A chromosome level genome assembly of Pseudoroegneria libanotica reveals a key Kcs gene involves in the cuticular wax elongation for drought resistance

Dandan Wu[1], Xingguang Zhai[1], Chen Chen[1], Xunzhe Yang[1], Shaobo Cheng[1], Lina Sha[1], Yiran Cheng[1], Xing Fan[1], Houyang Kang[1], Yi Wang[1], Dengcai Liu[1], yonghong zhou[1], and Haiqin Zhang[1]

[1]Sichuan Agricultural University - Chengdu Campus

May 23, 2023

Abstract

The St genome of Pseudoroegneria (Triticeae, Poaceae) not only accounts for more than 60% perennial speciation, but also prominent for forage and crop breeding. The diploid Pseudoroegneria libanotica with more ancient St genome is covered by cuticular wax on the aerial part, and exhibited strong drought resistance. To reveal the genetic relationship among Triticeae species and illustrate the mechanism between water deficit and cuticular wax formation, in this study, we (1) assembled the chromosome level St genome of Pse. libanotica with 2.99 Gb assembled into seven pseudochromosomes; less repeat clusters (TEs) is the main reason for the St genome with smaller genome size, and high genome heterozygosity might cause abundant St-containing speciation; (2) the genus Pseudoroegneria diverged during the middle and late Miocene, and unique genes, gene family expansion and contraction in Pse. libanotica were enriched in biotic and abiotic stresses, such as fatty acid biosynthesis which may be greatly contribute to the its drought adaption; (3) in total, 14 genes were involved in wax biosynthesis under 28 days drought treatment, more importantly, a new Kcs gene evm.TU.CTG175.54 plays a critical role in the very long chain fatty acid (VLCFA) elongation from C18 to C26 in Pse. libanotica. Our study lays a foundation for the genome diversification of Triticeae species and deciphers cuticular wax formation genes that have contributed to the drought resistance of Pse. libanotica.

A chromosome level genome assembly of *Pseudoroegneria libanotica* reveals a key *Kcs* gene involves in the cuticular wax elongation for drought resistance

Dandan Wu[1, 2#], Xingguang Zhai[2#], Chen Chen[2#], Xunzhe Yang[2], Shaobo Cheng[2], Lina Sha[1,3], Yiran Cheng[1], Xing Fan[1,2], Houyang Kang[1,2], Yi Wang[1,2], Dengcai Liu[1,2], Yonghong Zhou[1, 2*] and Haiqin Zhang[1, 3*]

[1] State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China,Sichuan Agricultural University, Chengdu 611130, Sichuan, China

[2] Triticeae Research Institute, Sichuan Agricultural University, Chengdu 611130, Sichuan, China

[3] College of Grassland Science and Technology, Sichuan Agricultural University, Chengdu 611130, Sichuan, China

*Correspondence

Yonghong Zhou (*zhouyh@sicau.edu.cn*),

Haiqin Zhang (haiqinzhang@163.com)

# These authors contributed equally to this work.

Abstract

The St genome of *Pseudoroegneria* (Triticeae, Poaceae) not only accounts for more than 60% perennial speciation, but also prominent for forage and crop breeding. The diploid *Pseudoroegneria libanotica* with more ancient St genome is covered by cuticular wax on the aerial part, and exhibited strong drought resistance. To reveal the genetic relationship among Triticeae species and illustrate the mechanism between water deficit and cuticular wax formation, in this study, we (1) assembled the chromosome level St genome of *Pse. libanotica* with 2.99 Gb assembled into seven pseudochromosomes; less repeat clusters (TEs) is the main reason for the St genome with smaller genome size, and high genome heterozygosity might cause abundant St-containing speciation; (2) the genus *Pseudoroegneria* diverged during the middle and late Miocene, and unique genes, gene family expansion and contraction in *Pse. libanotica* were enriched in biotic and abiotic stresses, such as fatty acid biosynthesis which may be greatly contribute to the its drought adaption; (3) in total, 14 genes were involved in wax biosynthesis under 28 days drought treatment, more importantly, a new *Kcs* gene *evm.TU.CTG175.54* plays a critical role in the very long chain fatty acid (VLCFA) elongation from C18 to C26 in *Pse. libanotica*. Our study lays a foundation for the genome diversification of Triticeae species and deciphers cuticular wax formation genes that have contributed to the drought resistance of *Pse. libanotica*.

KEYWORDS

genome assemble, St genome, Triticeae, comparative analyses, cuticular wax, drought resistance

1 | INTRODUCTION

The Triticeae tribe (Poaceae), includes economically important annual crops (e.g., wheat, rye, and barley) and crucial perennial forage grasses (e.g., *Roegneria* , *Agropyron* , *Leymus* , and *Pseudoroegneria*) (Dewey, 1984). The genus *Pseudoroegneria* consists of six diploid (2n = 2x = 14, StSt) and nine autotetraploid species (2n = 4x = 28, StStStSt, Yen & Yang, 2011). The diploid *Pseudoroegneria* species, with St genome, involved in more than 60% of perennial polyploid speciation of Triticeae as maternal donor (Yen & Yang, 2011; Chen et al., 2020). At present, the St-containing species were extensively utilized for wheat breeding material improvement such as wheat stripe rust resistance wheat-*Thinopyrum elongatum* (StStEEE) translocation line and Fusarium head blight (FHB) resistance wheat-*Roegneria ciliaris* (StY) disomic addition line (Kong et al., 2018; Yang et al., 2021). Besides, high grass production, excellent disease and saline-alkali resistance characters make the St-containing species to be the ideal germplasm for selecting forage varieties. Thus, the St genome not only account for extensively perennial speciation, but also prominent for forage and crop breeding.

More and more efforts involved the origin of the St genome and genetic relationships among Triticeae species. Molecular data estimated the genus *Pseudoroegneria* derived is more ancient than the *Triticum* / *Aegilops* group (8.0–8.3 Mya) (Zhang et al., 2022). Genomic relationships suggested that the St genome was closely related to the J (=E) genome of (*Thinopyrum elongatum*), and diverged from H (*Hordeum bogdanii*), D sub-genome, then followed by B and A sub-genome in common wheat (Liu & Wang, 1993; Liu et al., 2007). In recent years, Wang et al (2020) mapped 14 linkage groups (LGs) and identified seven homologous groups of the St genome, and revealed the genome shared-homology between the St and ABD, and H genome of 35% and 24%, respectively (Wang et al., 2020). Comparative cytogenetic karyotype result revealed highly conserved collinearity between St genome and common wheat genome, except a well-known 4A chromosome translocation (Wu et al., 2022). These studies provide partly St genome information, nevertheless, the whole genome homoeology and genetic relationship between the St genome and published Triticeae have not been studied.

Pseudoroegneria plants are predominately cool-season and drought tolerant grasses, which distributed in arid or semi-arid areas. The *Pse. libanotica* distributes in the lithoid slopes of Lebanon, Iraq, and the north of Iran (Yen & Yang, 2011). Molecular studies indicate that *Pse. libanotica* is placed at the base of the

phylogenetic tree, implying probably a more ancient than species from Eastern Europe (*Pse. strigosa*), East Asia (*Pse. stipifolia*), and North America (*Pse. spicata*) (Yan and Sun, 2011; Gamache et al., 2015). Morphologically, *Pse. libanotica* is covered by thicken cuticular wax on aerial part, which differs from other diploid *Pseudoroegneria* species. It is largely reported that the primary function of cuticular waxes is to preventing non-stomata water loss (Larson & Kiemnec, 2003; Fraser et al., 2009). Thus, characterizing the St genome of *Pse. libanotica* might reveal the regulation of cuticular wax biosynthesis in response to water deficit.

The regulation of cuticle deposition in response to drought stress was found in model species, crops and important economic species such as*Arabidopsis thaliana* , rice, wheat (*Triticum aestivum*), maize, soybean (*Glycine max*), sesame (*Sesamum indicum*), sorghum, oats (*Avena sativa*), *Medicago sativa* , cotton (*Gossypium hirsutum*), tree tobacco (*Nicotiana glauca*), and pine (*Pinus palustris*) (Li et al., 2019; Lewandowska et al., 2020; Lee & Suh, 2022). Multiple gene involved the cuticular wax biosynthesis have been investigated. The plastid Acetyl-CoA carboxylase are responsible for *de novo* fatty acid biosynthesis (up to C16:0, C16:1 and C18:1) (Roesler et al., 1997; Roudier et al., 2010). Further, the long-chain fatty acids (C16–C18) exported to the cytosol after hydrolysis of ACPs by acyl-ACP thioesterases (Bonaventure et al., 2003). Those 3-ketoacyl-CoA synthase (KCS) isoforms involved in fatty acid elongation (Joubès et al., 2008). Some cytochrome P450s have been reported to be involved in plant cuticle wax synthesis as well (Greer et al, 2007; Zhang et al., 2020). Eceriferum (CER) genes effect different steps of the wax biosynthesis pathway including primary alcohol forming, aldehydes, and alkanes synthesis (Bernard et al., 2012). Researchers investigated distinctive cuticular waxes among different plant species, different tissues and organs, even in different growth and developmental stages (Jetter et al., 2006). Novel genes involved in cuticular wax biosynthesis may be practically used as valuable genetic resource to improve crop drought tolerance in plant breeding (Xue et al., 2017). However, the drought induced wax biosynthesis have not been investigated in crucial wild germplasm in Triticeae.

Given the importance of the St genome in Triticeae and the excellent drought resistance traits of the *Pseudoroegneria* species, the diploid wild species *Pse. libanotica* with the more ancient St genome were performed whole-genome sequencing in this study. This study, we will assemble chromosome scale reference St genome by sequencing; identify the genetic relationship especially among Triticeae species using whole genome comparative analysis; elucidate mechanism of fatty acid biosynthesis in *Pse. libanotica* under drought stress. Those result will provide solid information for the evolution of Triticeae species, and facility better germplasm utilization.

2 | METHODS

2.1 | Plant materials

The diploid *Pse. libanotica* accession PI 228392 (2n=14) was used for genome sequencing in this study. The original seeds were collected from the northeast side of Kuhe Savalan, Azerbaijan, Iran (altitude: 47.85N, 38.30E). It was kindly provided by the National Plant Germplasm System (NPGS, United States). The voucher specimens were kept in the herbarium of Triticeae Research Institute, Sichuan Agricultural University, China.

2.2 | DNA extraction and library preparation

High-molecular-weight genomic DNA was extracted from fresh leaves of a single plant using a DNAsecure Plant Kit (TIANGEN, Beijing, China). For Nanopore sequencing, a ligation of the sequencing adapters library was prepared following the manufacturer's protocol (PromethION, CA). For Illumina (San Diego, CA) short-read sequencing, libraries were size-selected for PE150 sequencing. Sequencing libraries with insert sizes of 350 bp were constructed and sequenced using an Illumina HiSeq X Ten platform at the Novogene Bioinformatics Institute, Beijing.

2.3 | Genome assembly

We constructed a de novo assembly of the St genome of *Pse. libanotica* by combining sequences from three

different technologies: Illumina PE150 short-read sequencing, Nanopore long-read sequencing, and Hi-C conformational alignment.

The clean Nanopore reads after filtering and decontamination were assembled with wtdbg2. The iterative polishing was conducted using Pilon (v1.22) in which clean Illumina reads were aligned with the pre-assembled contigs and BWA with the default parameters (Li & Durbin, 2010; Walker et al., 2014). Further, we combined the final pre-assembled contig sequences from Nanopore sequencing and clean paired-read data from Illumina sequencing into scaffolds using SSPACE (v3.0) tool (Bolger et al., 2014). Genome assembly completeness was assessed using the plantae database of 1440 single-copy orthologues using BUSCO (v3) with a BLAST threshold E-value of 1×10^{-5} (Simão et al., 2015).

The Hi-C libraries were prepared as described previously (Lieberman-Aiden et al., 2009). Hi-C library sequence used a modified SNAP read mapper to align the draft input assembly (Zaharia et al., 2011) (http://snap.cs.berkeley.edu). HiRise was used to analyze the segregation of Hi-C read pairs mapped within draft scaffolds, and a likelihood model of the genomic distance between the read pairs was generated. The model was used to identify and break putative mis-joins, score prospective joins, and select joins above a threshold.

2.4 | Annotation of repetitive sequences

Both homology-based and de novo- based approaches were used to search for TEs. Tandem Repeat was extracted using TRF (http://tandem.bu.edu/trf/trf.html) by *ab initio* prediction. The homolog prediction used Repbase (http://www.girinst.org /repbase) database employing RepeatMasker (*http://www.repeatmasker.org/*) software and its in-house scripts (RepeatProteinMask) with default parameters to extract repeat regions. For the de novo-based approach, we used LTR_FINDER (*http://tlife.fudan.edu.cn/ltr_finder/*), RepeatScout (*http://www.repeatmasker.org/*), and RepeatModeler (http://www.repeatmasker.org/RepeatModeler.html) to build the de novo repeat library. All the repeats identified by different methods were combined into the final repeat annotation after removing the redundant repeats.

2.5 | Genome annotation

To predict protein-coding genes, three approaches were used: de novo gene prediction, homolog prediction, and RNA-sequencing annotation. For de novo prediction, Augustus (v3.2.3), Geneid (v1.4), Genescan (v1.0), GlimmerHMM (v3.04), and SNAP (*http://homepage.mac.com/iankorf/*) were applied to predict genes. For homolog prediction, the protein sequences of twelve published plant genomes (*A. tauschii* ,*B. distachyon* , *T. aestivum* , *T. durum* , *T. dicoccoides* , *T. urartu* , *H. vulgare* , *O. sativa* ,*S. cereale* , *Sorghum bicolor* , *Z. mays,* and*Arabidopsis thaliana*) were aligned to the genome using TblastN (v2.2.26; E-value [?]1e-5), and then used Gene-Wise (v2.4.1) (Birney et al., 2004) to predict gene structures. To optimize the genome annotation, the RNA-seq reads were aligned to the genome using TopHat (v2.0.11) (Trapnell et al., 2009), and the alignments were used as input for Cufflinks (v.2.2.1) (Trapnell et al., 2012). The non-redundant reference gene set was generated by merging genes predicted by three methods with EvidenceModeler (v1.1.1) using PASA (Program to Assemble Spliced Alignment) terminal exon support and including masked transposable elements as input into gene prediction (Haas et al., 2008).

Genes functions were assigned according to the best match by aligning the protein sequences to the Swiss-Prot (with a threshold of E-value [?]$1e^{-5}$) (Bairoch and Apweiler, 2000). The motifs and domains were annotated using InterProScan70 (v5.31) by searching against publicly available databases, including ProDom, PRINTS, Pfam, SMRT, PANTHER, and PROSITE (Mulder & Apweiler, 2008; Finn et al., 2014, 2015, 2017). The Gene Ontology (GO) IDs for each gene were assigned according to the corresponding InterPro entry.

2.6 | Constructing gene families

To construct the dataset for gene-family clustering, the protein sequences from the genomes of *Pse. libanotica* and 13 other plants (*A. tauschii* , *B. distachyon* , *T. aestivum* , *T. durum* , *T. dicoccoides* , *T. urartu* , *H. vulgare* ,*O. sativa* , *S. cereale* , *Sorghum bicolor* , *Z. mays* , *Dactylis glomerata* and *Arabidopsis thaliana*)

were used. In the included species, only the longest transcript in the coding region was retained for further analysis when multiple transcripts were present in a gene. Additionally, genes encoding proteins with fewer than 50 amino acids were filtered. The protein sequences of all species were filtered by BLASTP with an E-value of $1e^{-5}$ Protein sequences from all 14 species were clustered into paralogous and orthologous groups using OrthoMCL (http://orthomcl.org/orthomcl/) with an inflation parameter equal to 1.5.

2.7 | Phylogenetic tree reconstruction

Protein sequences of all single-copy gene families were aligned using MUSCLE (Edgar, 2004), and the alignments of each gene family were concatenated into a super-alignment matrix. These data matrices were used for maximum likelihood phylogenetic analyses by RAxML (*http://sco.h-its.org/exelixis/web/software/raxml/index.html*) with a bootstrap value of 100. *A. thaliana* was designated as outgroups. The Venn diagram was constructed to display the number of gene families that were shared among four Poaceae species (*A. tauschii* , *T. aestivum* , *H. vulgare* , and *S. cereale*) clustered into one group of the phylogenetic tree.

2.8 | Species divergence time estimation

Single-copy gene families among *Pse. libanotica* , *A. tauschii* , *B. distachyon* , *T. aestivum* , *T. durum* ,*T. dicoccoides, T. urartu* , *H. vulgare* , *O. sativa* , *S. cereale* , *Sorghum bicolor* , *Z. mays* ,*Dactylis glomerata* and *Arabidopsis thaliana* were selected using the MCMCTree program (*http://abacus.gene.ucl.ac.uk/software/*paml.html) in Phylogenetic Analysis with Maximum Likelihood (PAML) for an estimate the divergence time of the nodes on the phylogenetic tree. The MCMCTree parameters were as follows: a burn-in of 10,000 steps, sample number of 100,000, and sample frequency of 2. The calibration times of divergence were obtained from the TimeTree database (http://*www.timetree.org/*): 2.20–3.80 Mya for *T. durum* and *T. aestivum* , 3.60–4.40 Mya for *T. dicoccoides* and *T. urartu* , 10.0–11.40 Mya for*Pse. libanotica* and *A. tauschii* , 9.60–11.90 Mya for*S* . c*ereale* and *H. vulgare* , 31.4–47.2 Mya for*Dactylis glomerata* and *B. distachyon* , 9. 0–12.0 Mya for*S. bicolor* and *Z. mays* , 47.2–189 Mya for *O. sativa*and *A. thaliana* .

2.9 | Gene family expansion and contraction

The expansion and contraction of gene families were determined by comparing the cluster size differences between the ancestor and each species using the CAFE program (*http://sourceforge.net/projects/cafehahnlab/*). A random birth-and-death model was used to evaluate changes in gene families of the phylogenetic tree. A probabilistic graphical model (PGM) was used to calculate the transfer probability of each gene family from parent to child nodes in the phylogeny. The conditional likelihood was used as the test statistics to calculate the corresponding$P$ -value of each lineage, and a P -value of or below 0.05 was considered significant.

2.10 | Whole-genome duplication

The homologous search in the *Pse. libanotica* genome was performed using BLASTP (E-value $< 1e^{-5}$), and then MCScanX (*http://chibba.pgml.uga.edu/mcscan2/*) was used to identify syntenic blocks in the genome according to the gene location and blast results. For each gene pair in syntenic blocks, Ks values were calculated, and values of all gene pairs were plotted to identify putative whole-genome duplication events in *Pse. libanotica* .

2.11 | Drought experiments

The seeds of *Pse. libanotica* PI 228392 were germinated in a petri dish lined with a double layer of filter paper. Afterwards, young seedlings were transplanted to potted monocultures and cultivated in a greenhouse at a temperature of 20 with a light cycle of 16h/8h until sturdy stage and had a certain number of tillers.

For drought treatments, five month-old potted *Pse. libanotica* PI 228392 plants of uniform growth were selected to start the experiment. The experimental groups were not watered for 28 days (every seven days was taken as a treatment 7d, 14d, 21d, 28d, and each treatment has three biology repeats), and the control group was watered every 7 days for a total of 28 days. All leaves were taken at 9-10 am and stored at -80.

2.12 | Library preparation, Transcriptome sequencing and analysis

Accurate detection of RNA integrity and total volume with Agilent 2100 bioanalyzer. NEB general library building using NEBNext(r) Ultra RNA Library Prep Kit for Illumina(r) kit and strand-specific library building using NEBNext(r) Ultra Directional RNA Library Prep Kit for Illumina(r) kit. After passing the library test, the different libraries are pooled according to the effective concentration and the target downstream data volume required for Illumina sequencing. The basic principle of sequencing is sequencing while synthesizing. The library construction and Illumina sequencing were conducted at Novogene limited liability company (Beijing, China).

Clean data were obtained after quality clipping of the raw data and Q20, Q30, and GC content. All the downstream analyses were based on clean data with high quality. We selected the genome of *Pse. libanotica*as the reference genome. Hisat2 (v2.0.5) as the mapping tool for that Hisat2 can generate a database of splice junctions based on the gene model annotation file and thus produce a better mapping result than other non-splice mapping tools. The mapped reads of each sample were assembled by StringTie (v1.3.3b) (Pertea et al., 2015) in a reference-based approach.

FeatureCounts v1.5.0 was used to count the reads numbers mapped to each gene. And the FPKM values were then mapped back to read counts according to known gene lengths. Differential expression analysis of two conditions/groups was performed using the DESeq2 R package (1.20.0). The resulting P-values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted P-value <0.05 found by DESeq2 were assigned as differentially expressed.

Gene Ontology (GO) enrichment analysis of differentially expressed genes was implemented by the clusterProfiler R package, in which gene length bias was corrected. GO terms with corrected P-values less than 0.05 were considered significantly enriched by differential expressed genes. We used the clusterProfiler R package to test the statistical enrichment of differential expression genes in KEGG pathways.

3 | RESULTS

3.1 | Genome assembly, quality evaluation, and annotation

Pse. libanotica had an estimated genome size of 3,273.28 Mb (1C) and 3,048.18 Mb based on flow cytometry and k-mer statistics, respectively (Figure S1). By integrating ~191 Gb (64x) Illumina short paired-end reads, ~440 Gb (sequencing depth 147x) Nanopore sequencing data, and ~330 Gb (110x) high-throughput chromosome conformation capture (Hi-C) data, we generated a chromosome-level assembly of *Pse. libanotica* . The assembly sequence comprised 2.99 Gb of genome data, with a contig N50 of 920.96 Kb and a super-scaffold N50 of 380.09 Mb, accounting for 96.45% of the estimated genome size with 1.34% heterozygous (Table 1; Table S1, S2; Figure S2, S3). Of the 2.99 Gb scaffold sequences, 2.75 Gb (91.97%) was anchored to seven super-scaffolds (chromosomes) using the Hi-C platform (Figure 1; Table S3).

The integrity and base accuracy of the assembled *Pse. libanotica*genome was verified by CEGMA (Parra et al., 2007) and BUSCO (Simao et al., 2015). CEGMA showed that the assembled genome completely covered 228 (91.94%) of the 248 core genes, and partially covered 11 core genes. Less than 4% of the core genes were not detected. BUSCO displayed that 95.2% of the 1440 single-copy genes were homologous sequences in Triticeae species (Table S4). The draft assembly was further evaluated by mapping short high-quality reads into the assembled genome. The mapping rate was 98.95%, with 58.95% of the average sequencing depth (Table S5). In *Pse. libanotica* , 151,872 expressed sequence tag (EST) sequences were mapped to the genome with >95% identity, in which 132,240 (87.10%) were aligned to the reference genome with >87% coverage (Table S6). Collectively, these data showed the high coverage of the assembled St genome.

A total of 46,369 protein-coding genes were identified, of which 91.4% had functional annotations (Figure 1B; Table S7, S8). We also identified 1,483 transfer RNAs, 18,438 miRNAs, 1,427 small nuclear RNAs, and 473 ribosomal RNAs (Supplemental Table 9). Repeat sequences comprised 71.62% of the assembled genome, with transposable elements (TEs) being the major component (Figure 1B; Table 2; Table S10). The long terminal repeats (LTRs) were the most abundant repeat type, and the other retrotransposons,

short interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs) had the lowest proportion in the final assembly. In the St genome, 0.31% of repeat sequences could not be annotated (Table 2). Moreover, we predicted the centromere position of *Pse. libanotica* , in which 2St, 3St, 4St, 6St, and 7St were metacentric chromosomes, the 1St and 5St were submetacentric chromosomes (Figure 1B; Table S11).

3.2 | Whole-genome duplication, phylogenetic evolution, and genome synteny

To clarify whole-genome duplication in *Pse. libanotica* , synonymous substitutions (Ks) were calculated in *Pse. libanotica* ,*D. glomerat* , and *B. distachyon* . The peak Ks at 0.20 which occurred after the divergence peak at 0.05 between *Pse. libanotica* and *D. glomerat* and 0.05 between *Pse. libanotica* and *B. distachyon* (Figure 2B), indicating a whole-genome duplication event occurred in the common ancestor of Poaceae. To explore the phylogeny and divergence of the St genome, a unique set of gene families was identified in 14 Poaceae plants based on available genomic resources (Figure 2A). *Sorghum bicolor* and*Zea mays* were C4 plants and diverged firstly in the phylogenetic tree, followed by *Oryza sativa* , *Brachypodium distachyon,*and *Dactylis glomerat* . Triticeae species were clustered into one monophyletic group (*A. tauschii* , *H. vulgare* , *S. cereale* , *T. aestivum* , *T. durum* , *T. dicoccoides* ,*T. urartu*). The *Pse. libanotica* diverged ˜10.7 million years ago (Mya) after *H. vulgare*(˜11.9 Mya), the *S. cereale* and the genus*Triticum* separated ˜9.6 Mya.

Orthologous gene in *Pse. libanotica* , *T. aestivum* , *T. urartu* , *H. vulgare* , and *D. glomerat* were analyzed to illustrate chromosome collinearity and derivation (Figure S4).*Pse. libanotica* exhibited conserved chromosome homoeologous collinearity with A genome, B-subgenome, D-subgenome and H genome, and varied with *D. glomerat.* Interestingly, the 4St in *Pse. libanotica* and the 4A chromosome of *T. urartu* displayed highly consistency, nevertheless, a pericentric inversion was observed in the 4A chromosome, and a 4AL (long arm)-5AL-7BS (short arm) translocation in the common wheat. However, the St genome and the *D. glomerat*displayed chaotic chromosome collinearity. The 2St and 3St chromosomes mainly matched to the 1 and 5 chromosomes of *D. glomerat* , the 5St had syntenic to the 3 and 4 chromosomes of *D. glomerat* , and the 7 St was syntenic to the 3 and 7 chromosome groups of *D. glomerat*(Figure S4). Thus, gene order and location varied among Poaceae species after divergence, conserved among diploid Triticeae species, and partly rearranged during polyploidization of Triticeae.

3.3 | Unique gene families, contract and expanded gene families involved in cuticular wax biosynthesis in *Pse. libanotica*

To characterizing the St genome, unique gene families and contract and expanded genes familied were analyzed in the St genome. In total, 10,073 shared gene families were identified in published Triticeae species and*B. distachyon* , and 882 gene families were specific to *Pse. libanotica* (Figure 2A, 2C; Table S12, 13). Unique gene families in*Pse. libanotica* were mainly involved in ABC transport and cuticular wax biosynthetic pathways (Figure S5). Compared with six Triticeae species, *Pse. libanotica* contained 3,558 expanded and 7,007 contracted genes (Figure 2C). The expanded gens were enriched in five pathways: cutin, suberine, and wax biosynthesis, sesquiterpenoid and triterpenoid biosynthesis, flavone and flavonol biosynthesis, and benzoxazinoid biosynthesis (Table S14 and S15). In summary, unique gene family and expanded and contract gene family in the St genome are clustered in the cutin, suberine and wax biosynthesis pathway.

3.4 | Fatty acid related genes under drought stress in*Pse. libanotica*

To the identify candidate genetic regions associated with fatty acids and their derivatives metabolism pathway, we performed 28 d drought stress treatment without irrigation and conducted transcriptome sequencing on each treatment. Every seven days was taken as a treatment 7 d, 14 d, 21 d, 28 d, and each treatment has three biology repeats. In total, 1,007 differentially expressed genes (DEGs) were shared at four treatments (Figure S6). Differential genes were largely expressed at 14 d drought treatment. Gene Ontology (GO) enrichment analysis showed that 1,190 GO terms were assigned to the 2,300 DEGs that responded to drought treatment (Figure S7). KEGG enrichment analysis showed that DEGs were mainly involved in biosynthesis of unsaturated fatty acids, fatty acid metabolism and elongation process (Figure S8). Comparing the DEGs with expansion and contraction genes, 1,576 DEGs (9.25%) were found in the expansion genes, whereas 61 DEGs (29.35%) were found in the contraction genes.

In total, we identified 14 significantly different expression genes including 20 transcripts which directedly participated in fatty acid biosynthesis under drought stress (Figure 3). In fatty acid biosynthesis pathway of *Pse. libanotica* , the *Acc* and *Fatb* were upregulated during 21 day's water deficit. Except for two transcripts of*Kcs5* and a transcript *evm.TU.CTG175.54* , other genes*Kcs1* , *Kcs6* , *Kcs11* , *Kcs12* , and *Kcs20*were significantly downregulated. Genes catalyzed VCLFAs derivatives in alcohol-forming pathway (*Far* , *Far1* and *Far4*) and alkane-forming pathway (*Cer1* , *Cer3* /*Cytb5* ,*Cyp96a15* /*Mah1*) were upregulated apart from one of the*Far1* transcript-*evm.TU.CTG5611.2* . In summary, only*Kcs5* and *evm.TU.CTG175.54* were responsible for VCLFA elongation in *Pse. libanotica* .

3.5 | Characterization of the *evm.TU.CTG175.54*gene related to VCLFA elongation

To characterize the *evm.TU.CTG175.54* , we aligned the cDNA in the National Center for Biotechnology Information (https://blast.ncbi.nlm.nih.gov/Blast.cgi, NCBI database), the result showed that it was *Kcs5* like and/or*Kcs6* like in Poaceae species such as *T. aestivum* ,*Ae . tauschii* , *H. vulgare* , *Loliumrigidum* , *Z. mays* , *Oryza sativa* , and *Panicum virgatum* etc. It shared 73.05% and 77.22% DNA sequence similarity compared with *Kcs5* and *Kcs6* respectively. Phylogenetic tree based on more than 121 KCS proteins in Poaceae displayed that the*evm.TU.CTG175.54* clustered with KCS5 and KCS6, and formed a mixed group (Figure 4). Thus, we speculated that the*evm.TU.CTG175.54* might be a new *Kcs* gene during fatty acid elongation under drought stress in *Pse. libanotica* and was named as *PlKCS5/6.*

To investigate the growth of yeast cells exposed to stresses, the yeast cells harboring empty pYES2 (as a control) or pYES2-*PlKCS5/6*plasmid was evaluated after treated with 2M sorbitol (Figure 5). The results showed that the growth rate of yeast contained empty pYES2 (as a control) or pYES2-*PlKCS5/6* had no difference under the treatment of 0M sorbitol. However, the growth of PYES2-*PlKCS5/6* plasmid was worse than empty pYES2 under 2M sorbitol treatment. These results demonstrated that expression of *PlKCS5/6* gene in yeast related to the abiotic stress treatments.

4 | DISCUSSION

4.1 | Characterization of the reference genome of*Pse. libanotica*

In this study, we utilized assemble a highly contiguous reference genome sequence for *Pse. libanotica* with 2.99 Gb in size, by using Illumina sequencing, nanopore-based next-generation sequencing and chromosome-scale scaffolding by Hi-C. The St genome is smaller than other diploid Triticeae species reported at present, such as the D genome (4.3 Gb) of *A. tauschii* (Luo et al., 2017), the E genome (4.78 Gb) of *Th. elongata* (Wang et al., 2020), the R genome (7.74 Gb) of *S. cereale* (Li et al., 2021). Repeat sequences comprise 71.62% of the assembled *Pse. libanotica* genome which is lower than other Triticeae (with 80%˜90% repeats in reported Triticeae species genome) even though it displayed no composition difference (Wang et al., 2022).

From a standpoint of germplasm innovation for breeding, clarifying the genome relationships in Triticeae can provide some useful strategies for obtaining the germplasm with the high level genetic diversity and fitness by artificial hybridization (Chen et al., 2020). In this study, we found that the St genome closely related to the A, B, and D genomes of wheat, offering the potential to transfer gene(s) from St genome to wheat. Furthermore, we observed that *T. aestivum* 4A was syntenic to *Pse. libanotica* chromosome 4St, 5St, and 7St, and*T. urartu* 4A was syntenic to *Pse. libanotica* chromosome 4St and 5St., which were related to the most significant overall re-arrangement of chromosome 4A in wheat (Devos et al., 1995; Miftahudin et al., 2004).

The St genome combines with other basic genomes and involve more than 60% perennial Triticeae speciation. Those St-containing species are excellent forages for feeding livestock and contain abundant resistant genes for crop breeding and improvement. The *Pse. libanotica* is an often-cross-pollination plant, the estimated genome size with 1.34% heterozygous. Cytogenetic researches revealed that 5St and 7St displayed high-level genetic heterogeneity among seven *Pse. libanotica*populations (Wu et al., 2022). The high genomic heterozygosity could frequently occur from open pollination plants, which not only spurs the establishment of fitness and abundant genetic diversity, but also contributes to the genetic advantage in terms of the richness of

their polyploid offspring species in Triticeae. Many natural hybrids between various St-containing polyploid species have been reported (Zeng et al., 2012; Lu et al., 2019; Chen et al., 2022; Wu et al., 2023), further enlarging the richness of St-containing descendants.

The St genomic analysis revealed that unique gene families and expanded families displayed enrichment in cutin, suberine, and wax biosynthesis. Morphologically, the aerial parts of *Pse. libanotica* is covered by cuticular wax (Yen & Yang, 2011). It has been reported that plant cuticular wax, which acts as the first barrier against environmental threats, consistently serves a critical role in restricting nonstomatal water loss which were often related to drought resistance (Seo et al., 2011; Zhu & Xiong, 2013; Li et al., 2019; Lewandowska et al., 2020; Lee & Suh, 2022). Back to the origin of genus *Pseudoroegneria* , it diverged about 10.7 Mya during the middle and late Miocene. During that stage, the ice sheet growth, and the Tethyan Seaway went through gradually to completely closure leading to the low temperature and water deficit in Iranian Plateau (Sun et al., 2021). Thus, it is reasonable to speculated that*Pse. libanotica* diverged at during middle and late Miocene in Tran might develop cuticular wax resistance to cold and drought environment.

4.2 | A combination of cuticular wax and related genes might promote the excellent drought resistance of *Pse. libanotica*

Cuticular waxes mostly comprise very long chain fatty acids (VLCFAs) and their derivatives, including alkanes, wax esters, branched alkanes, primary alcohols, alkenes, secondary alcohols, aldehydes ketones, and unsaturated fatty alcohols, as well as cyclic compounds including terpenoids and metabolites such as sterols and flavonoids. In this research, 14 genes involved in cuticular wax biosynthesis were identified in leaves of *Pse. libanotica* during 28 days drought treatment. The *Acc* was reported to catalyze the first step and increased significantly the overall rate of fatty acid biosynthesis up to C16:0, C16:1 and C18:1 (Baud et al., 2003). The long-chain fatty acids are further hydrolyzed by *Fatb,* the FATB thoresterase encode a group of enzymes with more heterogeneous substrate specificity, but generally showing high activities towards saturated acyl-ACPs which is the major determinant of the chain length and level of saturated fatty acid found on most plant tissues (Bonaventure et al., 2003). In*Pse. libanotica* , *Acc* and *Fatb* , determined the*de novo* fatty acid biosynthesis, are significantly upregulated during 21 days water deficit which might indicate its excellent drought resistance.

Based on differences in VLCFA length and degree of unsaturation, 21 KCS genes in *Arabidopsis* showed substrate specificity and organ- or tissue-specific expression patterns during fatty acid elongation (Joubes et al., 2008; Zhao et al., 2019). It is proved that the *Kcs1* is the important gene and has broad substrate specificity for saturated and mono-unsaturated C16–C24 acyl CoAs formation (Trenkamp et al., 2004). The *Kcs20* and *Kcs2 /DAISY* are involved in the two-carbon elongation to C22 VLCFA, and required for cuticular wax and root suberin biosynthesis and expressed in stem epidermal peels (Lee et al., 2009). The *Kcs6* is essential and crucial for the production of epicuticular and pollen coat lipids that are longer than C28 (Fiebig et al., 2000; Haslam and Kunst, 2021). In this study, two *Kcs5*transcripts which is essential for C26–C30 formation were upregulated. However, genes responsible for the C18 to C26 is missing. The*evm.TU.CTG175.54* displayed less DNA and protein sequences similarity with published KCS gene families. Thus, we speculate that*evm.TU.CTG175.54* might be a new gene responsible for C18 to C26 biosynthesis during VLCFA elongation in *Pse. libanotica* .

The fatty acyl-CoA reductase (FAR) gene family is expressed in the epidermis of aerial tissues and in roots, which is responsible for the primary alcohol formatting pathway. In *Pse. libanotica* , the*Far1* was downregulated, while the *Far4* was upregulated. It has been reported that *Far4* catalyzes the reduction of VLC-acyl-CoAs to primary alcohols, presumably via an unreleased aldehyde intermediate which was produced through alkane forming pathway (Kosma et al., 2012). The *Cer1* , *Cer3* , *Cyp96a15 /Mah1*related to alkane forming pathway were significantly upregulated in*Pse. libanotica* . *Cer1* interacts with the wax-associated protein *Cer3* and endoplasmic reticulum–localized *Cytb5s*resulted in VLC alkane synthesis (Bernard et al., 2012). The alkane was further hydrolyzed by *Cyp96a15 /Mah1* , which leading from alkanes to secondary alcohols and corresponding ketones (Greer et al, 2007). To sum up, we assume that the aldehydes, alkanes, secondary alcohol and ketones accumulated in the alkane forming pathway are the main cuticular wax component in

Pse. libanotica leaves, thereby increasing wax deposition and restricting water loss.

AUTHOR CONTRIBUTIONS

D.W. X.Z., C.C. whole-genome sequencing data analysis and discussed the results and prepared the manuscript. X.Z. X.Y. and S.C. drought stress experiments and transcriptome data analysis. L.S., X.F., H.K., Y.W., and Y.C. provide experimental guidance. Y.W., Y.Z., and D.L. provide data analysis guidance. J.L., D.L., W.Z., and Y.Z. discussed the manuscript. D.W., H.Z. review and editing the manuscript. H.Z. funding acquisition. Y.Z. and H.Z. designed and supervised this study.

ACKNOWLEDGEMENTS

We would like to thank National Plant Germplasm System (NPGS, United States) for providing seeds of *Pse. libanotica* . We would like to thank Prof. Jizheng Jia at Chinese Academy of Agricultural Sciences for providing helpful comments on this article.

CONFLICT OF INTEREST STATEMENT

There are no conflicts of interest to declare.

FUNDING INFORMATION

This research was supported by the National Natural Science Foundation of China (Grant Nos. 32270388 and 32200180), the Science and Technology Bureau of Sichuan Province (23NSFSC1995) for their financial supports.

DATA AVAILABILITY STATEMENT

All data are available in the manuscript, the supplementary materials, or at publicly accessible repositories. These data in the public repositories include all raw reads and assembled sequence data for*Pse. libanotica* in NCBI under BioProjectID PRJNA940619, Nanopore sequencing data (SAMN33575612), Illumina sequencing data (SAMN33589869), Hi-C data (SAMN34127759), and RNA-seq data (SAMN33846863, SAMN33593293). The assembly and annotation data of *Pse. libanotica* in the Genome Warehouse in BIG Data Center under accession numbers WGS038485, which are accessible at*https://bigd.big.ac.cn/gwh*.

REFERENCES

1. Bairoch, A., & Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res* , 28(1), 45–48.
2. Baud, S., Guyon, V., Kronenberger, J., Wuilleme, S., Miquel, M., Caboche, M., Lepiniec, L., & Rochat, C. (2003). Multifunctional acetyl-CoA carboxylase 1 is essential for very long chain fatty acid elongation and embryo development in *Arabidopsis* . *Plant J* , 33(1), 75–86
3. Bernard, A., Domergue, F., Pascal, S., Jetter, R., Renne, C., Faure, J.D., Haslam, R.P., Napier, J.A., Lessire, R., & Joubes, J. (2012). Reconstitution of plant alkane biosynthesis in yeast demonstrates that*Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane syn- thesis complex. *Plant Cell* , 24(7), 3106–3118.
4. Birney, E., Clamp, M., & Durbin, R. (2004). GeneWise and genomewise.*Genome Res* , 14(5), 988–995.
5. Bolger, A.M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* , 30(15), 2114–2120.
6. Bonaventure, G., Salas, J.J., Pollard, M.R., & Ohlrogge, J.B. (2003). Disruption of the *FATB* gene in *Arabidopsis* demonstrates an essential role of saturated fatty acids in plant growth.*Plant Cell* , 15(4), 1020–1033.
7. Chen, C., Zheng, Z.L., Wu, D.D., Tan, L., Yang, C.R., Liu, S.Q., Lu, J.L., Cheng, Y.R., Sha, L.N., Wang, Y., Kang, H.Y., Fan, X., Zhou, Y.H., Zhang, C.B., & Zhang, H.Q. (2022). Morphological, cytological, and molecular evidences for natural hybridization between*Roegneria stricta* and *Roegneria turczaninovii*(Triticeae: Poaceae). *Ecol Evol* , 12(1), e8517.

8. Chen, N., Chen, W.J., Yan, H., Wang, Y., Kang, H.Y., Zhang, H.Q., Zhou, Y.H., Sun, G.L., Sha, L.N., & Fan, X. (2020). Evolutionary patterns of plastome uncover diploid-polyploid maternal relationships in Triticeae. *Mol Phylogenet Evol* , 149, 106838.
9. Devos, K., Dubcovsky, J., Dvořák, J., Chinoy, C., & Gale, M. (1995). Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor. *Appl Genet* , 91, 282–288.
10. Dewey, D.R. (1984). The genomic system of classification. A guide to intergeneric hybridization with the perennial Triticeae. Gene Manipulation in Plant Improvement. 209–279.
11. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* , 32(5), 1792–1797.
12. Fiebig, A., Mayfield, J.A., Miley, N., Chau, S., Fischer, R.L., & Preuss, D. (2000). Alterations in *CER6* , a gene identical to*CUT1* , differentially affect long-chain lipid content on the surface of pollen and stems. *Plant Cell* , 12(10), 2001–2008.
13. Finn, R.D., Attwood, T.K., Babbitt, P.C., Bateman, A., Bork, P., Bridge, A.J., Chang, H.Y., Dosztányi, Z., El-Gebali, S., & Fraser, M. (2017). InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res* , 45(D1), 190–199.
14. Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., & Mistry, J. (2014). Pfam: the protein families database. *Nucleic Acids Res* , 42(D1), 222–230.
15. Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F., Bateman, A., and Eddy, S.R. (2015). HMMER web server: 2015 update. *Nucleic Acids Res* , 43(W1), 30–38.
16. Fraser, L.H., Greenall, A., Carlyle, C., Turkington, R., & Friedman, C.R. (2009). Adaptive phenotypic plasticity of *Pseudoroegneria spicata* : response of stomatal density, leaf area and biomass to changes in water supply and increased temperature. *Ann Bot* , 103(5), 769–775.
17. Gamache, J., & Sun, G.L. (2015). Phylogenetic analysis of the genus*Pseudoroegneria* and the Triticeae tribe using the *rbcL*gene. *Biochem Syst Ecol* , 62, 73–81.
18. Greer, S., Wen, M., Bird, D., Wu, X., Samuels, L., Kunst, L., & Jetter, R. (2007). The cytochrome P450 enzyme CYP96A15 is the midchain alkane hydroxylase responsible for formation of secondary alcohols and ketones in stem cuticular wax of *Arabidopsis* . *Plant Physiol* , 145(3), 653–667.
19. Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., & Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biol* , 9, 1–22.
20. Haslam, T.M., & Kunst, L. (2021). *Arabidopsis*ECERIFERUM2-LIKEs are mediators of condensing enzyme function.*Plant Cell Physiol* , 61(12), 2126–2138.
21. Jetter, R., Kunst, L., & Samuels, A.L. (2006). Composition of plant cuticular waxes. In: Riederer M, Muller C, eds. *Biology of the plant cuticle* , Oxford: Blackwell 145–175.
22. Joubès, J.M., Raffaele, S., Bourdenx, B., Garcia, C., Laroche-Traineau, J., Moreau, P., Domergue, F., & Lessire, R. (2008). The VLCFA elongase gene family in *Arabidopsis thaliana* : phylogenetic analysis, 3D modelling and ex- pression profiling.*Plant Mol Biol* , 67, 547–566.
23. Kong, L., Song, X., Xiao, J., Sun, H., Dai, K., Lan, C., Singh, P., Yuan, C., Zhang, S., & Singh, R. (2018). Development and characterization of a complete set of *Triticum aestivum–Roegneria ciliaris* disomic addition lines. *Theor Appl Genet* , 131, 1793–1806.
24. Kosma, D.K., Molina, I., Ohlrogge, J.B., & Pollard, M. (2012). Identification of an *Arabidopsis* fatty alcohol: caffeoyl-Coenzyme A acyltransferase required for the synthesis of alkyl hydroxycinnamates in root waxes. *Plant Physiol* , 160(1), 237–248.
25. Larson, L., & Kiemnec, G. (2003). Seedling Growth and Interference of Diffuse Knapweed (*Centaurea diffusa*) and Bluebunch Wheatgrass (*Pseudoroegneria spicata*). *Weed Technol* , 17(1), 79–83.
26. Lee, S.B., Jung, S.J., Go, Y.S., Kim, H.U., Kim, J.K., Cho, H.J., Park, O.K., & Suh, M.C. (2009). Two *Arabidopsis* 3-ketoacyl CoA synthase genes, *KCS20* and *KCS2/DAISY* , are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress. *Plant J* , 60(3), 462–475.
27. Lee, S.B., & Suh, M.C. (2022). Regulatory mechanisms underlying cuticular wax biosynthesis. *J Exp*

Bot , 73(9), 2799–2816.

28. Lewandowska, M., Keyl, A., & Feussner, I. (2020). Wax biosynthesis in response to danger: its regulation upon abiotic and biotic stress.*New Phytol* , 227(3), 698–713.
29. Li, H., Mo, Y.L., Cui, Q., Yang, X.Z., Guo, Y.L., Wei, C.H., Yang, J., Zhang, Y., Ma, J.X., & Zhang, X. (2019). Transcriptomic and physiological analyses reveal drought adaptation strategies in drought-tolerant and -susceptible watermelon genotypes. *Plant Sci* , 278, 32–43.
30. Li, G., Wang, L., Yang, J., He, H., Jin, H., Li, X., Ren, T., Ren, Z., Li, F., & Han, X. (2021). A high-quality genome assembly highlights rye genomic characteristics and agronomically important genes.*Nat Genet* , 53(4), 574–584.
31. Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics, 26(5), 589–595.
32. Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., & Dorschner, M.O. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome.*Science* , 326(5950), 289–293.
33. Liu, Z., Li, D., & Zhang, X. (2007). Genetic relationships among five basic genomes St, E, A, B and D in Triticeae revealed by genomic southern and in situ hybridization. *J Integr Plant Biol* , 49(7), 1080–1086.
34. Liu, Z.W., & Wang, R.R.C. (1993). Genome analysis of *Elytrigia caespitosa* , *Lophopyrum nodosum* , *Pseudoroegneria geniculata* ssp. *scythica* , and *Thinopyrum intermedium*(Triticeae: Gramineae). *Genome* , 36(1), 102–111.
35. Lu, X.W., Liu, B., Liu, R.J., & Dou, Q.W. (2019). Cytogenetic identification on interspecific hybrids in genus *Elymus* L. of Qinghai Plateau. *Bull Bot Res* , 39(6), 846–852.
36. Luo, M.C., Gu, Y.Q., Puiu, D., Wang, H., Twardziok, S.O., Deal, K.R., Huo, N., Zhu, T., Wang, L., & Wang, Y. (2017). Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii* .*Nature* , 551(7681), 498–502.
37. Miftahudin, K.R., Ma, X.F., Mahmoud, A., Layton, J., Milla, M.R., Chikmawati, T., Ramalingam, J., Feril, O., Pathan, M., Momirovic, G.S., Kim, S., Chema, K., Fang, P., Haule, L., Struxness, H., Birkes, J., Yaghoubian, C., Skinner, R., McAllister, J., Nguyen, V., Qi, L.L., Echalier, B., Gill, B.S., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvorak, J., Dilbirligi, M., Gill, K.S., Peng, J.H., Lapitan, N.L.V., Bermudez-Kandianis, C.E., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Lazo, G.R., Chao, S., Anderson, O.D., Gonzalez-Hernandez, J., Conley, E.J., Anderson, J.A., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O., Nguyen, H.T., & Gustafson, J.P., (2004). Analysis of expressed sequence tag loci on wheat chromosome group 4. *Genetics* , 168(2), 651–663.
38. Mulder, N., & Apweiler, R. (2008). Interpro and interproscan.*Comparative Genomics* , 59–70.
39. Parra, G., Bradnam, K., & Korf, I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes.*Bioinformatics* , 23(9), 1061–1067.
40. Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., & Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* , 33(3), 290–295.
41. Roesler, K., Shintani, D., Savage, L., Boddupalli, S., & Ohlrogge, J. (1997). Targeting of the *Arabidopsis* homomeric acetyl-coenzyme A carboxylase to plastids of rapeseeds. *Plant Physiol* , 113(1), 75–81.
42. Roudier, F., Gissot, L., Beaudoin, F., Haslam, R., Michaelson, L., Marion, J., Molino, D., Lima, A., Bach, L., Morin, H., Tellier, F., Palauqui, J., Bellec, Y., Renne, C., Miquel, M., DaCosta, M., Vignard, J., Rochat, C., Markham, J.E., Moreau, P., Napier, J., & Faure, J.D. (2010). Very-long-chain fatty acids are involved in polar auxin transport and developmental patterning in *Arabidopsis* .*Plant Cell* , 22(2), 364–375.
43. Seo, P.J., Lee, S.B., Suh, M.C., Park, M.-J., Go, Y.S., & Park, C.M. (2011). The MYB96 transcription factor regulates cuticular wax biosynthesis under drought conditions in *Arabidopsis* .*Plant Cell* , 23(3), 1138–1152.

44. Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., & Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* , 31(19), 3210–3212.
45. Sun, J.M., Sheykh, M., Ahmadi, N., Cao, M.M., Zhang, Z.L., Tian, S.C., Sha, J., Jian, Z.M., Windley, B.F., & Talebian, M. (2021). Permanent closure of the Tethyan Seaway in the northwestern Iranian Plateau driven by cyclic sea-level fluctuations in the late Middle Miocene.*Palaeogeogr Palaeocl* , 564, 110172.
46. Trapnell, C., Pachter, L., & Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* , 25(9), 1105–1111.
47. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., & Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* , 7(3), 562–578.
48. Trenkamp, S., Martin, W., & Tietjen, K. (2004). Specific and differential inhibition of very-long-chain fatty acid elongases from*Arabidopsis thaliana* by different herbicides. *Proc Natl Acad* , Sci. 101(32), 11903–11908
49. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., & Young, S.K. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* , 9(11), e112963.
50. Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., Lyu, Z., Chen, L., Xu, S., & Guo, J. (2020). Horizontal gene transfer of *Fhb7*from fungus underlies Fusarium head blight resistance in wheat.*Science* , 368(6493), eaba5435.
51. Wang, R.R.C., Li, X.F., Robbins, M.D., Larson, S.R., Bushman, S., Jones, T.A., & Thomas, A. (2020). DNA sequence-based mapping and comparative genomics of the St genome of *Pseudoroegneria spicata* (Pursh) Á. Löve versus wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.). *Genome* , 63(9), 445–457.
52. Wang, X., Yan, X., Hu, Y., Qin, L., Wang, D., Jia, J., & Jiao, Y. (2022). A recent burst of gene duplications in Triticeae. *Plant Communications* , 3(2), 100268.
53. Wu, D.D., Yang, N.M,, Xiang, Q., Zhu, M.K., Fang, Z.Y., Zheng, W., Lu, J.L., Sha, L.N., Fan, X., Cheng, Y.R., Wang, Y., Kang, H.Y., Zhang, H.Q., & Zhou, Y.H. (2022). *Pseudorogneria libanotica*intraspecific genetic polymorphism revealed by fluorescence in situ hybridization with newly identified tandem repeats and wheat single-copy gene probes. *Int J Mol Sci* . 23(23), 14818.
54. Wu, D.D., Liu, X.Y., Yu, Z.H., Tan, L., Lu, J.L., Cheng, Y.R., Sha, L.N., Fan, X., Kang, H.Y., & Wang, Y., Zhou, Y.H., Zhang, C.B., & Zhang, H.Q. (2023). Recent natural hybridization in *Elymus* and*Campeiostachys* of Triticeae: evidence from morphological, cytological and molecular analyses. *Bo. J Linn Soc* , 201(4), 428–442.
55. Xue, D., Zhang, X., Lu, X., Chen, G., & Chen, Z.H. (2017). Molecular and evolutionary mechanisms of cuticular wax for plant drought tolerance. *Front Plant Sci* , 8, 621.
56. Yan, C., & Sun, G.L. (2011). Nucleotide divergence and genetic relationships of *Pseudoroegneria* species. *Biochem Syst Ecol* , 39(4-6), 309–319.
57. Yang, G.T., Boshoff, W.H.P., Li, H.W., Pretorius, Z.A., Luo, Q.L., Li, B., Li, Z.S., & Zheng, Q. (2021). Chromosomal composition analysis and molecular marker development for the novel *Ug99* -resistant wheat–*Thinopyrum ponticum* translocation line WTT34.*Theor Appl Genet* , 134, 1587–1599.
58. Yen, C., & Yang, J.L. (2011). Triticeae biosystematics, vol. 4, Chinese Agricultural Press, Beijing.
59. Zaharia, M., Bolosky, W.J., Curtis, K., Fox, A., Patterson, D., Shenker, S., Stoica, I., Karp, R.M., & Sittler, T. (2011). Faster and more accurate sequence alignment with SNAP. *ArXiv* , 1111.5572.
60. Zeng, J., Fan, X., Zhang, H.Q., Sha, L.N., Kang, H.Y., Zhang, L., Yang, R.W., Ding, C.B., & Zhou, Y.H. (2012). Molecular and cytological evidences for the natural wheat grass hybrids occurrence and origin in west China. *Genes Genom* , 34(5), 499–507.
61. Zhang, D., Yang, H., Wang, X., Qiu, Y., Tian, L., Qi, X., & Qu, L.Q. (2020). Cytochrome P450 family member *CYP96B5* hydroxylates alkanes to primary alcohols and is involved in rice leaf cuticular wax synthesis. *New Phytol* , 225(5), 2094–2107.

62. Zhang, L., Zhu, X., Zhao, Y., Guo, J., Zhang, T., Huang, W., Huang, J., Hu, Y., Huang, C.H., & Ma, H. (2022). Phylotranscriptomics resolves the phylogeny of pooideae and uncovers factors for their adaptive evolution. *Mol Biol Evol* , 39(2), msac026.
63. Zhao, L., Haslam, T.M., Sonntag, A., Molina, I., & Kunst, L. (2019). Functional overlap of long-chain acyl-CoA synthetases in*Arabidopsis* . *Plant Cell Physiol* , 60(5), 1041–1054.
64. Zhu, X., & Xiong, L. (2013). Putative megaenzyme *DWA1* plays essential roles in drought resistance by regulating stress-induced wax deposition in rice. *PANS* , 110(44), 17790–17795.

SUPPLEMENTAL INFORMATION

Figure S1 Flow cytometry results of *Pse. libanotica*

Figure S2 The workflow of *Pse. libanotica* genome assembling

Figure S3 K-mer frequency distributions in *Pse. libanotica* Axis means sequence depth (X), and y axis means frequency of K-mer

Figure S4 Synteny analysis of seven chromosomes from *Pse. libanotica* with *T. aestivum*, *T. urartu*, *H. vulgare*, and *D. glomerat*

Figure S5 The physiological changes of *Pse. libanotica*under drought conditions at 7d, 14d, 21d and 28d (A) Soil water content (B) POD activity (C) CAT activity (D) Pro content (E) MDA content (F) ABA content. The X-axis indicates the treatment time; the Y-axis indicates the physiological activities. The data were obtained from three independent experiments. Different letters represent the statistical significance analyzed by Duncan's multiple range test ($P < 0.05$).

Figure S6 Number of differentially expressed genes (DEGs) under drought conditions at 7d, 14d, 21d and 28d Venn diagrams showing the number of co-expressed DEGs in *Pse. libanotica* under drought stress.

Figure S7 GO classification of differentially expressed genes (DEGs) of *Pse. libanotica* under drought stress The ordinate in the figure is GO Term, and the abscissa is the number of genes enriched by GO Term.

Figure S8 KEGG pathway enrichment scatter diagram of DEGs The19 most strongly represented pathways are displayed in the diagram. The degree of KEGG pathway enrichment is represented by the GeneRatio, the padj, and the number of genes enriched in a KEGG pathway.

Table S1. Estimation of genome size

Table S2. Sequencing libraries and statistics of the data used for the genome assembly

Table S3. Characteristics of *Pse. libanotica* assembly containing 7 chromosome

Table S4. Evaluation of benchmarking universal single-copy orthologs (BUSCO) and gene space coverage using core eukaryotic gene mapping approach (CEGMA) in *Pse. libanotica* genome

Table S5. Statistics of paired-end reads mapping

Table S6. Assessment of *Pse. libanotica* genome using full length EST sequences

Table S7. Prediction of protein-coding genes in *Pse. libanotica*

Table S8. Summary for annotation of predicted protein-coding genes in the *Pse. libanotica* genome assembly

Table S9. Non-coding RNAs in the assembly of *Pse. libanotica*

Table S10. Statistics of repeat sequence in *Pse. libanotica* genome via different methods

Table S11. Centromere position of *Pse. libanotica*

Table S12. GO analysis for the unique gene families in

Pse. libanotica

Table S13. KEGG pathway of unique families in *Pse. libanotica*

Table S14. GO analysis for the expanded gene families in *Pse. libanotica*

Table S15. KEGG pathway of expanded families in *Pse. libanotica*

Figures and Tables

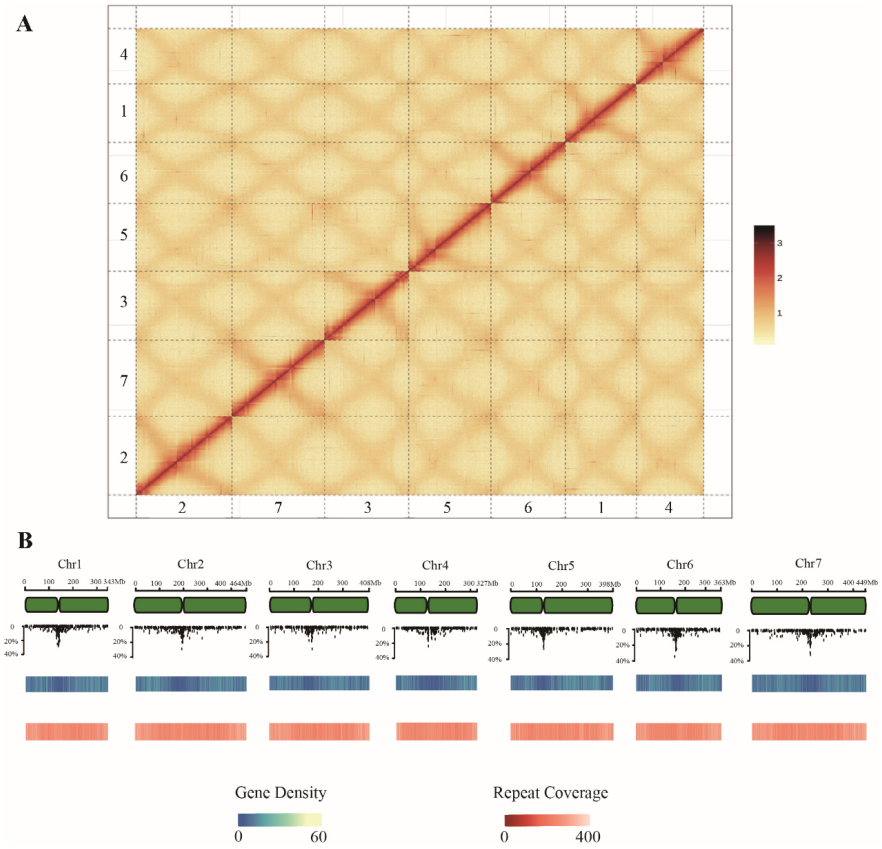


Figure 1. Genome assemble and structure of *Pse. libanotica* (A) Hi-C interaction matrix for genome assembly, the x and y axes indicate the mapping positions of the first and second read in the read pair respectively, grouped into bins. The color of each square gives the number of read pairs within that bin. Scaffolds less than 1 Mb are excluded. (B) From top to bottom are the predicted centromere position, gene density, and repeat coverage.

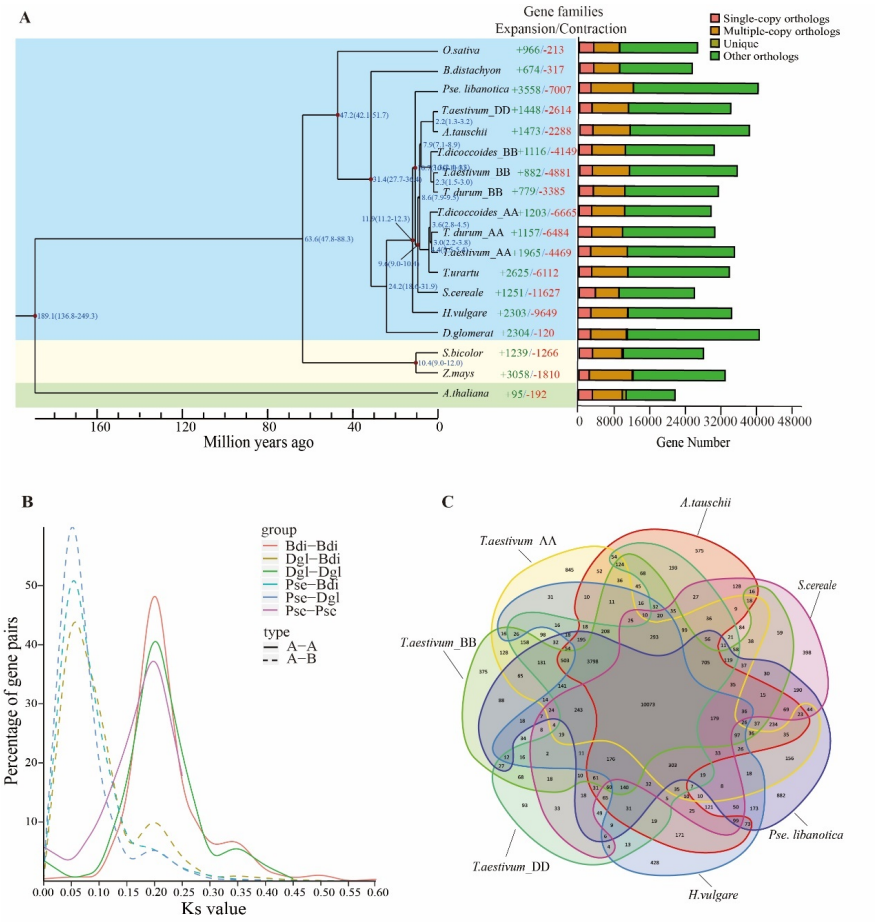


Figure 2. Gene family and genome evolution of *Pse. libanotica* (A) The left panel includes the estimation of divergence time of *Pse. libanotica* and *O. sativa* , *B. distachyon* , *T. aestivum* , *A. tauschii* , *T. durum* ,*T. dicoccoides* , *T. urartu* , *S. cereale* , *H. vulgare* , *Z. mays* and *A. thaliana* . The right panel displays the distribution of single-copy, multiple-copy, unique and other orthologues. (B) Distribution of the Ks values of the best reciprocal BLASTP hits in the genomes of *Pse. libanotica* (Pse),*D. glomerata* (Dgl) and *B. distachyon* (Bdi). (C) The number of gene families shared among seven Triticeae species genomes shown in Venn diagrams.

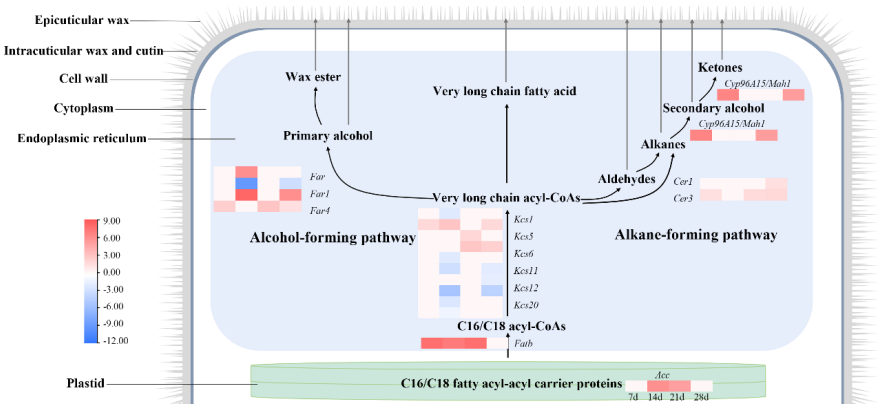


Figure 3. Cuticular wax biosynthetic pathway related genes in *Pse. libanotica* under drought stress Cuticular wax are synthesized and secreted in the endoplasmic reticulum (ER). First, long-chain fatty acids (C16–C18) are synthesized in the plastids. Second, the saturated C16 and C18 fatty acyl-CoA precursors are elongated to very long chain fatty acyl-coenzyme As (VLCFA-CoAs) in ER. Once the VLC-acyl-CoAs are synthesized, they can be further released as free VLCFAs that can either be directly exported as cuticular waxes, or undergo further modifications in the via either the alkane-forming pathway or the alcohol-forming pathway.

Figure 4. Maximum likelihood tree derived from 132 KCS protein sequences data from Poaceae KCS proteins with red color were produced in this research. Other KCS proteins were downloaded from

NCBI and named as 'KCS protein+species+protein ID'.

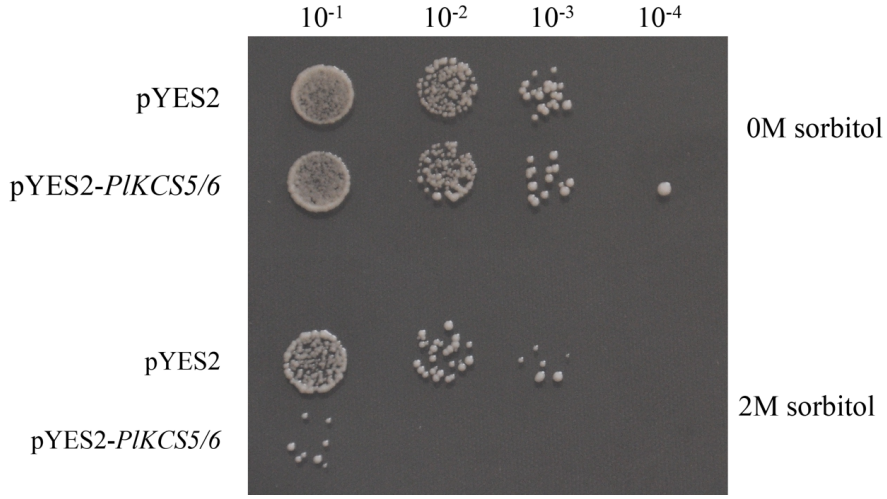


Figure 5. The growth activity of pYES2 and pYES2-PlKCS5/6 under different treatments. 5 μL of serial dilutions of pYES2 and pYES2-PlKCS5/6 stressed with 0M sorbitol and 2M sorbitol were cultivated on SC-Ura medium for 2 days.

Tables

Table 1 Overview of genome assembly and gene annotation for *Pse. libanotica*

Assembly characteristics	**Values**
Estimated genome size	3.10 Gb
Assembled genome size	2.99 Gb
Total length of contigs	2.99 Gb
N50 length of contigs	0.96 Mb
Total number of contigs	7,391
N50 length of super scaffolds	380.09 Mb
Number of annotated high-confidence genes	46,369
Percentage of repeat sequences	71.62 %
Complete BUSCOs	95.20 %
Fragmented BUSCOs	1.20 %
Missed BUSCOs	3.60 %

Table 2 Statistics of transposable elements in *Pse. libanotica*genome sequences

	Denovo+Repbase Length(bp)	**% in Genome**	**TE proteins Length(bp)**	**% in Genome**	**Combi**
DNA	80,757,639	2.70	49,637,759	1.66	128,642,
LINE	2,196,166	0.07	33,738,564	1.13	35,319,9
SINE	1,570,054	0.05	0	0	1,570,05
LTR	1,952,684,497	65.32	391,751,832	13.1	1,977,51
Unknown	9,318,886	0.31	120	0.000004	9,319,00
Total	2,045,135,021	68.41	475,109,220	15.89	2,105,45

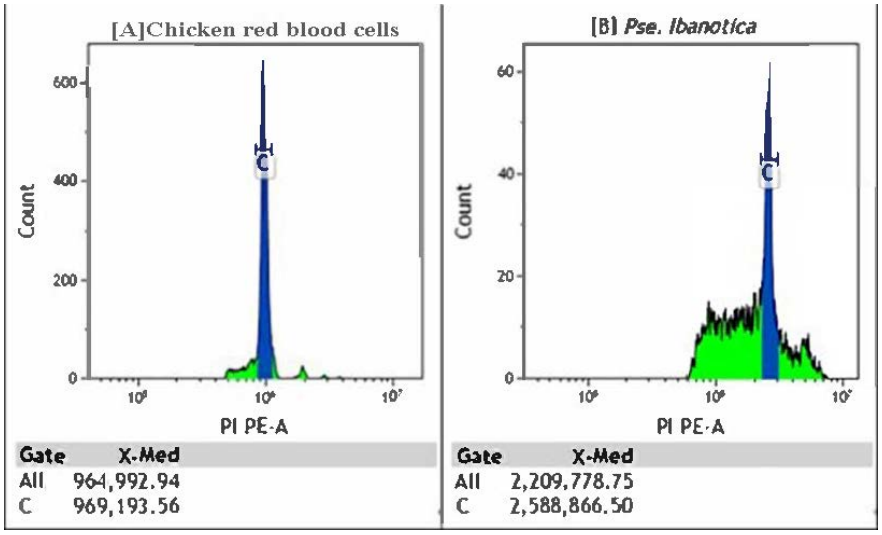


Figure S1

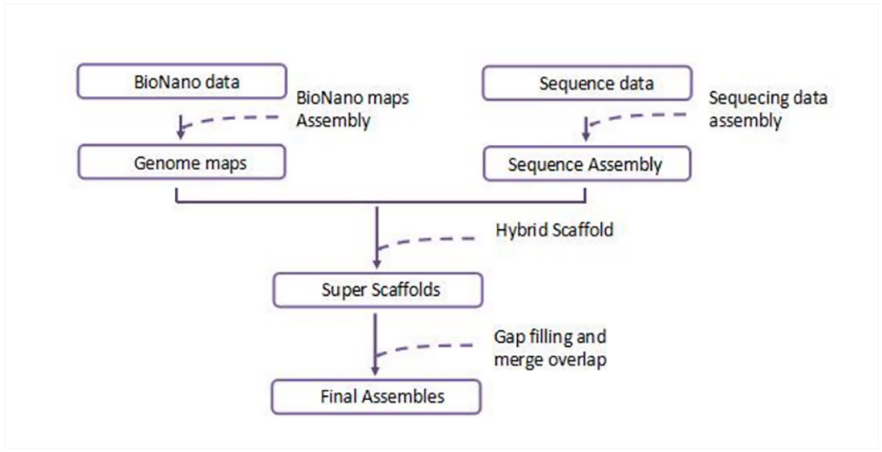


Figure S2

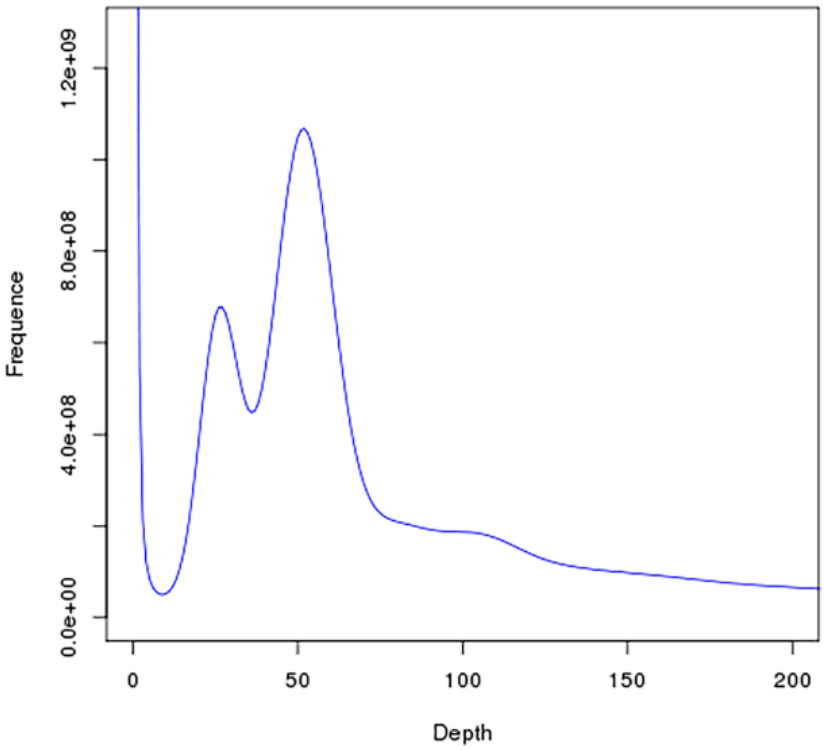


Figure S3

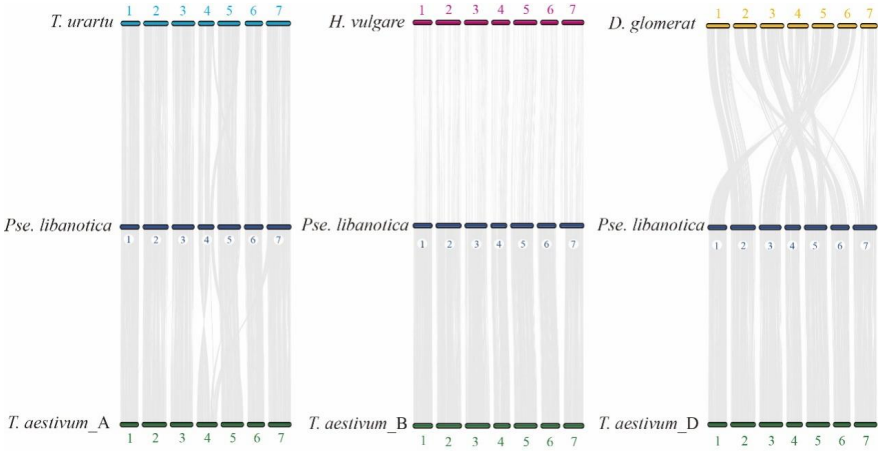


Figure S4

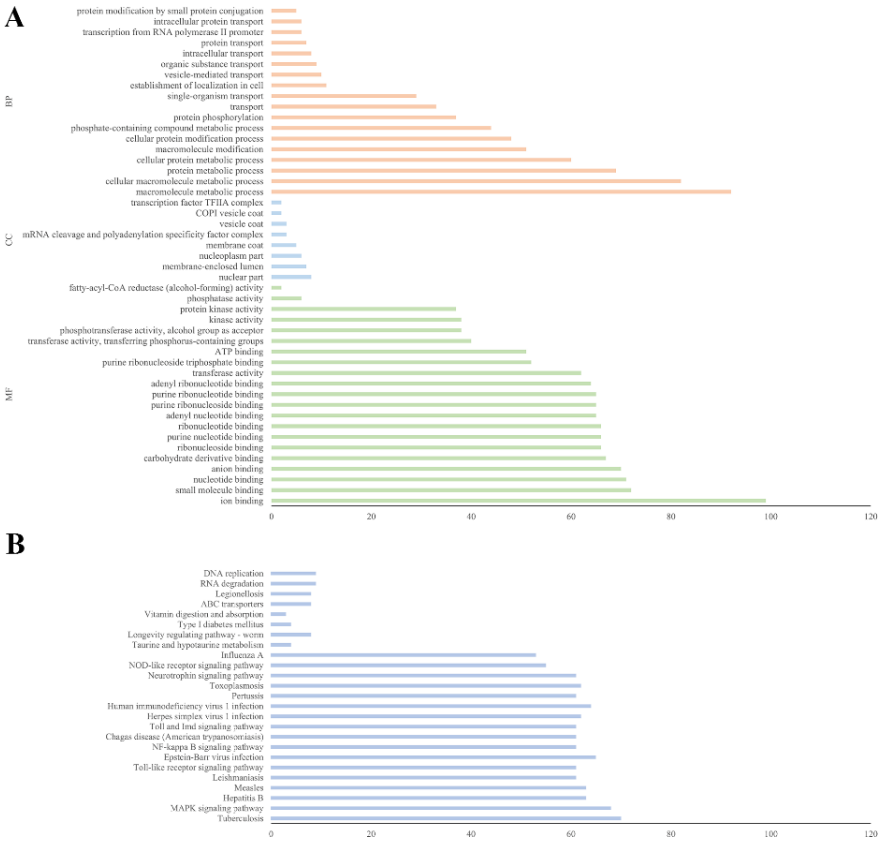


Figure S5

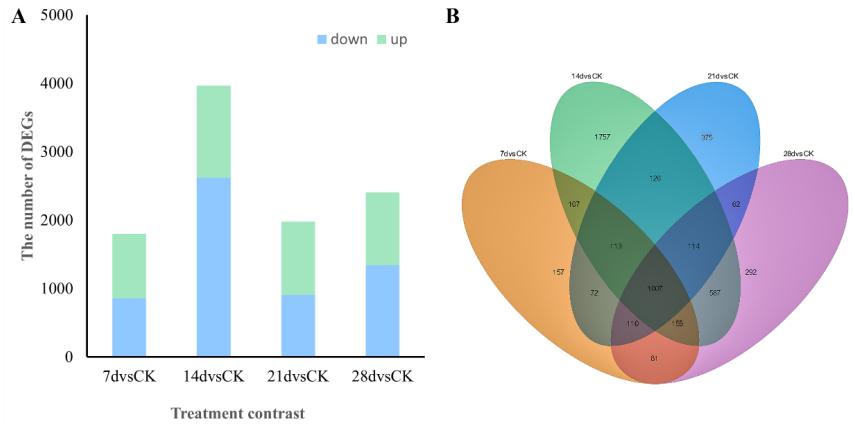
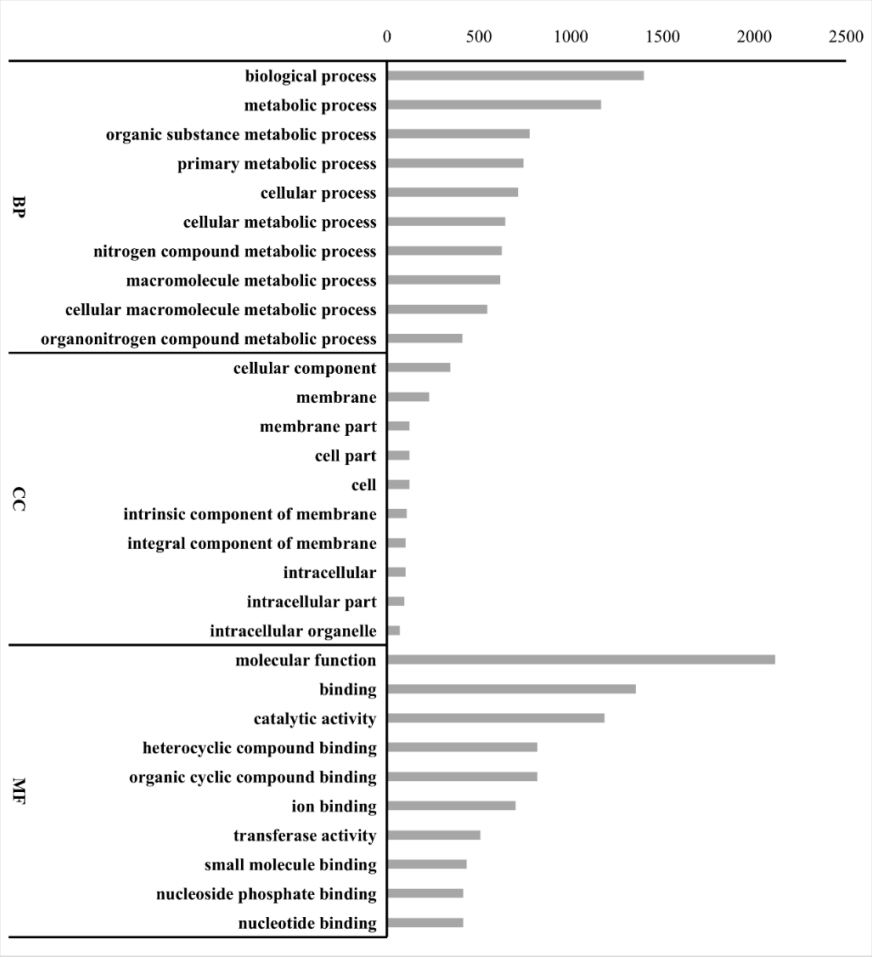
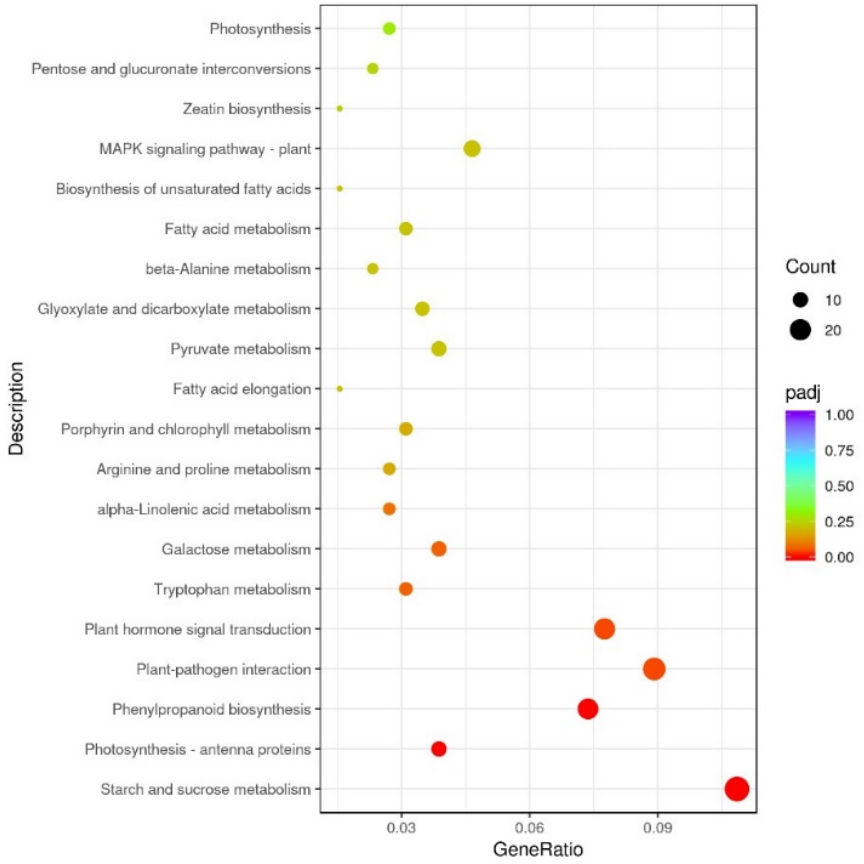


Figure S6



Supplemental Fig. 7



Supplemental Fig. 8