

# Recurrent single-gene duplication drives the expansion and expression diversification of the ADH gene family in pear and other Rosaceae species

Shaoling Zhang (✉ [slzhang@njau.edu.cn](mailto:slzhang@njau.edu.cn))

Nanjing Agricultural University

Weiwei Zeng

Nanjing Agricultural University

Xin Qiao

Nanjing Agricultural University

Qionghou Li

Nanjing Agricultural University

Chunxin Liu

Nanjing Agricultural University

Jun Wu

Nanjing Agricultural University

Hao Yin

Nanjing Agricultural University

---

## Research article

**Keywords:** Alcohol dehydrogenase, Pear, Rosaceae, Evolution, Expression pattern

**Posted Date:** October 30th, 2019

**DOI:** <https://doi.org/10.21203/rs.2.16577/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## Abstract

**Background** Alcohol dehydrogenases (ADHs) are essential to plant growth and the formation of aromatic compounds in fruits. However, the evolutionary history and characteristics of ADH gene expression remain largely unclear in Chinese white pear (*Pyrus bretschneideri*) and other fruit species from the family Rosaceae.

**Results** In this study, 464 ADH genes were identified in eight Rosaceae fruit species and 68 of the genes were from pear. Based on the analyses of phylogeny and conserved motifs, the pear ADH genes were classified into four subgroups (I, II, III, and IV). The chromosomal distribution of the genes was found to be uneven and numerous clusters of physically linked ADH genes were detected. Frequent single-gene duplication events were found to have contributed to the formation of ADH gene clusters and the expansion of the ADH gene family in these eight Rosaceae species. Purifying selection was the major force in ADH gene evolution. The younger genes derived from tandem and proximal duplications had evolved faster than those that derived from other types of duplication. RNA-sequencing and quantitative-real time-PCR analysis revealed that the expression levels of three ADH genes were closely correlated with the content of aromatic compounds that are found during fruit development.

**Conclusion** Comprehensive analyses were conducted in eight Rosaceae species and 464 ADH genes were identified. The results of this study provide new insights into the evolution and expression characteristics of ADH family genes in pear and other Rosaceae species.

## Background

Alcohol dehydrogenases (ADHs, EC 1.1.1.1) are members of the dehydrogenase superfamily and are found in a broad range of organisms [1–4]. For example, ADH has been investigated in humans, animals, bacteria and yeast [4–10]. The ADH enzyme is encoded by a multigene family in eukaryotes and prokaryotes and catalyzes reciprocal transformations between alcohols and aldehydes. ADH genes can be divided into three types including those that encode short- (~250 residues), medium- (~350 residues) and long-chain or Iron-ADH genes super-families (600–750 residues or approximately 385–900 residues) [1, 4, 11]. In plants, most ADHs contain zinc ligands and belong to the subfamily of medium-chain proteins [1]. These enzymes are widely involved in metabolic processes and have a positive function in improving resistance to biotic and abiotic stresses [12–18]. In addition, ADH also plays a vital role in catalyzing the synthesis of aromatic substances in plants, which involves the selective conversion of short linear alcohols and aldehydes to synthesize aromatic precursors. The process is regulated by amines such as ethylene [19–21]. *ADH* or *ADH*-like genes have been identified in certain plants such as mango (*Mangifera indica*), Arabidopsis (*Arabidopsis thaliana*), melon (*Cucumis melo*), tomato (*Solanum lycopersicum*), and grape (*Vitis vinifera*) [22–26]. However, information on *ADH* genes in Rosaceae species is limited.

Pear is one of the most widely grown commercial fruits in the global market. These fruits are cultivated in all temperate regions worldwide. The content of aromatic compounds is important to the quality of flavor in the pear and, hence, improving the aromatic content is a direct method of improving flavor. Fruits often contain more than 1,000 volatile compounds and the main aromatic components in pear include esters, alcohols, aldehydes, ketones, lactones, and terpenoids [27–30]. Alcohol and aldehyde substances form primary components of aromas in fruits and so these compounds are vital to fruit quality traits. The production of aromatic compounds has been investigated superficially in grape, apple, tomato, apricot, and peach [19, 26, 31–35] but fewer genome-wide annotation and evolutionary studies of *ADH* genes have been performed in pear or other Rosaceae fruit species. At present, the genome sequences of eight Rosaceae fruit species have been released and they include those of the Chinese white pear (*Pyrus bretschneideri* Rehd.) [36] and seven other Rosaceae species: apple (*Malus domestica*) [37], peach (*Prunus persica*) [38], sweet cherry (*Prunus avium*) [39], black raspberry (*Rubus occidentalis*) [40], strawberry (*Fragaria vesca*) [41, 42], Japanese apricot (*Prunus mume*) [43], and European pear (*Pyrus communis*) [44, 45]. These genomic resources lay a foundation for performing comparative analyses of the *ADH* gene family among different Rosaceae fruit species. In this study, we identified members of the *ADH* gene family in pear and the seven other Rosaceae fruit species, as described above, and unraveled the evolutionary history of *ADH* family genes based on comprehensive analyses of phylogeny, conserved domains, selective pressures, syntenic relationships, and gene duplication events. Moreover, we investigated expression patterns of *ADH* family genes based on transcriptome data from different pear tissues and quantitative-real time (qRT)-PCR analysis. Several candidate genes closely associated with alcohol and aldehyde biosynthesis were identified using correlation analysis of alcohol and aldehyde content changes and gene expression profiles. The results of this study provide insights into the evolution and functional roles of the *ADH* gene family.

## Results

### Non-random chromosomal distribution of *ADH* genes in pear and seven other Rosaceae species

A Hidden Markov Model (HMM) was used to identify *ADH* family genes in pear and seven other Rosaceae fruit species. A total of 464 *ADH* genes were identified, of which 68 were identified in Chinese white pear, 68 in European pear, 82 in apple, 68 in peach, 37 in strawberry, 61 in Japanese apricot, 37 in black raspberry, and 43 in sweet cherry (Fig. 1). Lineage-specific whole-genome duplication (WGD) was found to have occurred in the ancestor of Chinese white pear, European pear, and apple, which may have resulted in the higher number of *ADH* genes in these three species than in strawberry and black raspberry (Fig. 1). Indeed, we found more *ADH* genes located in syntenic blocks between apple (11 syntenic pairs) and pear (seven syntenic pairs). However, the number of *ADH* genes in peach and Japanese apricot is similar to that in apple and pear, although peach and Japanese apricot have not experienced a recent WGD.

To determine the mechanism for the expansion of *ADH* genes in peach and Japanese apricot in the absence of recent genome duplication, we investigated the chromosomal distribution of *ADH* genes in each species. The distribution of *ADH* genes is uneven on different chromosomes and homologous gene clusters were observed more often in each investigated species (Fig. 2). In peach, we found several gene clusters of homologous *ADH* genes on Chr3, Chr6, and Chr8 and the cluster size ranged from 3 to 7. A total of 22 genes were identified on Chr1, which is the highest number of *ADH* genes on all 17 chromosomes of apple. In Japanese apricot, eight *ADH* gene clusters were located on Chr1, Chr2, Chr4, and Chr6 and, as in peach, the size of the gene clusters ranged from 3 to 7. It is worth noting that a strong syntenic relationship was found between two gene clusters located on Chr1 and Chr7 in pear and apple. The explosion of gene clusters may account for the expansion of the *ADH* gene family in Rosaceae species.

## Single-gene duplication largely contributed to the expansion of the *ADH* gene family in Rosaceae species.

In addition to WGD, single-gene duplication events including tandem, proximal, transposed, and dispersed duplications also played important roles in the formation of local gene clusters and gene family expansion. We performed genome-wide identification of different modes of gene duplication in each of the eight Rosaceae species. To infer the evolutionary origins of *ADH* family genes, we searched for different types of duplicated gene pairs that contained *ADH* genes and classified them into five modes of gene duplication (Fig. 3, Table S2). We found that 90.5% of *ADH* genes were derived from single-gene duplications in pear, while 86.0% were derived in this way in apple, 98.9% in peach, 96.8% in strawberry, 98.3% in Japanese apricot, 97.7% in sweet cherry, 97.3% in black raspberry, and 96.3% in European pear. In contrast, only 1.1–14.0% of *ADH* genes were derived from WGD in the eight Rosaceae species and a relatively higher proportion of WGD-derived *ADH* genes were found in pear (9.5%) and apple (14.0%) due to lineage-specific genome duplication. *ADH* genes experienced a high frequency of tandem (12.2–20.2%) and proximal duplications (0–14.6%) in each of the investigated species, which contributed to the formation of the *ADH* gene clusters we have observed. In addition, dispersed duplications account for the highest number of derived genes (53.0–64.04%) in all the species investigated. However, the mechanism underlying dispersed gene duplication remains unclear.

## Recurrent tandem and proximal duplication occurred following whole-genome duplication

The  $K_s$  (synonymous substitutions per site) value is usually used to estimate the evolutionary dates of genome or gene duplication events [46, 47]. Here, the  $K_s$  value was estimated for each gene pair (Fig. 4 and Table S3). In pear, two WGD events were detected including an ancient WGD event which corresponds to the paleo-hexaploidization ( $\gamma$ ) event shared by core eudicots that took place ~140 Mya ( $K_s$  ~1.5–1.8) [37] and the recent WGD that is inferred to have occurred 30–45 Mya ( $K_s$  ~0.15–0.3) [36].

The  $K_s$  values of the majority of WGD-derived *ADH* gene pairs ranged from 0.14 to 0.28, suggesting that these genes may descend from the more recent WGD event. It is notable that the  $K_s$  values of *ADH* gene pairs derived from tandem and proximal duplications are much lower than those derived from the WGD, except for pear, suggesting that *ADH* genes experienced frequent small-scale gene duplications after the ancient or recent genome duplication events and are younger in age. *ADH* gene pairs derived from transposed duplications have high  $K_s$  values and the median  $K_s$  distribution is close to that of WGD-derived gene pairs. *ADH* gene pairs derived from WGD have higher  $K_s$  values in peach, sweet cherry, Japanese apricot, black raspberry, and strawberry than in pear and apple, suggesting that these *ADH* genes have been retained from the ancient eudicot  $\gamma$  duplication event.

## ***ADH* genes evolved under strong purifying selection**

$K_s$ ,  $K_a$ , and  $K_a/K_s$  values of paralogous *ADH* gene pairs in eight Rosaceae species were estimated for each gene pair (Fig. 4 and Table S3). The  $K_a/K_s$  ratio has been widely used as an index for measuring the strength and direction of selection pressure.  $K_a/K_s > 1$  indicated positive selection;  $K_a/K_s = 1$  demonstrated neutral evolution; and  $K_a/K_s < 1$  suggested negative (or purifying) selection [48]. Purifying selection can eliminate deleterious mutations and positive (Darwinian) selection can induce and fix advantageous mutations [49]. Our calculated  $K_a/K_s$  ratios for all paralogous *ADH* gene pairs in each species showed that they were less than one, indicating that purifying selection was the main force behind *ADH* family gene evolution in Rosaceae species. *ADH* genes derived from tandem and proximal duplications showed high  $K_a/K_s$  ratios in the species investigated, suggesting that these genes evolved at a faster evolutionary rate, which is a feature of new genes. In summary, the results from this and the aforementioned analysis supported the hypothesis that tandem and proximal duplications occurred more recently generated new *ADH* genes that contributed to the formation of homologous *ADH* gene clusters and supplied the expansion of the *ADH* gene family (Fig. 5 and Table S4).

## **Microsyntenic relationships among orthologous *ADH* genes from eight Rosaceae species**

In this study, we identified syntenic blocks among eight Rosaceae species by performing interspecies syntenic analysis. Nine *ADH* genes in pear were found to have orthologous syntenic genes in seven other species, while 16 had orthologous syntenic genes in one of the other seven species (Table S5). Surprisingly, good collinearity was detected among the eight Rosaceae species to the nine pear *ADH* genes, even after speciation and long-term evolution, which suggested that these genes originated before diversification of the Rosaceae species and may have conserved functional roles. For example, the orthologous syntenic gene cluster for pear *ADH* gene *Pbr012701.1* includes *MD07G1250800* (apple), *PCP001961.1* (European pear), *Pav\_sc0001102.1\_g440.1.mk* (sweet cherry), *Prupe.2G274500* (peach), *Pm019522* (Japanese apricot), *Ro07\_G12285* (black raspberry), and *FvH4\_7g27140.1* (strawberry). The genomic regions around *Pbr012701.1* also showed strong syntenic relationships with their counterparts

in the other seven Rosaceae species (Fig. 6 and Table S6). It is noteworthy that the directions of the *ADH* and surrounding genes within a distance of 500 kb on Chr5 in Japanese apricot were the inverse of the seven other species, which indicated that a chromosomal inversion occurred after the divergence of Japanese apricot from other Rosaceae species. In addition, seven *ADH* genes in pear were found to have no syntenic counterpart in the other seven species, suggesting that these genes were newly duplicated in the pear genome after the divergence of pear and apple. We found that five of seven *ADH* genes were derived from tandem or proximal duplications. This result supports the aforementioned finding that recent small-scale gene duplications have been important in the expansion of *ADH* gene family.

## Phylogenetic analysis of *ADH* family genes in pear

Based on phylogenetic analysis and conserved motif analysis, 68 *ADH* family genes in pear were classified into four subgroups (Fig. 7). Group-I and Group-II both contain 14 genes, Group-III contains 18 genes, and Group-IV contains 22 genes.

The Multiple EM for Motif Elicitation (MEME 5.05) motif search tool was used to predict conserved domains in *ADH* protein sequences (Fig. 2 and Table S7). The type and distribution of conserved motifs of *ADH* genes was similar within subgroups, supporting the classification results of the phylogenetic analysis. A total of 19 motifs were detected in all the *ADH* genes and the number of motifs contained in *ADH* protein sequences varied in different subgroups. The Group IV genes encoded more conserved motifs than other subgroups, while genes of Group III encoded fewer motifs. Motifs 1, 7, 2, and 9 were detected in almost all of the *ADH* genes, whereas motif 15 was only detected in group II; motif 17 was only detected in group III; and motif 4 was present only in group IV. Motifs 1, 3, 6, 7, 16, and 19 corresponded to the ADH\_N domain and motifs 8, 9, 11, 17, and 18 were identified in the ADH\_zinc\_N domain. In clade I, motifs 1, 3, and 19 represented the ADH\_N domain in all genes, except for Pbr033945.1, and motifs 9, 11, and 18 related to the ADH\_zinc\_N domain and were found in all *ADH* genes. In clade II, motifs 1 and 3 were identified in the ADH\_N domain and motifs 8 and 9 were identified in the ADH\_zinc\_N domain. Motifs 7, 6, and 16 were represented in the ADH\_N domain and motifs 17 and 19 were represented in the ADH\_zinc\_N domain and were found only in clade III.

The structures of *ADH* genes in pear were also compared among the different subgroups (Fig. S1). The number of exons varied from 2 to 13 and 21 genes were annotated in the untranslated region. The exon/intron structures varied in different subgroups. Group I members contained more exons and introns than other subgroups, whereas group II members contained fewer exons and introns than other subgroups. The exon/intron structures were similar between members of group III and group IV. We further investigated the genetic features of 68 *ADH* genes identified in Chinese white pear, including the CDS length, MW, and PI (Table S8). The lengths of the CDSs ranged from 900 to 2,261, the peptide length of *ADH* proteins ranged from 312 to 887 amino acids, the PI value ranged from 5.33 to 9.32, and the MW ranged from 32.32 to 69.83 kDa. Different subgroups showed distinct gene features (Fig. S2). For example, group I members had the longest CDSs, whereas group III members had the shortest. The PI

values varied greatly in each of the subgroups. Members of group III had the highest average PI values, whereas those of group I were the lowest.

## Transcriptome expression profiles and qRT-PCR analysis of *ADH* genes in different pear tissues

Based on the transcriptome data from pear fruits, pollen, leaves, petals, sepals, ovaries, stems, and buds, we investigated the expression patterns of the *ADH* gene family (Fig. 8 and Table S9). Transcripts per million (TPM) values were used to measure the gene expression level. We found that 12 *ADH* genes showed low or no expression in all investigated pear tissues. We further investigated the expression patterns of paralogous genes that corresponded to each of these 12 *ADH* genes. The results showed that 5 out of 12 *ADH* genes have diverged expression patterns compared to their highly expressed paralogous genes, suggesting that these five *ADH* genes (*Pbr024187.1*, *Pbr004679.1*, *Pbr003626.1*, *Pbr040240.1*, *Pbr040236.1*) may have undergone pseudogenization.

Seven *ADH* genes (*Pbr021220.1*, *Pbr039379.1*, *Pbr032936.1*, *Pbr022043.1*, *Pbr032775.1*, *Pbr000293.1*, and *Pbr042656.1*) are highly expressed in the four development stages of fruit, implying that these genes may play important roles during fruit development and ripening. The expression levels of four of the genes (*Pbr032936.1*, *Pbr032775.1*, *Pbr000293.1*, and *Pbr042656.1*) gradually increased during fruit development, whereas those of the two genes (*Pbr021220.1*, *Pbr022043.1*) decreased. The expression level of *Pbr039379.1* decreased at early stages and increased at later stages. Most *ADH* genes presented no or very low expression levels in pollen. Moreover, two *ADH* genes (*Pbr021220.1* and *Pbr039378.1*) showed high expression levels in leaf tissue. Six *ADH* genes (*Pbr022043.1*, *Pbr013915.1*, *Pbr032774.1*, *Pbr032777.1*, *Pbr013913.1*, and *Pbr000293.1*) were highly expressed in petals, sepals and ovaries.

As ADHs are suggested to be involved in the synthesis of aroma substances in fruit and function in the lipoxygenase pathway [3, 19, 50], we measured the contents of C6-C9 alcohols during pear fruit development and correlated the expression profiles of *ADH* genes and content changes in C6-C9 alcohols during the four development stages. The results indicated that the expression levels of eight *ADH* genes (*Pbr013912.1*, *Pbr026289.1*, *Pbr01252.1*, *Pbr028275.1*, *Pbr022043.1*, *Pbr014945.1*, *Pbr034883.1*, and *Pbr026287.1*) are closely correlated with the content of Z)-hex-3-en-1-ol. Furthermore, qRT-PCR analysis was performed to verify the RNA-Seq expression profiles of the aforementioned eight *ADH* genes. The RNA-Seq expression profiles of three *ADH* genes (*Pbr013912.1*, *Pbr026289.1*, and *Pbr01252.1*) were consistent with the results from the qRT-PCR analysis (Fig. 9 A-C). Expression patterns of three *ADH* genes were similar, with high expression at S1 and a sharp decline at S2, which is consistent with the previous report that ADH activity and/or alcohol contents are high at the early stages of fruit ripening, whereas derivative esters predominate at maturity [32, 34]

## Discussion

*ADH* genes have been widely studied in a range of plants and they have been reported to participate in plant growth, development, and stress responses [1–3]. ADHs can catalyze the reciprocal transformation between alcohols and aldehydes and are involved in the production of aromatic compounds during fruit ripening. The number of *ADH* genes in Chinese white pear, European pear, and apple is over 1.5-fold those in black raspberry, sweet cherry and strawberry. The number found in peach and Japanese apricot was similar to that found in pear and apple and larger than that in black raspberry, sweet cherry, and strawberry.

We reconstructed the phylogenetic tree of *ADH* family genes in pear using the neighbor-joining method and four distinct subfamilies were determined. The results from the analysis on conserved motifs, gene features, and gene structures of *ADH* family genes support the classification results obtained by the phylogenetic analysis. We found that the characteristics of *ADH* genes were similar within each subfamily and varied among different subfamilies.

Different types of gene duplication events including WGD, tandem, proximal, transposed and dispersed duplications are the main driving forces for gene family expansion in eukaryotes [51, 52]. WGD events can generate large numbers of duplicate genes in a very short period of time [53]. Pear and apple experienced a recent lineage-specific WGD, whereas strawberry, sweet cherry, Japanese apricot, black raspberry, and peach did not undergo this duplication event [36]. Indeed, the number of WGD-derived *ADH* gene pairs in pear and apple is far greater than in the other six species. In addition, single-gene duplications were also important to the expansion of the *ADH* gene family in Rosaceae species. *ADH* gene clusters occur frequently in Rosaceae species, which was largely attributed to the recently tandem and proximal duplication. The duplicated *ADH* genes generated by tandem and proximal duplications showed accelerated evolution. Evolutionary analysis suggested that purifying selection was the primary evolutionary force imposed on *ADH* family genes. This result is consistent with our previous observation in analyzing the evolution of the *Hsf* and *F-box* gene family [54, 55].

We identified nine orthologous syntenic gene clusters among eight Rosaceae species by searching interspecies syntenic gene pairs for pear *ADH* genes. For example, one of the syntenic gene clusters comprised *Pbr012701.1*, *MD07G1250800*, *PCP001961.1*, *Pav\_sc0001102.1\_g440.1.mk*, *Prupe.2G274500*, *Pm019522*, *Ro07\_G12285*, and *FvH4\_7g27140.1* and these genes were located in interspecies large syntenic blocks. This result suggested that some ancestral *ADH* genes and surrounding genes were retained in descendants during long-term evolution after speciation and diversification of the Rosaceae. In addition, seven pear *ADH* genes were found to have no syntenic counterparts in other Rosaceae species, and five out of the seven *ADH* genes derived from tandem or proximal duplication and two of them were originated from transposed or genome-wide duplications. This result implied that local small-scale gene duplications played important roles in generating new genes [56, 57].

Previous studies showed that ADH activity and/or alcohol levels are highest at an early stage in fruit ripening, whereas, derivative esters are dominant at maturity [33, 34]. It has been reported that alcohols play an important role in the process of fruit maturation in tomato, grape, and melon. Overexpressing

*ADH2* in tomato fruits significantly increased the content of C6-alcohol and has a little influence on the content of C6-aldehyde, and the flavor of mature fruits is closely related to the increase of  $\alpha$ -hex-3-en-1-ol [26]. In grape fruits, ADHs show increased activity at the later stage of fruit ripening when aldehydes are converted alcohols [33, 35]. The expression module of three *ADH* genes (*Pbr013912.1*, *Pbr026289.1*, *Pbr01252.1*) were consistent with the qRT-PCR results indicating that expression is high early in fruit development and decreases sharply later on, which is concomitant with previous findings [32, 34].

Following duplication, duplicated genes may undergo different evolutionary processes including subfunctionalization, neofunctionalization, conservation, or nonfunctionalization [58]. In this study, we found that the duplicated *ADH* gene pairs showed divergent expression in different pear tissues, suggesting that subfunctionalization occurred frequently after gene duplication. In addition, we found five *ADH* genes with no or few instances of expression compared to their paralogous genes in all investigated pear tissues, suggesting that these five *ADH* genes (*Pbr024187.1*, *Pbr004679.1*, *Pbr003626.1*, *Pbr040240.1*, and *Pbr040236.1*) may have undergone nonfunctionalization or pseudogenization. This result provides evidence for the hypothesis that two gene copies derived via gene duplication may evolve toward distinct evolutionary fates, with one of the two copies gradually losing function and undergoing pseudogenization [59].

## Conclusion

In summary, a total of 464 *ADH* genes were identified in eight Rosaceae genomes and 68 of these genes were from Chinese white pear. Based on phylogenetic, gene structure and conserved motif analyses, *ADH* family genes were divided into four subfamilies (groups I–IV). Single-gene duplication largely contributed to the expansion of the *ADH* gene family in Rosaceae species, although lineage-specific genome duplication events in the ancestor of pear and apple also supplied *ADH* family expansion. Nine orthologous syntenic gene clusters were found among eight Rosaceae species after long-term evolution following Rosaceae diversification, suggesting that ADHs have highly conserved functional roles. Purifying selection was the main evolutionary force imposed on *ADH* genes. *ADH* genes derived from different types of gene duplications showed divergent evolutionary rates, with tandem and proximally duplicated genes evolving faster than those derived from other types of duplications. Transcriptome and qRT-PCR analysis revealed that *ADH* genes play roles in promoting alcohol production at later stages in pear fruit development. The results of this study lay a foundation for further investigations into the molecular function of *ADH* genes in Chinese white pear and other Rosaceae species.

## Materials And Methods

### Identification of *ADH* genes

The Chinese white pear genome sequence was downloaded from the Pear Genome Project (<http://peargenome.njau.edu.cn/>) [36]. The genome sequences of apple, peach, and strawberry were downloaded from Phytozome V12 (<http://phytozome.jgi.doe.gov/pz/portal.html>) and the European pear

genome sequence was downloaded from the Genome Database for Rosaceae (GDR) (<http://www.Rosaceae.org/>) [60]. The Japanese apricot genome sequence was obtained from the *Prunus mume* Genome Project (<http://prunusmumegenome.bjfu.edu.cn/index.jsp>). Additionally, the seed alignment file for the *ADH* domain (PF00107 and PF08240) was obtained from the Pfam database [61] and a Hidden Markov Model (HMM) file was created using the HMMER3 software package [62]. A search using HMMsearch was then performed against the local protein databases of eight Rosaceae species: Chinese white pear (*Pyrus bretschneideri* Rehd.), apple (*Malus domestica*), peach (*Prunus persica*), sweet cherry (*Prunus avium*), black raspberry (*Rubus occidentalis*), strawberry (*Fragaria vesca*), Japanese apricot (*Prunus mume*), and European pear (*Pyrus communis*). Furthermore, all candidate ADH protein sequences were analyzed using the Pfam database (<https://pfam.xfam.org>) to verify the presence of GroES-like and zinc-binding domains. Any protein sequences lacking GroES-like and zinc-binding domains were removed.

## Analysis of conserved motifs and gene features of *ADH* genes

In order to identify conserved domains among pear *ADH* genes, all identified protein sequences were subjected to MEME (Multiple Em for Motif Elicitation; v5.0.5) [63]. The analyses were conducted using default parameters with the following exceptions: the occurrence of motifs was set at 0 or 1 per sequence; the number of motifs was set to 20; the optimum width of motifs was 6–50 residues; and the minimum and maximum numbers of motif sites were set to 2 and 68, respectively. The gene ID, coding sequence (CDS), length of coding sequence, and number of amino acids in the sequence were acquired from the Pear Genome Project (<http://peargenome.njau.edu.cn/>). The protein isoelectric point (PI) and molecular weight (MW) for all candidate family members were computed using the ExPASy website ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)).

## Chromosomal locations and structures of *ADH* genes

Information about chromosomal locations of *ADH* genes was obtained from genome annotation files and the data were visualized using Circos software [64]. In order to construct the gene structures displaying the intron-exon distributions, GSDS (Gene Structure Display Server v2.0; <http://gsds.cbi.pku.edu.cn/>) was used [65]. Genomic DNA as well as the CDSs of all the *ADH* genes were submitted to construct the gene structure map. A phylogenetic tree was constructed using the full-length protein sequences of ADH from pear using MEGA (version 7.0) [66] with the Neighbor-joining (NJ) method [67] using the Poisson model and bootstrap values for 1000 replicates.

## Syteny analysis

To identify collinear gene pairs and syntenic blocks in pear and other Rosaceae species, we used diamond software [68] to perform multiple alignments of protein sequences in the eight Rosaceae species (e-value <  $10^{-5}$ ) and then obtained genome annotation files using an in-house Perl script. Finally, MCScanX [69] was performed to produce orthologous gene pairs within and between each Rosaceae species. Five module gene duplication pairs from the eight Rosaceae species were downloaded from the Plant Duplicate Gene Database (<http://pdgd.njau.edu.cn:8080/>) [56].

## Calculation of Ka, Ks and Ka/Ks

The values of Ka (non-synonymous substitutions), Ks (synonymous substitutions), and the Ka/Ks ratio were calculated using the calculate\_Ka\_Ks\_pipeline ([https://github.com/qiaoxin/Scripts\\_for\\_GB/tree/master/calculate\\_Ka\\_Ks\\_pipeline](https://github.com/qiaoxin/Scripts_for_GB/tree/master/calculate_Ka_Ks_pipeline)) [56]. In brief, the coding sequence and gene pairs were prepared. Then `computing_Ka_Ks_pipe.pl` script was used to perform multiple alignments automatically using MAFFT software and convert them to AXT format for submission to the KaKs\_Calculator [70] in GMYN model. The readable results including Ka, Ks, Ka/Ks, and the P-value were generated.

## Quantitative real-time PCR analysis

Total RNA was extracted from pear fruits using the Plant RNA Isolation Kit plus (Fuji, China) with three biological replicates. The cDNA was used for qRT-PCR. Primer sequences (Table S1) were designed using the Primer-Blast ([https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK\\_LOC=BlastHome](https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome)) tool. The composition of the PCR mixture was as follows: 0.5  $\mu$ L of each primer, 5  $\mu$ L of 2 $\times$ SYBR Premix ExTaq<sup>TM</sup>, 1  $\mu$ L of cDNA, and 3  $\mu$ L of RNase-free water. The RT-PCR was performed on a Lightcycle-480 (Roche). The qRT-PCR sequence began with 10 min at 95°C, followed by 45 cycles of 95°C for 3 s and 60°C for 10 s, and 30 s of extension at 72°C. Relative expression levels were calculated using the  $2^{-\Delta\Delta Ct}$  method and normalized to the *ADH* genes.

## Fruit alcohol contents

We measured the content of C6-C9 alcohols (5-methyl-2-hexanol, (z)-hex-3-en-1-ol, 3-nonanol, n-heptanol, nonan-3-ol, nonan-3-ol, 3-hexanol, 3, 4-diethyl-3-hexanol, and heptanol) during pear fruit development and performed correlation analysis between the expression profiles of *ADH* genes and content changes in C6-C9 alcohols during the four development stages of fruit using SPSS Statistics tool and significance value was set  $P < 0.05$ . Stages S1-S4 were used to denote four different stages of fruit development: 15 days after flowering (DAF), 45 DAF, 90 DAF, and 120 DAF. In this study, S1 = 16 May, S2 = 1 July, S3 = 31 July, and S4 = 29 August.

## List Of Abbreviations

WGD: Whole-genome duplication; TD: tandem duplication; PD: proximal duplication; TRD: transposed duplication; DSD: dispersed duplication; MYA: Million years ago; HMM: Hidden Markov Model; NJ: Neighbor-joining; MEME: Multiple EM for Motif Elicitation; BFU: Beijing Forestry University; CDS: coding sequence; GDR: Genome database for Rosaceae; PI: protein isoelectric point; MW: molecular weight; GSDS: Gene Structure Display Server; TPM: Transcripts per million; DAF: days after flowering

## Declarations

### *Competing interests*

The authors declare that they have no competing interests.

### *Author s' contributions*

W. Z. carried out the experiments and data analysis and produced a draft of the manuscript; X. Q. and Q. L. participated in data analysis and helped modify the manuscript; C. L. participated in the qRT-PCR; J. W. provided funding support; H. Y. and S. Z. managed and designed the research and experiments. All authors have read and approved the final manuscript.

### *Acknowledgments*

We thank International Science Editing (<http://www.internationalscienceediting.com>) for editing this manuscript.

### *Funding*

This study was supported by the National Key Research and Development Program (2018YFD1000200); the Key Program of National Natural Science Foundation of China (31830081); the “Taishan Scholar” project from Shandong Province of China; and the Jiangsu Province Science and Technology Support Program (BE2018389).

### *Author details*

<sup>1</sup>State Key Laboratory of Crop Genetics and Germplasm Enhancement, Centre of Pear Engineering Technology Research, Nanjing Agricultural University, Nanjing 210095, China

## References

1. Chase T: *Alcohol Dehydrogenases: Identification and Names for Gene Families. Plant Molecular Biology Reporter* 1999(17):333–350.
2. Jornvall H, Hedlund J, Bergman T, Oppermann U, Persson B: *Superfamilies SDR and MDR: from early ancestry to present forms. Emergence of three lines, a Zn-metalloenzyme, and distinct variabilities. Biochemical and biophysical research communications* 2010, 396(1):125–130.

3. Strommer J: *The plant ADH gene family. The Plant journal: for cell and molecular biology* 2011, *66*(1):128–142.
4. Alka K, Windle HJ, Cornally D, Ryan BJ, Henehan GT: *A short chain NAD(H)-dependent alcohol dehydrogenase (HpSCADH) from Helicobacter pylori: a role in growth under neutral and acidic conditions. The international journal of biochemistry & cell biology* 2013, *45*(7):1347–1355.
5. Khan AJ, Husain Q, Choudhuri G, Parmar D: *Association of polymorphism in alcohol dehydrogenase and interaction with other genetic risk factors with alcoholic liver cirrhosis. Drug and alcohol dependence* 2010, *109*(1–3):190–197.
6. Kumar S, Sandell LL, Trainor PA, Koentgen F, Duester G: *Alcohol and aldehyde dehydrogenases: retinoid metabolic effects in mouse knockout models. Biochimica et biophysica acta* 2012, *1821*(1):198–205.
7. Çelik A, Aktaş F: *A new NADH-dependent, zinc containing alcohol dehydrogenase from Bacillus thuringiensis serovar israelensis involved in oxidations of short to medium chain primary alcohols. Journal of Molecular Catalysis B: Enzymatic* 2013, *89*:114–121.
8. Jornvall H, Hedlund J, Bergman T, Kallberg Y, Cederlund E, Persson B: *Origin and evolution of medium chain alcohol dehydrogenases. Chemico-biological interactions* 2013, *202*(1–3):91–96.
9. Plapp BV, Lee AT, Khanna A, Pryor JM: *Bradykinetic alcohol dehydrogenases make yeast fitter for growth in the presence of allyl alcohol. Chemico-biological interactions* 2013, *202*(1–3):104–110.
10. Quaglia D, Pori M, Galletti P, Emer E, Paradisi F, Giacomini D: *His-tagged Horse Liver Alcohol Dehydrogenase: Immobilization and application in the bio-based enantioselective synthesis of (S)-arylpropanols. Process Biochemistry* 2013, *48*(5–6):810–818.
11. Deng Y, Wang Z, Gu S, Ji C, Ying K, Xie Y, Mao Y: *Cloning and Characterization of a Novel Human Alcohol Dehydrogenase Gene (ADHFe1). DNA Sequence* 2009, *13*(5):301–306.
12. Liskens HFaS, J.: *Measurement of oxygen tension changes in the style during pollen tube growth. Planta* 1966(71):90–106.
13. Harberd NPae, K. J.R: *The effect of a mutation causing alcohol dehydrogenase deficiency of flooding tolerance in barley. The New phytologist* 1982(90):631–644.
14. Michel Jacobs IRD, ~ and Dirk Van Den Bossche: *Isolation and biochemical analysis of ethyl methanesulfonate-induced alcohol dehydrogenase null mutants of Arabidopsis thaliana (L.) Heynh. Biochemical Genetic* 1988(26):105–122.
15. Bailey-Serres J, Voeselek LACJ: *Flooding Stress: Acclimations and Genetic Diversity. Annual Review of Plant Biology* 2008, *59*(1):313–339.

16. Million Tadege IDaCK: *Ethanol fermentation: new functions for an old pathway*. *Plant Science* 1999(4):320–325.
17. Ming Zhang YM, Yasunaga Furihata, Yasuo Nakamaru and Yohji Esashi: *Enzymatic Conversion of Volatile Metabolites in Dry Seeds during Storage*. *Plant and Cell Physiology* 1994(4):49–56.
18. Garabagi F, Duns, G. and Strommer: *Selective recruitment of Adh genes for distinct enzymatic functions in Petunia hybrida*. *Plant Molecular Biology Reporter* 2005(58):283–294.
19. Gonçalves B, Oliveira I, Bacelar E, Morais MC, Aires A, Cosme F, Ventura-Cardoso J, Anjos R, Pinto T: *Aromas and Flavours of Fruits*. *IntechOpen* 2018.
20. Höög J-O, Strömberg P, Hedberg JJ, Griffiths WJ: *The mammalian alcohol dehydrogenases interact in several metabolic pathways*. *Chemico-biological interactions* 2003, 143–144:175–181.
21. Thompson CE, Salzano FM, de Souza ON, Freitas LB: *Sequence and structural aspects of the functional diversification of plant alcohol dehydrogenases*. *Gene* 2007, 396(1):108–115.
22. Singh RK, Sane VA, Misra A, Ali SA, Nath P: *Differential expression of the mango alcohol dehydrogenase gene family during ripening*. *Phytochemistry* 2010, 71(13):1485–1494.
23. Marcus A. Koch BH, and Thomas Mitchell-Olds: *Comparative Evolutionary Analysis of Chalcone Synthase and Alcohol*. *Mol Biol Evol* 2000, 17(10):1483–1498.
24. Manriquez D, El-Sharkawy I, Flores FB, El-Yahyaoui F, Regad F, Bouzayen M, Latche A, Pech JC: *Two highly divergent alcohol dehydrogenases of melon exhibit fruit ripening-specific expression and distinct biochemical characteristics*. *Plant molecular biology* 2006, 61(4–5):675–685.
25. Jin Y, Zhang C, Liu W, Tang Y, Qi H, Chen H, Cao S: *The Alcohol Dehydrogenase Gene Family in Melon (Cucumis melo L.): Bioinformatic Analysis and Expression Patterns*. *Frontiers in plant science* 2016, 7:670.
26. Speirs J LE, Holt K, Yong-Duk K, Steele S, Loveys B, Schuch W: *Genetic Manipulation of Alcohol Dehydrogenase Levels in Ripening Tomato Fruit Affects the Balance of Some Flavor Aldehydes and Alcohols*. *Plant physiology* 1998.
27. Stephen A. Goff<sup>1</sup>, Harry J. Klee<sup>2</sup>: *Plant Volatile Compounds: Sensory Cues for Health and Nutritional Value?* *Science* 2006, 311:815–819.
28. Chervin C SJ, Loveys B, Patterson BD: *Influence of low oxygen storage on aroma compounds of whole pears and crushed pear flesh*. *Postharvest Biology and Technology* 2000, 19(3):279–285.
29. Lara I, Miró RM, Fuentes T, Sayez G, Graell J, López ML: *Biosynthesis of volatile aroma compounds in pear fruit stored under long-term controlled-atmosphere conditions*. *Postharvest Biology and Technology*

2003, 29(1):29–39.

30. Qin G, Tao S, Zhang H, Huang W, Wu J, Xu Y, Zhang S: *Evolution of the aroma volatiles of pear fruits supplemented with fatty acid metabolic precursors. Molecules* 2014, 19(12):20183–20196.

31. Dixon J, Hewett EW: *Factors affecting apple aroma/flavour volatile concentration: A Review. New Zealand Journal of Crop and Horticultural Science* 2000, 28(3):155–173.

32. Gonzalez-Aguero M, Troncoso S, Gudenschwager O, Campos-Vargas R, Moya-Leon MA, Defilippi BG: *Differential expression levels of aroma-related genes during ripening of apricot (Prunus armeniaca L.). Plant physiology and biochemistry: PPB* 2009, 47(5):435–440.

33. Kalua CM, Boss PK: *Evolution of volatile compounds during the development of cabernet sauvignon grapes (Vitis vinifera L.). Journal of agricultural and food chemistry* 2009, 57(9):3818–3830.

34. Zhang B, Shen JY, Wei WW, Xi WP, Xu CJ, Ferguson I, Chen K: *Expression of genes associated with aroma formation derived from the fatty acid pathway during peach fruit ripening. Journal of agricultural and food chemistry* 2010, 58(10):6157–6165.

35. Catherine Tesnière CVs: *Molecular cloning and expression of cDNAs encoding alcohol dehydrogenases from Vitis vinifera L. during berry development. Plant Science* 2000(157):77–88.

36. Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, Khan MA, Tao S, Korban SS, Wang H *et al*: *The genome of the pear (Pyrus bretschneideri Rehd.). Genome research* 2013, 23(2):396–408.

37. Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D *et al*: *The genome of the domesticated apple (Malus × domestica Borkh.). Nature Genetics* 2010, 42(10):833–839.

38. International Peach Genome I, Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J *et al*: *The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet* 2013, 45(5):487–494.

39. Shirasawa K, Isuzugawa K, Ikenaga M, Saito Y, Yamamoto T, Hirakawa H, Isobe S: *The genome sequence of sweet cherry (Prunus avium) for use in genomics-assisted breeding. DNA Res* 2017, 24(5):499–508.

40. VanBuren R, Bryant D, Bushakra JM, Vining KJ, Edger PP, Rowley ER, Priest HD, Michael TP, Lyons E, Filichkin SA *et al*: *The genome of black raspberry (Rubus occidentalis). The Plant journal: for cell and molecular biology* 2016, 87(6):535–547.

41. Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP *et al*: *The genome of woodland strawberry (Fragaria vesca). Nature Genetics* 2010,

43(2):109–116.

42.Li Y, Pi M, Gao Q, Liu Z, Kang C: *Updated annotation of the wild strawberry *Fragaria vesca* V4 genome. Hort Res* 2019, 6:61.

43.Qixiang Zhang WC, Lidan Sun, Fangying Zhao, Bangqing Huang<sup>2,6</sup>, Weiru Yang, Ye Tao, Jia Wang, Zhiqiong Yuan, Guangyi Fan, Zhen Xing, Changlei Han, Huitang Pan, Xiao Zhong, Wenfang Shi,, Xinming Liang DD, Fengming Sun, Zongda Xu, Ruijie Hao, Tian Lv,Yingmin Lv, Zequn Zheng,Ming Sun, Le Luo, Ming Cai, Yike Gao, Junyi Wang, Ye Yin, Xun Xu, Tangren Cheng & Jun Wang: *The genome of *Prunus mume*. Nature communication* 2012(3):1318.

44.Chagne D, Crowhurst RN, Pindo M, Thrimawithana A, Deng C, Ireland H, Fiers M, Dzierzon H, Cestaro A, Fontana P *et al*: *The draft genome sequence of European pear (*Pyrus communis* L. 'Bartlett'). PloS one* 2014, 9(4):e92644.

45.Linsmith G, Rombauts S, Montanari S, Deng CH, Celton J-M, Guérif P, Liu C, Lohaus R, Zurn JD, Cestaro A *et al*: *Pseudo-chromosome length genome assembly of a double haploid 'Bartlett' pear (*Pyrus communis* L.)*. *bioRxiv* 2019.

46.Jun J, Mandoiu, Il, Nelson CE: *Identification of mammalian orthologs using local synteny. BMC Genomics* 2009, 10:630.

47.Lemoine F, Lespinet O, Labedan B: *Assessing the evolutionary rate of positional orthologous genes in prokaryotes using synteny data. BMC evolutionary biology* 2007, 7:237.

48.Yang Z: *PAML 4: Phylogenetic Analysis by Maximum Likelihood. Molecular Biology and Evolution* 2007, 24(8):1586–1591.

49.Starr TK, Jameson SC, Hogquist KA: *Positive and negative selection of T cells. Annual review of immunology* 2003, 21:139–176.

50.Meng Li LL, Jim M Dunwell, Xin Qiao, Xing Liu and Shaoling Zhang: *Characterization of the lipoxygenase (LOX) gene family in the Chinese white pear (*Pyrus bretschneideri*) and comparison with other members of the Rosaceae. BMC Genomics* 2014.

51.Doerks T, Copley RR, Schultz J, Ponting CP, Bork P: *Systematic identification of novel protein domain families associated with nuclear functions. Genome research* 2002, 12(1):47–56.

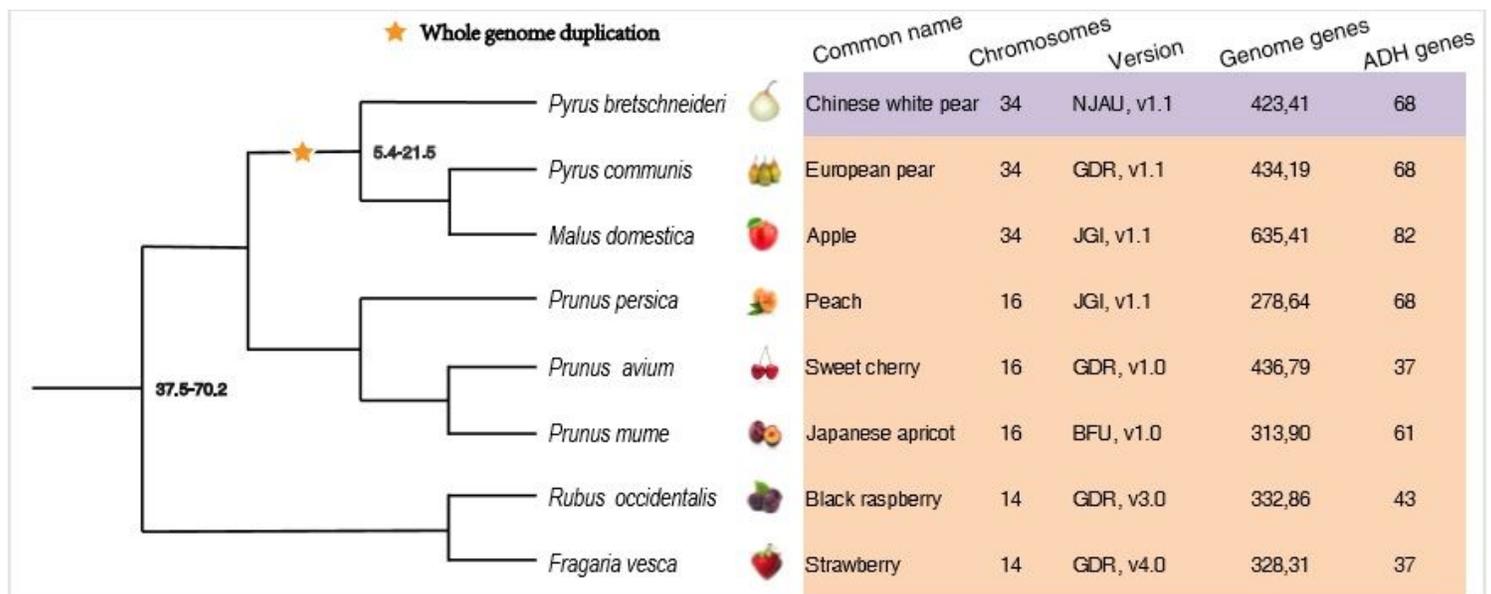
52.Moore RC, Purugganan MD: *The early stages of duplicate gene evolution. Proc Natl Acad Sci U S A* 2003, 100(26):15682–15687.

53.Wang Y, Wang X, Paterson AH: *Genome and gene duplications and gene expression divergence: a view from plants. Annals of the New York Academy of Sciences* 2012, 1256:1–14.

54. Qiao X, Li M, Li L, Yin H, Wu J, Zhang S: *Genome-wide identification and comparative analysis of the heat shock transcription factor family in Chinese white pear (Pyrus bretschneideri) and five other Rosaceae species. BMC plant biology* 2015, *15*:12.
55. Wang GM, Yin H, Qiao X, Tan X, Gu C, Wang BH, Cheng R, Wang YZ, Zhang SL: *F-box genes: Genome-wide expansion, evolution and their contribution to pollen growth in pear (Pyrus bretschneideri). Plant science: an international journal of experimental plant biology* 2016, *253*:164–175.
56. Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, Zhang S, Paterson AH: *Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. Genome Biol* 2019, *20*(1):38.
57. Freeling M: *Bias in Plant Gene Content Following Different Sorts of Duplication: Tandem, Whole-Genome, Segmental, or by Transposition. Annual Review of Plant Biology* 2009, *60*(1):433–453.
58. Hahn MW: *Distinguishing Among Evolutionary Models for the Maintenance of Gene Duplicates. Journal of Heredity* 2009, *100*(5):605–617.
59. Xie J, Li Y, Liu X, Zhao Y, Li B, Ingvarsson PK, Zhang D: *Evolutionary Origins of Pseudogenes and Their Association with Regulatory Sequences in Plants. The Plant cell* 2019, *31*(3):563–578.
60. Jung S, Lee T, Cheng CH, Buble K, Zheng P, Yu J, Humann J, Ficklin SP, Gasic K, Scott K *et al*: *15 years of GDR: New data and functionality in the Genome Database for Rosaceae. Nucleic acids research* 2019, *47*(D1):D1137-D1145.
61. Finn RD, Mistry J, Tate J, Coghill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K *et al*: *The Pfam protein families database. Nucleic acids research* 2010, *38*(Database issue):D211–222.
62. Eddy SR: *Accelerated Profile HMM Searches. PLoS computational biology* 2011, *7*(10):e1002195.
63. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: *MEME SUITE: tools for motif discovery and searching. Nucleic acids research* 2009, *37*(Web Server issue):W202–208.
64. Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA: *Circos: An information aesthetic for comparative genomics. Genome research* 2009.
65. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G: *GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics* 2015, *31*(8):1296–1297.
66. Kumar S, Stecher G, Tamura K: *MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol* 2016, *33*(7):1870–1874.
67. Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S: *MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol* 2013, *30*(12):2725–2729.

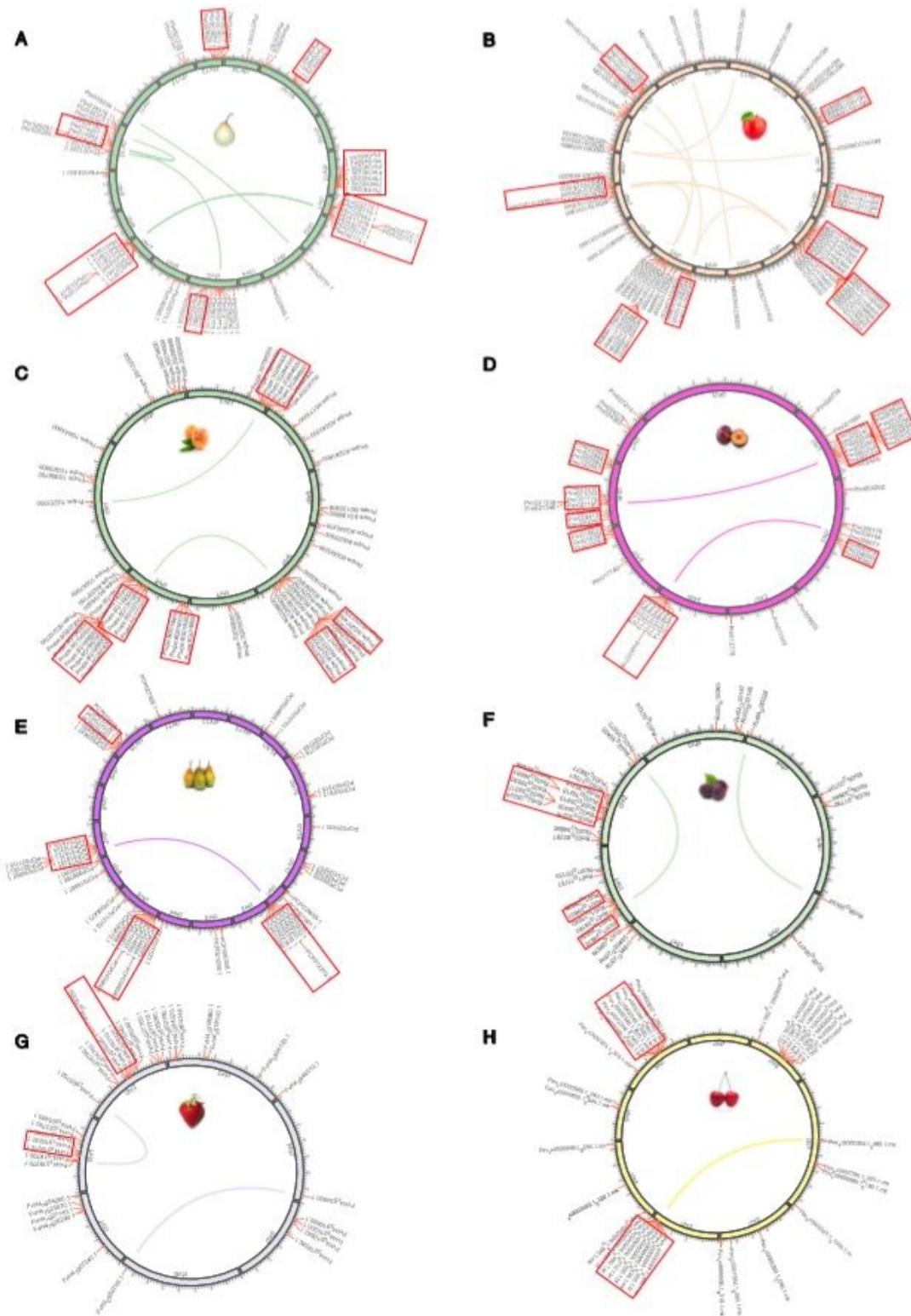
68. Buchfink B, Xie C, Huson DH: *Fast and sensitive protein alignment using DIAMOND*. *Nature Methods* 2014, 12(1):59–60.
69. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H *et al*: *MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity*. *Nucleic acids research* 2012, 40(7):e49.
70. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J: *KaKs\_Calculator 2.0: A Toolkit Incorporating Gamma-Series Methods and Sliding Window Strategies*. *Genomics, Proteomics & Bioinformatics* 2010, 8(1):77–80.
71. Chen C, Xia R, Chen H, He Y: *TBtools, a Toolkit for Biologists integrating various HTS-data handling tools with a user-friendly interface*. *bioRxiv* 2018.

## Figures



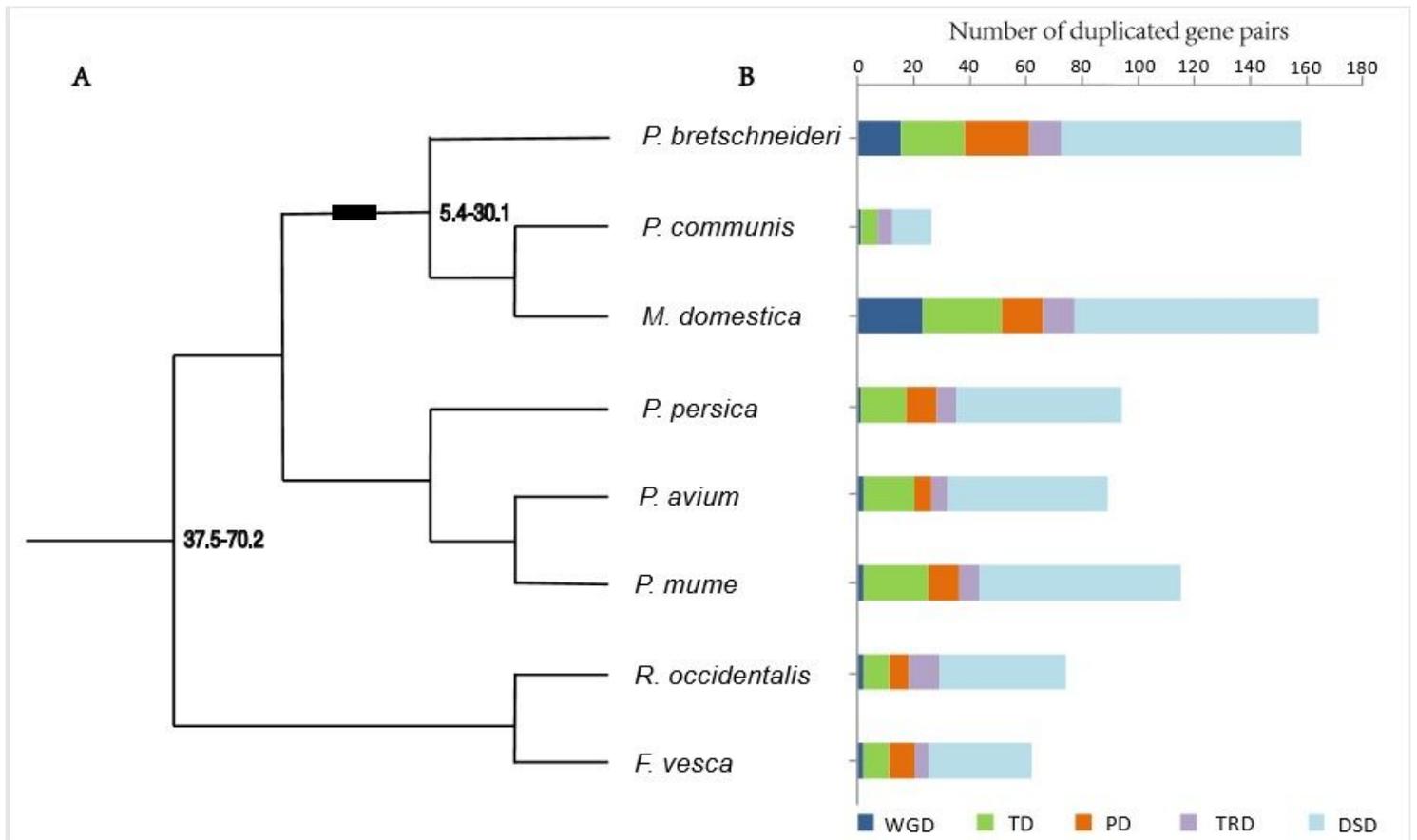
**Figure 1**

Phylogenetic tree and genome information for eight Rosaceae species. The pentagram (left) indicates the occurrence of WGD. The values on the left indicate species divergence time. Unit: MYA. The species tree was downloaded from NCBI taxonomy common tree (<http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi>) and the tree was visualized using MEGA 7 [66]. NJAU v1.1, Nanjing Agricultural University (<http://peargenome.njau.edu.cn/>); GDR, Genome Database for Rosaceae (<http://www.Rosaceae.org/>); JGI, Joint Genome Institute (<http://www.jgi.doe.gov/>); BFU, Beijing Forestry University (<http://prunusmumegenome.bjfu.edu.cn/index.jsp>).



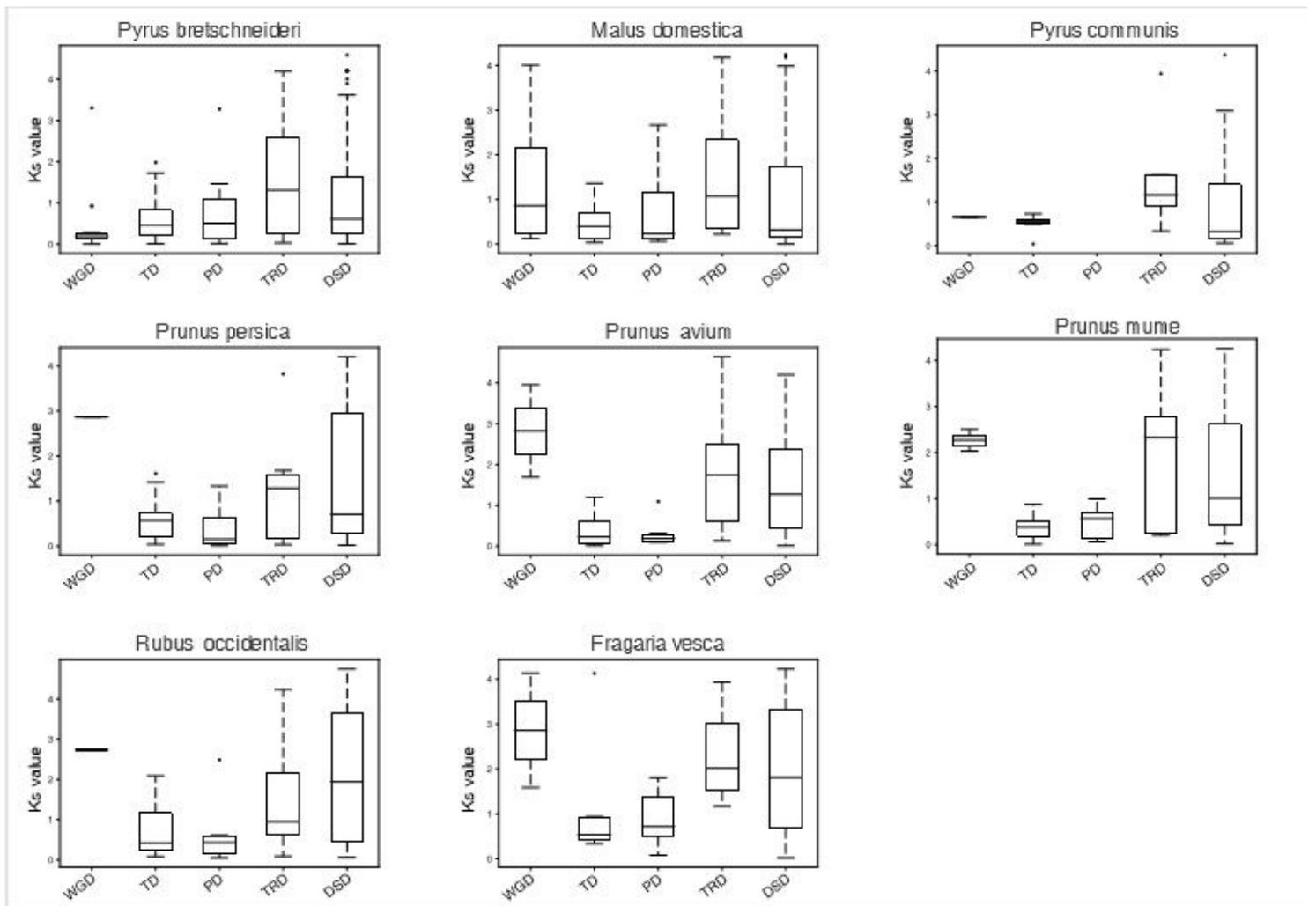
**Figure 2**

Chromosomal localization and syntenic relationships of ADH genes in eight Rosaceae species. (A) Chinese white pear, (B) apple, (C) peach, (D) Japanese apricot, (E) European pear, (F) black raspberry, (G) strawberry, and (H) sweet cherry. ADH genes are mapped on different chromosomes and syntenic gene pairs are linked by colored lines. The red rectangles represent tandem duplication events in each species.



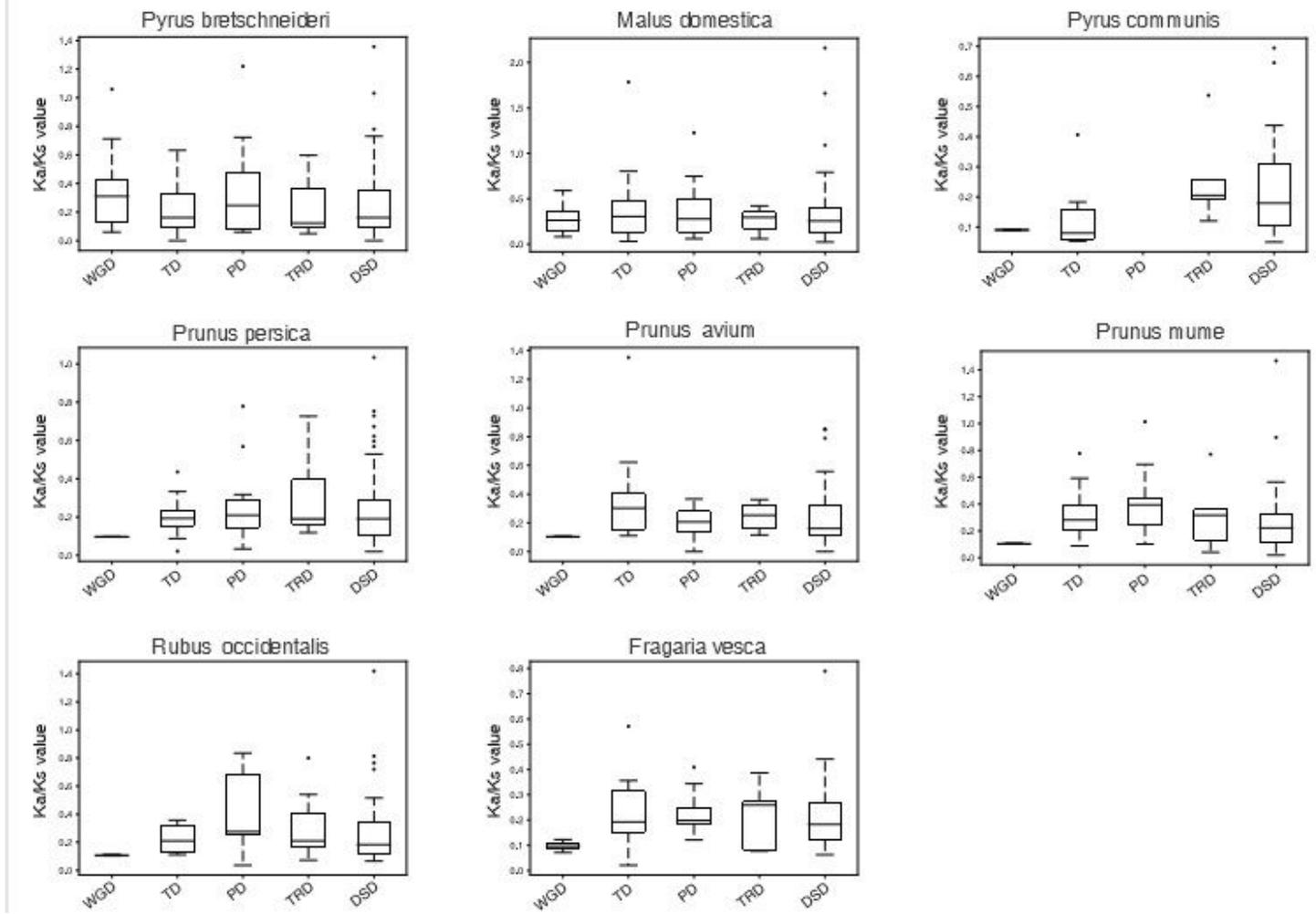
**Figure 3**

The number of ADH gene pairs derived from different modes of gene duplication in pear and seven other Rosaceae species. (A) The phylogenetic relationship among eight Rosaceae fruit species. (B) The number of different modes of duplicated gene pairs in each species. WGD: whole genome duplication, TD: tandem duplication, PD: proximal duplication, TRD: transposed duplication, DSD: dispersed duplication.



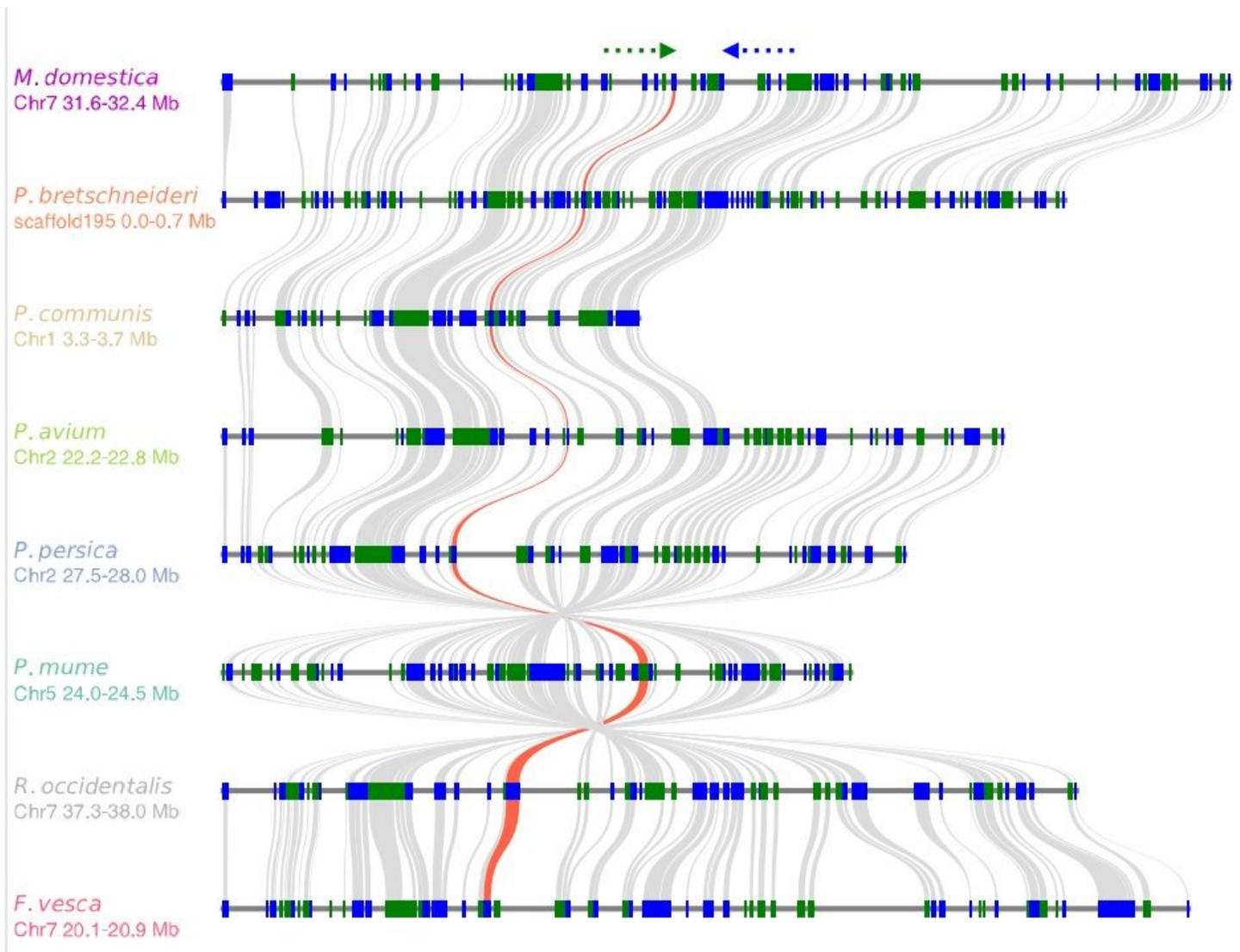
**Figure 4**

Ks distribution of different modes of duplicated ADH gene pairs in eight Rosaceae species.



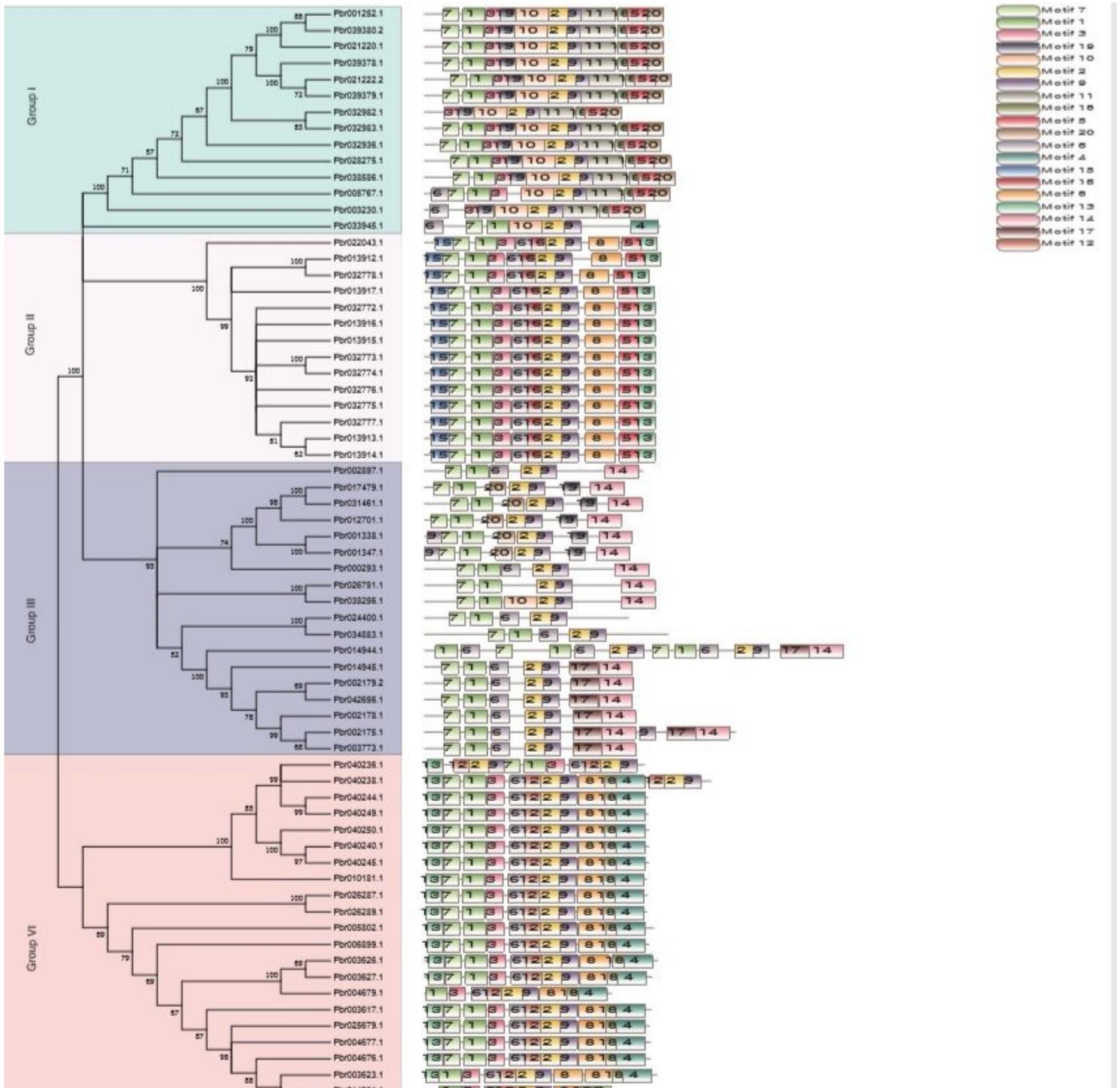
**Figure 5**

Ka/Ks distribution of different modes of duplicated ADH gene pairs in eight Rosaceae species.



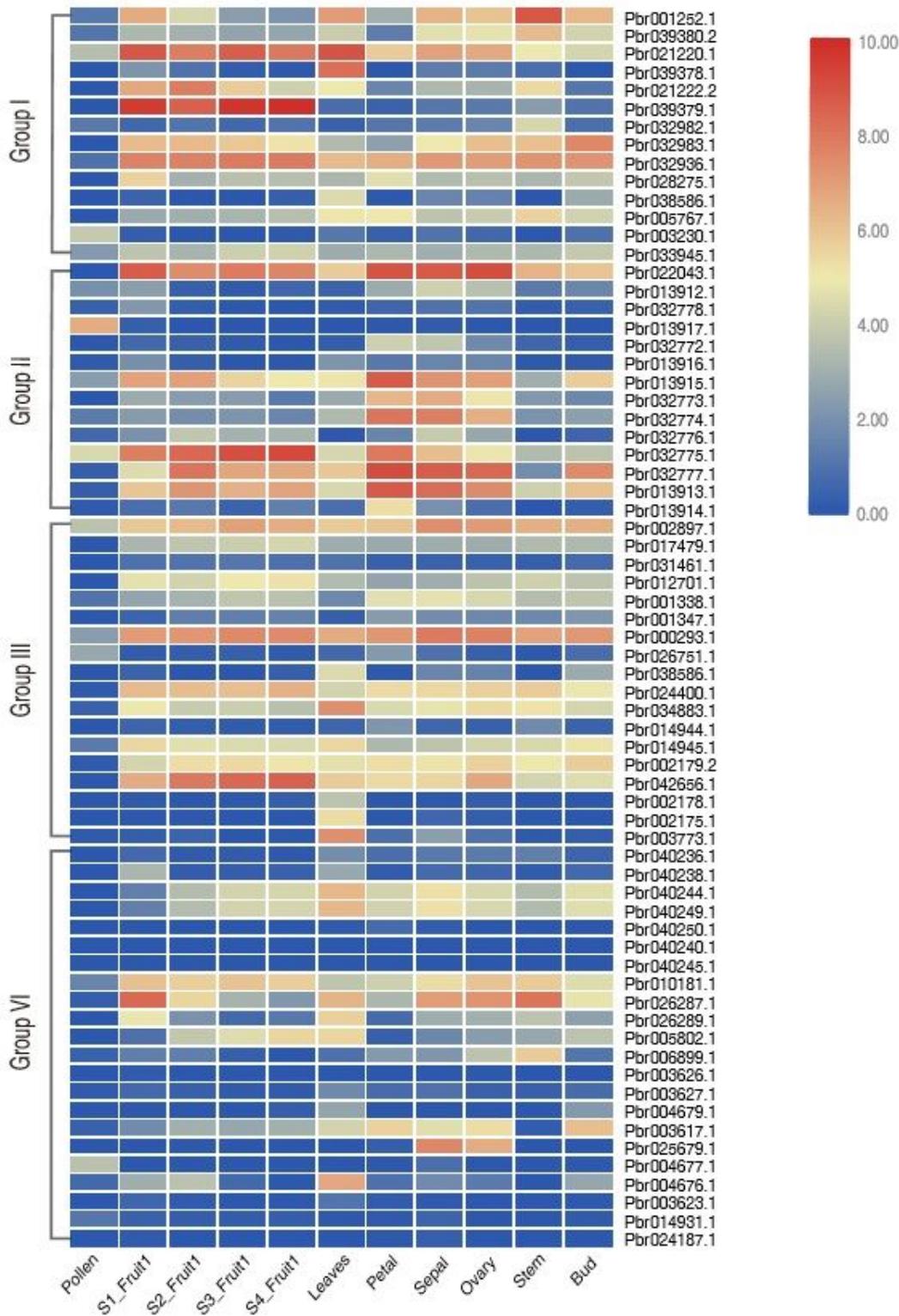
**Figure 6**

Microsyntenic analysis of ADH genes in eight Rosaceae species. The example ADH gene (Pbr012701.1) and its orthologous syntenic genes in seven other species of the Rosaceae are linked by the red line. The other syntenic gene pairs located in the syteny blocks containing ADH genes are linked by gray lines. Chromosome segments are represented by deep gray bars. The green bricks on the deep gray bars indicate genes that are transcribed in the forward direction and the blue bricks on the deep gray bars indicate genes that are transcribed in the reverse direction.



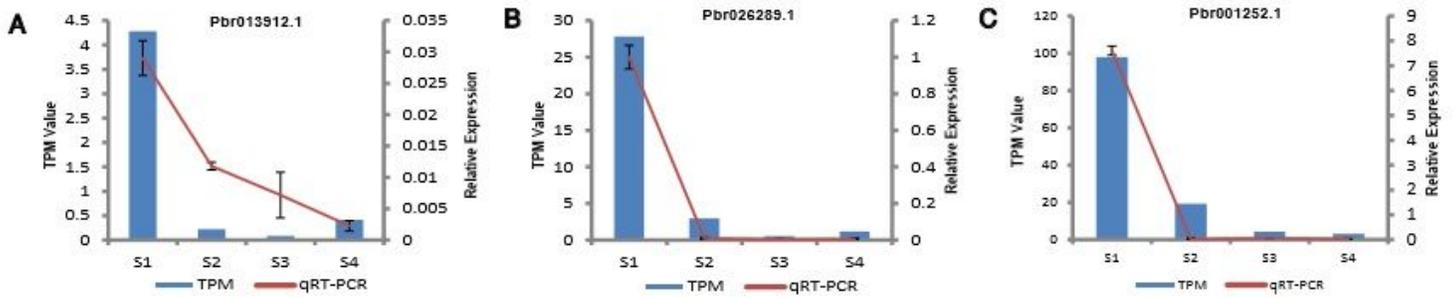
**Figure 7**

The phylogeny and conserved motifs of ADH genes in pear. (Left) Neighbor-joining phylogenetic tree of pear ADH genes and (right) the distribution of conserved motifs. The tree was reconstructed using MEGA 7.0 software and bootstrap analysis was conducted with 1,000 replicates. Twenty motifs (1 to 20) are indicated by different colors.



**Figure 8**

Expression patterns of pear ADH family genes in fruits, petals, sepals, stems, leaves, buds, ovaries, and root tissues. Groups I-IV represent four different subfamilies. S1-S4 correspond to four different developmental stages of fruit: S1 (16 May), S2 (1 July), S3 (31 July), and S4 (29 August). The color scale at the top represents log<sub>2</sub> transformed TPM values. Blue indicates a low expression level and red indicates a high expression level. The heatmap was generated using TBtools [71].



**Figure 9**

Expression patterns of three ADH genes verified by qRT-PCR analysis during the four fruit development stages of pear. (A-C) Expression profiles of three ADH genes (Pbr013912.1, Pbr026289.1, and Pbr01252.1). The left-hand y-axis represents the TPM values ascertained by RNA-seq analysis and the right-hand y-axis represents the relative expression levels obtained by qRT-PCR analysis. The stages (S1-S4) correspond to four different stages of fruit development: 15 days after flowering (DAF), 45 DAF, 90 DAF, and 120 DAF.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS3.xlsx](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS5.xlsx](#)
- [TableS7.xlsx](#)
- [TableS6.xlsx](#)
- [TableS9.xlsx](#)
- [TableS8.xlsx](#)
- [TableS4.xlsx](#)