BMC Plant Biology

RESEARCH

Open Access

Chromosome-level genome provides new insight into the overwintering process of Korla pear (*Pyrus sinkiangensis* Yu)

Wenwen Xia[1†], Saisai Wang[1†], Xiaoyan Liu[1†], Yifei Chen[1†], Caixia Lin[2], Ruina Liu[1], Hailiang Liu[1,3], Jin Li[1*] and Jianbo Zhu[1*]

Abstract

Korla pear has a unique taste and aroma and is a breeding parent of numerous pear varieties. It is susceptible to *Valsa mali* var. *pyri*, which invades bark wounded by freezing injury. Its genetic relationships have not been fully defined and could offer insight into the mechanism for freezing tolerance and disease resistance. We generated a high-quality, chromosome-level genome assembly for Korla pear via the Illumina and PacBio circular consensus sequencing (CCS) platforms and high-throughput chromosome conformation capture (Hi-C). The Korla pear genome is ~ 496.63 Mb, and 99.18% of it is assembled to 17 chromosomes. Collinearity and phylogenetic analyses indicated that Korla might be derived from *Pyrus pyrifolia* and that it diverged ~ 3.9-4.6 Mya. During domestication, seven late embryogenesis abundant (*LEA*), two dehydrin (*DHN*), and 54 disease resistance genes were lost from Korla pear compared with *P. betulifolia*. Moreover, 21 *LEA* and 31 disease resistance genes were common to the Korla pear and *P. betulifolia* genomes but were upregulated under overwintering only in *P. betulifolia* because key *cis* elements were missing in Korla pear. Gene deletion and downregulation during domestication reduced freezing tolerance and disease resistance in Korla pear. These results could facilitate the breeding of novel pear varieties with high biotic and abiotic stress resistance.

Keywords Disease, Genome, Overwintering, *Pyrus sinkiangensis* Yu

Introduction

Pear (family Rosaceae; subfamily Pomoideae) is the most important temperate fruit crop after apple and grape, and it has been cultivated worldwide for over 3,000 years [1]. Pears originated in the mountainous regions of southwestern China during the Tertiary period, 65–55 million years ago (Mya) and spread to other regions [2, 3]. There are thousands of cultivars, and they are generally classified as either occidental (European) or oriental (Asiatic). To date, only a few species, including *Pyrus bretschneideri*, *P. pyrifolia*, *P. ussuriensis*, *P. sinkiangensis*, and *P. communis* have been cultivated for fruit production. Of these, Korla pear (*P. sinkiangensis*), variety 'korla fragrant pear', is the most highly preferred and more frequently exported to overseas because of its sweet taste, pleasant aroma, and crisp flesh. Korla pear is a native plant

[†]Wenwen Xia, Saisai Wang, Xiaoyan Liu and Yifei Chen contributed equally to this article.

*Correspondence:
Jin Li
lijin@shzu.edu.cn
Jianbo Zhu
zjbshz@126.com
[1] Key Laboratory of Agricultural Biotechnology, College of Life Science, Shihezi University, Shihezi 832000, China
[2] Xinjiang Production and Construction Crops, Institute of Agricultural Sciences, Tiemenguan 841007, China
[3] Institute for Regenerative Medicine, Shanghai East Hospital, Tongji University School of Medicine, Shanghai 200123, China

of Xinjiang, China, and it is a primal parent for breeding novel pear varieties such as Xin 6, Xin 9, Zhongli 10, and others [4]. However, the genetic relationships of Korla pear remain in dispute. The pear originated in southwest China and spread across central and western Asia and Europe, with its dispersal eventually resulting in various new cultivars [5]. Preliminary studies on pear traits showed that Korla pear may be a hybrid of *P. communis* and *P. bretschneideri* [6, 7]. However, random amplified polymorphic DNA (RAPD) markers were used to assess the genetic relationships of the pear cultivars in Xinjiang and showed that *P. sinkiangensis* is actually a hybrid of *P. communis* and *P. pyrifolia* [8]. According to the minimum Euclidean distance between *P. sinkiangensis* and *P. bretschneideri* determined via the RAPD data, *P. sinkiangensis* was classified into the *P. bretschneideri* group [9]. Recently, Korla pear was identified as a hybrid of *P. communis* and *P. bretschneideri* based on 420 single-copy conserved genes from nine pear species [5]. Therefore, an understanding of the whole genome is essential and could be used to accurately identify the genetic relationships for Korla pear.

Pear trees are susceptible to a wide range of pathogens, such as *Valsa mali* var. *pyri*, which can cause an aggressive pear canker responsible for substantial crop and economic losses [10, 11]. This fungus invades the host tissue through bark wounds or girdling of the lateral and main stems and induces vascular tissue necrosis and tree death [12–15]. *Valsa mali* var. *pyri* secretes necrotizing and ethylene-inducing peptides and effector proteins that manipulate host immunity and enable successful fungal colonization in the plants [16, 17]. A total of 50–80% of Korla pear plants were found to be infected by *V. mali* in a recent study [18]. Since this disease is difficult to control without using chemical treatment, the common method used to prevent spreading of the disease is destruction of infected plants. Thus, understanding the reason that Korla pear is vulnerable to *V. mali* is vital to breeding new varieties with high resistance to the pathogen and will prevent sever yield loss.

To defend against microbial pathogens, a host may possess a robust innate immune system that includes pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) [19]. Genes regulating the immunity pathway in plants include NOD-like receptors (NLRs), ascorbate peroxidase 1 (APX1), BRI1-associated receptor kinase 1 (BAK1), botrytis-induced kinase 1 (BIK1), mitogen-activated protein kinases (MAPKs), UDP-glucose:phloretin 2'-O-glucosyltransferase (UGT88F1), and others [20–24]. However, the numbers and expression modes of various disease resistance genes differ among plant species and even cultivars. For example, the incidences of *Valsa mali* var. *pyri* in Yali (A kind of *Pyrus bretschneideri*), Crisp (Another kind of *Pyrus bretschneideri*), and Korla pear are 30%, 30–50%, and 50–80%, respectively [18]. *Pyrus betulifolia* Duli is highly resistant to *V. mali* var. *pyri* [18]. In general, *V. mali* var. *pyri* disease outbreaks mainly occur during spring and autumn. Moreover, extreme low wintertime temperatures exacerbate disease severity. Low temperatures were found to promote *V. mali* colony growth and increased disease incidence in apple [25]. Homeodomain-leucine zipper I and II (*HD-Zip* I and II), which regulate the response to fungal invasion, were identified in apple and are regulated by abscisic acid (ABA) and methyl jasmonate (MeJA) signaling [26]. We found that overexpress *PsHB*s could enhance freeze resistance in hybrid *Broussonetia papyrifera* [27]. Therefore, the disease induced by *V. mali* var. *pyri* is secondary to freezing injury, and both the molecular genetics of overwintering and a potential defense mechanism were areas of interest in our analysis.

Fruit tree dormancy is vital for ensuring survival during the cold season [28, 29]. It may be induced and maintained by ABA signaling [30–32]. Cellular dehydration and membrane injury are the two main mechanisms of freezing damage [33], so numerous osmoprotectants accumulate in the cytoplasm during acclimation and overwintering. Soluble proteins encoded by *LEA*s and *DHN*s help stabilize protein structures and membranes while preventing excessive cytoplasmic water loss. It was previously shown that *Betula platyphylla* dehydrated during seasonal cold acclimation and that *DHN* accumulation increased with ABA content [34, 35]. *LEA14* improves low-temperature stress tolerance in *P. communis.* Thus, *LEA* genes are associated with freezing tolerance in overwintering trees [36].

To achieve the purpose of revealing the genetic basis for evolution and the reduced stress resistance, whole genomic analysis is necessary for Korla pear. In this study, both genome and transcriptome sequencing were conducted on Korla pear to identify the genes which control its overwintering mechanism. A chromosome-level genome was produced through the PacBio circular consensus sequencing (CCS) platforms. The assembly was corrected using the high-throughput chromosome conformation capture (Hi-C) algorithm. The Korla pear genome was ~ 496.63 Mb and highly heterozygous (2.1%). Korla pear was found to be an Asiatic variety that diverged from *P. pyrifolia* ~ 3.9–4.6 Mya. During long-term evolution, numerous genes were deleted from the Korla pear genome compared with that of *P. betulifolia*. For instance, several genes regulating overwintering and disease resistance were not induced by freezing. The foregoing findings explained the low overwintering capacity and lack of disease resistance of Korla pear and could

provide theoretical and empirical guidance for breeding novel cold- and disease-tolerant Korla pear varieties.

Methods

Korla pear genome estimation

The Korla pear genome DNA was extracted from tissue culture seedlings via ultrasonic oscillation. The tissue culture seedings of Korla pear was provided by Agricultural Science Research Institute of the second division of Xinjiang Production and Construction Corps (Korla, Xinjiang, China). A library with 350-bp insert size was constructed for use in Illumina HiSeq 2500 paired-end 150 (PE150) sequencing (Illumina, San Diego, CA, USA). Raw data (29 Gb) were used to estimate sample DNA pollution, genome size, GC content, and heterozygosity. Sample DNA pollution was estimated via the NCBI Blast+v. 2.9.0 package (https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/) against the nucleotide (NT) database (https://www.ncbi.nlm.nih.gov/nucleotide/) using the parameters -num_descriptions 100, -num_alignments 100, and -evalue 1e-05 [37]. The k-mer frequency distribution was calculated and plotted. Extranuclear DNA was detected using SOAP v. 2.21 (https://www.soapui.org/downloads/soapui/) with cutoffs -m 260 and -x 440 [38]. The k-mer frequency distribution was used to estimate the Korla pear genome size.

De novo genome assembly using PacBio CCS reads

According to the Korla pear genome estimates, 31,327,856,661 bp with an estimated depth of coverage of ~ 59.65 X was obtained using PacBio circular consensus sequencing (PacBio CCS) technology (https://github.com/PacificBiosciences/ccs). The average CCS read length was ~ 13,469 bp and N50 = 13,563 bp. Highly accurate CCS data were used to assemble the Korla pear genome with hifiasm v. 0.12 (https://github.com/chhylp123/hifiasm). Redundant sequences were eliminated with purge_dups software (https://github.com/dfguan/purge_dups) [39]. De novo genome quality was estimated and properly mapped to the Illumina HiSeq data and the CEGMA v. 2.5 (https://github.com/KorfLab/CEGMA_v2/tree/v2.5) and BUSCO v. 4 (https://busco.ezlab.org/busco_v4_data.html) databases [40, 41]. The Illumina HiSeq reads were mapped to the Korla pear genome using bwa_mem (https://github.com/bwa-mem2/bwa-mem2) with its default parameters [42].

Chromosome-scale scaffold anchoring with Hi-C

High-throughput chromosome conformation capture (Hi-C) was used to categorize and order the contigs and erect chromosome-scale scaffolds [43]. Cells and tissues from a fresh pear sample were cross-linked with formaldehyde to establish and maintain interactions between the cellular DNA and the proteins and between DNA and RNA. Nuclear DNA was then digested with the restriction endonuclease *Hind*III. The sticky ends were filled with biotinylated nucleotides and the interacting DNA was cyclized with blunt ends. The DNA segments crosslinks were reversed, and the DNA segments were sheared to 300-700 bp fragments. Interacting DNA fragments were captured with streptomycin magnetic beads and used to construct an Illumina sequencing library. The quality of the latter was determined with a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) and an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and by qPCR.

A total of 51,120,244,986 bp raw data (~ 98.17 × of the estimated genome size) was generated on the Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA). Clean reads were mapped with Burrows-Wheeler Aligner (BWA v. 0.7.10-r789) (mapping mode; -aln and default cutoffs) (https://sourceforge.net/projects/bio-bwa/files/) to the Korla pear genome. A total of 171,078,949 read pairs were mapped to the genome, and 125,295,815 read pairs were uniquely mapped. Among the latter, the number of valid interacting read pairs was 78,369,912 (62.55%), and they were analyzed by HiC-Pro v. 2.10.0 (https://github.com/nservant/HiC-Pro/releases) [44]. The valid interacting read pairs data were used to correct the contig sequence of the de novo genome. Both the valid interacting read pairs and the corrected de novo genome contigs were divided into subgroups, sorted, and oriented with LACHESIS (https://github.com/shendurelab/LACHESIS) into chromosome scale scaffolds using the following parameters: CLUSTER_MIN_RE_SITES = 100; CLUSTER_MAX_LINK_DENSITY = 2; ORDER_MIN_N_RES_IN_TRUNK = 79; and ORDER_MIN_N_RES_IN_SHREDS = 86. The final genome assembly was confirmed by Hi-C contact heatmap and collinearity analysis. The genomes of *P. betulifolia* (ftp://download.big.ac.cn/gwh/Plants/Pyrus_betulifolia_Pbe-SD_GWHAAYT00000000), *P. communis* (https://www.rosaceae.org/organism/Pyrus/communis), *P. x bretschneideri* (https://www.ncbi.nlm.nih.gov/assembly/GCF_000auto315295.1/), *Vitis vinifera* (ftp://ftp.ensemblgenomes.org/pub/plants/release-45/fasta/vitis_vinifera/), *P. pyrifolia* Nijisseiki (https://www.rosaceae.org/rosaceae_downloads/Pyrus_pyrifolia/ppyrifolia_v1.0/), *P. pyrifolia* Cuiguan (https://download.cncb.ac.cn/gwh/Plants/Pyrus_pyrifolia_Pyrus_pyrifolia_cultivar_'Cuiguan'_GWHBAOS00000000/), *Malus domestica* (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/002/114/115/GCF_002114115.1_ASM211411v1), *Arabidopsis thaliana* (https://www.arabidopsis.org/download/index-auto.jsp?dir=%2Fdownload_files%2FGenes%2FTAIR10_genome_release%2FTAIR10_chromosome_files), *Betula*

pendula (https://genomevolution.org/coge/GenomeInfo.pl?gid=35080), and *Prunus persica* (https://www.ncbi.nlm.nih.gov/genome/?term=PEACH) were used to construct collinearity with the Korla pear genome. Diamond v. 0.9.29.130 with the parameters e < 1e-5 and C score > 0.5 (https://github.com/bbuchfink/diamond/releases?after=v0.9.21) was used to identify gene orthologs [45]. The adjacent relationships of the orthologous genes targeting in the chromosomes of two species were identified via MCScanX (https://github.com/wyp1125/MCScanX) [46]. The collinearity was plotted with JCVI v. 0.9.13 (https://pypi.org/project/jcvi/) and VGSC software (https://dvb.ac.cn/vgsc2/service/index.php?mode=VGSConline) [47, 48].

Genome annotation

The tandem and interspersed repeat sequences were categorized. The latter included transposable elements (TE), which are major annotations of repeat sequences. RepeatModeler2 v. 2.0.1 (http://www.repeatmasker.org/RepeatModeler/) was used for de novo prediction by invoking RECON v. 1.0.8 (http://repeatmasker.org/RECON-1.08.tar.gz) and RepeatScout v. 1.0.6 (https://anaconda.org/bioconda/repeatscout) [49–51]. RepeatClassifier (https://github.com/Dfam-consortium/RepeatModeler/blob/master/RepeatClassifier) classified predictions based on the repbase v. 19.06 (https://www.girinst.org/repbase/) [52], REXdb v. 3.0 (http://repeatexplorer.org/?page_id=918) [53], and Dfam v. 3.2 (https://dfam.org/home) [54] databases. Long terminal repeats (LTR) were annotated de novo with LTR_retriever v. 2.8 (https://github.com/oushujun/LTR_retriever) based on the LTRharvest v. 1.5.9 (https://github.com/oushujun/LTR_HARVEST_parallel) and LTR_FINDER v. 1.1 (https://github.com/xzhub/LTR_Finder) analyses [55–57]. The preceding results and databases were used to construct a specific redundancy-free repeat sequences database. The Korla pear genome TE was annotated with RepeatMasker v. 4.1.0 (http://repeatmasker.org/RepeatMasker/) against the specific repeat sequences database [58]. The tandem repeats were annotated with MIcroSAtellite Identification Tool v. 2.1 (https://webblast.ipk-gatersleben.de/misa/) and Tandem Repeats Finder v. 409 using the parameters 1 1 2 80 5 200 2000 -d -h (https://mybiosoftware.com/trf-4-04-tandem-repeats-finder.html) [59, 60].

Protein-coding genes were identified by de novo homolog- and transcriptome-based prediction. Transcriptome prediction was performed by Augustus v. 2.4 (https://github.com/Gaius-Augustus/Augustus) and SNAP v. 2006–07-28 (https://github.com/KorfLab/SNAp) [61, 62]. Homology prediction was conducted with GeMoMa v. 1.7 (https://anaconda.org/bioconda/gemoma), which aligned the gene orthologs of congenerics [63]. Fruit, leaf, and stem samples were mixed and used in transcriptome sequencing on the Illumina HiSeq 2500 platform. The transcriptome was assembled with the reference genome by HISAT v. 2.0.4 (http://daehwankimlab.github.io/hisat2/) and StringTie v. 1.2.3 (http://ccb.jhu.edu/software/stringtie/). GeneMarkS-T v. 5.1 (https://github.com/Magdoll/SQANTI2/blob/master/utilities/gmst/README.GeneMarkS-T) was used to predict the protein-coding genes in the genome [64–66]. The transcriptome assembly with no reference genome was acquired with Trinity v. 2.11 90 (https://www.osc.edu/resources/available_software/software_list/trinity) [67]. PASA v. 2.0.2 (https://sourceforge.net/projects/pasa/) was used to predict the coding genes in the genome [68]. The foregoing results were integrated with EVM v. 1.1.1 (https://github.com/EVidenceModeler/EVidenceModeler/releases) [69]. The primary gene set was modified with PASA v. 2.0.2 to produce the final gene set. The accuracy of protein-coding genes was qualified with BUSCO v. 4.0 against the BUSCO embryophyta database (http://busco.ezlab.org/v2/datasets/embryophyta_odb9.tar.gz). The percentage of RNA-Seq clean data mapped to the protein-coding genes was determined with HISAT2 software.

Non-coding RNA was predicted according to the RNA classification. The software tRNAscan-SE v. 1.3.1 (http://trna.ucsc.edu/tRNAscan-SE/) was used to predict the tRNA [70]. The rRNA was identified with the barrnap v. 0.9 (https://github.com/tseemann/barrnap) and the Rfam v. 12.0 (https://rfam.xfam.org) databases [71, 72]. The microRNA (miRNA) was predicted with miRbase (https://www.mirbase.org) [73]. The snoRNA and the snRNA were predicted with Infernal v. 1.1 (http://eddylab.org/infernal/) against the Rfam v. 12.0 database [71, 74]. The pseudogene sequences are usually similar to those of functional genes but might have lost their biological function because of certain genetic mutations such as insertions and deletions. GenBlastA v. 1.0.4 (https://mybiosoftware.com/genblasta-1-0-4-genblastg-1-39-homologous-gene-sequences.html) was used to scan the entire genomes after the predicted functional genes were masked. Putative candidates were analyzed by searching for non-mature and frameshift mutations with GeneWise v. 2.4.1 (https://www.ebi.ac.uk/Tools/psa/genewise/) [75, 76].

Functional annotations were assigned by blast searches against the NR (202,009, ftp://ftp.ncbi.nlm.nih.gov/blast/db), SWISS-PROT (http://ftp.ebi.ac.uk/pub/databases/swissprot; accessed May 2020) [77], and KEGG (http://www.genome.jp/kegg; accessed December 20, 2019) databases [78]. All genes were then subjected to eggNOG v. 5.0 (http://eggnog5.embl.de/download/eggnog_5.0/)

[79] and GO (http://geneontology.org; accessed June 15, 2020) annotations. Gene families were assigned with Pfam v. 33.1 (http://pfam.xfam.org) [80] and InterProScan v. 5.34-73.0) (https://www.ebi.ac.uk/about/news/updates-from-data-resources/InterPro-73.0/) [81]. Overall, 38,296 of the predicted genes (99.92%) were assigned to functional descriptions.

Comparative genomic analysis

The genomes of *P. betulifolia*, *V. vinifera*, *P. persica*, *A. thaliana*, *P. communis*, *M. domestica*, *P. bretschneideri*, *P. betulifolia*, *P. pyrifolia Nijisseiki*, *P. pyrifolia Cuiguan* and *B. pendula* were subjected to comparative genomic analysis. Protein sequence sets were collected from all of the aforementioned species. Orthofinder v. 2.4 with the parameters -align model diamond and -e 0.001 (http://www.stevekellylab.com/software/orthofinder) was used to assign the gene families [82]. The latter were then annotated with the PANTHER v. 15 database (pantherdb.org) [83] and subjected to GO and KEGG annotations.

A total of 1,857 single-copy gene sequences were extracted from all 10 species and aligned with MAFFT v. 7.205 using the parameters -localpair and -maxiterate 1000 (https://mafft.cbrc.jp/alignment/software/) [84]. Poorly aligned sequences and regions were estimated and deleted with Gblocks v. 0.91b using the parameter -b5 = h (https://bioweb.pasteur.fr/docs/modules/gblocks/0.91b/). In this manner, a supergene sequence was generated [85]. The optimal phylogenetic tree was JTT + F + I + G4, and it was determined with ModelFinder in IQ-TREE v. 1.6.11 (https://github.com/Cibiv/IQ-TREE/releases/tag/v1.6.11) [86, 87]. The latter program was then used to construct the maximum likelihood (ML) phylogenetic tree with 1,000 bootstraps. A rooted tree was constructed by setting *Vitis vinifera* as the outgroup. The divergence time was estimated with the MCMCTREE package in PAML v. 4.9 (http://abacus.gene.ucl.ac.uk/software/paml.html) [88]. The divergence times for *P. communis*_Vs_*M. domestica* (19-46 Mya) and *P. communis*_Vs_*V. vinifera* (105-115 Mya) were acquired from the TimeTree database (http://www.timetree.org/). The gradient and Hessian parameters were estimated with the MCMCTREE package in PAML v. 4.9 [88]. The ML phylogenetic tree was then constructed, and a correlated molecular clock and JC69 were used to estimate the divergence time. The numbers of Markov chain iterations were set to burnin = 5,000,000, sampfreq = 30, and nsample = 10,000,000. The correlation coefficient of two independent runs was determined to confirm the accuracy of divergence time estimation. The phylogenetic tree with divergence time was displayed with MCMCTreeR v. 1.1 (https://www.rdocumentation.org/packages/MCMCtreeR/versions/1.1) [89].

CAFE v. 4.2 (https://hahnlab.github.io/CAFE/) was used to detect gene family expansion and contraction based on the gene family clustering results and the estimated divergence times between species [90]. The standard for determining gene family expansion and contraction was based on $P < 0.05$.

The protein sequences of the single-copy families were extracted from *P. betulifolia*, *P. bretschneideri*, *P. sinkiangensis*, *P. communis*, *P. pyrifolia* Nijisseiki, *P. pyrifolia* Cuiguan and *M. domestica* were aligned with MAFFT using the parameters -localpair and -maxiterate 1000. The aligned sequences were then reversed into codon sequences with PAL2NAL software (http://www.bork.embl.de/pal2nal/). The chi2 program in PAML was used to perform likelihood ratio tests based on the branch-site model of CodeML and to identify significant differences ($P < 0.05$). Sites with $P > 0.95$ were considered positive gene selections based on the empirical Bayes method.

Whole-genome duplication (WGD) was identified from the distribution curve of synonymous substitutions per synonymous site (*Ks*) and four-fold synonymous (degenerative) third-codon transversion (4DTv). The *Ks* distribution was plotted with wgd v. 1.1.1 (https://wgd.readthedocs.io/en/latest/index.html) [91]. The *Ks* peak corresponded to a WGD event. The 4DTv distribution was plotted according to Scripts (https://github.com/JinfengChen/Scripts) and corrected through a HKY substitute model.

Transcriptome sequencing of Korla pear and *P. betulifolia* during overwintering

Two-year branches were harvested from Korla pear and *P. betulifolia* trees in Korla City, Xinjiang Province, China on November 11, 2020, December 25, 2020, and February 15, 2021. Each sample had three biological replicates. Phloem was scraped from the branches after they were frozen in liquid nitrogen and stored in an ultracold freezer. Total RNA was extracted with TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The RNA quality was evaluated by agarose gel electrophoresis and analysis using a Qubit 2.0 fluorometer and an Agilent 2100 bioanalyzer. RNA-Seq libraries were constructed with a TruSeq RNA sample preparation kit (Thermo Fisher Scientific) following the manufacturer's instructions. After raw read QC, ≥ 6 Gb data/sample were obtained on an Illumina HiSeq 2500 platform.

HISAT2 was used with its default parameters to map the clean reads to the reference genome. The transcript sequences were assembled with StringTie. The gene expression level was calculated per kilobase transcript per million fragments mapped (FPKM) using featureCounts (https://subread.sourceforge.net/featureCounts.

html) [92]. Differentially expressed genes (DEGs) were analyzed with DESeq2 (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) [93, 94]. P-values were calculated by the Benjamini-Hochberg method to obtain the false discovery rates (FDR). The DEGs were identified using the criteria $|\log_2(FC)| >= 1$, $P < 0.5$, and $FDR < 0.01$. All genes were annotated in the GO (geneontology.org), KEGG (https://www.genome.jp/kegg/), KOG (https://mycocosm.jgi.doe.gov/help/kogbrowser.jsf), NR (https://https.ncbi.nlm.nih.gov/refseq/about/nonredundantproteins/), and SwissProt (https://www.expasy.org/resources/uniprotkb-swiss-prot) databases. The transcription factors (TFs) were annotated using the hmmscan function in iTAK software (https://github.com/kentnf/iTAK) against the PlnTFDB and PlantTFDB (planttfdb.gao-lab.org) databases.

Results and discussion

Genome sequencing, assembly, annotation

We sequenced the Korla pear genome through the PacBio CCS sequencing platforms and corrected the assembly with the Hi-C algorithm. We obtained 31,327,856,661 base pairs (bp) (average read length = 13,469 bp) and 51,120,244,986 bp of the Hi-C paired-end reads (Table S1). We used the high-quality PacBio CCS data to assemble the Korla pear genome with hifiasm v. 0.12 [39] and generated an initial contig set with N50 = 30,533,258 bp (Table S2). We obtained 84 contigs with total length = 525,179,194 bp. To determine assembly quality, we generated 25.75 Gb clean reads on the Illumina HiSeq 2500 platform. The proportion of NGS sequencing short reads mapped to the assembly genome with the BWA software package was 98.19% (Table S3). Hence, the pear genome assembly had high accuracy [42]. We identified most of the core eukaryotic genes in the CEGMA v. 2.5 (98.91%) and BUSCO v. 4 databases (Tables S4 and S5). The foregoing results indicated that the initial contigs were properly assembled.

We used LACHESIS software to assemble the chromosome-level Korla pear genome. We integrated the Hi-C paired-end reads data (Table S6) [95], and we assigned 502,229,930 bp contigs data (99.18%) to 17 chromosomes (Fig. 1A; Table 1). The scaffold N50 was 29,545,866 bp, and the GC content was 37.67% (Table S7). According to the genome elucidated by Illumina HiSeq sequencing, the Korla pear genome was highly heterozygous (2.1%) and was 496.63 Mb in size (Tables S8-S10; Fig. S1). About 99.77% of the Korla pear genome was accurately assembled. To validate the results of the Korla pear genome assembly, we plotted a heatmap with Hi-C data and showed that all bins could be clearly assigned to all 17 chromosomes (Fig. S2). The foregoing analyses illustrated that the Korla pear genome was correctly and completely assembled at the chromosome scale.

We found that 52.32% of the Korla pear genome comprises repeat sequences including 21,860,188 bp tandem repeat sequences (4.32%) and 246,097,064 bp transposable elements (48.60%) (Tables S11 and S12).

The retroelement type had the highest proportion at 39.13% (Table S11). As for other plants, the long terminal repeats (LTRs) consisted mainly of *Gypsy*, *Unknown*, and *Copia* elements. However, the proportion of DNA transposon elements was higher in the Korla pear genome (9.47%) than it was in those of the other plants. The DIRS (*Dictyostelium* intermediate repeat sequence) and LARD (large retrotransposon derivative) proportions were lower in Korla pear than paper mulberry (5.11% and 8.68%, respectively), and accounted for 0.01% and 0% of the genome, respectively.

We identified 38,327 protein coding genes, 4,558 noncoding genes (corresponding to 850 tRNAs, 3,593 rRNAs, and 115 miRNAs), and 264 pseudogenes in the Korla pear genome (Tables S13 and S14). Of these, we identified 1,539 genes (95.35%) in the Embryophyta data of the BUSCO database (Table S15). Moreover, the gene length distribution trend of the Korla pear genome was similar to those of *P. betulifolia*, *P. communis*, and *P. bretschneideri* (Fig. S3). However, the Korla pear genome had fewer genes than that of *P. betulifolia* (59,520 genes) but more than those of *P. communis* (37,430) and *P. bretschneideri* (34,661) (Table S16). We annotated 38,296 genes (99.92%) in the GO, KEGG, KOG, SwissProt, TrEMBL, eggNOG, and NR databases (Table 2). The Unknown function term (10,132, 31.51%) was the most highly enriched in the Korla pear genome according to the eggNOG annotation (Fig. S4). Only 302 genes (0.94%) were assigned to the Defense mechanism term. Korla pear had a lower gene number than *P. betulifolia* in the Defense mechanism term [96].

Genomic evolution of Korla pear

We used IQ-TREE v. 1.6.11 software to plot a phylogenetic tree with the sequences of single-copy gene set and analyze the evolution of Korla pear. *Pyrus* spp. mentioned in this study and *Malus domestica* diverged from each other ~ 6.9–26.9 Mya (Figs. S5 and 6). The results indicated that Korla pear, *P. pyrifolia* (Cuiguan and Nijisseiki) and *P. bretschneideri*, the most closely related species, diverged ~ 3.9–4.6 Mya. The five Asiatic pear species [Korla pear, *P. bretschneideri*, *P. pyrifolia* (Cuiguan and Nijisseiki), and *P. betulifolia*] had more in

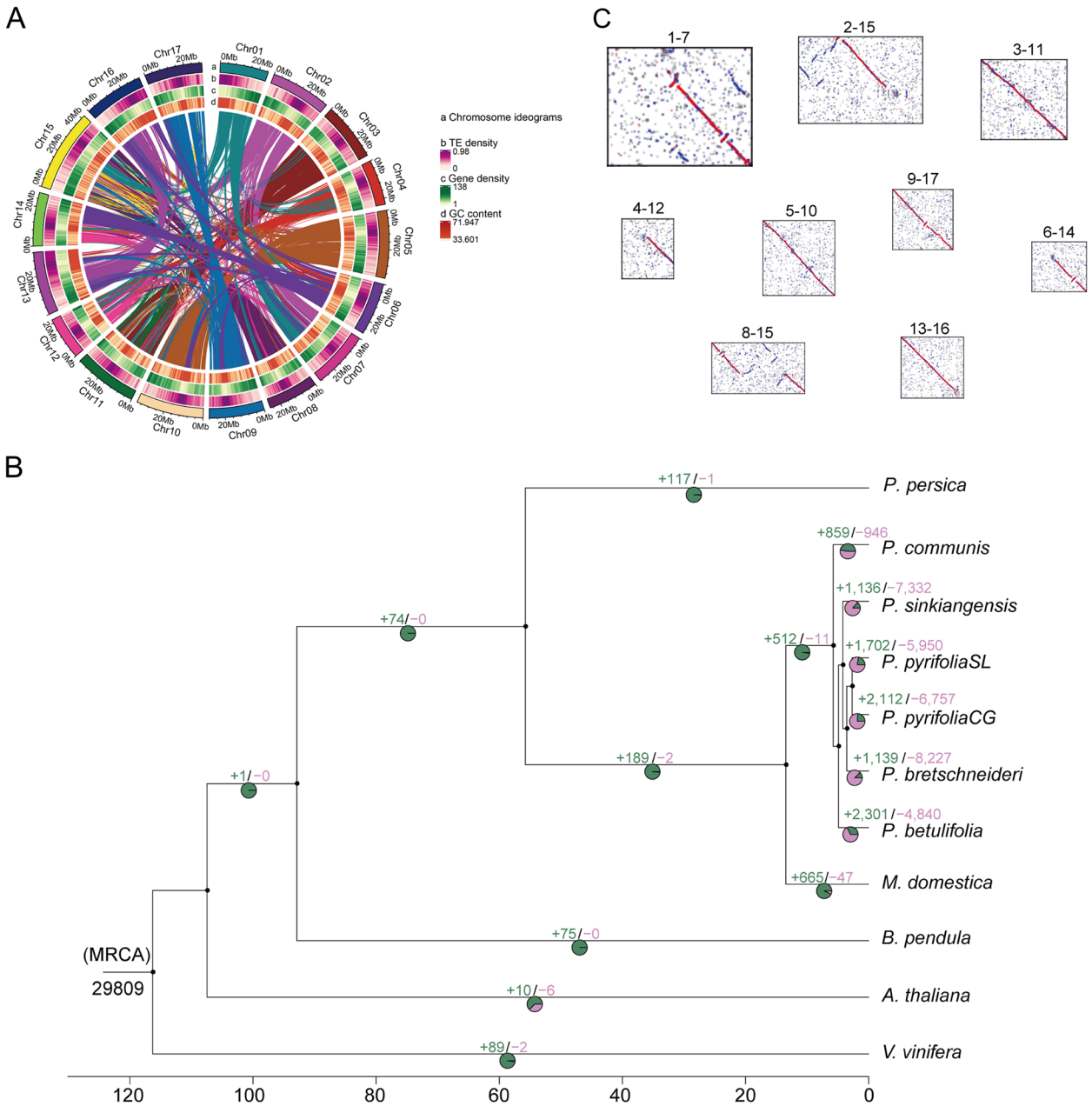

Fig. 1 The information of Korla pear chromosomes. **A** Korla pear pseudochromosomes. a, Chromosome ideograms. b, Transposable element (TE) content in each chromosome. c, Gene density in each chromosome. d, GC content in each chromosome. **B** Phylogenetic tree analysis of Korla pear genome with number of gene family expansion and contraction. +, gene family expansion. -, gene family contraction. The scale represents divergence times (million years ago). **C** Alignment of Korla pear chromosomes shown by pairwise dot plots

common than the *P. communis* that diverged ~ 4.9-6.8 Mya. All other phylogenetic analyses strongly corroborated this finding (Supplementary Fig. 6 and Fig. 1B). Moreover, the Korla pear genome had greater similarity to that of *P. pyrifolia* (89.74%) than that of *P. betulifolia* (89.11%), *P. bretschneideri* (84.59%) and *Pyrus communis* (76.76%) based on the number of common families. These discoveries were partially consistent with previously reported estimates that Korla pear was a hybrid of *Pyrus communis* and *P. bretschneideri* or *P. pyrifolia* [6, 7, 97]. According to our analysis, the genetic information of Korla pear is more closely

Table 1 Sequences distribution of Korla pear chromosome-level genome

Chromosomes	Cluster Number	Cluster Length (bp)	Order Number	Order Length (bp)
Chr01	3	24,866,863	1	24,564,898
Chr02	2	30,561,580	1	30,533,258
Chr03	18	30,115,825	1	29,545,866
Chr04	1	23,100,000	1	23,100,000
Chr05	1	34,388,895	1	34,388,895
Chr06	1	26,681,330	1	26,681,330
Chr07	1	28,997,902	1	28,997,902
Chr08	1	25,579,965	1	25,579,965
Chr09	2	28,084,273	2	28,084,273
Chr10	1	34,166,355	1	34,166,355
Chr11	1	32,444,340	1	32,444,340
Chr12	1	21,764,265	1	21,764,265
Chr13	3	32,553,751	2	32,304,445
Chr14	1	27,211,198	1	27,211,198
Chr15	3	41,898,571	3	41,898,571
Chr16	1	30,492,774	1	30,492,774
Chr17	1	29,322,043	1	29,322,043
Total (Ratio %)	42 (48.84)	502,229,930 (99.18)	21 (50.0)	501,080,378 (99.77)

Table 2 Statistic gene annotation of Korla pear genome

#Anno_Database	Annotated_Number	Annotated_Ratio
GO_Annotation	32,132	83.84
KEGG_Annotation	29,880	77.96
KOG_Annotation	22,276	58.12
Swissprot_Annotation	31,514	82.22
TrEMBL_Annotation	38,238	99.77
eggNOG_Annotation	32,157	83.9
nr_Annotation	38,295	99.92
All_Annotated	38,296	99.92

related to that of *P. pyrifolia* than that of *P. bretschneideri* and *P. communis*. There had been identified that the *P. bretschneideri* was diverged from the *P. pyrifolia* [5]. On the other hand, the Korla pear showed higher synteny with *P. pyrifolia* than *P. bretschneideri* (Fig. 2A). Accordingly, all these results illustrated that the Korla pear and *P. pyrifolia* have exhibited a close genetic relationship, suggest a common ancestral divergence.

Pairwise comparisons of the 17 Korla pear chromosomes showed strong collinearity between the large segments of chromosomes 1 and 7, 2 and 15, 3 and 11, 4 and 12, 5 and 10, 6 and 14, 8 and 15, 9 and 17, and 13 and 16 (Fig. 1C; Fig. S7). Hence, a whole-genome duplication (WGD) event occurred in Korla pear. The WGD event was clearly identified by the distribution of the synonymous substitutions per synonymous site (*Ks*) and by fourfold degenerate transversion rate (4DTv) analysis (Fig. 2B). A sharp peak (~0.15) in the *Ks* distribution indicated a recent WGD event in evolutionary history (Fig. 2B). The *Ks* peak values and the time sequence after the WGD event also showed that Korla pear diverged from *Malus domestica*, *P. communis*, *P. betulifolia*, and *P. bretschneideri*. The *Ks* peak values for apple (0.20) and *P. betulifolia* (0.17) were close to that of Korla pear. Therefore, our results indicated that the pear and apple species were derived from a common ancestor, which is in agreement with previous research [96, 98]. The time of divergence for Korla pear aligned with the phylogenetic tree.

Eudicots originated from a common ancestor wherein a karyotype consisting of seven (pre-γ ancestor) protochromosomes evolved into 21 (post-γ ancestor) chromosomes through a paleohexaploid event (WGT-γ) [99]. No remnant of any ancient duplication event was detected in Korla pear, according to the *Ks* distribution analysis. However, there was an obvious *Ks* peak value (~1.6) in apple [98]. Pairwise comparisons and syntenic blocks of the genome disclosed high collinearity between Korla pear and apple. Thus, the chromosome structures of both genomes were stable (Supplementary Figs. 8 and 9). Identical segments in chromosomes 9 and 17, 6 and 14, and 13 and 16 resembled those in apple and were collinear with chromosomes 1, 14, and 17 in grape, which is an ancient hexaploid [100]. The remnant of a paleohexaploid state of the eudicot progenitor revealed that Korla

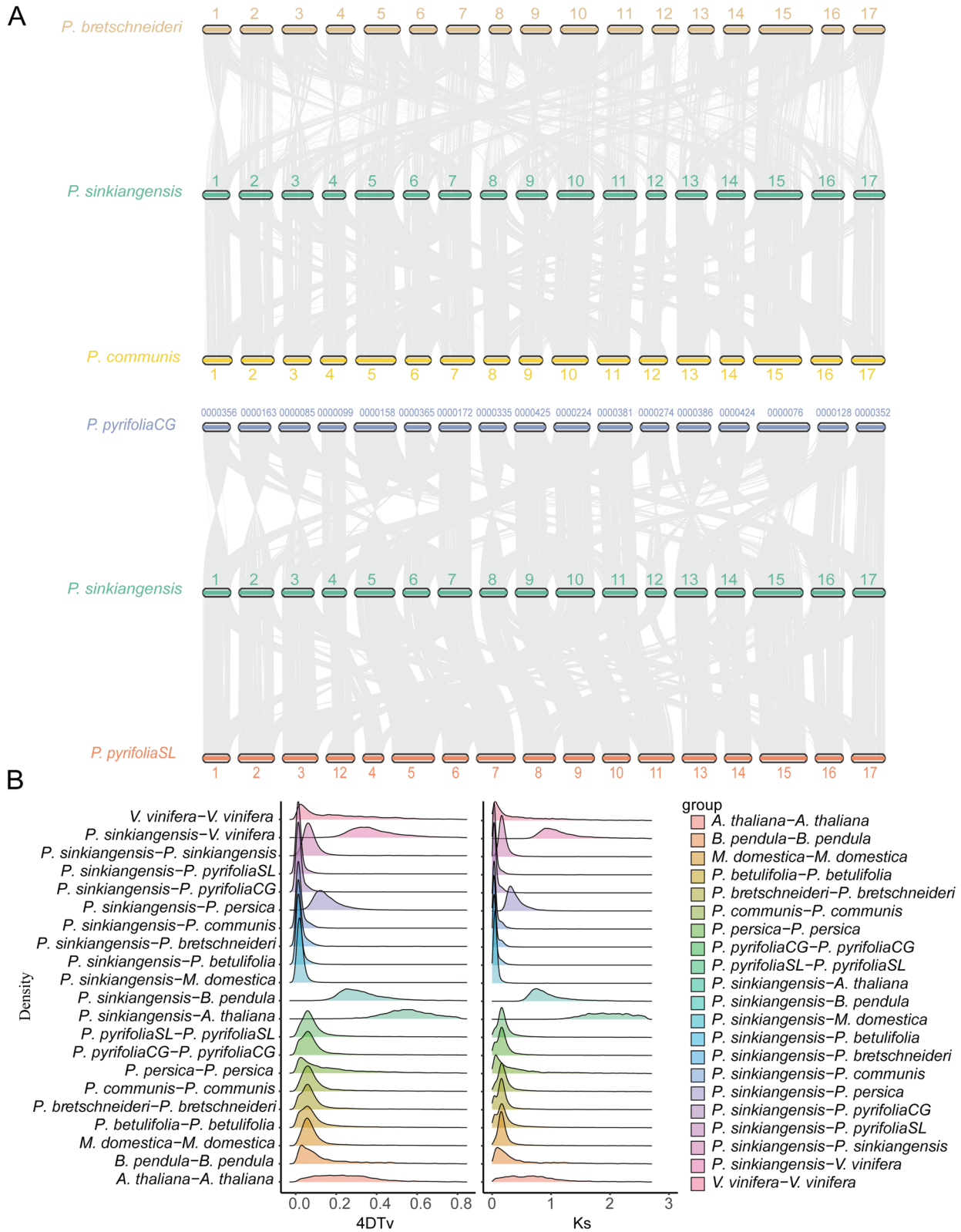

Fig. 2 Evolution of the Korla pear genome and comparative genomic analysis. **A** Collinearity analysis of Korla pear with White pear, European pear and Asian pear. **B** 4DTv and Ks distribution curve showing pairwise comparisons of Korla pear and other plant genomes

pear and apple experienced analogous paleoduplication events.

Comparative genomic analysis

We assigned 37,368 genes to 20,615 gene families in the Korla pear genome (Figure S10). A comparative analysis of Korla pear, *P. pyrifoliaSL*, *P. pyrifoliaCG*, *P. betulifolia*, and *P. bretschneideri* showed 14,595 gene families in all five species and 320 species-specific gene families in the Korla pear genome. There were 1,136 expansion and 7,332 contraction families through comparative analysis in Korla pear gene families(Fig. 1B). A GO enrichment analysis disclosed that the catalytic activity and transporter activity terms were the most abundant in the specific Korla pear gene families (Fig. S11). Most genes with catalytic activity were implicated in galactose metabolism, RNA polymerase, sphingolipid metabolism, pyrimidine metabolism, diterpenoid biosynthesis, amino sugar and nucleotide sugar metabolism, among other functions (Fig. S12). These genes might regulate the unique taste and aroma of Korla pear.

Pyrus betulifolia is a wild species with high tolerance to several abiotic and biotic stressors [101]. We found 6,058 gene families consisting of 12,761 genes that were specific to *P. betulifolia* (Fig. S10). The annotation revealed 151 genes in 85 orthogroups related to gene transcription, 353 genes in 188 orthogroups related to protein kinase activity, 54 genes in 23 orthogroups related to disease

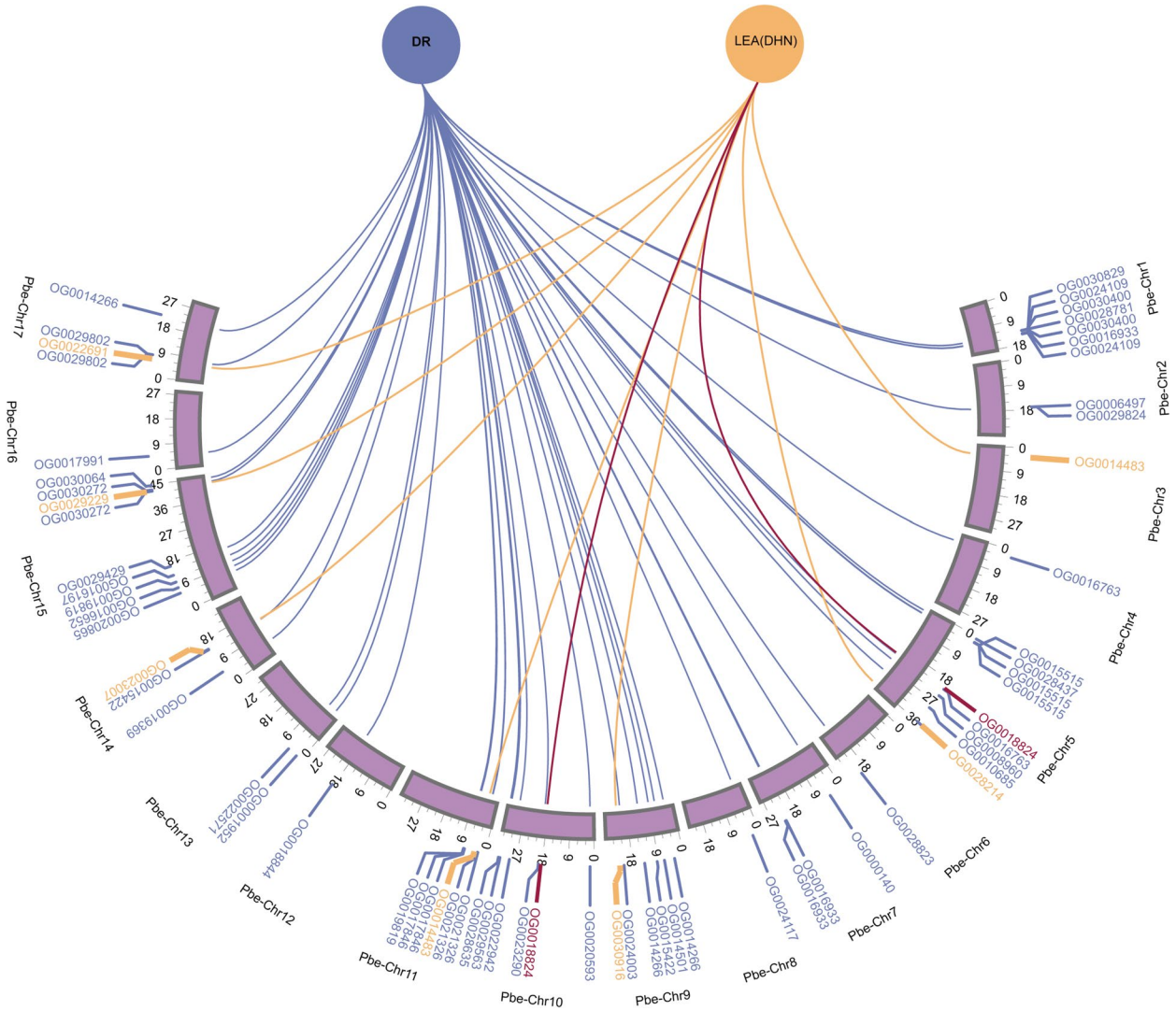

Fig. 3 LEA family and disease resistance genes mapped to *P. betulifolia* genome. Orange markers represent LEA genes deleted in the Korla pear genome. Red markers represent DHN genes deleted in the Korla pear genome. Blue markers represent disease resistance genes deleted in the Korla pear genome

resistance, and 9 genes in the LEA family possibly related to wintertime freezing resistance (Fig. 3).

Deciduous fruit trees have complex mechanisms enabling them to survive detrimental winter environments [102]. For example, phytohormone abscisic acid (ABA) induces dormancy and regulates the expression of genes encoding antifreeze proteins [103–105]. *LEA14* in *P. communis* encodes an antifreeze protein gene that confers wintertime freezing resistance. The expression patterns of this gene resemble those of *PcDREB1A* and *PcDREB1C* [36]. The *LEA* genes play important roles in freezing resistance under cold winter conditions. However, the Korla pear genome lost seven *LEA* and two *DHN* genes during long-term artificial selection. Hence,

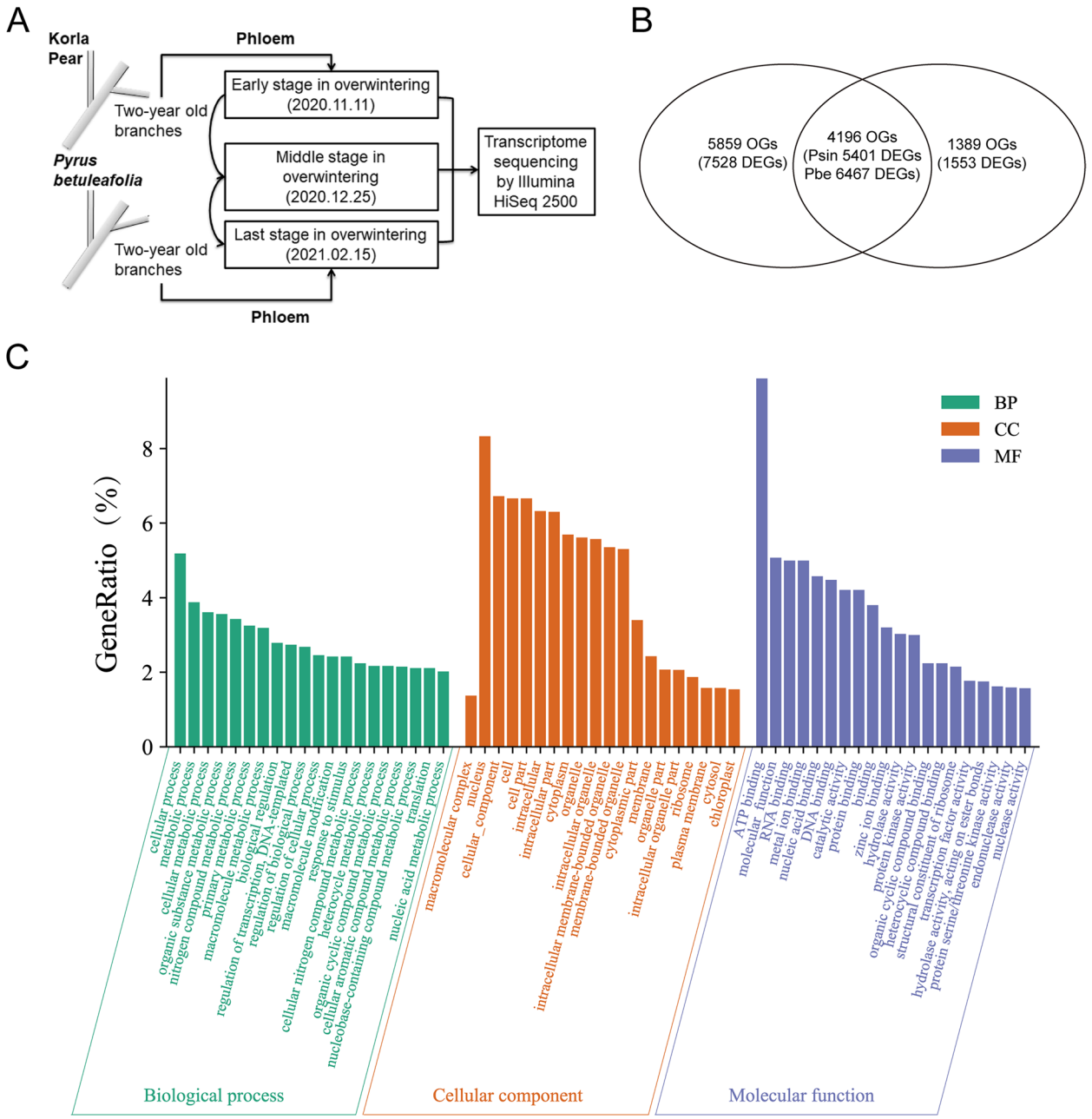

Fig. 4 Comparative transcriptomic analyses of Korla pear and *P. betulifolia*. **A** transcriptome sequencing sampling method for Korla pear and *P. betulifolia*. **B** Venn analysis of differentially expressed genes (DEGs) in Korla pear and *P. betulifolia*. OG: orthogroup ID. **C**, GO enrichment analysis of significantly differentially expressed genes in *P. betulifolia* but not Korla pear

Korla pear might have lower freezing resistance and higher disease susceptibility during winter.

Microorganism-associated molecular patterns (MAMPs) elicit vital pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) in plants. Nevertheless, 54 genes encoding disease resistance were depleted in the Korla pear genome compared with that of *P. betulifolia*. Most of these genes encode pattern

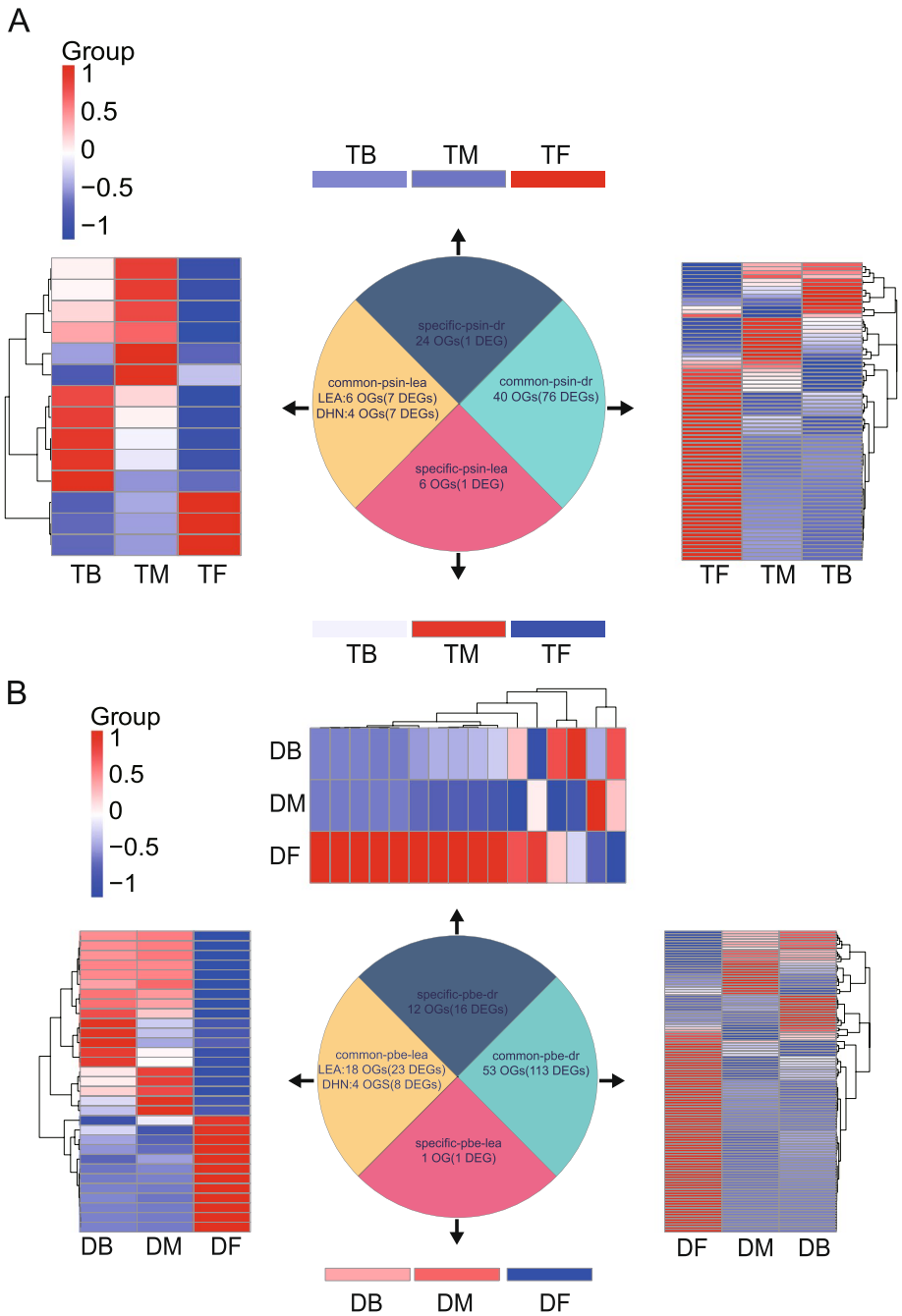

Fig. 5 Expression profiles of LEA family and disease resistance genes in overwintering pear species. **A** expression profiles of LEA family and disease resistance genes in Korla pear. **B** expression profiles of LEA family and disease resistance genes in *P. betulifolia*. Common-pbe/psin-lea: LEA genes common to Korla pear and *P. betulifolia*. Common-pbe/psin-dr: disease resistance genes common to Korla pear and *P. betulifolia*. Specific-pbe/psin-lea: LEA genes localized only to *P. betulifolia* or Korla pear genome. Specific-pbe/psin-dr: disease resistance genes localized only to *P. betulifolia* or Korla pear genome. OG: orthogroup. TB, TM, and TF: early, middle, and late overwintering stage, respectively, in Korla pear sampling. DB, DM, and DF: early, middle, and late overwintering stage, respectively, in *P. betulifolia* sampling

recognition receptors (PRRs) required to elicit PTI. Chitin elicitor receptor kinase 1 (CERK1) is a key immunoregulator in many plants [106–110]. No NOD-like receptors (NLRs) triggering ETI were identified among the 34 genes. However, a heat shock protein 90 (HSP90) required for nucleotide-binding leucine-rich domain (NLR) maturation was detected [111]. Therefore, both PTI and ETI immunity were attenuated in Korla pear compared with *P. betulifolia*.

Transcriptome sequencing of Korla pear and *P. betulifolia* during overwintering

Plant overwintering is highly complex and includes freezing and disease resistance during dormancy. We subjected early-, mid-, and late winter branch phloem samples to transcriptome sequencing to understand the differences in overwintering ability between Korla pear and *P. betulifolia* (Fig. 4A). We identified 7,526 DEGs by comparing the three wintertime stages of Korla pear (Table S17). By contrast, we found 15,676 DEGs in *P. betulifolia* (Table S18). A Venn analysis revealed 4,195 OGs (5,401 DEGs in Korla pear and 6,467 DEGs in *P. betulifolia*) for both Korla pear and *P. betulifolia* in winter, 2,125 DEGs (including 1,553 DEGs in 1,390 OGs) specific to Korla pear, and 9,209 DEGs (including 7,528 DEGs in 5,860 OGs) specific to *P. betulifolia* (Fig. 4B).

We performed GO enrichment analyses of the specific DEGs in *P. betulifolia* to demonstrate the differences between Korla pear and *P. betulifolia*. Twenty GO terms were assigned to the biological process category, including "cellular process" (5.18%), "response to stimulus" (2.42%), and "regulation of transcription, DNA-templated" (2.74%) (Fig. 4C). The most highly enriched GO term under the molecular function category was "ATP binding" (9.88%). The most highly enriched GO term under the cellular component category was "nucleus" (8.33%). The foregoing results showed that either numerous genes implicated in overwintering were downregulated or they were depleted in the Korla genome compared with that of *P. betulifolia*.

Expression analysis of genes regulating freezing and disease resistance

Certain LEA family genes were lost in the Korla pear genome compared with that of *P. betulifolia* (Fig. 3). There were 75 *LEA* genes and 10 *Dehydrin* genes (*DHNs*) in the Korla pear genome. Of these, seven *LEA* and eight *DHN* genes participated in overwintering (Fig. 5A). However, 24 *LEA* and eight *DHN* genes in *P. betulifolia* were differentially expressed during winter. There were 18 orthogroups (21 *LEA* genes) in the Korla pear genome, but they were not expressed during overwintering (Fig. 6). Hence, mutation of the Korla pear genome during long-term domestication resulted in downregulation of the genes induced by low temperatures and other signals. Seventy-seven disease resistance genes were differentially expressed in Korla pear during overwintering (Fig. 5B). Compared with *P. betulifolia*, there were 17 orthogroups with 31 disease resistance genes in the Korla pear genome that were not differentially expressed during overwintering (Fig. 6). Hence, wintertime disease resistance was markedly lower in Korla pear than it was in *P. betulifolia*.

Plants in various habitats have evolved different molecular mechanisms to adapt to their ambient environments. In *Arabidopsis,* core binding factor genes (*CBFs*) regulate cold acclimation required to resist freezing stress. The *CBF* regulon also controls various cold-responsive genes (CORs; *LEAs* and *DHNs*). However, tomato and *Arabidopsis CBF* overexpression did not improve freezing stress tolerance in tomato; the *CBF* regulons controlled only four genes in tomato. Certain genes detected in the Korla pear and *P. betulifolia* genomes were only differentially expressed in *P. betulifolia* during overwintering. Certain *cis* elements of these common gene promoters were deleted in Korla pear compared with those of *P. betulifolia* (Table 3). The *cis* elements such as GATA motif, G box, ABRE, and others associated with low temperature or ABA signal induction, were absent in nearly all these genes (87%) (Table 3). The most vital *cis* elements associated with

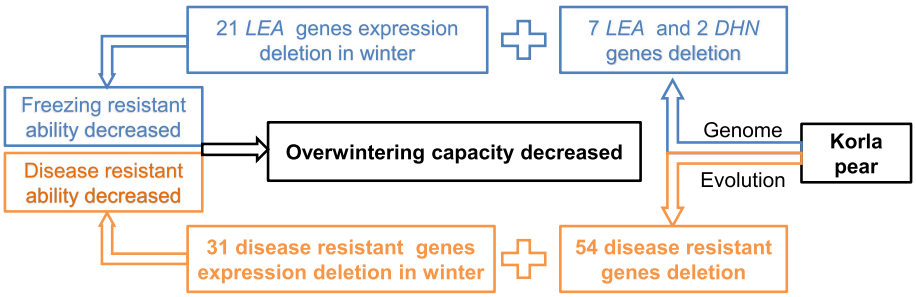

Fig. 6 Evolutionary mechanism by which disease resistance and winter hardness decreased in Korla pear

Table 3 Deletion Cis-element of Korla pear genes comparison with *P. betulifolia*

Orthogroups	Gene ID	Deletion Cis-element type
OG0000663	Psin03G015750	**G-box**, WUN-motif, MRE, **GATA-motif**, TCT-motif, MSA-like, ARE, **GARE-motif**, ATCT-motif
	Psin03G015770	CAT-box, **G-box**, **MYB**, MRE, **GATA-motif**, MSA-like, Box 4, **GARE-motif**, Myb-binding site, **LTR**, ATCT-motif
	Psin11G017180	**G-box**, ATCT-motif
OG0001765	Psin05G025970	**G-box**, WUN-motif, **circadian**, MRE, AE-box, TCT-motif, **ERE**, **TCA**, **TGA-element**, **TC-rich repeats**, **DRE core**
	Psin10G024390	MRE, AE-box, **ERE**, **LTR**, **TGA-element**, ATCT-motif
OG0002397	Psin02G003780	GT1-motif, TCT-motif, MBS, CARE, Box 4, **Myb**, GCN4_motif, **TCA-element**
	Psin15G014940	CARE, Box 4, GCN4_motif, **TCA-element**
OG0004580	Psin13G014220	AE-box, **I-box**, LAMP-element
	Psin14G003900	TATC-box, **ABRE4**, AE-box, GATA-motif, WRE3, ABRE3a, **I-box**, re2f-1, box S, LAMP-element
OG0006572	Psin04G011310	TATC-box, MRE, AE-box, TCT-motif, Box 4. **Myb**, **GARE-motif**, Gap-box, AACA_motif, chs-Unit 1 m1
	Psin12G013410	**Sp1**
OG0007775	Psin04G011330	**CGTCA-motif**, **ABRE4**, as-1, **W box**, AE-box, **ABRE3a**, TCT-motif, **TGACG-motif**, ERE, **AT ~ ABRE**, **TGA-box**
	Psin12G013430	**MYC**, AAGAA-motif, AC-I
OG0007905	Psin06G008070	**Circadian**, GTGGC-motif, O_2-site
	Psin14G009930	GTGGC-motif, O_2-site, **TCA**
OG0009066	Psin01G006360	GATA-motif, **MBS**, **Myb-binding site**, LAMP-element
OG0009468	Psin03G010340	Box 4
	Psin11G011430	
OG0012178	Psin06G008090	**TGA-element**, AT1-motif
	Psin14G009940	**P-box**, WRE3, **MYB recognition site**, CCAAT-box, Gap-box, **TC-rich repeats**, chs-CMA1a, AT1-motif
OG0017813	Psin17G018570	WUN-motif, **TGA-element**, Gap-box, chs-CMA1a
OG0019795	Psin14G008620	TATC-box, AT-rich sequence, MRE, MSA-like, Box 4, **MYB**, CCAAT-box, AuxRR-core, **LTR**, Box II -like sequence, TC-rich repeats, TGA-box
OG0000881	Psin11G006020	
	Psin11G006030	**G-Box**, **ABRE4**, WUN-motif, **ABRE3a**, **ABRE**
	Psin11G006040	
OG0002053	Psin01G001630	**MBS**, Myc, **MYB**, STRE, CCAAT-box, **ERE**, GTGGC-motif, **I-box**, **TCA-element**, **TGA-element**, ATCT-motif, GC-motif, CTAG-motif, **DRE1**, GA-motif, chs-CMA1a
	Psin02G006920	**G-Box**, **MBSI**, AAGAA-motif, MBS, **MYB**, CCAAT-box, GTGGC-motif, **TCA-element**, **TGA-element**, GC-motif, CTAG-motif, **DRE1**, GA-motif, chs-CMA1a
	Psin02G019970	**P-box**, MRE, MBSI, TCT-motif, Myc, **MYB**, CCAAT-box, GTGGC-motif, **TCA-element**, ATCT-motif, GC-motif, CTAG-motif, **DRE1**, GA-motif, chs-CMA1a
OG0002884	Psin01G017180	**ABRE4**, CAT-box, GT1-motif, W box, **MYB**, GATA-motif, **ABRE3a**, CCAAT-box, **I-box**, AT-rich element, O_2-site, **TCA**, TC-rich repeats
	Psin07G020990	**MYB**, GATA-motif, Box 4, CCAAT-box, **I-box**, **Sp1**, TC-rich repeats
OG0002938	Psin04G006730	AAGAA-motif, TCT-motif, **MYB**, O_2-site
	Psin04G006740	AAGAA-motif
OG0004845	Psin13G006630	WUN-motif, GATA-motif, MSA-like, **MYB**, CCAAT-box, **ERE**, ARE, AuxRR-core, ATCT-motif
	Psin16G006650	MRE, **MYB**, ATCT-motif
OG0004943	Psin13G005060	**CGTCA-motif**, **ABRE4**, **circadian**, as-1, W box, F-box, **ABRE3a**, TCCC-motif, **TGACG-motif**, **MBS**, **ERE**, **GARE-motif**, GTGGC-motif, **I-box**, AT-rich element, AuxRR-core, TC-rich repeats, 3-AF3 binding site
	Psin16G005030	GTGGC-motif, AT-rich element, **TCA-element**, AuxRR-core, TC-rich repeats
OG0005094	Psin13G000160	TCCC-motif, GTGGC-motif
	Psin16G000120	**CGTCA-motif**, **G-Box**, **ABRE4**, as-1, WRE3, **ABRE3a**, TCCC-motif, **TGACG-motif**, **MYB**, CCAAT-box, GTGGC-motif, GCN4_motif, **Sp1**, chs-CMA2a, AuxRE
OG0009377	Psin07G011590	AAAC-motif, W box, **MYB**
	Psin09G008440	AAAC-motif, **MYB**, GATA-motif, CCAAT-box, O_2-site
	Psin17G008000	TC-rich repeats
OG0009455	Psin15G011450	TATC-box, **P-box**, GATA-motif, TCCC-motif, TC-rich repeats, box S

Table 3 (continued)

Orthogroups	Gene ID	Deletion Cis-element type
OG0010430	Psin03G003220	MRE, GT1-motif, AE-box, **MYB**, Myc, **ERE**, **GARE-motif**, GATT-motif, **TGA-element**, Gap-box, ATCT-motif, 3-AF1 binding site, 3-AF3 binding site
	Psin11G003360	**G-Box**, MRE, **MBS**, Myc, **MYB**, **GARE-motif**, GATT-motif, **TCA-element**, **TGA-element**, Gap-box, ATCT-motif, 3-AF1 binding site, 3-AF3 binding site
OG0012298	Psin05G001200	**CGTCA-motif**, **G-Box**, AAAC-motif, CAT-box, as-1, **W box**, GATA-motif, WRE3, AAGAA-motif, **ABRE**, **TGACG-motif**, **MBS**, GTGGC-motif, **MYB**, AT-rich element, **TCA**, **LTR**, **Sp1**, **DRE1**
	Psin10G000970	AAAC-motif, CAT-box, W box, WRE3, TCT-motif, **ERE**, GTGGC-motif, AT-rich element, **Sp1**, TC-rich repeats, **DRE**
OG0012307	Psin13G002810	ACA-motif, AE-box, GATA-motif, MBS, CARE, Box 4, STRE, **GARE-motif**, **I-box**, chs-CMA2a, **DRE**
	Psin16G002870	ACA-motif, WUN-motif, TCT-motif, Box 4, **GARE-motif**, **I-box**, **LTR**, chs-CMA2a
OG0012332	Psin01G004080	**Circadian**, WRE3, **ABRE**, O_2-site, **TCA**, TC-rich repeats
OG0012474	Psin04G015630	**CGTCA-motif**, MRE, as-1, **TGACG-motif**, AuxRR-core, chs-CMA2a
OG0014546	Psin11G018420	**CGTCA-motif**, as-1, AAGAA-motif, **TGACG-motif**, **ERE**
OG0018032	Psin12G008700	WUN-motif
OG0021534	Psin09G008450	**ABRE4**, **MYB**, **ABRE3a**, **Myb**, GTGGC-motif, O_2-site, ATCT-motif, RY-element, CAG-motif

The Cis-elements with activities of inducing by low temperature or ABA signal are bolded

low temperature induction, including DRE1, ABRE, and LTR, were deleted from 20 genes (38%). W box is a key *cis* element that is bound by WRKY transcription factors (TFs) and is implicated in biotic stress resistance. It was absent in six disease resistance genes (19.4%). Therefore, the 53 DEGs in overwintering *P. betulifolia* but not Korla pear were attributed to the loss of key *cis* elements with low temperature tolerance and ABA signaling activity in Korla pear.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-05490-x.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements
Thanks to the National Natural Science Foundation of China and the Bingtuan Key Science and Technology Program of Xinjiang Province for the financial support of this experiment, and Professor Lin Zhongping of Peking University for his comments on the design of this experiment and the revision of this paper.

Authors' contributions
Jin Li and Jianbo Zhu conceived the study. Jin Li, Jianbo Zhu, Saisai Wang, and Wenwen Xia designed the experiments. Wenwen Xia, Saisai Wang, Xiaoyan Liu, Yifei Chen, and Ruina Liu performed the experiments and analyzed the data. Caixia Lin and Hailiang Liu helped in the experimentation. Jin Li, Jianbo Zhu, Wenwen Xia, and Saisai Wang wrote the paper. All authors read and approved the manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (32260726), Bingtuan Key Science and Technology Program of Xinjiang Province (2018AB035) and Natural Science Foundation of Shihezi University (RCZK202047 and CXFZ202006).

Availability of data and materials
Sequence data of Korla pear genome can be found in the China National Center for Bioinformation database under BioProject codes PRJCA007928 (https://ngdc.cncb.ac.cn/search/?dbId=&q=PRJCA007928) and GWH accession NO. GWHDTVS00000000. Other relevant data are within the manuscript and its additional files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 16 April 2024 Accepted: 6 August 2024
Published online: 14 August 2024

References

1. Rom RC, Carlson RFE. Rootstocks fruit crops. New York : John Wiley and Sons; 1987.
2. Rubtsov GA. Geographical Distribution of the genus Pyrus and trends and factors in its evolution. Am Nat. 1944;78(777):358–66.
3. Zhukovskiĭ PM. Dictionary of cultivated plants and their centres of diversity. Wageningen: Centre for Agricultural Publishing and Documentation; 1975.
4. Zhang S, Qian M, Yin H, Li X, Wu J, Qi K, Wu X. Pedigree analysis of pear varieties (lines) bred in China. Acta Horticulturae Sinica. 2018;45:2291–307.
5. Wu J, Wang Y, Xu J, Korban SS, Fei Z, Tao S, Ming R, Tai S, Khan AM, Postman JD, et al. Diversification and independent domestication of Asian and European pears. Genome Biol. 2018;19(1):77.
6. Pu F, Wang Y. Some problems in the classification of Pyrus in China. China Fruits. 1960;1(3):25–7.
7. Yu D, Zhang P. Xinjiang pears, a new series of cultivars of pears in China. Acta Horticulturae Sinica. 1979;6(1):27–32.

8. Teng Y, Tanabe K, Tamura F, Itai A. Genetic relationships of pear cultivars in Xinjiang, China, as measured by RAPD markers. J Hortic Sci Biotechnol. 2001;76(6):771–9.
9. Yz Ma, Yx Z. Classification position of Ku'erle aromatic pear Identified by RAPD. Southwest China J Agri Sci. 2009;22(3):743–5.
10. Chen C, Li B-H, Dong X-L, Wang C-X, Lian S, Liang W-X. Effects of temperature, humidity, and wound age on Valsa mali infection of apple shoot pruning wounds. Plant Dis. 2016;100(12):2394–401.
11. Zhai L, Zhang M, Lv G, Chen X, Jia N, Hong N, Wang G. Biological and molecular characterization of four botryosphaeria species isolated from pear plants showing stem wart and stem canker in China. Plant Dis. 2014;98(6):716–26.
12. Yin Z, Liu H, Li Z, Ke X, Dou D, Gao X, Song N, Dai Q, Wu Y, Xu JR, et al. Genome sequence of Valsa canker pathogens uncovers a potential adaptation of colonization of woody bark. New Phytol. 2015;208(4):1202–16.
13. Kange AM, Xia A, Si J, Li B, Zhang X, Ai G, He F, Dou D. The fungal-specific transcription factor VpFSTF1 is required for virulence in Valsa pyri. Front Microbiol. 2019;10:2945.
14. He F, Zhang X, Mafurah JJ, Zhang M, Qian G, Wang R, Safdar A, Yang X, Liu F, Dou D. The transcription factor VpCRZ1 is required for fruiting body formation and pathogenicity in Valsa pyri. Microb Pathog. 2016;95:101–10.
15. Li Z, Gao X, Kang Z, Huang L, Fan D, Yan X. Saccharothrix yanglingensis strain Hhs.015 is a promising biocontrol agent on apple Valsa canker. Plant Disease. 2016;100(2):510–4.
16. Liu J, Nie J, Chang Y, Huang L. Nep1-like proteins from Valsa mali differentially regulate pathogen virulence and response to abiotic stresses. J Fungi. 2021;7:830.
17. Wang W, Nie J, Lv L, Gong W, Wang S, Yang M, Xu L, Li M, Du H, Huang L. A Valsa mali effector protein 1 targets apple (Malus domestica) pathogenesis-related 10 protein to promote virulence. Front Plant Sci. 2021;12:741342.
18. Zhang M, Zhai L, Hu H, Zhang B, Wang G. Laboratory evaluation for the resistance of Pyrus germplasm accession to Valsa canker. Acta Horticulturae Sinica. 2014;41(7):1297–306.
19. Yuan M, Jiang Z, Bi G, Nomura K, Liu M, Wang Y, Cai B, Zhou J-M, He SY, Xin X-F. Pattern-recognition receptors are required for NLR-mediated plant immunity. Nature. 2021;592(7852):105–9.
20. Ma X, Xu G, He P, Shan L. SERKing coreceptors for receptors. Trends Plant Sci. 2016;21(12):1017–33.
21. Ma X, Claus LAN, Leslie ME, Tao K, Wu Z, Liu J, Yu X, Li B, Zhou J, Savatin DV, et al. Ligand-induced monoubiquitination of BIK1 regulates plant immunity. Nature. 2020;581(7807):199–203.
22. Zhang M, Feng H, Zhao Y, Song L, Gao C, Xu X, Huang L. Valsa mali pathogenic effector VmPxE1 contributes to full virulence and interacts with the host peroxidase MdAPX1 as a potential target. Front Microbiol. 2018;9:821.
23. Zhang J, Shao F, Li Y, Cui H, Chen L, Li H, Zou Y, Long C, Lan L, Chai J, et al. A pseudomonas syringae effector inactivates MAPKs to suppress PAMP-induced immunity in plants. Cell Host Microbe. 2007;1(3):175–85.
24. Zhou K, Hu L, Li Y, Chen X, Zhang Z, Liu B, Li P, Gong X, Ma F. MdUGT88F1-Mediated phloridzin biosynthesis regulates apple development and Valsa canker resistance. Plant Physiol. 2019;180(4):2290–305.
25. Meng X, Yang R, Liu A-T, Hu T, Wang YN, Cao K, Wang S. The influence of lower temperature induction of Valsa mali on the infection of apple trees. Plant Dis. 2021;105(10):2776–80.
26. Liu K, Han X, Liang Z, Yan J, Cong P, Zhang C. Enome-wide identification, classification, and expression analysis of the HD-Zip transcription factor family in apple (Malus domestica Borkh.). Int J Mol Sci. 2022;23(5):2632.
27. Liu X, Wang L, Zhang X, Li A, Xia W, Lin C, Li J, Zhu J. Expression of the Pyrus sinkiangensis HD-Zip **I** transcription factor PsHB7 and PsHB12 in hybrid Broussonetia papyrifera regulates its natural overwintering. Environ Exp Bot. 2023;216:105534.
28. Lang GA, Early JD, Darnell R, Martin GC. Endo-, Para-, and Eco-dormancy: physiological terminology and classification for dormancy research. HortScience. 1987;22:371–7.
29. Samish RM. Dormancy in woody plants. Annu Rev Plant Physiol. 1954;5(1):183–204.
30. Tuan PA, 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. 2017;58(8):1378–90.
31. Michailidis M, Karagiannis E, Tanou G, Sarrou E, Adamakis I-D, Karamanoli K, Martens S, Molassiotis A. Metabolic mechanisms underpinning vegetative bud dormancy release and shoot development in sweet cherry. Environ Exp Bot. 2018;155:1–11.
32. Tylewicz S, Petterle A, Marttila S, Miskolczi P, Azeez A, Singh RK, Immanen J, Mähler N, Hvidsten TR, Eklund DM, et al. Photoperiodic control of seasonal growth is mediated by ABA acting on cell-cell communication. Science. 2018;360(6385):212–5.
33. Thomashow MF. Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annu Rev Plant Physiol Plant Mol Biol. 1999;50:571–99.
34. Welling A, Kaikuranta P, Rinne P. Photoperiodic induction of dormancy and freezing tolerance in Betula pubescens. Involvement of ABA and dehydrins. Physiologia Plantarum. 1997;100(1):119–25.
35. Welling A, Palva ET. Molecular control of cold acclimation in trees. Physiol Plant. 2006;127(2):167–81.
36. Shibuya T, Itai R, Maeda M, Kitashiba H, Isuzugawa K, Kato K, Kanayama Y. Characterization of PcLEA14, a group 5 late embryogenesis abundant protein gene from pear (Pyrus communis). Plants (Basel. 2020;9(9):1138.
37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.
38. Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. Bioinformatics. 2008;24(5):713–4.
39. Cheng H, Concepcion G, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly with phased assembly graphs. 2020.
40. Parra G, Bradnam K, Korf I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics. 2007;23(9):1061–7.
41. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210–2.
42. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics. 2009;25(14):1754–60.
43. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D Map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159(7):1665–80.
44. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, Heard E, Dekker J, Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 2015;16(1):259.
45. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using diamond. Nat Methods. 2015;12(1):59–60.
46. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-H, Jin H, Marler B, Guo H, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49–e49.
47. Xu Y, Bi C, Wu G, Wei S, Dai X, Yin T, Ye N. A web-based vector graph toolkit of genome synteny and collinearity. Biomed Res Int. 2016;2016:7823429.
48. Tang H, Krishnakumar V, Zeng X, Xu Z, Taranto A, Lomas JS, Zhang Y, Huang Y, Wang Y, Yim WC et al. JCVI: A versatile toolkit for comparative genomics analysis. iMeta. 2024;3(4):e211.
49. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 2020;117(17):9451–7.
50. Bao Z, Eddy SR. Automated de novo identification of repeat sequence families in sequenced genomes. Genome Res. 2002;12(8):1269–76.
51. Price AL, Jones NC, Pevzner PA. De novo identification of repeat families in large genomes. Bioinformatics. 2005;21(suppl_1):i351–8.
52. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase update, a database of eukaryotic repetitive elements. Cytogenet Genome Res. 2005;110(1–4):462–7.
53. Neumann P, Novák P, Hoštáková N, Macas J. Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. Mob DNA. 2019;10(1):1.
54. Wheeler TJ, Clements J, Eddy SR, Hubley R, Jones TA, Jurka J, Smit AFA, Finn RD. Dfam: a database of repetitive DNA based on profile hidden Markov models. Nucleic Acids Res. 2013;41(D1):D70–82.

55. Ou S, Jiang N. LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. Plant Physiology. 2017;176(2):1410–22.
56. Ellinghaus D, Kurtz S, Willhoeft U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinformatics. 2008;9:18.
57. Xu Z, Wang H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. 2007;35(Web Server issue):W265–8.
58. Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics. 2009;25(1):4.10.1-4.10.14
59. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573–80.
60. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33(16):2583–5.
61. Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics. 2008;24(5):637–44.
62. Korf I. Gene finding in novel genomes. BMC Bioinformatics. 2004;5(1):59.
63. Keilwagen J, Wenk M, Erickson JL, Schattat MH, Grau J, Hartung F. Using intron position conservation for homology-based gene prediction. Nucleic Acids Res. 2016;44(9).
64. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–60.
65. Tang S, Lomsadze A, Borodovsky M. Identification of protein coding regions in RNA transcripts. Nucleic Acids Res. 2015;43(12):e78.
66. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5.
67. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29(7):644–52.
68. Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK Jr, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD, et al. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res. 2003;31(19):5654–66.
69. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. Genome Biol. 2008;9(1):R7.
70. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25(5):955–64.
71. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res. 2005;33(Database issue):D121-124.
72. Loman T. A novel method for predicting ribosomal RNA genes in prokaryotic genomes. 2017.
73. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Research. 2006;34(suppl_1):D140–4.
74. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics. 2013;29(22):2933–5.
75. Birney E, Clamp M, Durbin R. GeneWise and genomewise. Genome Res. 2004;14(5):988–95.
76. She R, Chu JS, Wang K, Pei J, Chen N. GenBlastA: enabling BLAST to identify homologous gene sequences. Genome Res. 2009;19(1):143–9.
77. Boeckmann B, Bairoch A, Apweiler R, Blatter M-C, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, et al. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. 2003;31(1):365–70.
78. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k -mer weighting and repeat separation. Genome Res. 2017;27(5):722–36.
79. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund S, Cook H, Mende D, Letunic I, Rattei T, Jensen L, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2018;47(D1):D309–14.
80. Finn R, Mistry J, Schuster-Böckler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Eddy S, et al. PFAM: clans, web tools and services. Nucleic Acids Res. 2006;34:D247–51.
81. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236–40.
82. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20(1):1–14.
83. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2018;47(D1):D419–26.
84. Katoh K, Asimenos G, Toh H. Multiple alignment of DNA sequences with MAFFT. In: Posada D, editor. Bioinformatics for DNA sequence analysis. Totowa, NJ: Humana Press; 2009. p. 39–64.
85. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007;56(4):564–77.
86. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74.
87. Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587–9.
88. Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. Bioinformatics. 1997;13(5):555–6.
89. Puttick MN. MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. Bioinformatics. 2019;35(24):5321–2.
90. Han MV, Thomas GWC, Lugo-Martinez J, Hahn MW. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Mol Biol Evol. 2013;30(8):1987–97.
91. Zwaenepoel A, Van de Peer Y. wgd—simple command line tools for the analysis of ancient whole-genome duplications. Bioinformatics. 2019;35(12):2153–5.
92. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.
93. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
94. Varet H, Brillet-Gueguen L, Coppee JY, Dillies MA. SARTools: A DESeq2- and EdgeR-based R pipeline for comprehensive differential analysis of RNA-seq data. PLoS ONE. 2016;11(6):e0157022.
95. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nat Biotechnol. 2013;31(12):1119–25.
96. Dong X, Wang Z, Tian L, Zhang Y, Qi D, Huo H, Xu J, Li Z, Liao R, Shi M, et al. De novo assembly of a wild pear (Pyrus betuleafolia) genome. Plant Biotechnol J. 2020;18(2):581–95.
97. Yu D. Geographical distribution of fruit tree resources in China. Acta Horticulturae Sinica. 1962;1(1):2–14.
98. 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 x domestica Borkh.). Nat Genet. 2010;42(10):833–9.
99. Salse J. Ancestors of modern plant crops. Curr Opin Plant Biol. 2016;30:134–42.
100. Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449(7161):463–7.
101. Pu F, Wang Y. Chinese Journal of fruit Chi, vol. 3. Shanghai: Shanghai Science and Technology Press; 1963.
102. Faust M, Erez A, Rowland LJ, Wang SY, Norman HA. Bud dormancy in perennial fruit trees: physiological basis for dormancy induction, maintenance, and release. HortScience. 1997;32:623–9.
103. Seeley SD, Powell LE. Seasonal changes of free and hydrolyzable abscisic acid in vegetative apple buds. J Amer Soc Hort Sci. 1981;106:405–9.

104. Rodriguez A, Sanchez Tames R. Dormancy and seasonal changes of plant growth regulators in hazel corylus avellana cultivar negreta buds. Physiol Plant. 1986;66(2):288–92.
105. Agarwal PK, Agarwal P, Reddy MK, Sopory SK. Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant Cell Rep. 2006;25(12):1263–74.
106. Miya A, Albert P, Shinya T, Desaki Y, Ichimura K, Shirasu K, Narusaka Y, Kawakami N, Kaku H, Shibuya N. CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. Proc Natl Acad Sci. 2007;104(49):19613–8.
107. Miyata K, Kozaki T, Kouzai Y, Ozawa K, Ishii K, Asamizu E, Okabe Y, Umehara Y, Miyamoto A, Kobae Y, et al. The bifunctional plant receptor, OsCERK1, regulates both chitin-triggered immunity and arbuscular mycorrhizal symbiosis in rice. Plant Cell Physiol. 2014;55(11):1864–72.
108. Lee W-S, Rudd JJ, Hammond-Kosack KE, Kanyuka K. Mycosphaerella graminicola LysM effector-mediated stealth pathogenesis subverts recognition through both CERK1 and CEBiP homologues in wheat. Mole Plant-Microbe Interact®. 2014;27(3):236–43.
109. Liao D, Sun X, Wang N, Song F, Liang Y. Tomato LysM receptor-like kinase SlLYK12 Is involved in arbuscular mycorrhizal symbiosis. Front Plant Sci. 2018;9:1004.
110. Chen Q, Dong C, Sun X, Zhang Y, Dai H, Bai S. Overexpression of an apple LysM-containing protein gene, MdCERK1–2, confers improved resistance to the pathogenic fungus, alternaria alternata, in nicotiana benthamiana. BMC Plant Biol. 2020;20(1):146.
111. Kadota Y, Shirasu K. The HSP90 complex of plants. Biochim Biophys Acta (BBA) - Mol Cell Res. 2012;1823(3):689–97.

Publisher's Note