,**
429 **Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K.,**
430 **Redaschi, N., Rossier, G., Xenarios, I., Stockinger, H.** (2012). ExPASy: SIB bioinformatics resource portal. *NUCLEIC.*
431 *ACIDS. RES.* **40**, 597-603.

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

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

437 **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
438 pathway of lignin in stone cells of pear. *Sci Hortic-Amsterdam.* **125**, 374-379.

439 **Cao, Y. P., Han, Y. H., Li, D. H., Lin, Y., Cai, Y. P.** (2016). Systematic Analysis of the 4-Coumarate: Coenzyme A Ligase
440 (4CL) Related Genes and Expression Profiling during Fruit Development in the Chinese Pear. *GENES.* **7**, 89-106.

441 **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).
442 Arabidopsis C3H14 and C3H15 have overlapping roles in the regulation of secondary wall thickening and anther
443 development. *J. EXP. BOT.* **66**, 2595-2609.

444   **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
445   CCR and CAD gene families at the whole-genome level for lignin synthesis of stone cells in pear (Pyrus bretschneideri)
446   fruit. *BIOL. OPEN.* **6**, 1602-1613.

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

449   **Chen, J., Wang, J. N., Kuang, J. F., Shan, W., Chen, J. Y., Lu, W. J.** (2014). Molecular characterisation of a fruit ripening-
450   related INDETERMINATE DOMAIN transcription factor in banana (*Musa acuminata*). *J. HORTIC. SCI. BIOTECH.* **89**,
451   373-380.

452   **Coelho, C. P., Huang, P., Lee, D. Y., Brutnell, T. P.** (2018). Making Roots, Shoots, and Seeds: IDD Gene Family
453   Diversification in Plants. *TRENDS. PLANT. SCI.* **23,** 66-78.

454   **Colasanti, J., Tremblay, R., Wong, A. Y. M., Coneva, V., Kozaki, A., Mable, B. K.** (2006). The maize INDETERMINATE1
455   flowering time regulator defines a highly conserved zinc finger protein family in higher plants. *BMC. GENOMICS.* **7**, 158-
456   175.

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

459   **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,
460   IDD15, and IDD16 Cooperatively Regulate Lateral Organ Morphogenesis and Gravitropism by Promoting Auxin
461   Biosynthesis and Transport. *PLOS. GENET.* **9**, e1003759.

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

464   **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).
465   Phylogenetic analysis of IDD gene family and characterization of its expression in response to flower induction in Malus.
466   *MOL. GENET. GENOMICS.* **292**, 755-771.

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

470   **Galvao V. C., Horrer D., Kuttner F., Schmid M.** (2012). Spatial control of flowering by DELLA proteins in Arabidopsis
471   thaliana. *Development.* **139**, 4072–4082.

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

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

475   **Guo, H., Lee, T. H., Wang, X. Y., Paterson, A. H.** (2013). Function Relaxation Followed by Diversifying Selection after
476   Whole-Genome Duplication in Flowering Plants. *Plant. Physiol.* **162**, 769-778.

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

479   **Hamauzu, Y., Forest, F., Hiramatsu, K., Sugimoto, M.** (2007). Effect of pear (Pyrus communis, L.) procyanidins on gastric
480   lesions induced by HCl/ethanol in rats. *FOOD. CHEM.* **99**, 255-263.

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

483   **Hu, L. F., Liu, S. Q.** (2011). Genome-wide identification and phylogenetic analysis of the ERF gene family in cucumbers.
484   *GENET. MOL. BIOL.* **34**, 624-U201.

485   **Huang, L. J., Gao, W. Y., Li, X., Zhao, W. S., Huang, L. Q., Liu, C. X.** (2010). valuation of the in Vivo Anti-inflammatory
486   Effects of Extracts from Pyrus bretschneideri Rehd. *J. AGR. FOOD. CHEM.* **58**, 8983-8987.

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

490   **Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C.,**
491   **Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyere, C., Billault, A.,**
492   **Segurens, B., Gouyvenoux, M., Ugarte, E., Cattonaro, F., Anthouard, V., Vico, V ., Del Fabbro, C., Alaux, M., Di**
493   **Gaspero, G., Dumas, V., Felice, N., Paillard, S., Juman, I., Moroldo, M., Scalabrin, S., Canaguier, A., Le Clainche, I.,**
494   **Malacrida, G., Durand, E., Pesole, G., Laucou, V., Chatelet, P., Merdinoglu, D., Delledonne, M., Pezzotti, M.,**
495   **Lecharny, A., Scarpelli, C., Artiguenave, F., Pe, M., Valle, G., Morgante, M., Caboche, M., Adam-Blondon, A. F.,**
496   **Weissenbach, J., Quetier, F., Wincker, P.** (2007). The grapevine genome sequence suggests ancestral hexaploidization in
497   major angiosperm phyla. *Nature.* **449**, 463-U5.

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

500   **Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., Sternberg, M.** (2015). The Phyre2 web portal for protein modeling,
501   prediction and analysis. *NAT. PROTOC.* **10**, 845-858.

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

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

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

508   **Lin, Y. X., Jiang, H. Y., Chu, Z. X., Tang, X. L., Zhu, S. W., Cheng, B. J.** (2011). Genome-wide identification, classification
509   and analysis of heat shock transcription factor family in maize. *BMC. GENOMICS.* **12**, 76.

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

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

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

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

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

519 **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.,**
520   **Colasanti, J., An, G., Han, C. D.** (2008). Rice Indeterminate 1 (*OsId1*) is necessary for the expression of Ehd1 (Early
521   heading date 1) regardless of photoperiod. *PLANT. J.* **56**, 1018-1029.

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

524 **Paterson, A. H., Bowers, J. E., Chapman, B. A.** (2004). Ancient polyploidization predating divergence of the cereals, and its
525   consequences for comparative genomics. *P. NATL. ACAD. SCI. USA.* **101**, 9903-9908.

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

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

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

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

535 **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
536   *JcZFP8*, a C2H2 zinc finger protein gene from Jatropha curcas L. *GENE.* **658**, 47-53.

537 **Shirasawa, K., Isuzugawa, K., Ikenaga, M., Saito, Y., Yamamoto, T., Hirakawa, H., Isobe, S.** (2017). The genome sequence
538   of sweet cherry (*Prunus avium*) for use in genomics-assisted breeding. *DNA. RES.* **24**, 499-508.

539 **Shulaev, V., Sargent, D. J., Crowhurst, R. N., Mockler, T. C., Folkerts, O., Delcher, A. L., Jaiswal, P., Mockaitis, K.,**
540   **Liston, A., Mane, S. P., Burns, P., Davis, T. M., Slovin, J. P., Bassil, N., Hellens, R. P., Evans, C., Harkins, T., Kodira,**
541   **C., Desany, B., Crasta, O. R., Jensen, R. V., Allan, A. C., Michael, T. P., Setubal, J. C., Celton, J. M., Rees, D. J. G.,**

542   **Williams, K. P., Holt, S. H., Ruiz Rojas, J. J. R., Chatterjee, M., Liu, B., Silva, H., Meisel, L., Adato, A., Filichkin, S.**
543   **A., Troggio, M., Viola, R., Ashman, T. L., Wang, H., Dharmawardhana, P., Elser, J., Raja, R., Priest, H. D., Bryant,**
544   **D. W., Fox, S. E., Givan, S. A., Wilhelm, L. J., Naithani, S., Christoffels, A., Salama, D., Carter, J., Girona, E. L.,**
545   **Zdepski, A., Wang, W., Kerstetter, R. A., Schwab, W., Korban, S. S., Davik, J., Monfort, A., Denoyes-Rothan, B.,**
546   **Arus, P., Mittler, R., Flinn, B., Aharoni, A., Bennetzen, J., Salzberg, S. L., Dickerman, A. W., Velasco, R.,**
547   **Borodovsky, M., Veilleux, R. E., Folta, K. M.** (2011). The genome of woodland strawberry (*Fragaria vesca*). *NAT.*
548   *GENET.* **43**, 109-116.

549   **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).
550   Comparative genomic analysis of the PKS genes in five species and expression analysis in upland cotton. *PEERJ.* **5**, e3974.

551   **Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R.,**
552   **Ploetz, L., Radenbaugh, A., Singh, S., Swing, V., Tissier, C., Zhang, P., Huala, E.** (2008). The Arabidopsis Information
553   Resource (TAIR): gene structure and function annotation. *NUCLEIC. ACIDS. RES.* **36**, 1009-1014.

554   **Tang, H. B., Bowers, J. E., Wang, X. Y., Ming, R., Alam, M., Paterson, A. H.** (2008). Synteny and Collinearity in Plant
555   Genomes. *Science.* **320**, 486-488.

556   **Tang, R. M., Zhu, W. J., Song, X. Y., Lin, X. Z., Cai, J. H., Wang, M., Yang, Q.** (2016). Genome-Wide Identification and
557   Function Analyses of Heat Shock Transcription Factors in Potato. *FRONT. PLANT. SCI.* **7**, 490.

558   **Tavaud, M., Zanetto, A., David, J. L., Laigret, F., Dirlewanger, E.** (2004). Genetic relationships between diploid and
559   allotetraploid cherry species (*Prunus avium*, *Prunus x gondouinii* and *Prunus cerasus*). *HEREDITY.* **93**, 631-638.

560   **Tian, Z. D., Zhang, Y., Liu, J., Xie, C. H.,** (2010). Novel potato C2H2-type zinc finger protein gene, *StZFP1*, which responds to
561   biotic and abiotic stress, plays a role in salt tolerance. *PLANT. BIOLOGY.* **12**, 689-697.

562   **VanBuren, R., Bryant, D., Bushakra, J. M., Vining, K. J., Edger, P. P., Rowley, E. R., Priest, H. D., Michael, T. P., Lyons,**
563   **E., Filichkin, S. A., Dossett, M., Finn, C. E., Bassil, N. V., Mockler, T. C.** (2016). The genome of black raspberry (*Rubus*
564   *occidentalis*). *PLANT. J.* **87**, 535-547.

565   **Vavilov, N.I.** (1951). The Origin, Variation, Immunity and Breeding of Cultivated Plants. *Soil Science*, **72**, 482.

566   **Wei, K., Pan, S., Li, Y.** (2016). Functional Characterization of Maize C2H2, Zinc-Finger Gene Family. *PLANT. MOL. BIOL.*
567   *REP.* **34**, 761-776.

568   **Wong, A. Y., Colasanti, J.** (2007). Maize floral regulator protein INDETERMINATE1 is localized to developing leaves and is
569   not altered by light or the sink/source transition. *J. EXP. BOT.* **58**, 403-414.

570   **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
571   zinc finger transcription factor, acts as a master switch from vegetative to floral development in rice. *P. NATL. ACAD. SCI.*
572   *USA.* **105**, 12915-12920.

573   **Wu, J., Wang, Z., Shi, Z., Zhang, S., Ming, R., Zhu, S., Khan, M. A., Tao, S., Korban, S. S., Wang, H., Chen, N. J., Nishio,**
574   **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.,**

575   **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,**
576   **X., Zhang, H., Zeng, L., Zheng, D., Wang, C., Chen, M., Wang, G., Xie, L., Sovero, V., Sha, S., Huang, W., Zhang, S.,**
577   **Zhang, M., Sun, J., Xu, L., Li, Y., Liu, X., Li, Q., Shen, J., Wang, J., Paull, R. E., Bennetzen, J. L., Wang, J., Zhang, S.**
578   (2013). The genome of the pear (Pyrus bretschneideri Rehd). *Genome. Res.* 23, 396–408.

579   **Wu, T., Zhang, R. P., Gu, C., Wu, J. Y., Wan, H. J., Zhang, S. J., Zhang, S. L.** (2010). Evaluation of candidate reference
580   genes for real time quantitative PCR normalization in pear fruit. *Afr. J. Agric. Res.* **7**, 3701-3704.

581   **Xue, C., Yao, J. L., Qin, M. F., Zhang, M. Y., Allan, A. C., Wang, D. Fu., Wu, J.** (2018). *PbrmiR397a* regulates lignification
582   during stone cell development in pear fruit. *PLANT. BIOTECHNOL. J.* **17**. 103-117.

583   **Yan, C. C., Yin, M., Zhang, N., Jin, Q., Fang, Z., Lin, Y., Cai, Y.P.** (2014). Stone cell distribution and lignin structure in
584   various pear varieties. *SCI. HORTIC-AMSTERDAM.* **174**, 142-150.

585   **Yi, G., Neelakandan, A. K., Gontarek, B. C., Vollbrecht, E., Becraft, P. W.** (2015). The naked endosperm Genes Encode
586   Duplicate INDETERMINATE Domain Transcription Factors Required for Maize Endosperm Cell Patterning and
587   Differentiation. *PLANT. PHYSIOL.* **167**, 443-456.

588   **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
589   Uncanonical CCCH-Tandem Zinc-Finger Protein Represses Secondary Wall Synthesis and Controls Mechanical Strength in
590   Rice. *MOL. PLANT.* **11**, 163-174.

591   **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).
592   Comparison of the transcriptomic analysis between two Chinese white pear (Pyrus bretschneideri Rehd.) genotypes of
593   different stone cells contents. *PLOS. ONE.* **12**, e0187114.

594   **Zhang, Q., Chen, W., Sun, L., Zhao, F., Huang, B., Yang, W., Tao, Y., Wang, J., Yuan, Z., Fan, G., Xing, Z., Han, C., Pan,**
595   **H., Zhong, X., Shi, W., Liang, X., Du, D., Sun, F., Xu, Z., Hao, R., Lv, T., Lv, Y., Zheng, Z., Sun, M., Luo, L., Cai, M.,**
596   **Gao, Y., Wang, J., Yin, Y., Xu, X., Cheng, T., Wang, J.** (2012). The genome of Prunus mume. *Nat. Commun.* **3**, 1318.

597   **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
598   exogenous GA(3) and its inhibitor paclobutrazol on floral formation, endogenous hormones, and flowering-associated genes
599   in 'Fuji' apple (*Malus domestica Borkh.*). *PLANT. PHYSIOL. BIOCH.* **107**, 178-186.

600   **Zhu, G. R., Chen, G. X., Zhu, J. T., Zhu, Y., Lu, X. B., Li, X. H., Hu, Y. K., Yan, Y. M.** (2015). Molecular Characterization
601   and Expression Profiling of NAC Transcription Factors in Brachypodium distachyon L. *PLOS. ONE.* **10**, e0139794.

602

603   **Table 1.** The IDD genes of *Pyrus bretschneideri* identified in this study are listed.

604   **Figure 1.** Phylogenetic relationships and subfamily designations in IDD proteins from *Pyrus bretschneideri*, *Fragaria vesca*,
605   *Prunus mume*, *Rubus occidentalis*, *Prunus avium*, *Malus domestica*, *Arabidopsis thaliana*, *Oryza sativa* and *Zea mays*. Groups 1-
606   4 are represented by shades of red, green, blue and cyan, respectively.

607   **Figure 2.** Predicted *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis* and *Prunus avium* IDD protein

608 conserved motifs and exon-intron structures. (A) Gene structures of the IDD genes. Black wedge indicates exon, black line
609 indicates intron and red wedge indicates UTR. (B) Distribution of MEME motifs in IDD genes. (C) The colour and
610 corresponding number of each motif box.

611 **Figure 3.** Conserved ID domain composition. All 68 IDD genes had a characteristic ID domain. The ID domain consists of four
612 zinc fingers (Z1, Z2, Z3 and Z4). Alignment analysis of the 68 IDD gene sequences using the ClustalW tool in MEGA 7.0
613 software. These domain diagrams were plotted using the online WebLogo tool.

614 **Figure 4.** Chromosomal locations of IDD genes in *Pyrus bretschneideri* (A), *Fragaria vesca* (B), *Rubus occidentalis* (C), *Prunus*
615 *mume* (D) and *Prunus avium* (E). Duplicated gene pairs are connected with coloured lines.

616 **Figure 5.** Microsynteny of regions among *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis*, *Prunus*
617 *avium* and *Malus domestica*. The chromosome numbers are indicated by differently coloured boxes and are labelled by Pb, Fv,
618 Pm, Ro, Pa, and Md. The differently coloured boxes also represent the sequence lengths of chromosomes in megabases. The
619 black line indicates the syntenic relationship among the IDD regions.

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

621 **Figure 7.** Expression patterns of IDD genes of Chinese white pear in different organs and in fruit at different developmental
622 stages. Expression patterns of IDD genes in Chinese white pear in different organs (A-O). Expression patterns of IDD genes in
623 Chinese white pear at different developmental stages (P-DD). *significant difference at P < 0.05, **significant difference at P <
624 0.01.

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

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

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

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

641 **Supplemental Figure S5.** *Prunus avium* IDD protein sequence alignment. Black underline indicates zinc finger domain (Z1, Z2,
642 Z3 and Z4). Red triangle indicates a conserved C residue, and blue triangle indicates a conserved H residue. The yellow underline

643   indicates the NLS sequence in the N-terminal region of the IDD gene. Green box means the MSATALLQKAA domain, and

644   purple box indicates the TRDFLG domain.

645   **Supplemental Figure S6.** N-terminal region of the ID-domain shows the putative NLS sequence. The yellow underline indicates

646   the NLS sequence in the N-terminal region of the IDD gene.

647   **Supplemental Figure S7.** Sliding window plots of duplicated IDD genes in Chinese white pear. The grey shaded portion

648   indicates conserved ID domain. The X-axis indicates the synonymous distance within each gene.

649   **Supplemental Figure S8.** A phylogenetic tree of the 12 species genomes (A). Phylogenetic relationships and subfamily

650   designations in IDD proteins from 12 species (B).

651   **Supplemental Figure S9.** Expression modes of candidate *PbIDD2*, *5*, *6*, *8*, *9*, *12*, *16* in Chinese white Pear buds treated with

652   gibberellin (A-G) and sucrose (H-N). *significant difference at P < 0.05, **significant difference at P < 0.01.

653   **Supplemental Figure S10.** A hypothetical evolutionary model map of IDD genes.

654   **Supplemental Figure S11.** Predicted three-dimensional structures of *PbIDD3, 5* and *OsIDD2*. *OsIDD2* have been proven to be

655   responsible for SCW formation and lignin biosynthesis.

656   **Supplementary Table S1.** Gene sequence involved in the article.

657   **Supplemental Table S2.** The IDD genes of *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis* and *Prunus avium* identified in

658   this study are listed.

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

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

661   **Supplemental Table S5.** Synteny data in *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis*, *Prunus*

662   *avium*, *Malus domestica*.

663   **Supplementary Table S6.** Promoter sequence of 16 IDD genes in Chinese white pear.

664   **Supplementary Table S7.** Complete information on *C*is-acting elements of the 16 *PbIDD* genes.

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

Basic Information of IDD Gene in *Pyrus bretschneideri.*

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

Figure 1

Phylogenetic relationships and subfamily designations in IDD proteins from *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis*, *Prunus avium*, *Malus domestica*, *Arabidopsis thaliana*, Oryza sativ an

Phylogenetic relationships and subfamily designations in IDD proteins from *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis*, *Prunus avium*, *Malus domestica*, *Arabidopsis thaliana*, *Oryza sativa* and *Zea mays*. Groups 1-4 are represented by shades of red, green, blue and cyan, respectively.
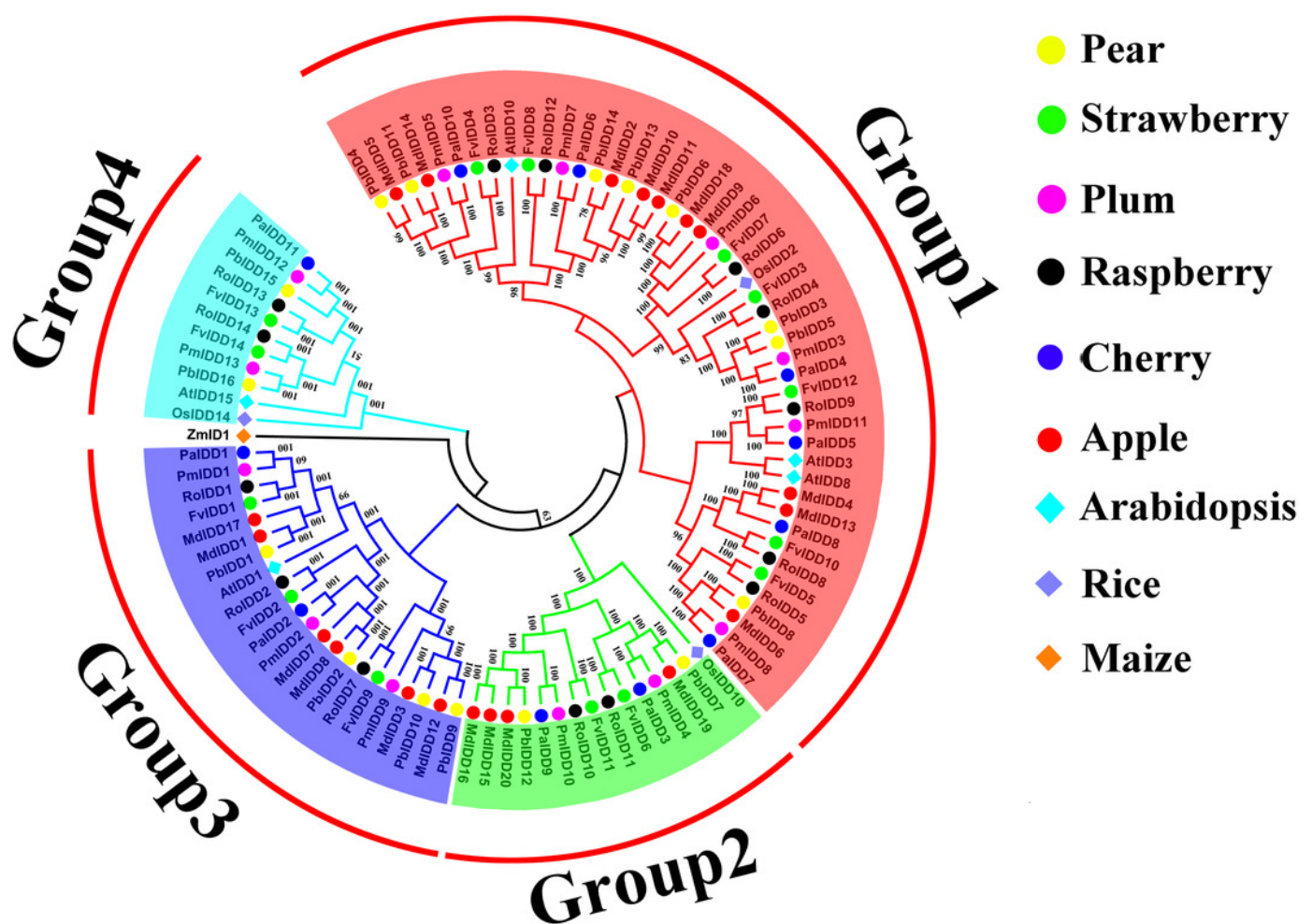
Figure 2

Predicted *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis* and *Prunus avium* IDD protein conserved motifs and exon-intron structures.

Predicted *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis* and *Prunus avium* 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.
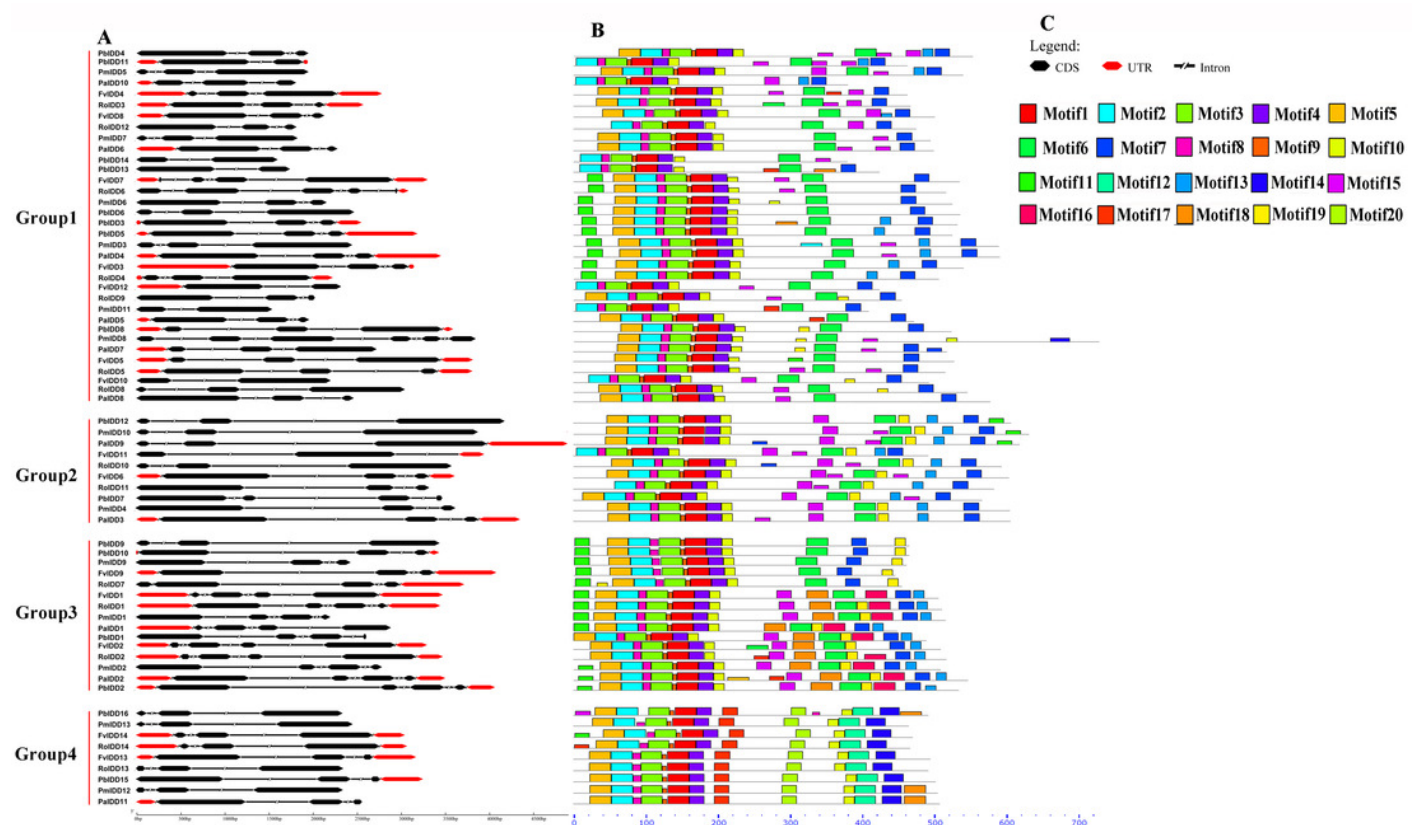
Figure 3

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.
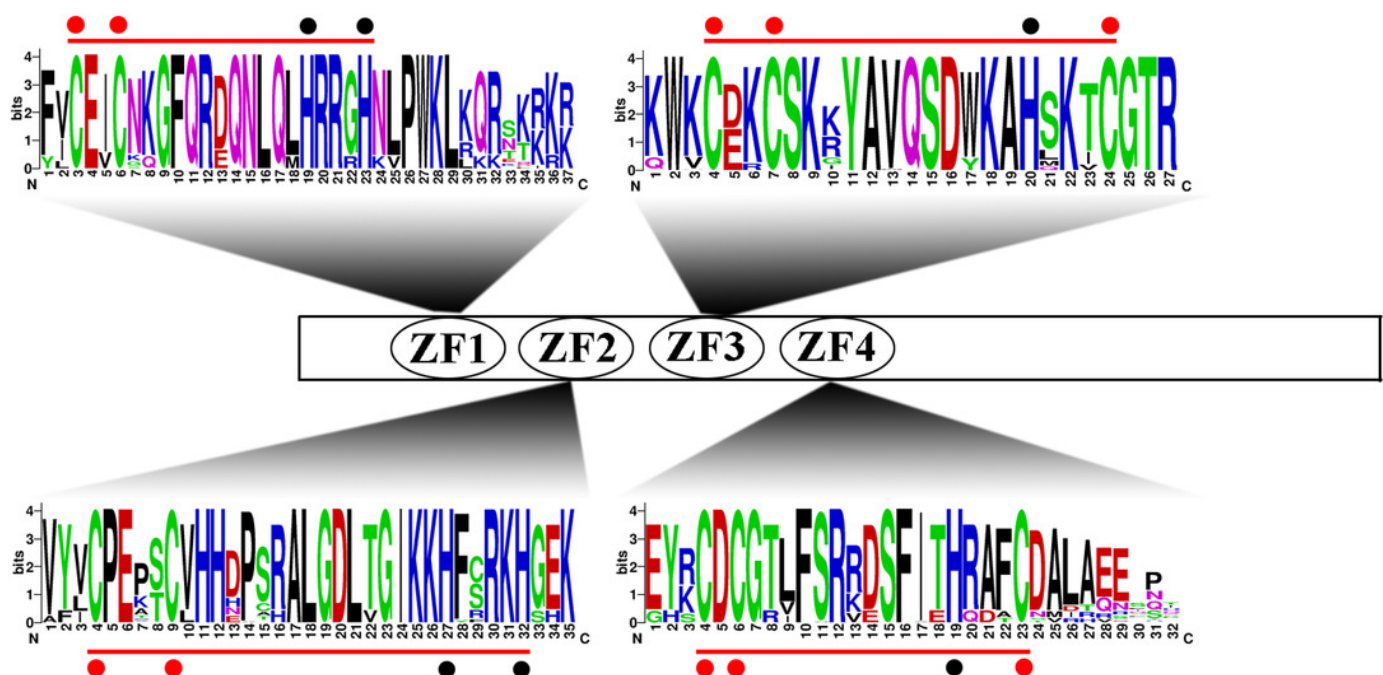
Figure 4

Chromosomal locations of five Rosaceae species.

Chromosomal locations of IDD genes in *Pyrus bretschneideri* (A), *Fragaria vesca* (B), *Rubus occidentalis* (C), *Prunus mume* (D) and *Prunus avium* (E). Duplicated gene pairs are connected with coloured lines.
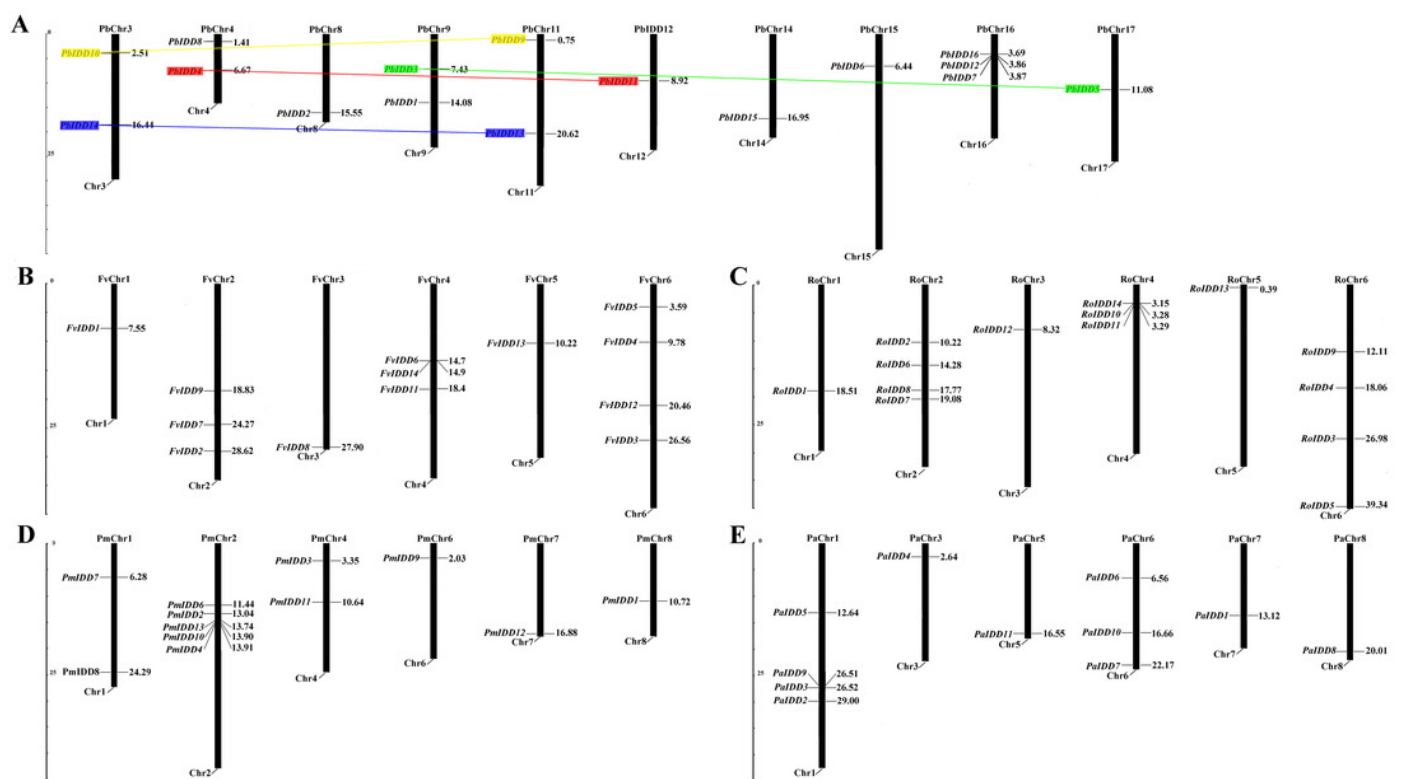
Figure 5

Microsynteny of regions among *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis*, *Prunus avium* and *Malus domestica*.

Microsynteny of regions among *Pyrus bretschneideri*, *Fragaria vesca*, *Prunus mume*, *Rubus occidentalis*, *Prunus avium* and *Malus domestica*. 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.
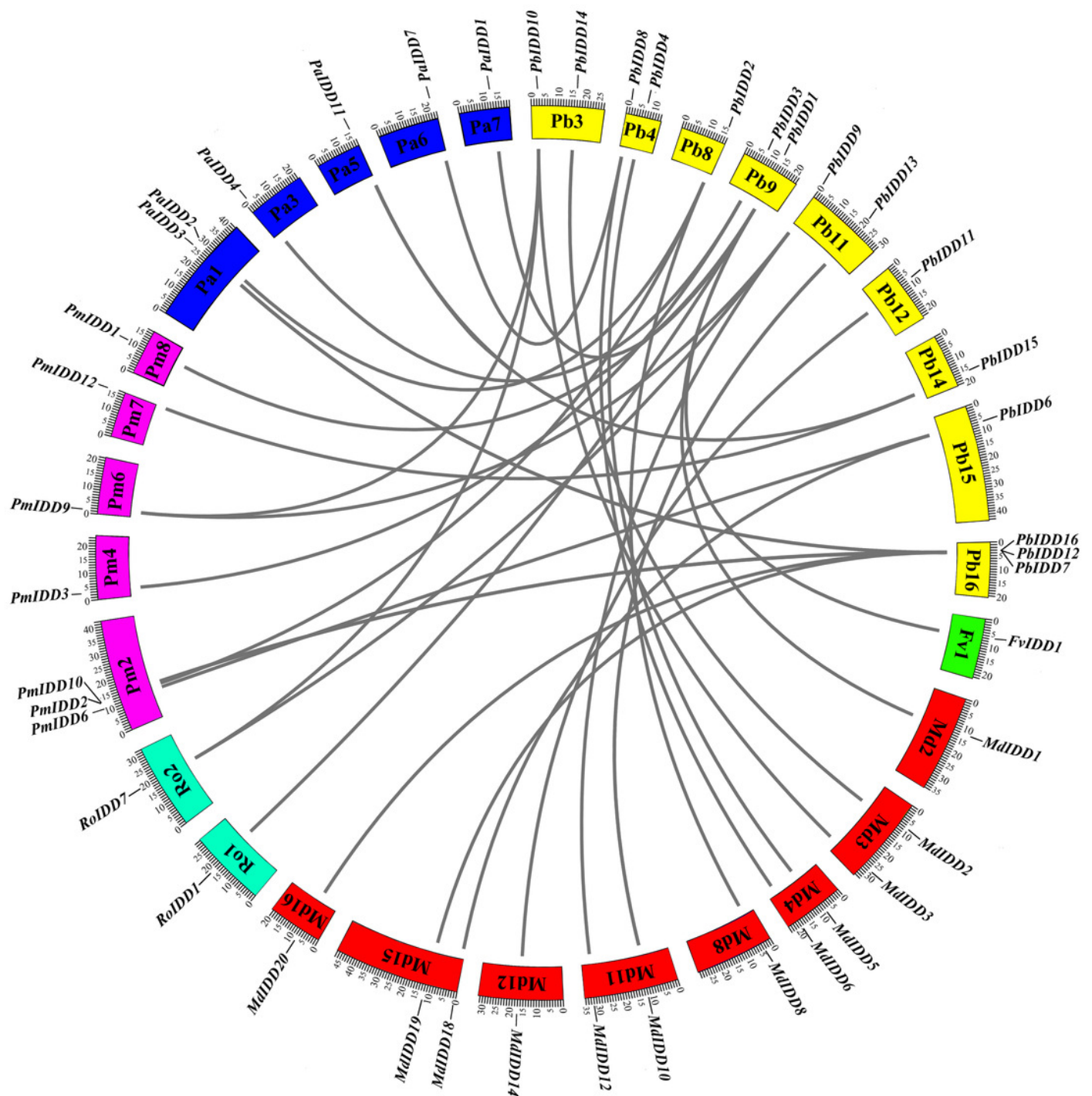
Figure 6

Promoter *C*is-elements of the 16 *PbIDDs.*

Potential *C*is-elements in the 5' regulatory sequences of the 16 *PbIDDs.*

Figure 7

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

Expression patterns of IDD genes of Chinese white pear in different organs and in fruit at different developmental stages. Expression patterns of IDD genes in Chinese white pear in different organs (A-O). Expression patterns of IDD genes in Chinese white pear at different developmental stages (P-DD). *significant difference at P < 0.05, **significant difference at P < 0.01.