The genome of low-chill Chinese plum 'Sanyueli' (*Prunus salicina* Lindl.) provides insights into the regulation of chilling requirement of flower bud

Zhi-Zhen Fang[1,2], Kui Lin-Wang[3], He Dai[4], Dan-Rong Zhou[1,2], Cui-Cui Jiang[1,2], Richard V. Espley[3], Yan-Juan Lin[1,2], Shao-Lin Pan[1,2], Xin-Fu Ye[1,2*]

[1] Fruit Research Institute, Fujian Academy of Agricultural Sciences, Fuzhou, Fujian, China

[2] Fujian Engineering and Technology Research Center for Deciduous Fruit Trees, Fujian Academy of Agricultural Sciences, Fuzhou, Fujian, China

[3]The New Zealand Institute for Plant & Food Research Ltd, Mt Albert Research Centre, Private Bag, Auckland 92169, New Zealand;

[4] Biomarker Technologies Corporation, Beijing, China

[*]Corresponding author

Tel: +86-591 83700095

E-mail: specialfruit@126.com, yexinfu@126.com

Fruit Research Institute

Fujian Academy of Agricultural Sciences

Hualin Road 247,

Fuzhou 350003, China

Abstract

Chinese plum (*Prunus salicina* Lindl.) is a stone fruit that belongs to the *Prunus* genus and plays an important role in the global production of plum. In this study, we report the genome sequence of the Chinese plum 'Sanyueli' (*P. salicina* Lindl.), which is known to have a low-chill requirement. The assembled genome size was 308.06 Mb, with a contig N50 of 815.7 kb. A total of 30,159 protein-coding genes was predicted from the genome and 56.42% (173.39 Mb) of the genome was annotated as repetitive sequence. Bud dormancy is influenced by chilling requirement in plum and partly controlled by *DORMANCY ASSOCIATED MADS-box* (*DAM*) genes. Six tandemly arrayed *PsDAM* genes were identified in the assembled genome. Sequence analysis of *PsDAM6* in 'Sanyueli'revealed the presence of large insertions in the intron and exon regions. Transcriptome analysis indicated that the expression of *PsDAM6* in the dormant flower buds of 'Sanyueli' was significantly lower than that in the dormant flower buds of the high chill requiring 'Furongli' plum. In addition, the expression of *PsDAM6* was repressed by chilling treatment. The genome sequence of 'Sanyueli' plum provides a valuable resource for elucidating the molecular mechanisms responsible for the regulation of chilling requirements, and it also useful for the identification of the genes involved in the control of other important agronomic traits and molecular breeding in plum.

Introduction

Plums are temperate fruit trees, which belong to *Prunus* genus in the *Rosaceae* family, and have been grown throughout the world for centuries. Plums provide the second largest stone fruit production after peaches and nectarines. Chinese plum (also known as Japanese plum, *P. salicina* Lindl.) and European plum (*Prunus domestica* L.) are the most commercially significant species[1]. Chinese plum is native to China[1]. The discovery of fossil plum stone dated to the Neolithic or the Warring States period implicated that plum fruits have been used as food for over 5,000-6,000 years in China [2]. The discoveries of plum stones in tombs of ancient Chinese people and an in-depth description of plum in the Book of Odes[2,3], a collection of poems dating from 1,100 BC to 600 BC, suggested that plum has been cultivated and popular in China for over 3,000 years. Chinese plum was transported to Japan over 2,000 years ago and was introduced into the USA in the 1870s[1,4]. It was estimated that Chinese plum and its hybrids account for over 70% of the world's global plum production[1]. China is the largest plum producer in the world and almost all of the production is Chinese plum[1]. According to Food and Agriculture Organization of the United Nations (FAO), plum fruit production in China was over 680 million tons in 2017.

Similar to other temperate/deciduous fruit trees, plum trees require a certain amount of cool temperature during the dormancy period to fulfill their chilling requirements and allow for the release of dormancy. Most Chinese plum varieties require 500-800 hours of chilling and European plums need even more (often >1,000 hours)[5]. Lack of winter chill results in delayed and abnormal flowering and negatively affects the yield and quality of fruits[6,7]. Plums are mainly grown in temperate zones[8]. In China, the production of plum is mainly concentrated in subtropical regions, including Guangdong, Guangxi and Fujian[9]. In the context of global warming, it is particularly pressing to develop low-chill plum cultivars[8,10]. A low chilling requirement has long been one of the principal objectives in plum breeding and breeding programmes to develop low-chill cultivars are evident in many countries, especially in the southern hemisphere[1,4,11-13]. Understanding the molecular mechanisms responsible for the regulation of bud dormancy and chilling requirements in plums is essential for the breeding of low-chill cultivars and ensuring consistent plum production in the changing environment.

Bud dormancy has attracted a great amount of attention in recent decades due to the agronomic disorders caused by warm winters[14-16]. Considerable efforts have been made to uncover the genetic basis of bud dormancy in Rosaceae species, including apple[17-24], peach[25-35], mume[36-41], pear[42-50], sweet cherry[51-54], almond[55,56] and apricot[57-61]. These studies as well as the publications on other species have emphasized the role of DORMANCY ASSOCIATED MADS-box (DAM) genes in regulation of bud dormancy and shed light on the transcriptional regulation of these genes[14]. However, to date the mechanisms involved in modulation of bud dormancy and chilling requirement in plum is still unknown.

To date the genome sequences of several *Prunus* species of *Rosaceae* family, including mume[62], peach[63], sweet cherry[64], flowering cherry[65], almond[66,67], apricot[68], have been published. Recently, the genome sequence of European plum has become available in the Genome Database for Rosaceae[69]. Genomes of commercial crops are valuable resources for identification of genetic loci regulating important agronomic traits and molecular breeding in the future. The availability of these genomes has enable the identification of genes for certain traits[70]. Currently, genome information for Chinese plum and its hybrids is not available. In this study, we sequenced and assembled the genome of Chinese plum ‘Sanyueli’, a

low-chilling requirement cultivar. We also conducted transcriptomic analyses to identify candidate genes underlying chilling requirement in plum. The results extend our understanding of the molecular control of dormancy and chilling requirement in plum. This genome sequence will be help facilitate the genetic research of the agronomic traits, such as chilling requirement and development of novel plum cultivars.

Materials and methods

Plant materials

Young leaves of 'Sanyueli' were used as sample for genome sequencing. Leaves collected in the plum repository in Fruit Research Institute of Fujian Academy of Agricultural Sciences (Jinan District, Fuzhou, Fujian province, China).

Annual branches of 8-year-old 'Sanyueli' were collected from the plum repository in Fruit Research Institute of Fujian Academy of Agricultural Sciences on December 7, 2017. Annual branches of 15-year-old 'Furongli' were collected from a commercial orchard in Gutian County, Fujian Province, China on December 5, 2017. Then, the collected branches were transported to the laboratory immediately and placed in glass jars containing fresh water and kept under 2-6□ in dark. Flower buds of 'Sanyueli' were collected at 0 h (S1), 170 h (S2) and 450 h (S3) and flower buds of 'Furongli' at 0 h (F1), 290 h (F2) and 530 h (F3). For determine the state of flower buds, the branches of 'Sanyueli' and 'Furongli' were transferred to and kept at 25 ± 1°C with white light (150 mol m^{-2} s^{-1}) under a 14 h light/10 h dark photoperiod at 75±5% humidity for 20 days. The chilling requirement of 'Sanyueli' flower bud was fulfilled after treated with chilling temperature for 50 h (Supplementary Fig. 2A). However, nearly 800 h is need for breaking dormancy in flower bud of 'Furongli' (Supplementary Fig. 2B). Three biological replicates for each sample were collected. Flower bud samples were immediately frozen in liquid nitrogen and then stored at -80°C until RNA extraction.

Genome Sequencing

Genomic DNA was extracted from young leaves of 'Sanyueli' using a modified CTAB method. For short-read sequencing, a library with insert size of 270bp was constructed using TrueLib DNA Library Rapid Prep Kit for Illumina (Genetimes Technology Inc. shanghai, China) and sequenced using Illumina HiSeq 4000 system (Illumina, San Diego, CA, USA). All raw reads were processed using a local perl script (developed by Biomarker Technologies Corporations) to remove adapter sequence and reads with length less than 100 bp or Q30<85%. For PacBio sequencing, genomic DNA was sheared into 20Kb fragments using a g-TUBE device (Covaris, Woburn, MA, USA), then the sheared DNA was purified and concentrated with AmpureXP beads (Agencourt, Beverly, MA, USA). SMRTbell library was constructed with SMRTbell kits (Pacific Biosciences) according to manufacturer's protocol and sequenced on a PacBio Sequel system.

Genome size estimation and heterozygosity

The filtered Illumina data were used to estimate the genome size, heterozygosity and the content of repetitive sequences through the k-mer depth frequency distribution analysis. This analysis was carried out using "kmer freq stat" software (developed by Biomarker Technologies Corporation, Beijing, China).

Genome size G=K_num/Peak_depth, where the K_num is the total number of K-mer(k =19), and Peak_depth is the expected value of K-mer depth.

Genome assembly

Raw PacBio reads were filtered using a local perl script (developed by Biomarker Technnologies Corporations). Only subreads that equal to or longer than 500bp were used for subsequent genome assembly. PacBio long reads were corrected and assembled using the Canu program[71]. In the correction step, Canu selects longer seed reads, then detects overlapping raw reads with MHAP (mhap-2.1.2), and finally performs an error correction through falcon_sense method. Error-corrected reads are trimmed to obtain the longest supported range with the default parameters. Finally, a draft assembly was generated using the longest trimmed reads. In addition, we also used WTDBG program (https://github.com/ruanjue/wtdbg) to assemble the Canu corrected reads. To improve the assembly contiguity, the assemblies generated by Canu and WTDBG were merged by Quickmerge (v0.2). Canu-generated contigs were used as query input and WTDBG -generated contigs were used as ref input. Finally, we employed Pilon (v1.22) to the merged assembly using high-quality cleaned Illumina reads.

Evaluation of Genome Quality

To evaluate the coverage of the assembled plum genome, the Illumina paired-end reads were aligned to the genome assembly using the BWA-MEM (version 0.7.10-r789)[72]. To evaluate the completeness of the assembly, 1,440 Benchmarking Universal Single-Copy Orthologs (BUSCOs) and 458 Core Eukaryotic Genes (CEGs) were mapped to plum genome using BUSCO v2.0[73] and CEGMA v2.5[74], respectively.

Genome Annotation

The *de novo* repetitive sequence database of *P. salicina* was constructed employing LTR-FINDER v1.05[75], MITE-Hunter[76], RepeatScout v1.0.5[77] and PILER-DF v1.0[78]. The database was classified using PASTEClassifier v1.0[79] and merged with the Repbase database v20.01[80] to create the final repeat library. Repeat sequences in *P. salicina* were finally identified and classified using RepeatMasker program v4.0.6[81].

We used *Ab initio-based*, protein homology-based and RNA-seq-based approaches to predict protein-coding genes. *Ab initio-based* gene prediction was performed using Genscan v3.1[82], Augustus v3.1[83], GlimmerHMM v3.0.4[84], GeneID v1.4[85] and SNAP[86]; GeMoMa v1.3.1[87] was employed to carried out homology-based prediction with the homologous peptides from *Oryza sativa*, *Prunus persica*, *Malus* × *domestica* and *Pyrus bretschneideri*. For RNA-seq-based prediction, the RNA-seq reads generated from leaf, stem and flower were assembled into unigenes using Hisat v2.0.4[88] and Stringtie v1.2.3[88]. The unigenes were aligned to the assembly using BLAT[89] and gene structures of BLAT alignment results were modeled using PASA v2.0.2[90]. The prediction of protein-coding regions were performed with TransDecoder v3.0.1 (Haas, http://transdecoder.github.io) and GeneMarkS-Tv5.1[91], respecitively. Finally, the *de novo* predictions, protein alignments and transcripts data were integrated using EVM v1.1.1[92]. Annotations of the predicted genes were performed by blasting their sequences against a series of nucleotide and protein sequence databases, including KOG[93], KEGG[94], Pfam[95], Swissprot[96], NCBI-NT, NCBI-NR and TrEMBL[96] with an E-value cutoff of $1e^{-5}$. Gene ontology (GO) for each gene were assigned by the Blast2GO[97] based

on NCBI databases.

The rRNA fragments were identified by aligning the rRNA template sequences (Pfam database v32.0) against the assembled plum genome using BLAST with E-value at $1e^{-10}$ and identity cutoff at 95% or more. The tRNAs were predicted using tRNAScan-SE v1.3.1 algorithms[98] with default parameters. The miRNAs were predicted by INFERNAL v1.1 software[99,100] against the Rfam database v14.0[101] with cutoff score at 30 or more. The minimum cutoff score was based on the settings which yield a false positive rate of 30 bits.

Comparative genomic analysis

Orthologous groups among *P. salicina* and other 14 species (including *Vitis vinifera*, *Rosa chinensis*, *Populus trichocarpa*, *P. persica*, *Prunus mume*, *Prunus dulcis*, *Pyrus bretschneideri*, *Prunus avium*, *Prunus armeniaca*, *O. sativa*, *M. domestica*, *Fragaria vesca*, *Amborella trichopoda* and *Arabidopsis thaliana*) were constructed using OrthoMCL v2.0.9[102] based on an all-to-all BLASTP strategy (with an E-value of 1e-5). We extracted 436 single-copy genes from the 15 species and aligned proteins for each gene. All the alignments were combined to one supergene to construct a phylogenetic tree using RAxML v7.2.8[103] with 1,000 rapid bootstraps followed by a search of the best-scoring maximum likelihood (ML) tree in one single run. Divergence time was estimated using the MCMCTree[104] program in PAML v4.9 under the relaxed clock model. Several calibrated time points were used to date the divergence time in the unit of Ma.

A GF (gene family) was defined as a group of similar genes that descended from a single gene in the last common ancestor. Expansion and contraction of GFs were determined using CAFÉ v4.0[105] based on changes in GF size. The cluster size of each branch was compared with the cluster size of the ancestral node. P value was calculated using Viterbi method under the hidden Markov model, with $p < 0.01$ defining significant expansion or contraction. Genes belonging to expanded GFs were subjected to GO and KEGG enrichment.

All-against-all BLASTP analyses of protein sequences were performed between *P. salicina*, *P. mume* and *P. armeniaca* using E-value cut-off of $1e^{-10}$. Syntenic regions between species were identified using MCScan[106] based on the BLASTP results. A syntenic region was identified if it contained a minimum of 10 and a maximum of 25 genes in the identified gene pairs. Protein sequences of homologous gene pairs in the identified syntenic regions were aligned by MUSCLE v3.8.31[107] and the protein alignments were then converted to coding sequence (CDS) alignments. The synonymous substitution (Ks) value of each syntenic gene pair was calculated using the Yn00 program in the PAML package[104].

Identification and analysis of *PsDAM*s

To identify *DAM* genes in the genome of ‘Sanyueli’, the deduced amino acid sequences of all six peach *DAM* genes[32] and six *P. mume PmDAM* genes[61] were used as queries for Blast analyses using TBtools[108]. The schematic gene structure of DAM genes and structural alignments of the PsDAM6 gene from 'Sanyueli' and ‘Furongli’ plum was displayed using TBtools[108]. Phylogenetic analyses were performed using the neighbor-joining method by MEGA-X software with 1,000 bootstrap replicates. The predicted amino acid sequences of DAM genes in other Rosaceae species were downloaded from Genbank or Genome Database for Rosaceae, including *M. domestica* MdDAM1 (KT582786), MdDAM2 (KT582787),

MdDAM3 (MDP0000527190), MdDAM4 (KT582789), MdJa(KT582788), MdJb (LC004730); *P. bretschneideri* Rehd. PbrDAM1 (KP164027), PbrDAM2 (KP164026) and PbrDAM3 (KP164028); *P. persica* PpDAM1 (DQ863253), PpDAM2 (DQ863255), PpDAM3 (DQ863256), PpDAM4 (DQ863250), PpDAM5 (DQ863251) and PpDAM5 (DQ863252); *P. armeniaca* ParDAM1 (PARG08688m02), ParDAM2 (PARG08688m01), ParDAM3 (PARG08689m02), ParDAM4 (PARG08689m04), ParDAM5 (PARG08690m03) and ParDAM6 (PARG08690m02); *P. avium* PavDAM1(LC544139), PavDAM2 (LC544140), PavDAM3 (LC544141), PavDAM4 (LC544142), PavDAM5 (LC544143) and PavDAM6 (LC544144); *P. mume* PmDAM1 (XR_001677199.1), PmDAM2 (XM_008221048.2), PmDAM3 (XM_008221049.2), PmDAM4 (XM_008221050.2), PmDAM5 (NM_001293268.1) and PmDAM6 (NM_001293262.1); DAM genes were predicted from the genome sequence of *P. dulcis* in Genome Database for Rosaceae according to the sequence of peach DAM genes, PduDAM1 (Pd01:39769437-39776175), PduDAM2 (Pd01:39778066-39783736), PduDAM3 (Pd01:39795512-39804580), PduDAM4 (Pd01:39805806-39814186), PduDAM5 (Pd01:39815116-39823593) and PduDAM6 (Pd01:39824719-39831140). Alignments of DAM proteins were generated with ClustalW and displayed using ESPript 3.0[109].

Transcriptome sequencing and analysis of flower bud

Total RNA was extracted from approximately100 mg flower buds using the RNAprep Pure Plant Kit (Tiangen, Beijing, China). The quality of RNA was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The concentration of RNA was determined using Qubit 2.0 Fluorometer (Invitrogen, Life Technologies, CA, USA). High quality RNA was subjected to library preparation using NEBNext Ultra^{TM} RNA Library Prep Kit for Illumina (NEB, USA). The libraries were sequenced on a HiSeq X Ten sequencer (Illumina). The RNAseq reads have been deposited into the NCBI Short Read Archive and are accessible under project PRJNA645255. The RNA-seq reads were mapped to the ‘Sanyueli’ plum genome sequences using HISAT2[110]. Reads were assembled into transcripts and the expression level of transcripts was calculated as the FPKM values using the StringTie software package[111]. Principal component analyses (PCA) were performed on FPKM values from different datasets using the GenAlEx version 6.5 program. Differential gene expression analysis was carried out using DESeq2[112]. Genes with fold change greater than 2 and false discovery rate (FDR) below 0.05 were considered significantly differentially expressed. Gene Ontology (GO) enrichment analysis of differentially expressed genes (DEGs) was performed with the GOseq R package. GO terms with corrected p-values < 0.05 were considered to be significantly enriched by differentially expressed genes[113]. KOBAS software was used to test the statistical enrichment of DEGs in KEGG pathways[114].

The PlantTFDB v5.0 (http://planttfdb.cbi.pku.edu.cn/prediction.php) was used to predicted transcription factors from all assembled transcripts. The correlation coefficient between the expression of *PsDAM6* and differentially expressed genes in buds of 'Sanyueli' and 'Furongli' plum was calculating using Microsoft Excel software.

Results

Genome assembly

‘Sanyueli’ is an early-maturing and high-yielding plum variety with red skin and yellow flesh (Fig. 1A-C).

‘Sanyueli’ trees are able to bloom and fruit without exposure to large amount of chilling temperatures (Fig. 1D-E). Daily maximum temperature and minimum temperature from November 2015 to February 2016 was indicated in Supplementary Fig. 1. He, et al.[115] reported that ‘Sanyueli’ has the lowest chilling requirement among the varieties cultivated in Guangzhou, Guangdong Province, China. Later, they demonstrated that 28 h of chilling temperatures (≤7.2□) is sufficient to fulfill the chilling requirement of ‘Sanyueli’ flower bud[116].

We sequenced and assembled the genome of ‘Sanyueli’ using a combination of short-read sequencing from Illumina HiSeq 4000, and SMRT from Pacific Biosciences (PacBio, Menlo Park, CA). First, we generated 28.08 Gb Illumina paired-end (PE) reads with insert sizes of 270 bp (Supplementary Table 1) and used them to estimate the genome size and heterozygosity ratio of ‘Sanyueli’ genome. Based on a 19-mer analysis, we evaluated the genome size to be 308.06 Mb, with a heterozygosity of 0.33%, and the estimated repeat sequence content is 52.63% (Supplementary Fig.3). Then 14.89 Gb of PacBio reads (Supplementary Table 2) were produced and assembled into 1,752 contigs (longest of 7,898,106 bp) with an N50 of 815.7 kb. The size of the final assembled genome was 307.29 Mb, which is very close to the estimated size. The GC content of the assembled plum genome was 37.78% (Table 1).

The quality of the assembled genome was evaluated using three strategies. First, BWA alignment result indicated that 97.85% of the Illumina paired-end reads were successfully aligned to the genome (Supplementary Table 1). Second, according to BUSCO (version 2)[117], our assembly contained 97.01% (1397 of 1440) of the core eukaryotic genes, including 1292 single-copy orthologs and 197 duplicated orthologs (Supplementary Table 3). Third, CEGMA[74] results indicated that 451 of 458 highly conserved core genes were detected in our assembly (Supplementary Table 4), suggesting the high completeness of our genome assembly. These results indicated that our plum genome sequence was almost complete.

Genome annotation

Annotation of repeat sequences

56.42% (173.39 Mb) of the assembled genome was predicted to be repetitive, which is higher than the repeat content observed in mume (44.92%), sweet cherry (43.8%), apricot (38.28%), peach (37.14%) and almond (34.6%). The repetitive sequences including retrotransposons (Class I elements, 48.34%), DNA transposons (Class II elements, 11.99%), potential host genes (1.61%), simple sequence repeats (0.02%) and unclassified elements (3.84%) (Supplementary Table 5). The proportion of Gypsy retrotransposon (24.38%) appears to have expanded considerably in the ‘Sanyueli’ genome compared with that of the peach (9.97%) and mume (8.6%) and is comparable with that of apple (25.2%). The proportion of full-length long terminal repeats (LTR)/Copia repeats (9.32%) in the ‘Sanyueli’ genome was similar to that in mume (10.0%) and higher than that in peach (8.60%) and apple (5.0%). The PLE/LARD retrotransposon derivative repeat elements represented 9.63% of the genome, similar to the proportion of the genome represented by LTR/ Copia elements (Supplementary Table 5).

Gene prediction and functional annotation

A combination of *ab initio*, homology-based and RNA-Seq based prediction methods was used to predict gene models from plum genome sequence. A total of 30,159 protein-coding genes were predicted

(representing 35.26% of the genome assembly), with an average gene length of 3593 bp, an average coding region sequence size of 1297 bp (Supplementary Table 6). The average gene density of plum was 98 genes per Mb, which is lower than in apricot (137 genes per Mb), mume (132 genes per Mb), peach (122 genes per Mb) and almond (112 genes per Mb), but is higher than in sweet cherry (87 genes per Mb). 21939 genes (72.74%) were supported by RNA-sequencing data, 29190 genes (96.79%) were supported by homology to known proteins. A total of 19486 genes (64.61%) were supported by all three methods (Supplementary Fig. 4), and these genes were annotated with high confidence. A total of 29,817 (98.87%) protein-coding genes were annotated based on the Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), EuKaryotic Orthologous Groups (KOG), Pfam, Swissprot, TrEMBL, Nr and Nt databases (Supplementary Table 7). In addition, 85 microRNAs (miRNAs), 602 ribosomal RNAs, 525 transfer RNAs, 2210 small nucleolar RNAs (snoRNAs), 101 small nuclear RNAs (snRNAs) and 1419 pseudogenes were identified in the 'Sanyueli' plum genome (Supplementary Table 8).

Comparative genomic and genome evolutionary analysis

We performed gene family cluster analysis on genome of *P. salicina*, *P. armeniaca*, *P. persica*, *P. avium*, *P. mume*, *P. dulcis*, *P. bretschneideri*, *Malus* x *domestica*, *Fragaria vesca*, *Rosa chinensis*, *Vitis vinifera*, *Populus trichocarpa*, *O. sativa*, *Arabidopsis thaliana* and *Amborella trichopoda.* The number of single-copy gene in *P. salicina* is less than other *Prunus* species but more than *P. bretschneideri* and *M. domestica*, while the number of multi-copy gene in *P. salicina* is more than other *Prunus* species and less than *P. bretschneideri* and *M. domestica* (Figure 2A). A total of 28006 genes from the plum genome were grouped into 17622 gene clusters and 219 gene families containing 553 genes (Supplementary Table 9 and 10) were identified as plum specific. We compared the gene numbers among the six *Prunus* species. 11053 gene families were shared by all *Prunus* species, and 288 gene families were specific to *P. salicina*, which is more than the number found in genomes *P. persica* and *P. mume* but less than that of *P.avium*, *P.armeniaca* and *P.dulcis* (Fig. 2B).

A phylogenetic tree of plum, five *Prunus* species and nine other sequenced species was constructed using single-copy genes. The phylogenetic tree showed that *P. salicina* is relatively closely related to *P. mume* and *P. armeniaca* (Figure 2A). It also indicated that *P. salicina* diverged from *P. mume* and *P. armeniaca* approximately 7.45 million years ago (Mya) and from *P. dulcis* and *P. persica* 10.32 Mya, after the divergence of *P. avium* at 15.14 Mya (Figure 2A). Genome collinearity analysis demonstrated that *P. salicina* genome showed high collinear relationships with the genome of *P. armeniaca* and *P. mume*, which are closely related to *P. salicina*. (Supplementary Fig. 5).

The synonymous nucleotide substitutions (Ks) analysis in plum, *M. domestica* and *Vitis vinifera* indicated that there has been an ancient WGD (ancient γ whole-genome duplication) in the plum genome before the divergence with *M. domestica* and *V. vinifera* (Figure 2C). In addition, 1248 and 2550 gene families were found to have expanded and contracted in the plum genome, which is more than the number of expanded and contracted gene families in the genomes of *P. mume* and *P.armeniaca*. KEGG enrichment analysis of the expanded genes showed that plant-pathogen interaction, starch and sucrose metabolism and phenylpropanoid biosynthesis were the most enriched pathways (Supplementary Table 11). Among the 74 genes assigned to starch and sucrose metabolism pathway, 17 were annotated as polygalacturonase and two of them (PsSY0006672 and PsSY0008876) were significantly upregulated in the flesh during fruit ripening.

GO analysis of the expanded orthogroups revealed that the oxidation-reduction process (GO:0055114), regulation of transcription, DNA-templated (GO:0006355), metabolic process (GO:0008152) and protein phosphorylation (GO:0006468) are significantly enriched (Supplementary Table 12).

A, Phylogenetic analysis and gene family cluster analysis of plum and 14 other species. The phylogenetic tree (left panel) was constructed using 436 single-copy genes from plum and 14 other species, including *P. armeniaca*, *P. persica*, *P. avium*, *P. mume*, *P. dulcis*, *P. bretschneideri*, *M. domestica*, *F. vesca*, *R. chinensis*, *V. vinifera*, *P. trichocarpa*, *O. sativa*, *A. thaliana* and *A. trichopoda.* The numbers indicate the divergence time. Gene family expansions and contractions are indicated by numbers in yellow and blue boxes. The distribution of paralogous genes in the analyzed plants is indicated in the right panel. B, Venn diagram comparison of gene families in six *Prunus* species. C, Distribution of synonymous nucleotide subsitutions (Ks) among *P. salicina*, *M. domestica* and *V. vinifera*. Duplication components were identified by MCScan without tandem duplicated pairs. Colored curves superimposed on the Ks distribution represent different component identified. Ks distributions of paralogous gene pairs were identified from syntenic blocks between species, and orthologous gene pairs were identified from duplication component within one genome.

DORMANCY ASSOCIATED MADS-box (*DAM*) genes in plum genome

DAM genes have been reported to be the major regulatory factors involved in the control of dormancy induction and release in flower bud of relative species[61,118,119]. Blast analyses against the genome of 'Sanyueli' plum enabled us to identify six *PsDAM* genes, including *PsDAM1* (PsSY0028611), *PsDAM2* (PsSY0012373), *PsDAM3* (PsSY0007206), *PsDAM4* (PsSY0000114), *PsDAM5* (PsSY0007447) and *PsDAM6* (PsSY0011722), which are tandemly arrayed in the genome 'Sanyueli' plum (Figure 3A). Genomic structure analysis revealed that *PsDAM1-PsDAM5* have similar genomic structures consisting of eight exons and seven introns to their homologs from other *Prunus* species (Figure 3B). However, *PsDAM6* only have six exons, which was different from *DAM6* genes in other *Prunus* species. In addition, *PsDAM6* is much longer due to the large size of the fourth intron, when compared to other *Prunus* species (Figure 3B). Sequence analysis of *PsDAM6* gene from 'Sanyueli' and 'Furongli' plum indicated there were several insertions, including two large insertions (1327bp and 6493bp) (Figure 3C and Supplementary Fig. 6). The 6493bp-fragment insertion was detected in exon5 of *PsDAM6* gene from 'Sanyueli', which causes the loss of exon5 (Figure 3C and Supplementary Fig. 6). Furthermore, a fragment in the seventh intron of *PsDAM6* from 'Furongli' was reverse-inserted into the first intron of *PsDAM6* from 'Sanyueli' (Figure 3C and Supplementary Fig. 6).

Phylogenetic analysis indicated that PsDAMs group into a clade with DAMs from other *Prunus* species and DAM1-DAM6 sequences formed a subgroup with the respective sequences from other *Prunus* species (Figure 4). Multiple sequence alignment of plum PsDAMs with DAMs from Prunus species indicated that all *Prunus* DAM proteins contain conserved MADS domain, I domain and K domain(Figure 5). A conserved EAR motif, which acts as a repression domain, was also detected in the C-terminal of these DAM proteins (Figure 5). The deletion of exon5 led to the lack of 14 amino acid residues, which belong to K3 helix in the K domain (Figure 5). These results suggested that DAMs are conserved in *Prunus* species and plum PsDAMs may have similar function with their homologs from *Prunus* species.

Transcriptome of Dormancy Release in plum flower bud

To investigate the molecular processes and genes involved in the regulation of flower bud dormancy, we

compared the RNA-seq data derived from flower buds of ‘Sanyueli’ plum (low-chill) and ‘Furongli’ plum (high-chill) during chilling treatment. In total, 153.60Gb clean data were obtained and 85.38% to 92.02% of the clean reads from the libraries were successfully mapped to the genome of ‘Sanyueli’ (Supplementary Table 13).

Principal component analyses (PCA) showed that for each RNA-seq experiment, biological replicate samples tended to cluster together (Supplementary Fig.7). Gene differential expression analysis identified 7,782 DEGs with 3844 DEGs in F1 vs S1, 1,313 DEGs in F1 vs F2, 1,096 DEGs in F2 vs F3, 2,213 DEGs in F1 vs F3, 1,961 DEGs in S1 vs S2 and 1595 DEGs in S1 vs S3, respectively (Supplementary Table 14). However, only 288 DEGs were detected in S2 vs S3. KEGG enrichment analysis indicated that phenylpropanoid biosynthesis was one of the most enriched pathway in all comparisons except S2 vs S3 (Supplementary Fig.8). Starch and sucrose metabolism and plant hormone signal transduction pathways were also enriched in F2 vs F3 and F1 vs F3 (Supplementary Fig.8).

The dramatic changes in transcriptome of flower bud during chilling treatment suggested a number of differentially-expressed genes indicateding that the fulfillment of chilling requirement is a complex process that involves a series of transcription regulatory events. To identify candidate transcription factors that may participate in this process, we employed PlantTFDB v5.0 to predict transcription factors from the expressed transcripts. In total, 1,687 transcripts were predicted to encode transcription factors (Supplementary Table 15) and 511 of them were differentially expressed (Supplementary Table 16). All six *PsDAMs* were identified as DEGs (Fig. 6). *PsDAM1* and *PsDAM3* showed a similar expression pattern, their expression only significantly downregulated in the buds of ‘Furongli’ at F3 stage and unchanged in the buds of ‘Sanyueli’. The expression profile of *PsDAM2*, *PsDAM4* and *PsDAM5* was similar in both cultivars during cold treatment. The expression of *PsDAM2* decreased in both buds of ‘Furongli’ and ‘Sanyueli’. Both the transcript levels of *PsDAM4* and *PsDAM5* were increased by cold treatment and then decreased. However, the expression of *PsDAM4* is much higher than that of *PsDAM5*. The expssion pattern of *PsDAM6* in flower buds of ‘Furongli’ during cold treatment was similar to that of *PsDAM4* and *PsDAM5*, but it is noteworthy that its expression in flower buds of ‘Sanyueli’ was much lower than that in flower buds of ‘Furongli’ and was significantly repressed after treated with chilling temperature for 170h (Fig. 6). Thus, the expression of *PsDAM6* coincides with release of dormancy in flower buds of ‘Sanyueli’ and ‘Furongli’ plum.

To search for genes that may regulate the expression of *PsDAM6* or regulated by *PsDAM6*, the correlation between the expression of *PsDAM6* and other differentially expressed genes was calculated. The expression of 190 differentially expressed genes showed a high significant correlation ($|r| > 0.90$) with that of *PsDAM6* (Supplementary Table 17). Some of them was low expressed or not expressed in the flower bud of ‘Sanyueli’ (Fig. 7), including three transcription factors NAC (PsSY0001238), TT2-like (PsSY0001884) and LOB domain-containing protein (PsSY0011216). Interestingly, a transcriptional activator DEMETER-like protein encoding gene (PsSY0014977) was expressed well in flower buds of ‘Furongli’, but it was barely expressed in flower buds of ‘Sanyueli’.

Discussion

Plums are widely grown for their fruits, which have excellent taste, nutritive value and processing compatibility[1]. Although Chinese plum is one of the most economically important stone fruits, its genome sequence has not been reported to date. In the present study, a 308.06 Mb genome of low-chill Chinese plum ‘Sanyueli’ with N50

of 815.7 kb was generated. The presence of the majority of BUSCO genes (97.01%) in our assambly indicated that the quality of 'Sanyueli' plum genome is comparable to that of sweet cherry (96.0%)[64], almond(96.0%)[66] and apricot(98.0%)[68]. The genome size of 'Sanyueli' plum is larger than that of mume[62], peach[63], almond[66] and apricot (38.28%)[68]. 56.42% of the genome was predicted to be repetitive, which is much higher than that observed in other *Prunus* species, including flowering cherry (47.2%)[65], mume (45.0%)[62], sweet cherry (43.8%)[64], almond (34.6%)[66], apricot (38.28%)[68] and peach (29.6%)[63]. The higher proportion of repetitive sequences (especially LTR insertions and expansions) may contribute to the larger genome size of 'Sanyueli' plum. 30,159 genes were predicted from the assembled genome of 'Sanyueli' plum, which is similar to that in apricot (30,436)[68] and mume (31,390)[62], but more than that in almond (27,969)[66] and peach (27,852)[63] and less than that in and flowering cherry (41,294)[65] and sweet cherry (43,349)[64]. The assembled 'Sanyueli' plum genome provides important information for understanding the evolution of *Prunus* species and the molecular basis of important agronomic traits and molecular assisted breeding of plum.

As global warming increases, unveiling the mechanism of flower bud dormancy and chilling requirement and the development of low-chill cultivars is of great importance for sustained production of plum. Low-chill cultivars of deciduous fruit trees have been used to identify genes related to bud domancy regulation[31,32,39,47,55,120,121]. Among the identified candidates, *DAM* genes were proposed as the central regulators controlling bud dormancy[14,33]. In the present study, *PsDAM* genes were found to be arranged in six tandem repeats in the genome of plum similar to *DAM* genes in peach and Japanese apricot[32,62]. Recently, six *PavDAM* genes was identifed in sweet cherry by Masuda, et al.[122]. Blast analysis results indicated there are six *DAM* genes which were also arranged in six tandem repeats in the genome of *P. armeniaca* and *P. dulcis.* Sequence analysis showed that *DAM* genes from *Prunus* species shared similar gene structure and their predicted protein sequences were highly homologous to each other. These results suggested that *DAM* genes were conserved during the evolution of *Prunus* family and they may shared similar function. *PsDAM2*, *PsDAM4* and *PsDAM5* showed similar expression patterns in high-chill cultivar 'Furongli' and low-chill cultivar 'Sanyueli'. Although the expression pattern of *PsDAM1* and *PsDAM3* was different between 'Furongli' and 'Sanyueli', their expression in 'Sanyueli' was not repressed by cold temperature. These results suggested that *PsDAM1- PsDAM5* are not responsible for the difference of chilling requirement between the cultivars. In peach, the expression of *DAM5* and *DAM6* is well correlated with the dormancy status of buds, highly expressed in dormant buds but lowly expressed at buds of which the chilling requirements has been fulfilled[33,123]. In addition, the expression of *PpDAM6* in buds of low-chill 'Okinawa (Tsukuba)' was lower and decreased earlier than that in high-chill 'Akatsuki'[124]. Consistent with this result, our results demonstrated that the transcript level of *PsDAM6* was extremely low in buds of 'Sanyueli' and signifcantly decreased after 170h of cold treatment, while its expression in buds of 'Furongli' was approximately ten times higher than that in buds of 'Sanyueli'. Furthermore, *PsDAM6* was upregulated at 290h and decreased to a level comparable to untreated control at 530h after cold treatment. Fan, et al.[31] proposed that PpDAM5 and PpDAM6 could act as a dose-dependent inhibitor to peach bud break. Yamane, et al.[125] indicated that break competency of dormant buds was significantly repressed in transgenic apple lines overexpressing *PmDAM6* from Japanese apricot. These results suggested that *PsDAM6* was the most significant candidate responsible for chilling requirement in plum.

Our resuslts indicated that a large insertion in exon5 of *PsDAM6* in 'Sanyueli' resulted in partial deletion of the K domain, which is important for protein-protein interaction[126,127]. In Japanese apricot, Zhao, et al.[36] demonstrated that PmDAM1 and PmDAM6 were shown to form heteromeric complexes with the cold response C-repeat binding factor PmCBF5. In addition, apricot PmuDAM6 was shown to interact with PmuSOC1[41].

Recently, similar interactions between PavDAM1/5 and the dormany-associated SUPPRESSOR OF OVEREXPRESSION OF CO1 PavSOC1 were reported by Wang, et al.[118] in sweet cherry. The authors suggested these interactions between DAMs and SOCs could play a role in bud dormancy regulantion. However, it remains unclear whether the the lack of exon5 affects the function of *PsDAM6* gene in 'Sanyueli'.

Activation of *DAM* genes should be fine regulated in consideration of their role as dose-dependent inhibitor in the regulation of bud break[31]. Recently, transcription factors, such as peach PpTCP20 and pear PpyABF3 has been shown to play roles in control of bud dormancy by regulating the expression of *DAM* genes[128,129]. Our results showed that the transcript abundance of three genes predicted to encode transcription factors, NAC, TT2-like and LOB domain-containing protein, was positively correlated with that of *PsDAM6*. Conrad, et al.[58] suggested that the phenylpropanoid pathway was associated with dormancy in apricot. Consistent with their results, transcriptome analysis results showed that DEGs were enriched in phenylpropanoid biosynthesis. TT2 is an R2R3 MYB activator involved in the regulation ofproanthocyanidin accumulation in arabidopsis[130,131]. The differentially expressed TT2-like identified in this study could be responsible for transcriptional regulation of structural genes in phenylpropanoid biosynthesis pathway. *NACs* have been demonstrated to be associated with dormancy release in seveval species, including peach[132], pear[49,133], grapre[25,134] and peony[135-137]. Kumar, et al.[24] reported that a NAC domain containing protein encoding gene was demethylated during the chilling acquisition. In addition Tuan, et al.[138] indicated that NAC, as a cofactor of PpAREB1 repressed expression of *PpDAM1* in endodormancy release of pear buds. The differential expression of LOB domain-containing protein encoding genes has been reported during bud break of grape[25,139]. In sweet cherry, *PavLOB* was found to be highly expressed during endodormancy and around the time of dormancy release[53]. Epigenetic mechanisms were shown to participate in regulation of bud dormancy[55,140]. Several studies have reported that histone modification and DNA methylation are involved in transcriptional regulation of *DAM* genes during bud dormancy[28,48,51,119,140]. A gene predicted to encode DEMETER-like proetin was found to be highly expressed in the flower buds of 'Furongli' and expressed at an extremely low level in the flower buds of 'Sanyueli'. Conde, et al.[141] indicated that overexpression of DNA demethylase *CsDML*, a homolog of *PtaDML6*, significantly enhanced flavonoid accumulation through activating the expression of flavonoid biosynthesis genes and accelerated short□day□induced bud formation in poplar. Later, they demonstrated that a reduction of gDNA methylation in apex tissue during bud break is accompanied by the chilling-induced expression of *PtaDML10* and knock-down of *PtaDML8/10* delayed bud break in poplar. In addition, they further showed that the gene targets of DML-dependent DNA demethylation are genetically associated with bud break[142]. These results suggested DML participated in the control of bud dormancy through modulating the methylation level of related genes. Further study will be required to determine wether the DEMETER-like gene is involved in regulating the expression of dormancy-related genes, such as *PsDAM6*, and flower bud dormancy in plum.

Several insertions was found in the introns of PsDAM6 in low-chill plum 'Sanyueli'. Falavigna, et al.[14] proposed that intronic regions of *DAM* genes may function in regulation of their transcription. Large insertions was observed in the first intron of both *PpDAM5* and *PpDAM6* in low-chill peach and was suggested to be linked to lower chilling requirements for dormancy release[124,143]. Saito, et al.[144] reported a 3218 bp insertion in the first intron of *DAM* gene *MADS13-1*from low-chill pear 'Hengshanli'. However, they found that the insertion exists in in a high-chill pear. Whether these insertions in the intron and exon region affect the expression and function of *PsDAM6* in flower buds of 'Sanyueli' requires further studies.

Conclusions

In summary, we first report the sequencing, assembly and annotation of the genome of chinese plum ‘Sanyueli’, which is an extremely low-chilling requirement plum cultivar. Six PsDAM genes were identified in plum genome. Transcriptome analysis suggested that *PsDAM6* was the key candidate responsible for the low-chilling requirement in plum. The genome of ‘Sanyueli’ plum provides a valuable resource for further research on genetic basis of agronomic traits, such as chilling requirement and the genetic improvement of plum.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (31801916), the Basic Scientific Research Funds of Public Welfare Scientific Research Institutes of Fujian Province (2018R1013-1), the China Scholarship Council (CSC) (201909350001) and Projects of Fujian Academy of Agricultural Sciences(YC2015-9, STIT2017-1-4 and AGY2018-3).

Table 1 Statistics of 'Sanyueli' plum genome assembly and annotation.

Genome size (Mb)	307.29
Number of contigs (≥1 kb)	1752
N50 contig length (bp)	815,770
N90 contig length (bp)	50,776
Largest contig (bp)	7,898,106
GC content (%)	37.78
Number of gene models	30,159
Gene length (bp)	108,356,597
Mean gene length (bp)	3592.8
Total coding region length (bp)	39,130,040
Mean coding region length (bp)	1297.4
Total intron length (bp)	58,765,475
Mean intron length (bp)	437.5

Fig. 1 Images of 'Sanyueli' plum used in this study. A, B and C indicate fruit of 'Sanyueli'. D, a 'Sanyueli' plum tree with flowers (E) and fruits (F), the photo was took in Zhangpu county (Zhangzhou, Fujian province, China) in 5th, February, 2016.

Fig. 2 Comparative genomic analysis of plum and other species.

A, Phylogenetic analysis and gene family cluster analysis of plum and and 14 other species. The phylogenetic tree (left panel) was constructed using 436 single-copy genes from plum and 14 other species, including *P. armeniaca*, *P. persica*, *P. avium*, *P. mume*, *P. dulcis*, *P. bretschneideri* , *Malus x domestica*, *Fragaria vesca*、*Rosa chinensis*、*Vitis vinifera*, *Populus trichocarpa*, *Oryza sativa*, *Arabidopsis thaliana* and *Amborella trichopoda*. The numbers indicate the divergence time. Gene family expansions and contractions are indicated by numbers in yellow and blue boxes. The distribution of paralogous genes in the analyzed plants is indicated in the right panel. B, Distribution of synonymous nucleotide subsitutions (Ks) among *P. salicina*, *Malus x domestica* and *Vitis vinifera*. Duplication components were identified by MCScan without tandem duplicated pairs. Colored curves superimposed on the Ks distribution represent different component identified. Ks distributions of paralogous gene pairs were identified from syntenic blocks between species, and orthologous gene pairs were identified from duplication component within one genome. C, Venn diagram comparison of gene families in six *Prunus* species.

Fig. 3 *PsDAM* genes in plum and other Rosaceae species. A, Overview of the *DAM* locus in the genome of 'Sanyueli' plum and other Rosaceae species. *DAM* genes show synteny are indicated in same color. B, Schematic overview of introns and exons in *DAM* genes from plum and other Rosaceae species. C, Structural alignments of the *PsDAM6* gene from 'Sanyueli' and 'Furongli' plum. Boxes and lines represent exons and introns, respectively.

Fig. 4 Phylogenetic analysis of predicted amino acid sequences of DAM genes from plum and other Rosaceae species. DAMs in plum are colored by red. PsDAM6-sy and PsDAM6-fr represent PsDAM6 from 'Sanyueli' and 'Furongli' plum, respectively.

Fig. 5 Multiple alignment of predicted amino acid sequences of DAM genes from plum and other *Prunus* species. Plum PsDAMs are indicated in red font. PsDAM6-sy and PsDAM6-fr represent PsDAM6 from 'Sanyueli' and 'Furongli' plum, respectively. The MADS domain, I domain, and K domain are highlighted in grey, blue and green colors on bottom of the alignment. Three putative amphipathic a-helices, K1, K2, and K3, are indicated by arrows. EAR (ethylene-responsive element-binding factor-associated amphiphilic repression) motif is shown by lines on bottom of the alignment. Amino acids lost in PsDAM6-sy was indicated in red box.

Fig. 6 The expression profile of PsDAMs in dormant flower bud of 'Sanyueli' and 'Furongli' plum treated with chilling temperature. S and F represent 'Sanyueli' and 'Furongli', respectively.

Fig. 7 Expression profiles of selected differentially expressed genes correlated with PsDAM6. The expression levels of differentially expressed genes are shown as FPKM values.

Supplementary Fig. 1 Daily maximum temperature and minimum temperature from November 2015 to February 2016.

Supplementary Fig. 2 Bud break of cuttings of plum trees after treated with chilling temperature. A,‘Sanyueli’. B, 'Furongli’. The pictures were taken at 20 days after being transferred to and kept at 25 ± 1°C with white light (150 mol m^{-2} s^{-1}) under a 14 h light/10 h dark photoperiod at 75% humidity.

Supplementary Fig. 3 Distribution frequency of the 19-kmer graph for genome size estimation. Density plot of the frequency of unique 19-kmer for each kmer depth (x axis) is plotted.

Supplementary Fig. 4 Distribution map of genes from the three prediction methods

Supplementary Fig. 5 Collinear analysis of P. salicina, P. armeniaca and P. mume genome

Supplementary Fig. 6 Alignment of genomic DNA sequences of the *PsDAM6* from ‘Sanyueli’ and ‘Furongli’ plum. Exons are highlighted in green shading and different bases between the *PsDAM6* from ‘Sanyueli’ and ‘Furongli’ plum are indicated in red font. The genomic DNA sequence of *PsDAM6* in ‘Furongli’ plum was extracted from the genome of ‘Furongli’.

Supplementary Fig. 7 Principal component analyses of the RNA-Seq samples.

Supplementary Fig. 8 KEGG pathway enrichment analysis of the annotated DEGs.

References

1. Milošević, T. & Milošević, N. in *Advances in Plant Breeding Strategies: Fruits: Volume 3* (eds Jameel M. Al-Khayri, Shri Mohan Jain, & Dennis V. Johnson) 165-215 (Springer International Publishing, 2018).
2. Zhang, J. & zhou, E. *China Fruit-plant Monograph Chinese Plum*. (China Forestry Press, 1998).
3. Jiang, H., Yang, J., Ferguson, D. K., Li, Y., Wang, C., Li, C.-S. & Liu, C. Fruit stones from Tiao Lei's tomb of Jiangxi in China, and their palaeoethnobotanical significance. *J. Archaeol. Sci.* **40**, 1911-1917, (2013).
4. Topp, B. L., Russell, D. M., Neumüller, M., Dalbó, M. A. & Liu, W. in *Fruit Breeding* (eds Marisa Luisa Badenes & David H. Byrne) 571-621 (Springer US, 2012).
5. Okie, W. R. & Hancock, J. F. in *Temperate Fruit Crop Breeding: Germplasm to Genomics* (ed James F. Hancock) 337-358 (Springer Netherlands, 2008).
6. Atkinson, C. J., Brennan, R. M. & Jones, H. G. Declining chilling and its impact on temperate perennial crops. *Environ. Exp. Bot.* **91**, 48-62, (2013).
7. Wenden, B., Campoy, J. A., Lecourt, J., López Ortega, G., Blanke, M., Radičević, S., Schüller, E., Spornberger, A., Christen, D., Magein, H., Giovannini, D., Campillo, C., Malchev, S., Peris, J. M., Meland, M., Stehr, R., Charlot, G. & Quero-García, J. A collection of European sweet cherry phenology data for assessing climate change. *Sci. Data* **3**, 160108, (2016).
8. Ruiz, D., Egea, J., Salazar, J. A. & Campoy, J. A. Chilling and heat requirements of Japanese plum cultivars for flowering. *Sci. Hortic.* **242**, 164-169, (2018).
9. Liu, W. in *VIII International Symposium on Plum and Prune Genetics, Breeding and Pomology 734*. 89-92.
10. Campoy, J. A., Ruiz, D. & Egea, J. Dormancy in temperate fruit trees in a global warming context: A review. *Sci. Hortic.* **130**, 357-372, (2011).
11. Okie, W. R. & Ramming, D. W. Plum Breeding Worldwide. *HortTechnology* **9**, 162-176, (1999).
12. Sherman, W. B., Topp, B. L. & Lyrene, P. M. 317 edn 149-154 (International Society for Horticultural Science (ISHS), Leuven, Belgium).
13. Sherman, W. B. & Lyrene, P. M. 622 edn 599-605 (International Society for Horticultural Science (ISHS), Leuven, Belgium).
14. Falavigna, V. d. S., Guitton, B., Costes, E. & Andrés, F. I want to (bud) break free: the potential role of dam and *SVP*-like genes in regulating dormancy cycle in temperate fruit trees. *Front. Plant Sci.* **9**, (2019).
15. Arora, R., Rowland, L. J. & Tanino, K. Induction and release of bud dormancy in woody perennials: a science comes of age. *HortScience* **38**, 911-921, (2003).
16. Lundell, R., Hänninen, H., Saarinen, T., Åström, H. & Zhang, R. Beyond rest and quiescence (endodormancy and ecodormancy): A novel model for quantifying plant-environment interaction in bud dormancy release. *Plant Cell Environ.* **43**, 40-54, (2019).
17. Miotto, Y. E., Tessele, C., Czermainski, A. B. C., Porto, D. D., Falavigna, V. d. S., Sartor, T., Cattani, A. M., Delatorre, C. A., de Alencar, S. A., da Silva-Junior, O. B., Togawa, R. C., Costa, M. M. d. C., Pappas, G. J., Grynberg, P., de Oliveira, P. R. D., Kvitschal, M. V., Denardi, F., Buffon, V. & Revers, L. F. Spring is coming: genetic analyses of the bud break date locus reveal candidate genes from the cold perception pathway to dormancy release in apple (*Malus* × *domestica* Borkh.). *Front. Plant Sci.* **10**, 33, (2019).
18. Artlip, T., McDermaid, A., Ma, Q. & Wisniewski, M. Differential gene expression in non-transgenic and transgenic "M.26" apple overexpressing a peach *CBF* gene during the transition from eco-dormancy to bud break. *Hort. Res.* **6**, 86, (2019).
19. Wu, R., Tomes, S., Karunairetnam, S., Tustin, S. D., Hellens, R. P., Allan, A. C., Macknight, R. C. & Varkonyi-Gasic, E. *SVP*-like MADS box genes control dormancy and budbreak in apple. *Front. Plant Sci.* **8**,

477, (2017).

20. Porto, D. D., da Silveira Falavigna, V., Arenhart, R. A., Perini, P., Buffon, V., Anzanello, R., dos Santos, H. P., Fialho, F. B., de Oliveira, P. R. D. & Revers, L. F. Structural genomics and transcriptional characterization of the Dormancy-Associated MADS-box genes during bud dormancy progression in apple. *Tree Genet. Genomes* **12**, 46, (2016).
21. Allard, A., Bink, M. C. A. M., Martinez, S., Kelner, J.-J., Legave, J.-M., di Guardo, M., Di Pierro, E. A., Laurens, F., van de Weg, E. W. & Costes, E. Detecting QTLs and putative candidate genes involved in budbreak and flowering time in an apple multiparental population. *J. Exp. Bot.* **67**, 2875-2888, (2016).
22. Porto, D. D., Bruneau, M., Perini, P., Anzanello, R., Renou, J.-P., Santos, H. P. d., Fialho, F. B. & Revers, L. F. Transcription profiling of the chilling requirement for bud break in apples: a putative role for *FLC*-like genes. *J. Exp. Bot.* **66**, 2659-2672, (2015).
23. Celton, J. M., Martinez, S., Jammes, M. J., Bechti, A., Salvi, S., Legave, J. M. & Costes, E. Deciphering the genetic determinism of bud phenology in apple progenies: a new insight into chilling and heat requirement effects on flowering dates and positional candidate genes. *New Phytol.* **192**, 378-392, (2011).
24. Kumar, G., Rattan, U. K. & Singh, A. K. Chilling-mediated dna methylation changes during dormancy and its release reveal the importance of epigenetic regulation during winter dormancy in apple (*Malus x domestica* Borkh.). *PLOS ONE* **11**, e0149934, (2016).
25. Tang, L., Chhajed, S., Vashisth, T., Olmstead, M. A., Olmstead, J. W. & Colquhoun, T. A. Transcriptomic study of early responses to the bud dormancy-breaking agent hydrogen cyanamide in 'TropicBeauty' peach. *J. Am. Soc. Hortic. Sci.* **144**, 244, (2019).
26. Wang, D., Gao, Z., Du, P., Xiao, W., Tan, Q., Chen, X., Li, L. & Gao, D. Expression of ABA metabolism-related genes suggests similarities and differences between seed dormancy and bud dormancy of peach (*Prunus persica*). *Front. Plant Sci.* **6**, 1248, (2016).
27. Sun, M.-Y., Fu, X.-L., Tan, Q.-P., Liu, L., Chen, M., Zhu, C.-Y., Li, L., Chen, X.-D. & Gao, D.-S. Analysis of basic leucine zipper genes and their expression during bud dormancy in peach (*Prunus persica*). *Plant Physio. Bioch.* **104**, 54-70, (2016).
28. de la Fuente, L., Conesa, A., Lloret, A., Badenes, M. L., Ríos, G. J. T. G. & Genomes. Genome-wide changes in histone H3 lysine 27 trimethylation associated with bud dormancy release in peach. *Tree Genet. Genomes* **11**, 45, (2015).
29. Leida, C., Romeu, J. F., García-Brunton, J., Ríos, G. & Badenes, M. L. Gene expression analysis of chilling requirements for flower bud break in peach. *Plant Breed.* **131**, 329-334, (2012).
30. Leida, C., Conejero, A., Arbona, V., Gómez-Cadenas, A., Llácer, G., Badenes, M. L. & Ríos, G. Chilling-dependent release of seed and bud dormancy in peach associates to common changes in gene expression. *PLoS ONE* **7**, e35777, (2012).
31. Fan, S., Bielenberg, D. G., Zhebentyayeva, T. N., Reighard, G. L., Okie, W. R., Holland, D. & Abbott, A. G. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). *New Phytol.* **185**, 917-930, (2010).
32. Bielenberg, D. G., Wang, Y., Li, Z., Zhebentyayeva, T., Fan, S., Reighard, G. L., Scorza, R., Abbott, A. G. J. T. G. & Genomes. Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. *Tree Genet. Genomes* **4**, 495-507, (2008).
33. Yamane, H., Ooka, T., Jotatsu, H., Hosaka, Y., Sasaki, R. & Tao, R. Expressional regulation of *PpDAM5* and *PpDAM6*, peach (*Prunus persica*) dormancy-associated MADS-box genes, by low temperature and dormancy-breaking reagent treatment. *J. Exp. Bot.* **62**, 3481-3488, (2011).

34. Li, Z., Reighard, G. L., Abbott, A. G. & Bielenberg, D. G. Dormancy-associated MADS genes from the *EVG* locus of peach [*Prunus persica* (L.) Batsch] have distinct seasonal and photoperiodic expression patterns. *J. Exp. Bot.* **60**, 3521-3530, (2009).
35. Ito, A., Tuan, P. A., Saito, T., Bai, S., Kita, M. & Moriguchi, T. Changes in phytohormone content and associated gene expression throughout the stages of pear (*Pyrus pyrifolia* Nakai) dormancy. *Tree Physiol.*, (2019).
36. Zhao, K., Zhou, Y., Ahmad, S., Yong, X., Xie, X., Han, Y., Li, Y., Sun, L. & Zhang, Q. PmCBFs synthetically affect *PmDAM6* by alternative promoter binding and protein complexes towards the dormancy of bud for *Prunus mume*. *Sci. Rep.* **8**, 4527, (2018).
37. Zhao, K., Zhou, Y., Ahmad, S., Xu, Z., Li, Y., Yang, W., Cheng, T., Wang, J. & Zhang, Q. Comprehensive cloning of *Prunus mume* dormancy associated MADS-box genes and their response in flower bud development and dormancy. *Front. Plant Sci.* **9**, 17, (2018).
38. Zhang, Z., Zhuo, X., Zhao, K., Zheng, T., Han, Y., Yuan, C. & Zhang, Q. Transcriptome profiles reveal the crucial roles of hormone and sugar in the bud dormancy of *Prunus mume*. *Sci. Rep.* **8**, 5090, (2018).
39. Kitamura, Y., Habu, T., Yamane, H., Nishiyama, S., Kajita, K., Sobue, T., Kawai, T., Numaguchi, K., Nakazaki, T., Kitajima, A., Tao, R. J. T. G. & Genomes. Identification of QTLs controlling chilling and heat requirements for dormancy release and bud break in Japanese apricot (*Prunus mume*). *Tree Genet. Genomes* **14**, 33, (2018).
40. Wen, L. H., Zhong, W. J., Huo, X. M., Zhuang, W. B., Ni, Z. J. & Gao, Z. H. Expression analysis of ABA- and GA-related genes during four stages of bud dormancy in Japanese apricot (*Prunus mume* Sieb. et Zucc). *J. Hortic. Sci. Biotech.* **91**, 362-369, (2016).
41. Kitamura, Y., Takeuchi, T., Yamane, H. & Tao, R. Simultaneous down-regulation of *DORMANCY-ASSOCIATED MADS-box6* and *SOC1* during dormancy release in Japanese apricot (*Prunus mume*) flower buds. *J. Hortic. Sci. Biotech.* **91**, 476-482, (2016).
42. Yang, Q., Niu, Q., Tang, Y., Ma, Y., Yan, X., Li, J., Tian, J., Bai, S. & Teng, Y. *PpyGAST1* is potentially involved in bud dormancy release by integrating the GA biosynthesis and ABA signaling in 'Suli' pear (*Pyrus pyrifolia* White Pear Group). *Environ. Exp. Bot.* **162**, 302-312, (2019).
43. Yang, Q., Niu, Q., Li, J., Zheng, X., Ma, Y., Bai, S. & Teng, Y. PpHB22, a member of HD-Zip proteins, activates *PpDAM1* to regulate bud dormancy transition in 'Suli' pear (*Pyrus pyrifolia* White Pear Group). *Plant Physio. Bioch.* **127**, 355-365, (2018).
44. Gabay, G., Faigenboim, A., Dahan, Y., Izhaki, Y., Itkin, M., Malitsky, S., Elkind, Y. & Flaishman, M. A. Transcriptome analysis and metabolic profiling reveal the key role of α-linolenic acid in dormancy regulation of European pear. *J. Exp. Bot.*, ery405-ery405, (2018).
45. Niu, Q., Li, J., Cai, D., Qian, M., Jia, H., Bai, S., Hussain, S., Liu, G., Teng, Y. & Zheng, X. Dormancy-associated MADS-box genes and microRNAs jointly control dormancy transition in pear (*Pyrus pyrifolia* white pear group) flower bud. *J. Exp. Bot.* **67**, 239-257, (2016).
46. Anh Tuan, P., Bai, S., Saito, T., Imai, T., Ito, A. & Moriguchi, T. Involvement of *EARLY BUD-BREAK*, an AP2/ERF transcription factor gene, in bud break in Japanese pear (*Pyrus pyrifolia* Nakai) lateral flower buds: expression, histone modifications and possible target genes. *Plant Cell Physiol.* **57**, 1038-1047, (2016).
47. Takemura, Y., Kuroki, K., Shida, Y., Araki, S., Takeuchi, Y., Tanaka, K., Ishige, T., Yajima, S. & Tamura, F. Comparative transcriptome analysis of the less-dormant Taiwanese pear and the dormant japanese pear during winter season. *PLoS ONE* **10**, e0139595, (2015).
48. Saito, T., Bai, S., Imai, T., Ito, A., Nakajima, I. & Moriguchi, T. Histone modification and signalling

cascade of the dormancy-associated MADS-box gene, *PpMADS13-1*, in Japanese pear (*Pyrus pyrifolia*) during endodormancy. *Plant Cell Environ.* **38**, 1157-1166, (2015).

49. Liu, G., Li, W., Zheng, P., Xu, T., Chen, L., Liu, D., Hussain, S. & Teng, Y. Transcriptomic analysis of 'Suli' pear (*Pyrus pyrifolia* white pear group) buds during the dormancy by RNA-Seq. *BMC Genomics* **13**, 700, (2012).
50. Gabay, G., Dahan, Y., Izhaki, Y., Isaacson, T., Elkind, Y., Ben-Ari, G. & Flaishman, M. A. Identification of QTLs associated with spring vegetative budbreak time after dormancy release in pear (*Pyrus communis* L.). *Plant Breed.* **136**, 749-758, (2017).
51. Rothkegel, K., Sánchez, E., Montes, C., Greve, M., Tapia, S., Bravo, S., Prieto, H. & Almeida, A. M. DNA methylation and small interference RNAs participate in the regulation of MADS-box genes involved in dormancy in sweet cherry (*Prunus avium* L.). *Tree Physiol.* **37**, 1739-1751, (2017).
52. Castède, S., Campoy, J. A., Le Dantec, L., Quero-García, J., Barreneche, T., Wenden, B. & Dirlewanger, E. Mapping of candidate genes involved in bud dormancy and flowering time in sweet cherry (*Prunus avium*). *PLoS ONE* **10**, e0143250, (2015).
53. Vimont, N., Fouché, M., Campoy, J. A., Tong, M., Arkoun, M., Yvin, J.-C., Wigge, P. A., Dirlewanger, E., Cortijo, S. & Wenden, B. From bud formation to flowering: transcriptomic state defines the cherry developmental phases of sweet cherry bud dormancy. *BMC Genomics* **20**, 974, (2019).
54. Vimont, N., Quah, F. X., Schöepfer, D. G., Roudier, F., Dirlewanger, E., Wigge, P. A., Wenden, B. & Cortijo, S. ChIP-seq and RNA-seq for complex and low-abundance tree buds reveal chromatin and expression co-dynamics during sweet cherry bud dormancy. *Tree Genet. Genomes* **16**, 9, (2019).
55. Prudencio, Á. S., Werner, O., Martínez-García, P. J., Dicenta, F., Ros, R. M. & Martínez-Gómez, P. DNA methylation analysis of dormancy release in almond (*Prunus dulcis*) flower buds using epi-genotyping by sequencing. *Int. J. Mol. Sci.* **19**, 3542, (2018).
56. Prudencio, Á. S., Dicenta, F. & Martínez-Gómez, P. Monitoring dormancy transition in almond [*Prunus Dulcis* (Miller) Webb] during cold and warm mediterranean seasons through the analysis of a *DAM* (*Dormancy-Associated MADS-Box*) gene. *Horticulturae* **4**, 41, (2018).
57. Olukolu, B. A., Trainin, T., Fan, S., Kole, C., Bielenberg, D. G., Reighard, G. L., Abbott, A. G. & Holland, D. Genetic linkage mapping for molecular dissection of chilling requirement and budbreak in apricot (*Prunus armeniaca* L.). *Genome* **52**, 819-828, (2009).
58. Conrad, A. O., Yu, J., Staton, M. E., Audergon, J.-M., Roch, G., Decroocq, V., Knagge, K., Chen, H., Zhebentyayeva, T., Liu, Z., Dardick, C., Nelson, C. D. & Abbott, A. G. Association of the phenylpropanoid pathway with dormancy and adaptive trait variation in apricot (*Prunus armeniaca*). *Tree Physiol.* **39**, 1136-1148, (2019).
59. Balogh, E., Halász, J., Soltész, A., Erös-Honti, Z., Gutermuth, Á., Szalay, L., Höhn, M., Vágújfalvi, A., Galiba, G. & Hegedüs, A. Identification, structural and functional characterization of dormancy regulator genes in apricot (*Prunus armeniaca* L.). *Front. Plant Sci.* **10**, 402, (2019).
60. Zhuang, W., Gao, Z., Wang, L., Zhong, W., Ni, Z. & Zhang, Z. Comparative proteomic and transcriptomic approaches to address the active role of GA4 in Japanese apricot flower bud dormancy release. *J. Exp. Bot.* **64**, 4953-4966, (2013).
61. Sasaki, R., Yamane, H., Ooka, T., Jotatsu, H., Kitamura, Y., Akagi, T. & Tao, R. Functional and expressional analyses of *PmDAM* genes associated with endodormancy in Japanese apricot. *Plant Physiol.* **157**, 485-497, (2011).
62. 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, 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., Gao, Y., Wang, J., Yin, Y., Xu, X., Cheng, T. & Wang, J. The genome of *Prunus mume*. *Nat Commun* **3**, 1318, (2012).

63. Verde, I., Abbott, A. G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M. T., Grimwood, J., Cattonaro, F., Zuccolo, A., Rossini, L., Jenkins, J., Vendramin, E., Meisel, L. A., Decroocq, V., Sosinski, B., Prochnik, S., Mitros, T., Policriti, A., Cipriani, G., Dondini, L., Ficklin, S., Goodstein, D. M., Xuan, P., Del Fabbro, C., Aramini, V., Copetti, D., Gonzalez, S., Horner, D. S., Falchi, R., Lucas, S., Mica, E., Maldonado, J., Lazzari, B., Bielenberg, D., Pirona, R., Miculan, M., Barakat, A., Testolin, R., Stella, A., Tartarini, S., Tonutti, P., Arus, P., Orellana, A., Wells, C., Main, D., Vizzotto, G., Silva, H., Salamini, F., Schmutz, J., Morgante, M. & Rokhsar, D. S. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet* **45**, 487-494, (2013).
64. 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.* **24**, 499-508, (2017).
65. Baek, S., Choi, K., Kim, G.-B., Yu, H.-J., Cho, A., Jang, H., Kim, C., Kim, H.-J., Chang, K. S., Kim, J.-H. & Mun, J.-H. Draft genome sequence of wild *Prunus yedoensis* reveals massive inter-specific hybridization between sympatric flowering cherries. *Genome Biol*. **19**, 127, (2018).
66. Alioto, T., Alexiou, K. G., Bardil, A., Barteri, F., Castanera, R., Cruz, F., Dhingra, A., Duval, H., Fernández i Martí, Á., Frias, L., Galán, B., García, J. L., Howad, W., Gómez-Garrido, J., Gut, M., Julca, I., Morata, J., Puigdomènech, P., Ribeca, P., Rubio Cabetas, M. J., Vlasova, A., Wirthensohn, M., Garcia-Mas, J., Gabaldón, T., Casacuberta, J. M. & Arús, P. Transposons played a major role in the diversification between the closely related almond and peach genomes: results from the almond genome sequence. *Plant J.* **101**, 455-472, (2020).
67. Sánchez-Pérez, R., Pavan, S., Mazzeo, R., Moldovan, C., Aiese Cigliano, R., Del Cueto, J., Ricciardi, F., Lotti, C., Ricciardi, L., Dicenta, F., López-Marqués, R. L. & Møller, B. L. Mutation of a bHLH transcription factor allowed almond domestication. *Science* **364**, 1095-1098, (2019).
68. Jiang, F., Zhang, J., Wang, S., Yang, L., Luo, Y., Gao, S., Zhang, M., Wu, S., Hu, S., Sun, H. & Wang, Y. The apricot (*Prunus armeniaca* L.) genome elucidates Rosaceae evolution and beta-carotenoid synthesis. *Hort. Res.* **6**, 128, (2019).
69. Jung, S., Lee, T., Cheng, C.-H., Buble, K., Zheng, P., Yu, J., Humann, J., Ficklin, S. P., Gasic, K., Scott, K., Frank, M., Ru, S., Hough, H., Evans, K., Peace, C., Olmstead, M., DeVetter, L. W., McFerson, J., Coe, M., Wegrzyn, J. L., Staton, M. E., Abbott, A. G. & Main, D. 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. *Nucleic Acids Res.* **47**, D1137-D1145, (2019).
70. Aranzana, M. J., Decroocq, V., Dirlewanger, E., Eduardo, I., Gao, Z. S., Gasic, K., Iezzoni, A., Jung, S., Peace, C., Prieto, H., Tao, R., Verde, I., Abbott, A. G. & Arús, P. *Prunus* genetics and applications after de novo genome sequencing: achievements and prospects. *Hort. Res.* **6**, 58, (2019).
71. Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H. & Phillippy, A. M. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*. **27**, 722-736, (2017).
72. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760, (2009).
73. Whelan, S. & Goldman, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691-699, (2001).
74. Parra, G., Bradnam, K. & Korf, I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic

genomes. *Bioinformatics* **23**, 1061-1067, (2007).

75. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265-W268, (2007).
76. Han, Y. & Wessler, S. R. MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Res.* **38**, e199-e199, (2010).
77. Price, A. L., Jones, N. C. & Pevzner, P. A. De novo identification of repeat families in large genomes. *Bioinformatics* **21**, i351-i358, (2005).
78. Edgar, R. C. & Myers, E. W. PILER: identification and classification of genomic repeats. *Bioinformatics* **21**, i152-i158, (2005).
79. Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P. & Schulman, A. H. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **8**, 973-982, (2007).
80. Bao, W., Kojima, K. K. & Kohany, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* **6**, 11, (2015).
81. Tarailo-Graovac, M. & Chen, N. Using repeatmasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* **25**, 4.10.11-14.10.14, (2009).
82. Burge, C. & Karlin, S. Prediction of complete gene structures in human genomic DNA11Edited by F. E. Cohen. *J. Mol. Bio.* **268**, 78-94, (1997).
83. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, ii215-ii225, (2003).
84. Majoros, W. H., Pertea, M. & Salzberg, S. L. TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20**, 2878-2879, (2004).
85. Blanco, E., Parra, G. & Guigó, R. Using geneid to Identify Genes. **18**, 4.3.1-4.3.28, (2007).
86. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59, (2004).
87. Keilwagen, J., Wenk, M., Erickson, J. L., Schattat, M. H., Grau, J. & Hartung, F. Using intron position conservation for homology-based gene prediction. *Nucleic Acids Res.* **44**, e89-e89, (2016).
88. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* **11**, 1650-1667, (2016).
89. Kent, W. J. BLAT—The BLAST-Like Alignment Tool. **12**, 656-664, (2002).
90. Haas, B. J., Delcher, A. L., Mount, S. M., Wortman, J. R., Smith Jr, R. K., Hannick, L. I., Maiti, R., Ronning, C. M., Rusch, D. B., Town, C. D., Salzberg, S. L. & White, O. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654-5666, (2003).
91. Tang, S., Lomsadze, A. & Borodovsky, M. Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res.* **43**, e78-e78, (2015).
92. Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., White, O., Buell, C. R. & Wortman, J. R. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **9**, R7, (2008).
93. Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., Kiryutin, B., Galperin, M. Y., Fedorova, N. D. & Koonin, E. V. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* **29**, 22-28, (2001).
94. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27-30, (2000).
95. Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric,

G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L. L., Eddy, S. R., Bateman, A. & Finn, R. D. The Pfam protein families database. *Nucleic Acids Res.* **40**, D290-D301, (2011).

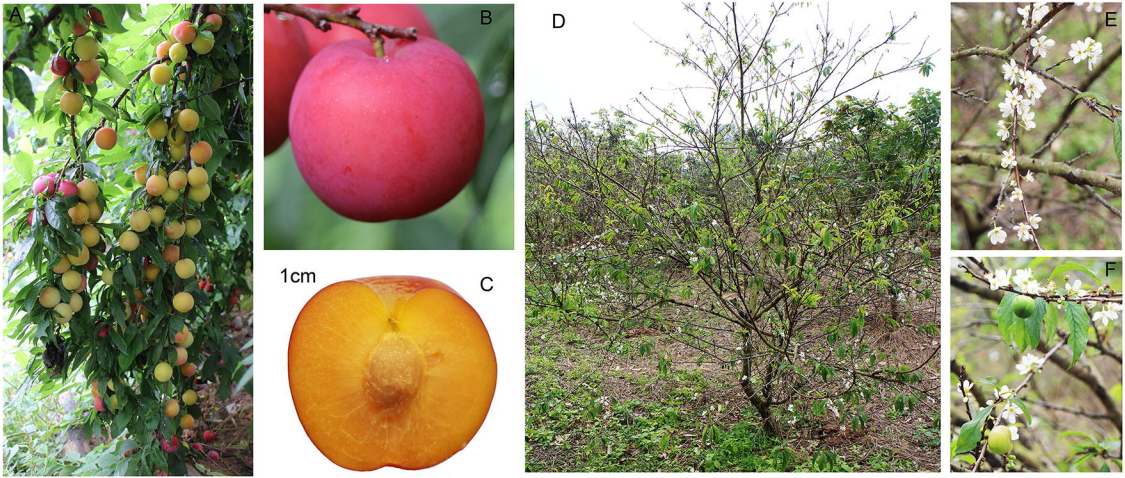
96. Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S. & Schneider, M. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **31**, 365-370, (2003).
97. Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M. & Robles, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674-3676, (2005).
98. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer rna genes in genomic sequence. *Nucleic Acids Res.* **25**, 955-964, (1997).
99. Nawrocki, E. P. & Eddy, S. R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933-2935, (2013).
100. Nawrocki, E. P., Kolbe, D. L. & Eddy, S. R. Infernal 1.0: inference of RNA alignments. *Bioinformatics* **25**, 1335-1337, (2009).
101. Kalvari, I., Argasinska, J., Quinones-Olvera, N., Nawrocki, E. P., Rivas, E., Eddy, S. R., Bateman, A., Finn, R. D. & Petrov, A. I. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* **46**, D335-D342, (2017).
102. Li, L., Stoeckert, C. J. & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178-2189, (2003).
103. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313, (2014).
104. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586-1591, (2007).
105. De Bie, T., Cristianini, N., Demuth, J. P. & Hahn, M. W. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269-1271, (2006).
106. Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., Lee, T.-h., Jin, H., Marler, B., Guo, H., Kissinger, J. C. & Paterson, A. H. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*. **40**, e49-e49, (2012).
107. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792-1797, (2004).
108. Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y. & Xia, R. TBtools - an integrative toolkit developed for interactive analyses of big biological data. *bioRxiv*, 289660, (2020).
109. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-W324, (2014).
110. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357-360, (2015).
111. Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T. & Salzberg, S. L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*. **33**, 290-295, (2015).
112. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550, (2014).
113. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. J. G. b. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol*. **11**, R14, (2010).
114. Mao, X., Cai, T., Olyarchuk, J. G. & Wei, L. J. B. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* **21**, 3787-3793,

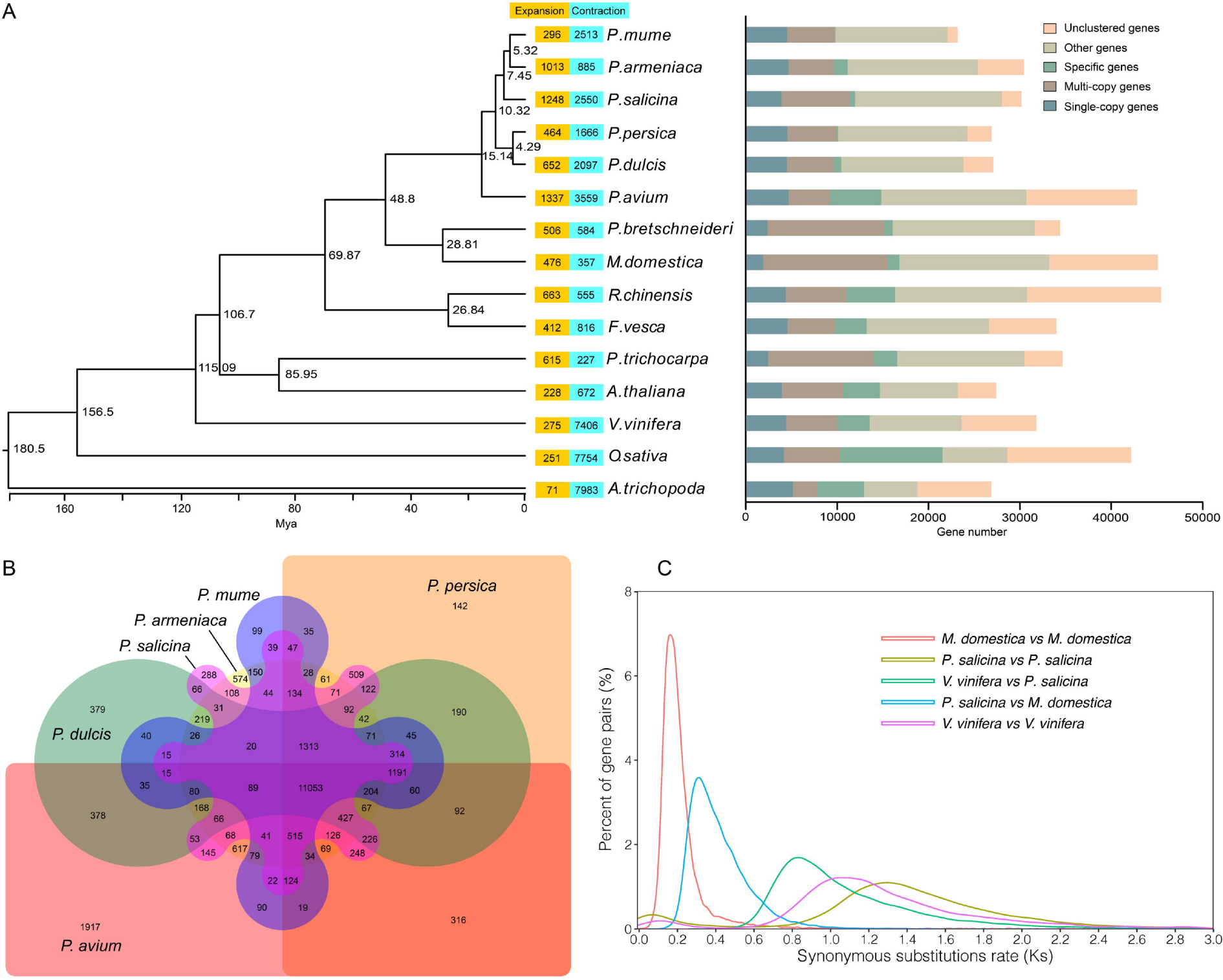
(2005).

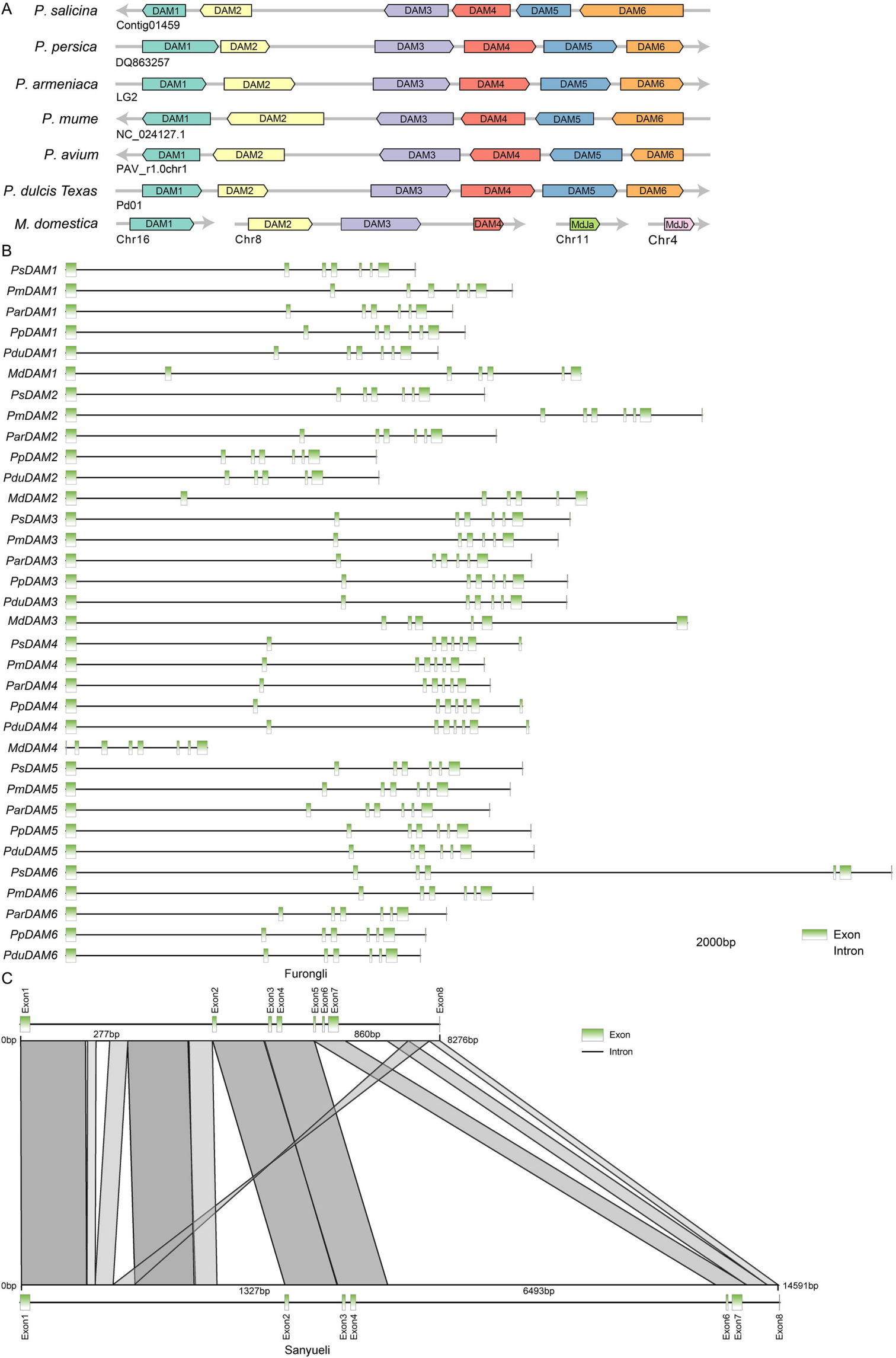
115. He, Y., Peng, W., Li, C., Yang, X., Liu, C., Hao, J. & Xie, T. in *The 2018 Annual Conference of the Chinese Society for Horticultural Science.*
116. Hao, J., YeHua, H., Yang, X., Pan, J., Xu, R., Peng, W. & Liu, C. in *The 2019 Annual Conference and the Celebration of the 90th Anniversary of the Chinese Society for Horticultural Science.* (Chinese Society for Horticultural Science).
117. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212, (2015).
118. Wang, J., Gao, Z., Li, H., Jiu, S., Qu, Y., Wang, L., Ma, C., Xu, W., Wang, S. & Zhang, C. Dormancy-associated MADS-Box (*DAM*) genes influence chilling requirement of sweet cherries and co-regulate flower development with *SOC1* gene. *Int. J. Mol. Sci.* **21**, 921, (2020).
119. Leida, C., Conesa, A., Llácer, G., Badenes, M. L. & Ríos, G. Histone modifications and expression of *DAM6* gene in peach are modulated during bud dormancy release in a cultivar-dependent manner. *New Phytol*. **193**, 67-80, (2012).
120. Calle, A., Cai, L., Iezzoni, A. & Wünsch, A. Genetic dissection of bloom time in low chilling sweet cherry (*Prunus avium* L.) using a multi-family QTL approach. *Front. Plant Sci*. **10**, 1647, (2020).
121. Castède, S., Campoy, J. A., García, J. Q., Le Dantec, L., Lafargue, M., Barreneche, T., Wenden, B. & Dirlewanger, E. Genetic determinism of phenological traits highly affected by climate change in *Prunus avium*: flowering date dissected into chilling and heat requirements. *New Phytol*. **202**, 703-715, (2014).
122. Masuda, K., Yamane, H., Ikeda, K., Tetsumura, T., Takai, M. & Tao, R. 1235 edn 421-426 (International Society for Horticultural Science (ISHS), Leuven, Belgium).
123. Jiménez, S., Reighard, G. L. & Bielenberg, D. G. Gene expression of *DAM5* and *DAM6* is suppressed by chilling temperatures and inversely correlated with bud break rate. *Plant Mol. Biol.* **73**, 157-167, (2010).
124. Yamane, H., Tao, R., Ooka, T., Jotatsu, H., Sasaki, R. & Yonemori, K. J. J. o. t. J. S. f. H. S. Comparative analyses of dormancy-associated MADS-box genes, *PpDAM5* and *PpDAM6*, in low-and high-chill peaches (*Prunus persica* L.). *J. Japan. Soc. Hort. Sci*. **80**, 276-283, (2011).
125. Yamane, H., Wada, M., Honda, C., Matsuura, T., Ikeda, Y., Hirayama, T., Osako, Y., Gao-Takai, M., Kojima, M. & Sakakibara, H. J. P. o. Overexpression of *Prunus DAM6* inhibits growth, represses bud break competency of dormant buds and delays bud outgrowth in apple plants. *PLoS One* **14**, e0214788, (2019).
126. Yang, Y., Fanning, L. & Jack, T. The K domain mediates heterodimerization of the Arabidopsis floral organ identity proteins, APETALA3 and PISTILLATA. *Plant J*. **33**, 47-59, (2003).
127. Yang, Y. & Jack, T. J. P. m. b. Defining subdomains of the K domain important for protein-protein interactions of plant MADS proteins. *Plant Mol. Biol.* **55**, 45, (2004).
128. Yang, Q., Yang, B., Li, J., Wang, Y., Tao, R., Yang, F., Wu, X., Yan, X., Ahmad, M., Shen, J., Bai, S. & Teng, Y. ABA-responsive ABRE-BINDING FACTOR3 activates *DAM3* expression to promote bud dormancy in Asian pear. *Plant Cell Environ.* **43**, 1360-1375, (2020).
129. Wang, Q., Xu, G., Zhao, X., Zhang, Z., Wang, X., Liu, X., Xiao, W., Fu, X., Chen, X., Gao, D., Li, D. & Li, L. Transcription factor TCP20 regulates peach bud endodormancy by inhibiting *DAM5/DAM6* and interacting with ABF2. *J. Exp. Bot.* **71**, 1585-1597, (2019).
130. Nesi, N., Jond, C., Debeaujon, I., Caboche, M. & Lepiniec, L. The Arabidopsis *TT2* gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant cell* **13**, (2001).
131. Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B. & Lepiniec, L. TT2, TT8, and TTG1

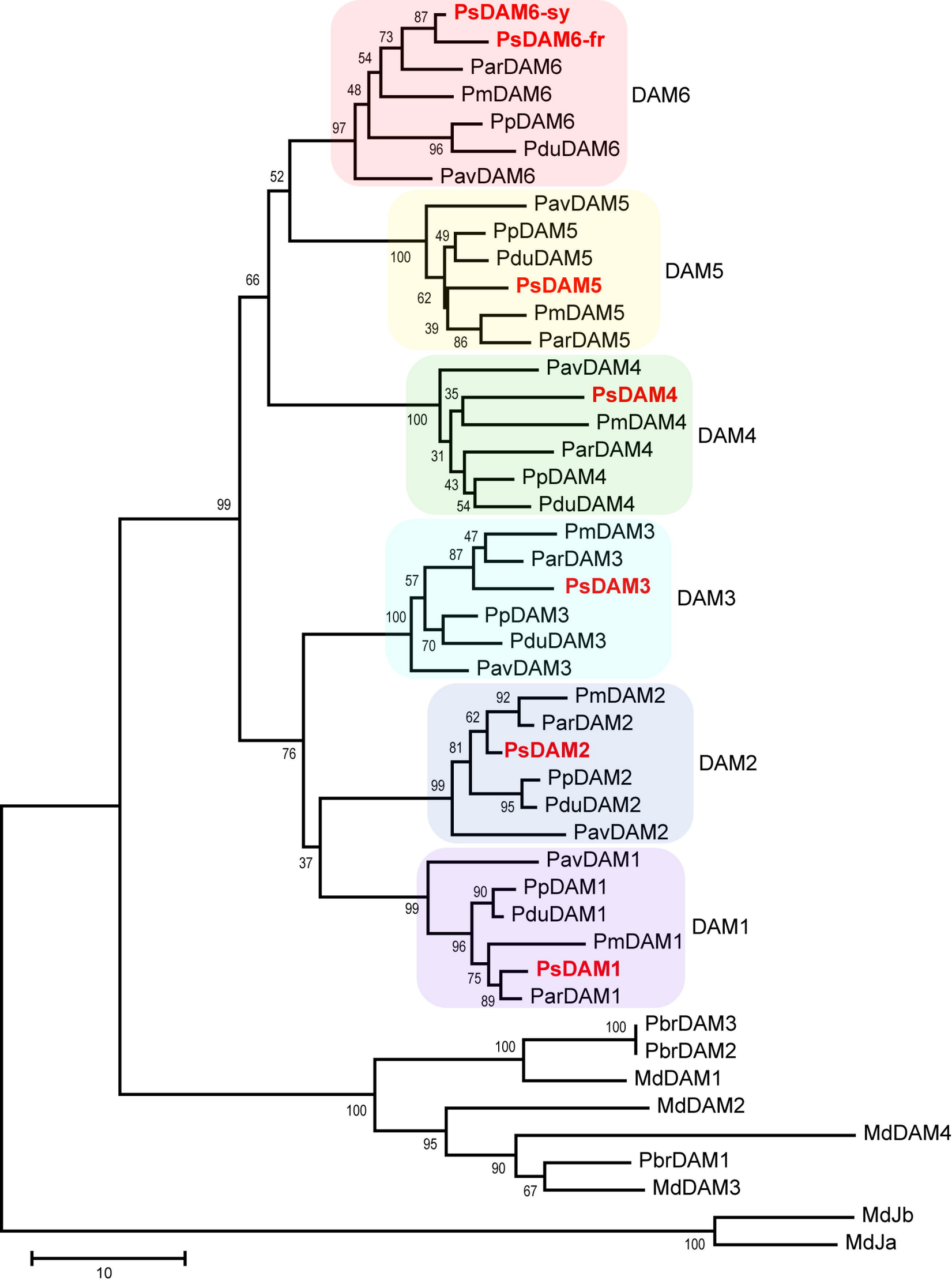
synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J*. **39**, 366-380, (2004).

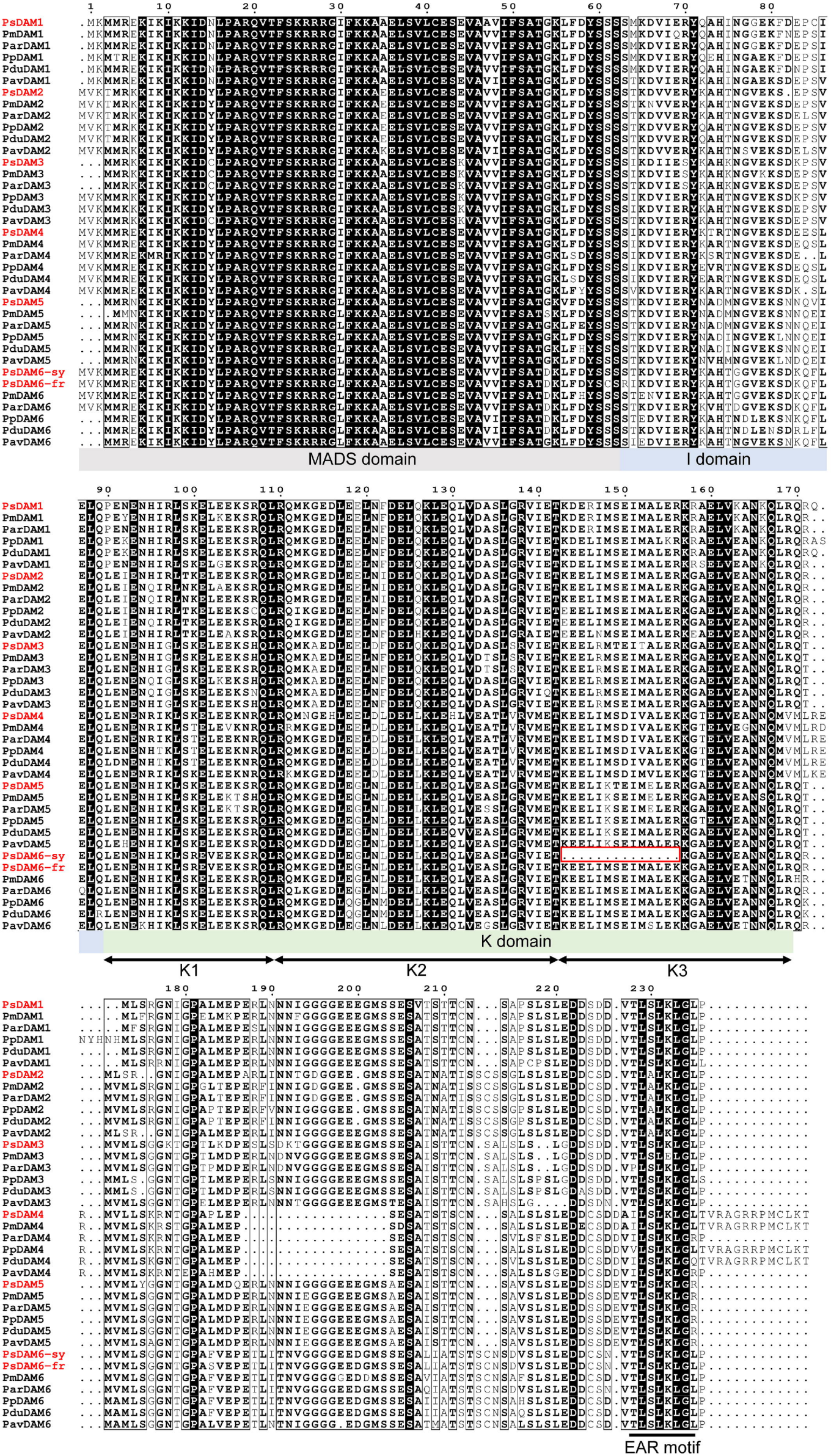
132. Leida, C., Terol, J., Martí, G., Agustí, M., Llácer, G., Badenes, M. L. & Ríos, G. Identification of genes associated with bud dormancy release in *Prunus persica* by suppression subtractive hybridization. *Tree Physiol*. **30**, 655-666, (2010).
133. Nishitani, C., Saito, T., Ubi, B. E., Shimizu, T., Itai, A., Saito, T., Yamamoto, T. & Moriguchi, T. Transcriptome analysis of *Pyrus pyrifolia* leaf buds during transition from endodormancy to ecodormancy. *Sci. Hortic*. **147**, 49-55, (2012).
134. Sudawan, B., Chang, C.-S., Chao, H.-f., Ku, M. S. B. & Yen, Y.-f. Hydrogen cyanamide breaks grapevine bud dormancy in the summer through transient activation of gene expression and accumulation of reactive oxygen and nitrogen species. *BMC Plant Biol.* **16**, 202, (2016).
135. Zhang, J., Wu, Y., Li, D., Wang, G., Li, X. & Xia, Y. Transcriptomic analysis of the underground renewal buds during dormancy transition and release in 'Hangbaishao' peony (*Paeonia lactiflora*). *PLoS ONE* **10**, e0119118, (2015).
136. Gai, S., Zhang, Y., Liu, C., Zhang, Y. & Zheng, G. Transcript profiling of *Paoenia ostii* during artificial chilling induced dormancy release identifies activation of ga pathway and carbohydrate metabolism. *PLOS ONE* **8**, e55297, (2013).
137. Gai, S., Zhang, Y., Mu, P., Liu, C., Liu, S., Dong, L. & Zheng, G. Transcriptome analysis of tree peony during chilling requirement fulfillment: Assembling, annotation and markers discovering. *Gene* **497**, 256-262, (2012).
138. Tuan, P. A., Bai, S., Saito, T., Ito, A. & Moriguchi, T. Dormancy-associated MADS-Box (*DAM*) and the abscisic acid pathway regulate pear endodormancy through a feedback mechanism. *Plant Cell Physiol.* **58**, 1378-1390, (2017).
139. Sreekantan, L., Mathiason, K., Grimplet, J., Schlauch, K., Dickerson, J. A. & Fennell, A. Y. Differential floral development and gene expression in grapevines during long and short photoperiods suggests a role for floral genes in dormancy transitioning. *Plant Mol. Biol.* **73**, 191-205, (2010).
140. Ríos, G., Leida, C., Conejero, A. & Badenes, M. L. Epigenetic regulation of bud dormancy events in perennial plants. *Front. Plant Sci.* **5**, 247, (2014).
141. Conde, D., Moreno-Cortés, A., Dervinis, C., Ramos-Sánchez, J. M., Kirst, M., Perales, M., González-Melendi, P. & Allona, I. Overexpression of DEMETER, a DNA demethylase, promotes early apical bud maturation in poplar. *Plant Cell Environ.* **40**, 2806-2819, (2017).
142. Conde, D., Le Gac, A.-L., Perales, M., Dervinis, C., Kirst, M., Maury, S., González-Melendi, P. & Allona, I. Chilling-responsive DEMETER-LIKE DNA demethylase mediates in poplar bud break. *Plant Cell Environ.* **40**, 2236-2249, (2017).
143. Zhebentyayeva, T. N., Fan, S., Chandra, A., Bielenberg, D. G., Reighard, G. L., Okie, W. R. & Abbott, A. G. Dissection of chilling requirement and bloom date QTLs in peach using a whole genome sequencing of sibling trees from an F2 mapping population. *Tree Genet. Genomes* **10**, 35-51, (2014).
144. Saito, T., Bai, S., Ito, A., Sakamoto, D., Saito, T., Ubi, B. E., Imai, T. & Moriguchi, T. Expression and genomic structure of the dormancy-associated *MADS box* genes *MADS13* in Japanese pears (*Pyrus pyrifolia* Nakai) that differ in their chilling requirement for endodormancy release. *Tree Physiol.* **33**, 654-667, (2013).

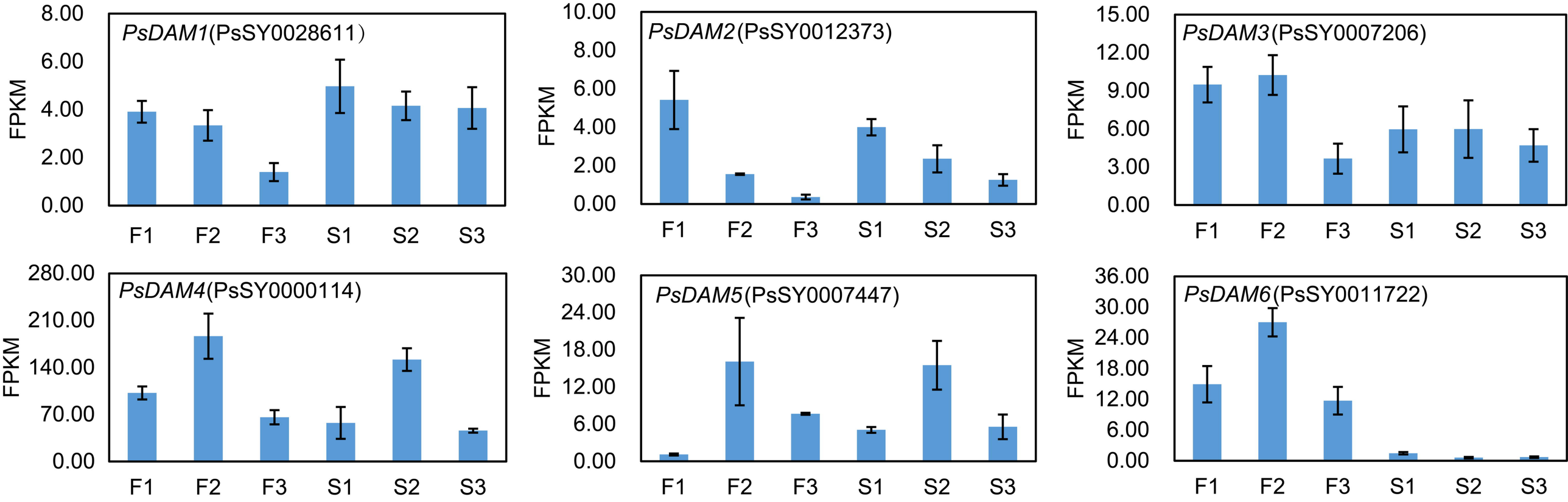

A
B
1cm
C
D
E
F

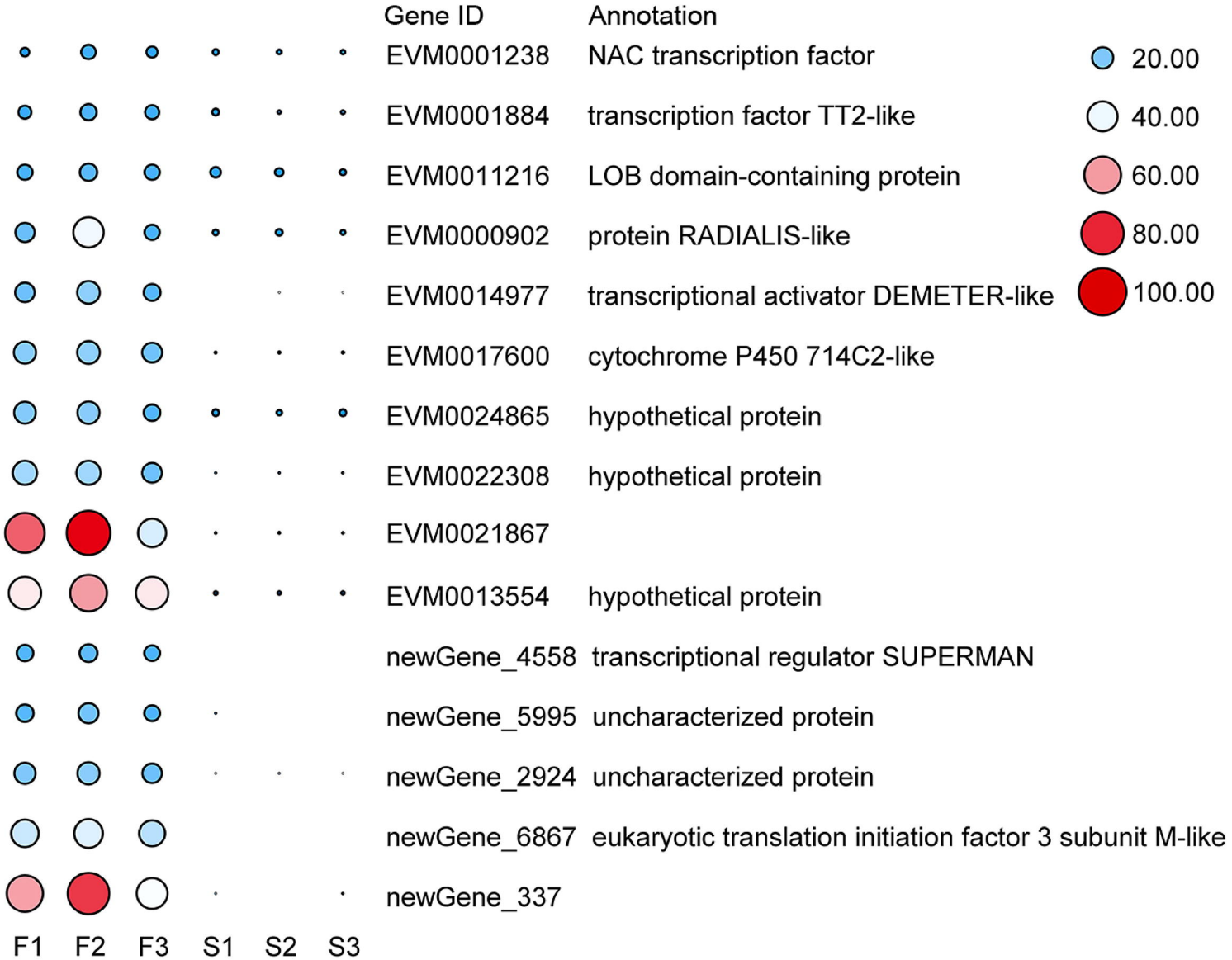

A
Expansion
Contraction
296 2513 P.mume
1013 885 P.armeniaca
1248 2550 P.salicina
464 1666 P.persica
652 2097 P.dulcis
1337 3559 P.avium
506 584 P.bretschneideri
476 357 M.domestica
663 555 R.chinensis
412 816 F.vesca
615 227 P.trichocarpa
228 672 A.thaliana
275 7406 V.vinifera
251 7754 O.sativa
71 7983 A.trichopoda
5.32
7.45
10.32
4.29
15.14
48.8
28.81
69.87
26.84
106.7
115.09
85.95
156.5
180.5
160
120
80
40
0
Mya
Unclustered genes
Other genes
Specific genes
Multi-copy genes
Single-copy genes
0
10000
20000
30000
40000
50000
Gene number
B
P. mume
P. persica
P. armeniaca
P. salicina
P. dulcis
P. avium
142
379
190
378
92
1917
316
C
M. domestica vs M. domestica
P. salicina vs P. salicina
V. vinifera vs P. salicina
P. salicina vs M. domestica
V. vinifera vs V. vinifera
Percent of gene pairs (%)
Synonymous substitutions rate (Ks)


A
P. salicina
Contig01459
P. persica
DQ863257
P. armeniaca
LG2
P. mume
NC_024127.1
P. avium
PAV_r1.0chr1
P. dulcis Texas
Pd01
M. domestica
Chr16
Chr8
Chr11
Chr4
DAM1
DAM2
DAM3
DAM4
DAM5
DAM6
MdJa
MdJb
B
PsDAM1
PmDAM1
ParDAM1
PpDAM1
PduDAM1
MdDAM1
PsDAM2
PmDAM2
ParDAM2
PpDAM2
PduDAM2
MdDAM2
PsDAM3
PmDAM3
ParDAM3
PpDAM3
PduDAM3
MdDAM3
PsDAM4
PmDAM4
ParDAM4
PpDAM4
PduDAM4
MdDAM4
PsDAM5
PmDAM5
ParDAM5
PpDAM5
PduDAM5
PsDAM6
PmDAM6
ParDAM6
PpDAM6
PduDAM6
2000bp
Exon
Intron
Furongli
C
Exon1
Exon2
Exon3
Exon4
Exon5
Exon6
Exon7
Exon8
0bp
277bp
860bp
8276bp
Exon
Intron
0bp
1327bp
6493bp
14591bp
Exon1
Exon2
Exon3
Exon4
Exon6
Exon7
Exon8
Sanyueli


PsDAM6-sy
PsDAM6-fr
ParDAM6
PmDAM6
PpDAM6
PduDAM6
PavDAM6
DAM6
PavDAM5
PpDAM5
PduDAM5
PsDAM5
PmDAM5
ParDAM5
DAM5
PavDAM4
PsDAM4
PmDAM4
ParDAM4
PpDAM4
PduDAM4
DAM4
PmDAM3
ParDAM3
PsDAM3
PpDAM3
PduDAM3
PavDAM3
DAM3
PmDAM2
ParDAM2
PsDAM2
PpDAM2
PduDAM2
PavDAM2
DAM2
PavDAM1
PpDAM1
PduDAM1
PmDAM1
PsDAM1
ParDAM1
DAM1
PbrDAM3
PbrDAM2
MdDAM1
MdDAM2
MdDAM4
PbrDAM1
MdDAM3
MdJb
MdJa
10


PsDAM1
PmDAM1
ParDAM1
PpDAM1
PduDAM1
PavDAM1
PsDAM2
PmDAM2
ParDAM2
PpDAM2
PduDAM2
PavDAM2
PsDAM3
PmDAM3
ParDAM3
PpDAM3
PduDAM3
PavDAM3
PsDAM4
PmDAM4
ParDAM4
PpDAM4
PduDAM4
PavDAM4
PsDAM5
PmDAM5
ParDAM5
PpDAM5
PduDAM5
PavDAM5
PsDAM6-sy
PsDAM6-fr
PmDAM6
ParDAM6
PpDAM6
PduDAM6
PavDAM6
MADS domain
I domain
K domain
K1
K2
K3
EAR motif


PsDAM1(PsSY0028611)
PsDAM2(PsSY0012373)
PsDAM3(PsSY0007206)
PsDAM4(PsSY0000114)
PsDAM5(PsSY0007447)
PsDAM6(PsSY0011722)
FPKM
F1
F2
F3
S1
S2
S3


Gene ID
Annotation
EVM0001238 NAC transcription factor
EVM0001884 transcription factor TT2-like
EVM0011216 LOB domain-containing protein
EVM0000902 protein RADIALIS-like
EVM0014977 transcriptional activator DEMETER-like
EVM0017600 cytochrome P450 714C2-like
EVM0024865 hypothetical protein
EVM0022308 hypothetical protein
EVM0021867
EVM0013554 hypothetical protein
newGene_4558 transcriptional regulator SUPERMAN
newGene_5995 uncharacterized protein
newGene_2924 uncharacterized protein
newGene_6867 eukaryotic translation initiation factor 3 subunit M-like
newGene_337
F1
F2
F3
S1
S2
S3
20.00
40.00
60.00
80.00
100.00