Transcriptomic analysis reveals candidate genes for male sterility in *Prunus sibirica*

Jianhua Chen [1] , **Hao Xu** [1] , **Jian Zhang** [1] , **Shengjun Dong** Corresp., [1] , **Quangang Liu** Corresp., [1] , **Ruoxi Wang** [1]

[1] College of Forestry, Shenyang Agricultural University, Shenyang, Liaoning, China

Corresponding Authors: Shengjun Dong, Quangang Liu
Email address: dsj928@163.com, liuquangang007@126.com

Background. The phenomenon of male sterility widely occurs in *Prunus sibirica* and has seriously affected the production yield. We identified the key stage and cause of male sterility and found differentially expressed genes related to male sterility in *Prunus sibirica*, and we analyzed the expression pattern of these genes. This work aimed to provide valuable reference and theoretical basis for the study of reproductive development and the mechanisms of male sterility in *Prunus sibirica.* **Method**. The microstructures of male sterile flower buds and male fertile flower buds were observed by paraffin section. Transcriptome sequencing was used to screen genes related to male sterility in *Prunus sibirica.* Quantitative real-time PCR analysis was performed to verify the transcriptome data. **Results**. Anther development was divided into the sporogenous cell stage, tetrad stage, microspore stage, and pollen maturity stage. Compared with male fertile flower buds, in the microspore stage, the pollen sac wall tissue in the male sterile flower buds showed no signs of degeneration. In the pollen maturity stage, the tapetum and middle layer were not fully degraded, and the anther development stopped. Therefore, the microspore stage was the key stage for anther abortion , and the pollen maturity stage was the post stage for anther abortion. A total of 4108 differentially expressed genes were screened out by transcriptome analysis. Among them, 1899 were up-regulated, and 2209 were down-regulated in the transcript of male sterile flower buds. We found that “protein kinase activity”, “apoptosis process”, “calcium binding”, “cell death”, “cytochrome c oxidase activity”, “aspartate peptidase activity”, “cysteine peptidase activity” and other biological pathways such as “starch and sucrose metabolism” and “proteasome” were closely related to male sterility in *Prunus sibirica*. A total of 332 key genes were preliminarily screened. **Conclusion**. The occurrence of male sterility in *Prunus sibirica* involved many biological processes and metabolic pathways. According to the results of microstructure observations, related physiological indexes determination and transcriptome analysis, we inferred that the occurrence of male sterility in *Prunus sibirica* may be caused by abnormal metabolic processes such as the release of cytochrome c in

the male sterile flower buds, the imbalance of the antioxidant system was destroyed, and inability of macromolecular substances such as starch to be converted into soluble small molecules in time, resulting in energy loss. The tapetum cannot be fully degraded, thereby blocking anther development, which eventually led to the formation of male sterility.

Transcriptomic analysis reveals candidate genes for male sterility in *Prunus sibirica*

Jianhua Chen[1], Hao Xu[1], Jian Zhang[1], Shengjun Dong[1], Quangang Liu[1], Ruoxi Wang[1]

[1] College of Forestry, Shenyang Agricultural University, Shenyang, Liaoning, China

Corresponding authors:
Shengjun Dong, Quangang Liu
No.120 Dongling Road, Shenhe District, Shenyang, Liaoning, 110866, China
Email address: dsj928@163.com, liuquangang007@126.com

Abstract

Background. The phenomenon of male sterility widely occurs in *Prunus sibirica* and has seriously affected the production yield. We identified the key stage and cause of male sterility and found differentially expressed genes related to male sterility in *Prunus sibirica*, and we analyzed the expression pattern of these genes. This work aimed to provide valuable reference and theoretical basis for the study of reproductive development and the mechanisms of male sterility in *Prunus sibirica.*

Method. The microstructures of male sterile flower buds and male fertile flower buds were observed by paraffin section. Transcriptome sequencing was used to screen genes related to male sterility in *Prunus sibirica.* Quantitative real-time PCR analysis was performed to verify the transcriptome data.

Results. Anther development was divided into the sporogenous cell stage, tetrad stage, microspore stage, and pollen maturity stage. Compared with male fertile flower buds, in the microspore stage, the pollen sac wall tissue in the male sterile flower buds showed no signs of degeneration. In the pollen maturity stage, the tapetum and middle layer were not fully degraded, and the anther development stopped. Therefore, the microspore stage was the key stage for anther abortion, and the pollen maturity stage was the post stage for anther abortion. A total of 4108 differentially expressed genes were screened out by transcriptome analysis. Among them, 1899 were up-regulated, and 2209 were down-regulated in the transcript of male sterile flower buds. We found that “protein kinase activity”, “apoptosis process”, “calcium binding”, “cell death”, “cytochrome c oxidase activity”, “aspartate peptidase activity”, “cysteine peptidase activity” and other biological pathways such as “starch and sucrose metabolism” and “proteasome” were closely related to male sterility in *Prunus sibirica*. A total of 332 key genes were preliminarily screened.

Conclusion. The occurrence of male sterility in *Prunus sibirica* involved many biological processes and metabolic pathways. According to the results of microstructure observations, related physiological indexes determination and transcriptome analysis, we inferred that the occurrence of male sterility in *Prunus sibirica* may be caused by abnormal metabolic processes such as the release of cytochrome c in the male sterile flower buds, the imbalance of the antioxidant system was destroyed, and inability of macromolecular substances such as starch to be converted into soluble small molecules in time, resulting in energy loss. The tapetum cannot be fully degraded, thereby blocking anther development, which eventually led to the formation of male sterility.

Introduction

Prunus sibirica is a shrub or small tree that belongs to the family Rosaceae. It is mostly found in the wild and semi-wild state, and most of *Prunus sibirica* are self-incompatible. There is a common phenomenon of interspecific hybridization in *Prunus sibirica*, so the germplasm resources are abundant. *Prunus sibirica* is widely cultivated in China, and its economic benefits are generally recognized by those who reside in the areas where this species commercially cultivated (*Wan et al*., *2015*). However, the low and unstable yield, which is caused by many factors, such as frost damage during flowering stage, low fruit setting rate, self-incompatibility, male sterility and pistil abortion, has become the main bottleneck hindering the rapid development of the *Prunus sibirica* industry (*Wang et al., 2014*). Among them, the differentiation of male organs directly affects the yield of *Prunus sibirica*. At present, the key stage and cause of male sterility in *Prunus sibirica* have not been determined.

Plant male sterility is widespread in nature. At present, domestic and foreign scholars have conducted studies on different plants, such as apricot (*Badenes et al*., *2000*), *Prunus mume* (*Yaegaki et al., 2003*), *Prunus salicina* (*Radice et al*., *2008*), poplar (*Liu et al*., *2019*), *Camellia crassicolumna* (*Jiang et al*., *2020*), *Brassica napus* (*Du et al*., *2016*), and tobacco (*Liu et al*., *2020*), from the aspects of cytology, physiology and biochemistry, and molecular biology to explore the mechanism of plant male sterility. The plant male sterility can usually be classified into several types such as abnormal meiosis of pollen mother cells (*Nonomura et al., 2003*; *Zhou et al., 2011*), abnormal metabolism of callose (Wan *et al*., *2011*), early or late degeneration of tapetum cells (*Jung et al., 2005*; *Li et al., 2006*), abnormal development of pollen wall (*Shi et al., 2011*), failure of anther dehiscence (*Steiner-Lange et al., 2003*). The tapetum plays an important role in pollen development, which provides nutrients for the developing pollen (*Pacini 2010*; *Gómez et al., 2015*). The middle layer cells are located between the tapetum and endothecium, and play an important role in the differentiation and function of tapetum cells (*Ma et al., 2007*; *Roque et al., 2007*). In recent years, with the rapid development of high-throughput sequencing technology, transcriptome sequencing has been widely used in the study of male sterility in herbs or cereals such as *Salvia miltiorrhiza* Bunge (*Yu et al*., *2021*), wheat (*Liu et al*., *2020*), *Cucumis melo* (Dai *et al*., *2019*), onion (*Yuan et al*., *2018*) and woody plant such as *Vernicia fordii* (Liu *et al*., *2019*), and *Citrus suavissima* (*Zhang et al*.,*2018*). However, to our knowledge, no study has used transcriptome sequencing technology to investigate male sterility in *Prunus*.

By using paraffin section, transcriptome sequencing and quantitative real-time PCR (RT-qPCR), this study clarified the critical stage and cause of male sterility in *Prunus sibirica*. The differentially expressed genes in male sterile flower buds and male fertile flower buds were analyzed, and the mechanism of male sterility was explored at the level of microstructure and gene transcription level. This work aims to provide a scientific basis for the study of the mechanism of male sterility in *Prunus sibirica*, and also lays the foundation for the development of molecular markers to select sterile materials and create high-yield and stable product varieties through use of genetic engineering technology to.

Materials & Methods

Plant materials

The eight-year-old *Prunus sibirica* clones were selected as the experimental material, which were cultivated in the *Prunus sibirica* germplasm resource nursery of Shenyang Agricultural University (Beipiao, Liaoning, China). Based on our research team's multi-year investigation of the inflorescence, florescence, male sterile flower buds of clone No. 1 were selected as the experimental group, and male fertile flower buds of clone No. 60 were selected as the control group. The anthers clone No. 1 were aborted thoroughly, there was no phenomenon of pollen dispersal, and the sterility was stable. The stamens of clone No. 60 developed completely, and can release abundant mature pollen after flowering, and they can pollinate and bear fruits well. Flower bud samples of male sterile clone and male fertile clone were collected from the end of July in 2018 to the beginning of April in 2019. Samples were taken every 10 to 15 days. Each time, 10~20 complete flower buds free of any observable disease symptoms or insect pests were selected from the upper part of the fruiting branches outside the canopy on the sunny side. They were collected and quickly stored in FAA fixative to create paraffin sections of flower buds. At the same time, three biological replicates of flower buds were collected and rapidly transferred into the liquid nitrogen (-196°C). The samples were taken back to the laboratory and stored in an ultra-low temperature freezer(-80°C) for RNA extraction. MSFB and MFFB were used to represent male sterile clone No. 1 and male fertile clone No. 60, respectively. MSFB_1, MSFB_2 and MSFB_3 were used to represent the experimental group, and MFFB_1, MFFB_2 and MFFB_3 were used to represent the control group.

Microstructure observation

Paraffin section technology was used to make sections (*Xu et al*., *2008*). The microstructure of male sterile flower buds and male fertile flower buds in different sampling periods were observed and photographed with an inverted fluorescent microscope (Zeiss Axio Vert.A1).

Determination of physiological indicators

The anthrone colorimetric method was used to determine the contents of soluble sugar and starch (*Gao*, *2005*). Coomassie Brilliant Blue G-250 staining was used to determine the contents of soluble protein (*Qu*, *2006*).

RNA extraction and detection

Total RNA was extracted from male sterile flower buds and fertile flower buds by using an RNA extraction kit (Tiangen Biotech Co., Ltd., Beijing) according to the manufacturer's instructions.

Then, 1.0 % (w/v) agarose gel electrophoresis was used to detect whether the extracted RNA was degraded or of appropriate quality. A NanoPhotometer spectrophotometer (IMPLEN, CA, USA) was used to detect the purity of RNA. An Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA) was used to detect the concentration and integrity of RNA.

Construction of sequencing library and Illumina sequencing

The sequencing library was constructed by using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA). First of all, poly (A) mRNA was purified from total RNA with Oligo(dT) magnetic beads. Fragmentation was carried out using divalent cations in NEBNext First Strand Synthesis Reaction Buffer (5X). Fragmented mRNA was taken as a template. The first strand cDNA was synthesized using a random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). The second cDNA synthesis was subsequently performed using DNA polymerase I and RNase H. The obtained double-stranded cDNAs were end-repaired. A tail was added, and a sequencing connector was added. In order to select cDNA fragments with 250~300 bp in length, the library fragments were purified with AMPure XP beads (Beckman Coulter, Beverly, USA). The purified cDNA was amplified by PCR, the PCR products were purified again by AMPure XP beads, and finally the cDNA libraries were obtained. The constructed libraries were quantified by using a Qubit 2.0 Fluorometer, and then the library quality was detected by using an Agilent 2100 bioanalyzer. To ensure the quality of the libraries, RT-qPCR was used to accurately quantify the effective concentration of the libraries. Meanwhile, the clustering of the samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS. After cluster generation, the library preparations were sequenced on an Illumina HiSeq platform and 150 bp paired-end reads were generated.

Transcriptome sequencing data analysis

The raw image data obtained from sequencing were converted to sequence reads using CASAVA. Then, the raw data were filtered by R language, and finally the clean reads for subsequent analysis were obtained. Trinity v2.4.0 program was used to stitch and assemble clean reads of all samples (*Grabherr et al., 2011*). Redundancy was removed by clustering with Corset hierarchy. The longest transcript of each gene was selected as unigene for subsequent analysis. The transcript sequence obtained by splicing with Trinity was used as a reference sequence. RSEM software (version v1.2.15) (bowtie2, mismatch=0) was used to compare clean reads of each repeated sample with reference sequences (*Li & Dewey, 2011*), and the readcount of each gene was counted and compared. Fragments per kilobase of exon per million mapped fragments (FPKM) was used to standardize the readcount of genes.

Differentially expressed gene analysis

The DEGseq R package (1.12.0) was used for differential expression analyses. The P-value was calculated on the basis of a negative binomial distribution model. P-values were adjusted using the Benjamini-Hochberg method. Genes with an adjusted P-value < 0.05 and log2(Fold change) > 1 were considered as differentially expressed.

Gene Ontology (GO) and KEGG enrichment analysis

Based on the hypergeometric distribution, GOseq R package (1.10.0) was used for GO enrichment analysis, and KOBAS (v2.0.12) was used for KEGG pathway enrichment analysis.

Quantitative Real Time PCR (RT-qPCR) analysis

RNA was extracted as described above. The first-strand cDNA was synthesized using EasyScript One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotech, China). The 18SrRNA gene was used as an internal reference gene (Table S1). RT-qPCR was performed with SYBR Green I method. The reaction system was 20 μL, including 10 μL of 2 × SuperReal PreMix Plus, 0.6 μL of 0.3 μmol/L forward primer, 0.6 μL of 0.3 μmol/L reverse primer, 1 μL of cDNA template, 2 μL of 50 × ROX Reference Dye, and 5.8 μL ddH_2O. The RT-qPCR was performed on the Applied Biosystems Step One Plus system, and the experiments were carried out in three replications. The PCR program was as follows: initial denaturation at 95°C for 15 min, denaturation at 95 °C for 10 s, annealing at 60 °C and extension for 32 s, for 40 cycles. The $2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression. The correlation coefficient between transcriptome sequencing and RT-qPCR was analyzed using SPSS 22.0 software.

Results

Microstructural characteristics of male sterile flower buds and male fertile flower buds of *Prunus sibirica*

To identify the key period and characteristics of male sterility in *Prunus sibirica*, the microstructures of male sterile flower buds and male fertile flower buds at different developmental periods were observed by paraffin section (Fig. 1). The results showed that anther development can be divided into four stages, namely sporogenous cell stage, tetrad stage, microspore stage, and pollen maturity stage. At the sporogenous cell stage (Fig. 1A, E) and tetrad stage (Fig. 1B, F), no significant difference was observed in anther development between male sterile flower buds and male fertile flower buds of *Prunus sibirica*. The sporogenous cells located in the four corners of the anther differentiate into multiple microspore mother cells. These cells produce microspores through meiosis, and they are surrounded by callose to form tetrads. When flower buds reach the microspore stage, microspores are released from tetrads with the degradation of the callose. In male fertile flower buds, the tapetum and middle layer cells in the pollen sac wall were degraded, and rendering the border between the two difficult to observe. At the pollen maturity stage, the tapetum and middle layer cells disappeared completely. Only the endothecium and epidermis were left in the pollen sac wall (Fig. 1G, H). In male sterile flower buds, however, the pollen sac wall tissue was complete at the microspore stage, with any signs of degradation failing to be observed. At the pollen maturity stage, the tapetum and middle layer cells were not degraded sufficiently, and the development of pollen was blocked, which eventually led to male sterility (Fig. 1C, D). Therefore, the abnormal anther development was the cause of male sterility in *Prunus sibirica*. The microspore stage was the key stage of anther abortion, and the pollen maturity stage was the post stage of anther abortion.

Total RNA extraction and detection

The total RNA concentrations of the six samples ranged from 452 ng·μL^{-1} to 670 ng·μL^{-1}. The RNA integrity number values were all close to 10 (Table S2). These results indicated that the extracted RNA had good integrity, high purity, and no obvious degradation, which met the quality requirements of sequencing library construction.

Six cDNA libraries of male sterile flower buds and male fertile flower buds at the microspore stage of abortion were sequenced on an Illumina sequencing platform using double-ended sequencing. After filtering and quality control, a total of 40 030,500~61,826,716 clean reads were obtained. The clean reads rate was between 98.02% and 98.45%. The total number of clean bases was 45.80 Gb, the sequencing quality represented by Q30 percentage was over 94%, and the GC content was between 45.65% and 45.88%. The transcripts assembled by Trinity were used as the reference transcriptome, and the clean reads of each sample were mapped to the reference sequences. The alignment proportion of each sample was greater than 80% (Table S3). All the above results showed that the quality and accuracy of the sequencing data were sufficient for further analysis.

Trinity software was used to assemble the clean reads and obtain non-redundant unigenes, and a total of 34,377 unigenes were obtained. The gene function of the obtained unigenes were annotated in seven databases (Nr, Nt, KEGG, SwissProt, Pfam, GO, and KOG), and it was found that a total of 27,798 unigenes were annotated, accounting for 80.86% of the number of originally identified unigenes. The number of unigenes that were annotated successfully in all seven databases was 3,283, accounting for 9.54% (Table S4).

Analysis of the differentially expressed genes

The expressed genes with false discovery rate (FDR) < 0.05 and fold change (FC) > 2 were designated as differentially expressed genes. Differences in gene expression in male sterile flower buds and male fertile flower buds were compared and analyzed. A total of 4,108 genes were differentially expressed with 1,899 up-regulated genes and 2,209 down-regulated genes in male sterile flower buds (Fig. S1).

GO functional enrichment analysis of differentially expressed genes

GO functional classification analysis of the differentially expressed genes between male sterile flower buds and male fertile flower buds was conducted. The results are shown in Fig. 2. Among the biological processes, differential genes were mainly enriched in the processes of “cellular process”, “metabolic process”, and “single-organism process”. In cellular component category, the differential genes were mainly distributed in “cells”, “cell part”, “cell composition”. In the molecular function category, the differential genes were mainly enriched in “binding”, “catalytic activity”, and “heterocyclic compound binding”.

Based on the results of GO enrichment analysis, eight GO functional subclasses with significant enrichment (corrected *P*-value <0.5) were selected (Fig. 3), including "DNA integration" in biological process and “ADP binding”, “heme binding”, “tetrapyrrole binding”, “iron ion binding”, “oxidoreductase activity”, “acting on paired donors with incorporation or reduction of molecular oxygen”, “oxidoreductase activity”, and “terpene synthase activity” in molecular function. The results indicated that these functional categories played a central role in the occurrence of male sterility in *Prunus sibirica*.

According to microstructure observations, the insufficient degradation of tapetal cells at the pollen maturity stage was one of the causes of male sterility in *Prunus sibirica*. The degradation process of tapetal cells belongs to programmed cell death. Therefore, the genes involved in the categories of “protein kinase activity”, “apoptosis process”, “calcium binding”, “cell death”, “cytochrome c oxidase activity”, “aspartate peptidase activity” and “cysteine peptidase activity” may be the key genes that regulate the male sterility of *Prunus sibirica*. Among them, a total of 297 related genes were enriched, and the detailed information of the 297 genes is listed in Table S5.

Hierarchical clustering was performed on key differentially expressed genes of male sterility in *Prunus sibirica*, which were screened on the basis of GO function enrichment analysis, and the heat map was drawn (Fig. S2). The column represents the expression of the same gene in different samples, and the horizontal row represents the expression of different genes in the same sample. The differences between individual samples at the transcriptome level can be preliminarily depicted by the use of a cluster map, and the differentially expressed genes can be divided into four categories (Category 1, Category 2, Category 3 and Category 4). The detailed information of differentially expressed genes in GO functional enrichment analysis is listed in Table S5.

KEGG pathway enrichment analysis of differentially expressed genes

To explore the main metabolic pathways involved in differentially expressed genes, KEGG pathway cluster analysis was conducted on differentially expressed genes between male sterile flower buds and male fertile flower buds of *Prunus sibirica*. The differentially expressed genes were mapped to 215 biological pathways, and 20 of them were significantly enriched (Fig. 4). They mainly included “plant-pathogen interaction”, “monoterpenoid biosynthesis”, “protein processing in endoplasmic reticulum”, “flavonoid biosynthesis”, “Stilbenoid, diarylheptanoid and gingerol biosynthesis”, “phenylpropanoid biosynthesis”, “Vitamin B6 metabolism”, “glycolysis / gluconeogenesis”, and “diterpenoid biosynthesis”. These pathways might play an important role played a key role in regulating the formation of male sterility in *Prunus sibirica*.

The metabolism of macromolecular nutrients plays a key role in the formation of male sterility in plants. In this study, the differences of soluble sugar content, starch content, and soluble protein content between male sterile flower buds and male fertile flower buds were analyzed. The results showed that during the critical stage of anther abortion, the soluble sugar content, starch content and soluble protein content in male sterile flower buds were significantly lower than those in male fertile flower buds at the same stage. During the post-abortion stage, the soluble sugar content and soluble protein content in male sterile flower buds were significantly lower than those in male fertile flower buds, whereas the starch content was significantly higher than that of male fertile flower buds (Fig. 5).

Combined with the results of KEGG pathway enrichment analysis, we found that “starch and sucrose metabolism” and “proteasome” pathways were enriched separately, and a total of 35 differentially expressed genes were involved (Fig. S3). The result suggested that these differentially expressed genes played an important regulatory role in the occurrence of male sterility in *Prunus sibirica*. The detailed information of differentially expressed genes in KEGG pathway enrichment analysis is listed in Table S6.

RT-qPCR analysis

To verify the accuracy of the transcriptome sequencing data, 14 DEGs were selected and used for RT-qPCR verification (Table S7). The RT-qPCR results were largely consistent with the transcriptome data (Fig. 6), and the correlation coefficient was 0.964. The results showed that the accuracy of transcriptome sequencing was high, which can be used to analyze the dynamic changes to the expression of genes related to male sterility in the flowers of *Prunus sibirica*, thus revealing the molecular mechanism of male sterility.

Discussion

In the studies of male sterility of *Prunus*, Lillecrapp *et al*. (1999) found that the anthers of `Trevatt Blue' apricot contained degraded microspores, with some failure in tapetal breakdown. Badenes et al. (2000) found a marker (M4-650) linked to male fertility traits in apricot using RAPD markers combined with the BSA method. *Yaegaki et al.* (2003) found that genes related to the male sterility of *Prunus mume* belong to the cytoplasmic type.

According to the microstructural characteristics of male sterile flower buds and fertile flower buds at different development stages of *Prunus sibirica*, anther development can be divided into four stages, namely sporogenous cell stage, tetrad stage, microspore stage and pollen maturity stage. This was similar to the findings reported by *Zhang et al*. (*2018*). The reason for the male sterility in *Prunus sibirica* was because the tapetum and middle layer cells failed to degrade sufficiently. The microspore stage was the key stage of anther abortion, whereas the pollen maturity stage was the post-abortion stage.

The soluble sugar content in male sterile flower buds of *Prunus sibirica* was significantly lower than that in fertile flower buds during the key stage of anther abortion and post-abortion stage, and the difference reached the maximum at the post-abortion stage. This result indicated that the degree of carbohydrate metabolism in male sterile flower buds was lower, which possibly led to a deficiency in the supply of required energy, and therefore, anther abortion (*Han et al*., *2020*). The content of soluble protein in male sterile flower buds of *Prunus sibirica* was significantly lower than that of fertile flower buds during the key stage of abortion and post-abortion stage, and the difference reached the maximum during the key stage of abortion. The starch content in male sterile flower buds was significantly higher than that in fertile flower buds during the key stage of abortion, and significantly lower than that in fertile flower buds during the post-abortion stage. This result indicated that anther abortion was closely related to the inability of macromolecular nutrients to produce soluble small molecular substances in time, which was consistent with the results of *Liu et al*. (*2014*) on male sterility in *Sesamum indicum*.

In recent years, many studies have used transcriptome sequencing to investigate male sterility, including studies in pepper (*Lv et al*., *2020*), *Catalpa bungei* (*Mao et al*., *2017*), maize (*Xue et al*., *2019*), and soybean (*Li et al*., *2019*). However, no study has investigated male sterility of *Prunus sibirica* using transcriptome sequencing. The filtered clean reads account for more than 80% of raw reads (*Ge et al*., *2014*). The GC content can reflect the structure of nucleic acid sequence within a certain range and can be used as an important feature of transcriptome sequencing (*Xu et al*., *2020*). In our study, the proportion of clean reads obtained from each repeated sample was more than 98%, and the GC content was between 45.65% and 45.88%. The results showed that the reliability of the sequencing data was high.

Male sterility is regulated by a series of genes related to fertility, which can be broadly divided into abnormal meiosis genes (*Nonomura et al*., *2004*), abnormal tapetum development genes (*Li et al*., *2006*), abnormal callose metabolism genes (*Wan et al*., *2011*), abnormal pollen wall formation genes (*Shi et al*., *2011*) and abnormal anther cracking genes (*Steiner-Lange et al*., *2003*). The development of the tapetum is closely related to male sterility, and the tapetum is the innermost tissue of the pollen sac wall, which secretes enzymes needed to degrade the callose component of the wall (*Cui et al*., *2017*). Furthermore, the tapetum provides essential nutrients for the development of microspore mother cells (*Li et al*., *2020*), and it is important for normal pollen development. Through the enrichment analysis of differentially expressed genes, the biological functions and metabolic pathways related to male sterility of *Prunus sibirica* were screened. We found that “protein kinases activity” (*Mizuno et al*., *2007*), “apoptotic process” (*Love et al*., *2008*), “calcium ion binding” (*Cao et al*., *2012*), “cell death” (*Zhu et al*., *2006*), “cytochrome c oxidase activity” (*Luo et al*., *2013*), “aspartic peptidase activity” (*Li et al*., *2006*) and “cysteine- peptidase activity” (*Niu et al*., *2013*) and other terms are closely related to insufficient degradation of tapetum. Biological pathways such as “starch and sucrose metabolism” and “proteasome” also participate in the formation of male sterility of *Prunus sibirica*. These analyses therefore provided strong initial evidence for elucidating the regulatory pathways of male sterility in *Prunus sibirica*.

Conclusions

We inferred that during anther development in *Prunus sibirica*, starch and other macromolecular substances cannot be converted into soluble small molecules in time due to abnormal metabolic process such as the release of cytochrome c, resulting in the abnormal development of tapetum cells and abnormal pollen development to form male sterility. The stage, characteristics and related metabolic pathways of male sterility in *Prunus sibirica* were analyzed at the microscopic and transcriptome level in our study, which has important reference value for the study of reproductive development and the mechanism of male sterility in *Prunus sibirica*. At the same time, this study provides scientific reference for breeding male sterile germplasms with better comprehensive traits, and the selected germplasms can be used as a worthful female parent in the hybridization breeding. It is of great significance for giving full play to heterosis and utilizing reasonably the *Prunus sibirica* germplasm resources in future hybrid breeding.

Acknowledgements

We thank Professor Xiujun Lu and Lecturer Xiaolin Zhang from Shenyang Agricultural University for their helpful comments and suggestions to improve our manuscript.

References

Badenes ML,Hurtado MA, Archelos DM, Burgos L, Egea J, Llácerl G. 2000. Searching for molecular markers linked to male sterility and self-compatibility in apricot, *Plant Breeding* **119(2)**:157-160 DOI 10.1046/j.1439-0523.2000.00463.x.

Cao Y, Kang XY, Zhang ZY, Jing YP. 2012. Anomalous changes in Ca^{2+} -ATPase distribution during the process of pollen abortion in *Populus tomentosa* Carr. *Journal of Beijing Forestry University* **34(4)**:10-17 DOI 10.13332/j.1000-1522.2012.04.020.

Cui HF, Zhang F, Yin JL, Guo YQ, Yue YL. 2017. Callose deposition and pollen development. *Journal of Yunnan Agricultural University (Natural Science)* **32(3)**:551-557 DOI 10.16211/j.issn.1004-390X(n).2017.03.022.

Dai DY, Xiong AP, Yuan LW, Sheng YY, Ji P, Jin YZ, Li DD, Wang YH, Luan FS. 2019. Transcriptome analysis of differentially expressed genes during anther development stages on male sterility and fertility in *Cucumis melo* L. line. *Gene* **707**:65-77. DOI 10.1016/j.gene.2019.04.089.

Du K, Liu Q, Wu XY, Jiang JJ, Wu J, Fang YJ, Li AM, Wang YP. 2016. Morphological structure and transcriptome comparison of the cytoplasmic male sterility line in *Brassica napus* (SaNa-1A) derived from somatic hybridization and its maintainer line SaNa-1B. *Frontiers in Plant Science* **7**: 1313 DOI 10.3389/fpls.2016.01313.

Gao JF. 2005. Experimental Guidance of Plant Physiology. Beijing: Higher Education Press.

Ge SJ, Sun AQ, Liu P, Zhang JD, Dong ST. 2014. *In silico* Expression profile of maize genes in response to osmotic stress. *Acta Agronomica Sinica* **40(7)**:1164-1173 DOI 10.3724/SP.J.1006.2014.01164.

Gómez JF, Talle B, Wilson ZA. 2015. Anther and pollen development: A conserved developmental pathway. *Journal of Integrative Plant Biology* **57(011)**:876-891. DOI 10.1111/jipb.12425.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, Palma FD, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29(7)**:644-652 DOI 10.1038/nbt.1883.

Han PA, Bai CH, Wu XR, Wang L, Zhang ZQ,Sun RF, Zhang BZ, Li XD. 2020. Differential Protein Analysis between Male Sterile Line and Maintainer Line in Bud Stage of Sugarbeet. *Molecular Plant Breeding* **18(16)**:5230-5236 DOI 10.13271/j.mpb.018.005230.

Jiang HB, Sun YN, Xu Y, Song WX, Li YY, Tian YP. 2020. Physiological and biochemical changes during bud development in male sterile plant of *Camellia crassicolumna*. *Plant Physiology Journal* **56(9)**:1807-1817 DOI 10.13592/j.cnki.ppj.2020.0209.

Jung KH, Han MJ, Lee YS, Kim YW, Hwang I, Kim MJ, Kim YK, Nahm BH, An G. 2005. Rice Undeveloped Tapetum1 is a major regulator of early tapetum development. *Plant Cell* **17(10)**:2705-2722. DOI 10.1105/tpc.105.034090.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**:323 DOI 10.1186/1471-2105-12-323.

Li N, Zhang DS, Liu HS, Yin CS, Li XX, Liang WQ, Yuan Z, Xu B, Chu HW, Wang J, Wen TQ, Huang H, Luo D, Ma H, Zhang DB. 2006. The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. *The Plant Cell* **18(11)**:2999-3014 DOI 10.1105/tpc.106.044107.

Li R, Lin CJ, Peng B, Ding XY, Li YK, Zhao GL, Zhao LM, Zhang CB. 2019. Transcriptomic analysis of soybean cytoplasmic male sterile lines with different outcrossing rate. *Chinese Journal of Oil Crop Sciences* **41(5)**:696-704 DOI 10.19802/j.issn.1007-9084.2019053.

Li YQ, Wang QL, Wei CY, Dong T, Chen QJ, Zhou RY. 2020. Morphological observation and physiological and biochemical characteristics of ccotton cytoplasmic male sterile cells. *Southwest China Journal of Agricultural Sciences* **33(1)**:58-63 DOI 10.16213/j.cnki.scjas.2020.1.010.

Lillecrapp A M, Wallwork M A, Sedgley M. 1999. Female and male sterility cause low fruit set in a clone of the `Trevatt' variety of apricot (*Prunus armeniaca*). *Scientia Horticulturae* **82(3-4):**255-263. DOI 10.1016/S0304-4238(99)00061-8.

Liu HY, Tan MP, Yu HJ, Li L, Zhou F, Yang MM, Zhou T, Zhao YZ. 2016. Comparative transcriptome profiling of the fertile and sterile flower buds of a dominant genic male sterile line in seasame (*Sesamum indicum* L). *BMC Plant Biology* **16(1)**:250 DOI 10.1186/s12870-016-0934-x.

Liu HY, Wu K, Yang MM, Zhou XA, Zhao YZ. 2014. Variation of soluble sugar, starch and plant hormones contents in sesame dominant genic male sterile line during bud development. *Chinese Journal of Oil Crop Sciences* **36(02)**:175-180 DOI 10.7505/j.issn.1007-9084.2014.02.006.

Liu ML, Li WY, Zhao G, Fan XM, Long HX, Fan YR, Shi MW, Tan XF, Zhang L. 2019. New Insights of Salicylic Acid Into Stamen Abortion of Female Flowers in Tung Tree (*Vernicia fordii*). *Frontiers in Genetics* **10** DOI 10.3389/fgene.2019.00316.

Liu WS, Han LZ, Zhu S, Pan HX. 2019. Cytological observation on pollen development of the male sterile poplar variety. *Journal of Nanjing Forestry University (Natural Sciences Edition)* **43(1)**:198-203 DOI 10.3969/j.issn.1000-2006.201805034.

Liu ZH, Li S, Li W, Liu Q, Zhang LL, Song XY. 2020. Comparative transcriptome analysis indicates that a core transcriptional network mediates isonuclear alloplasmic male sterility in wheat (*Triticum aestivum* L.). *BMC Plant Biology* **20(1)**:10 DOI 10.1186/s12870-019-2196-x.

Liu ZW, Liu YF, Sun YH, Yang AG, Li FX. 2020. Comparative transcriptome analysis reveals the potential mechanism of abortion in Tobacco *sua*-Cytoplasmic male sterility. *International Journal of Molecular Sciences* **21(7)**:2445 DOI 10.3390/ijms21072445.

Love AJ, Milner JJ, Sadanandom A. 2008. Timing is everything: regulatory overlap in plant cell death. *Trends in Plant Science* **13(11)**:589-595 DOI 10.1016/j.tplants.2008.08.006.

Luo DP, Xu H, Liu ZL, Guo JX, Li HY, Chen LT, Fang C, Zhang QY, Bai M, Yao N, Wu H, Wu H, Ji CH, Zheng HQ, Chen YL, Ye S, Li XY, Zhao XC, Li RQ, Liu YG. 2013. A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. *Nature genetics* **45(5)**:573-U157 DOI 10.1038/ng.2570.

Lv JH, Liu ZB, Liu YH, Ou LJ, Deng MH, Wang J, Song JS, Ma YQ Chen WC, Zhang ZQ, Dai XZ, Zou XX**. 2020.** Comparative transcriptome analysis between Cytoplasmic male-sterile line and its maintainer during the floral bud development of pepper. *Horticultural Plant Journal* **6(2)**:89-98 DOI 10.1016/j.hpj.2020.01.004.

Ma J, Duncan D, Morrow DJ, Fernandes J, Walbot V. 2007. Transcriptome profiling of maize anthers using genetic ablation to analyze pre-meiotic and tapetal cell types. *The Plant Journal* **50(4)**:637-648. DOI 10.1111/j.1365-313X.2007.03074.x.

Mao WB, Chen FJ, Wang CL, Liang HW. 2017. Transcriptome Sequencing and Analysis of Male Sterile Flower Buds in *Catalpa bungei*. Scientia Silvae Sinicae **53(6)**:141-150 DOI 10.11707/j.1001-7488.20170617.

Mizuno S, Osakabe Y, Maruyama K, Ito T, Osakabe K, Sato T, Shinozaki K, Yamaguchi-Shinozaki K. 2007. Receptor-like protein kinase 2 (RPK 2) is a novel factor controlling anther development in *Arabidopsis thaliana*. *The Plant Journal* **50(5)**:751-766 DOI 10.1111/j.1365-313X.2007.03083.

Nonomura K, Miyoshi K, Eiguchi M, Suzuki T, Miyao A, Hirochika H, Kurata N. 2003. The MSP1 Gene Is Necessary to Restrict the Number of Cells Entering into Male and Female Sporogenesis and to Initiate Anther Wall Formation in Rice. *Plant Cell* **15(8)**:1728-1739. DOI 10.1105/tpc.012401.

Nonomura KI, Nakano M, Fukuda T, Eiguchi M, Miyao A, Hirochika H, Kurata N. 2004. The novel gene *HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS1* of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis. *The Plant Cell* **16(4)** :1008-1020 DOI 10.1105/tpc.020701.

Niu NN, Liang WQ, Yang XJ, Jin WL, Wilson AZ, Hu JP, Zhang DB. 2013. EAT1 promotes tapetal cell death by regulating aspartic proteases during male reproductive development in rice. *Nature Communications* **4**:1445 DOI 10.1038/ncomms2396.

Pacini E. 2010. Relationships between tapetum, loculus, and pollen during development. *International Journal of Plant Sciences* **171(1)**:1-11. DOI 10.1086/647923.

Qu CX, Shen SD, Wang XF, Cui YH, Song WP. 2006. Method research of measuring soluble protein contents of plant rough extraction using Coomassie Brilliant Blue. *Journal of Suzhou University (Natural Science Edition)* **22(2)**: 82-85

Radice S, **Ontivero M**, **Giordani E, Bellini E. 2008.** Anatomical differences on development of fertile and sterile pollen grains of *Prunus salicina* Lindl. *Plant Systematics & Evolution* **273(1-2)**:63-69 DOI 10.1007/s00606-008-0011-5.

Roque E, MD Gómez, Ellul P, Wallbraun M, Madueño F, Beltrán J, Cañas LA. 2007. The PsEND1 promoter: a novel tool to produce genetically engineered male-sterile plants by early anther ablation. *Plant Cell Reports* **26(3)**:313-325. DOI 10.1007/s00299-006-0237-z.

Shi J, Tan HX, Yu XH, Liu YY, Liang WQ, Ranathunge K, Franke RB, Schreiber L, Wang YJ, Kai GY, Shanklin J, Ma H, Zhang DB. 2011. *Defective pollen wall* is required for anther and microspore development in rice and encodes a fatty acyl carrier protein reductase. *The Plant Cell* **23(6)**:2225-2246 DOI 10.1105/tpc.111.087528.

Steiner-Lange S, Unte US, Eckstein L, Yang CY, Wilson ZA, Schmelzer E, Dekker K, Saedler H. 2003. Disruption of *Arabidopsis thaliana MYB26* results in male sterility due to non-dehiscent anthers. *The Plant Journal* **34(4)**:519-528 DOI 10.1046/j.1365-313x.2003.01745.x.

Wan LL, Zha WJ, Cheng XY, Liu C, Lv L, Liu CX, Wang ZQ, Du B, Chen RZ, Zhu LL, He GC. 2011. A rice β-1,3-glucanase gene *Osg1* is required for callose degradation in pollen development. *Planta* **233(2)**:309-323 DOI 10.1007/s00425-010-1301-z.

Wan DM, Wu YL, Wang MH, Dong SJ, Liu MG. 2015. Genetic Diversity of *Armeniaca sibirica* Germplasm Resources Revealed by SRAP Analysis. *Molecular Plant Breeding* **13(11)**:2537-2541 DOI 10.13271/j.mpb.013.002537.

Wang Z, Kang M, Liu H, Gao J, Zhang ZD, Li YY, Wu RL, Pang XM. 2014. High-level genetic diversity and complex population structure of Siberian apricot (*Prunus sibirica* L.) in China as revealed by nuclear SSR markers. *PloS one* **9(2)**:e87381 DOI 10.1371/journal.pone.0087381

Xu YL, Jin JJ, Zhao YZ, Wei P, Xi JQ, Yang J, Cao PJ, Zhang JF. 2020. Genome survey of *Lasioderma serricorne* based on high-throughput sequencing. *Tobacco Science & Technology* **53(11)**:1-6 DOI 10.16135/j.issn1002-0861.2020.0211.

Xu YT, Luo SP, Li J, Yang WY, Zeng B. 2008. Morphological Differentiation of Flower Bud in *Amygdalus Ledebouriana* Schleche. *Nonwood Forest Research* **26(4)**:86-89 DOI 10.14067/j.cnki.1003-8981.2008.04.017.

Xue YD, Yang L, Yang HL, Li B, Lin YN, Zhang HS, Guo ZY, Tang JH. 2019. Comparative transcriptome analysis among the three line of cytoplasmic male sterility in maize. *Scientia Agricultura Sinica* **52(8)**:1308-1323 DOI 10.3864/j.issn.0578-1752.2019.08.002.

Yaegaki H, Miyake M, Haji T, Yamaguchi M. 2003. Inheritance of male sterility in Japanese apricot (*Prunus mume*). *Hortscience* **38(7)**:1422-1423 DOI 10.21273/HORTSCI.38.7.1422.

Yu Y, Jiang Y, Wang L, Wu L, Wu YC, Liao JQ, Zhong MZ, Yang RW, Chen XF, Li QM, Zhang L. 2021. Comparative transcriptome analysis reveals key insights into male sterility in *Salvia miltiorrhiza* Bunge. *PeerJ* **9**:e11326 DOI 10.7717/peerj.11326.

Yuan QL, Song C, Gao LY, Zhang HH, Yang CC, Sheng J, Ren J, Chen D, Wang Y. 2018. Transcriptome de novo assembly and analysis of differentially expressed genes related to cytoplasmic male sterility in onion. *Plant Physiology and Biochemistry* **125**:35 DOI 10.1016/j.plaphy.2018.01.015.

Zhang C, Yu DH, Ke FZ, Zhu MM, Xu JG, Zhang M. 2018. Seedless mutant 'Wuzi Ougan' (*Citrus suavissima* Hort. ex Tanaka 'seedless') and the wild type were compared by iTRAQ-based quantitative proteomics and integratedly analyzed with transcriptome to improve understanding of male sterility. *BMC Genomics*, **19**:106 DOI 10.1186/s12863-018-0693-9.

Zhang CY, Guo SF, Sun Q, Zhang GZ, Xu JJ, Liu L. 2018. Embryological characterization of apricot pollen and anther wall development. *Acta Horticulturae Sinica* **45(12)**:2427-2436 DOI 10.16420/j.issn.0513-353x.2018-0271.

Zhou SR, Wang Y, Li WC, Zhao ZG, Ren YL, Wang Y, Gu SH, Lin QB, Wang D, Jiang L, Su N, Zhang X, Liu LL, Cheng ZJ, Lei CL, Wang JL, Guo XP, Wu FQ, Ikehashi H, Wang HY, Wan JM. 2011. *Pollen semi-sterility1* encodes a kinesin-1-like protein important for male meiosis, anther dehiscence, and fertility in rice. *Plant Cell* **23(1)**:111-129 DOI 10.1105/tpc.109.073692.

Zhu BB, Li CQ, Lu JJ, Peng M. 2006. The progresses of research on programmed cell death of plants. *Molecular Plant Breeding* **4(S1)**:11-15.

Figure 1

The microstructure of male sterile flower buds (A-D) and male fertile flower buds (E-H) at different developmental stages in *Prunus sibirica*

A, E: sporogenous cell stage; B, F: tetrad stage; C, G: microspore stage; D, H: anther maturity stage. Sp: sporogenous cells; Ps: pollen sac; T: tapetum; Tds: tetrads; Ep: epidermis; En: endothecium; ML: middle layer; MSp: microspores; A, D: sporogenous cell stage; B, F: tapetum stage.

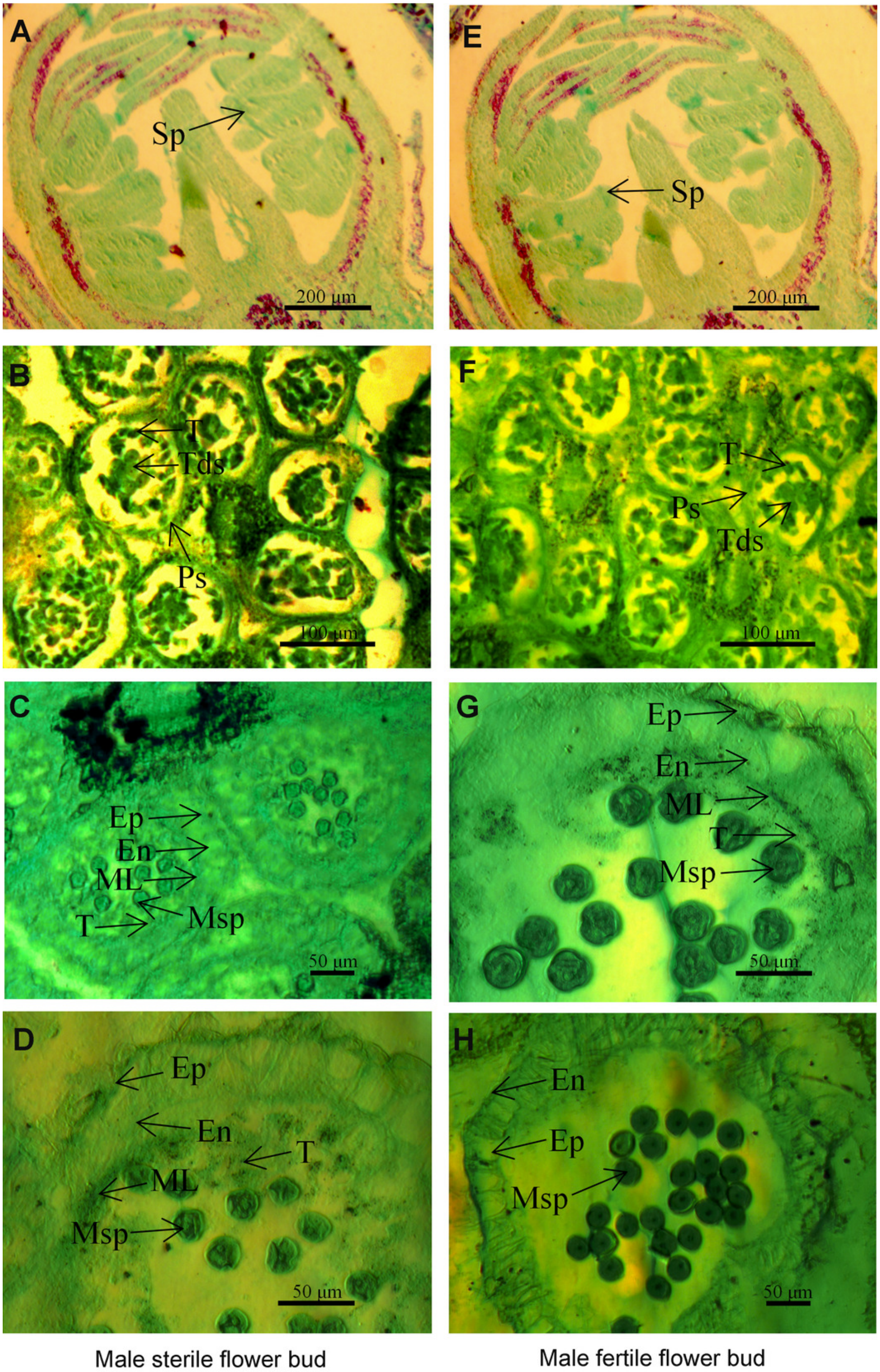

A
Sp
200 µm
E
Sp
200 µm
B
T
Tds
Ps
100 µm
F
T
Ps
Tds
100 µm
C
Ep
En
ML
Msp
T
50 µm
G
Ep
En
ML
T
Msp
50 µm
D
Ep
En
T
ML
Msp
50 µm
H
En
Ep
Msp
50 µm
Male sterile flower bud
Male fertile flower bud

Figure 2

Figure 2 GO functional classification of differentially expressed genes

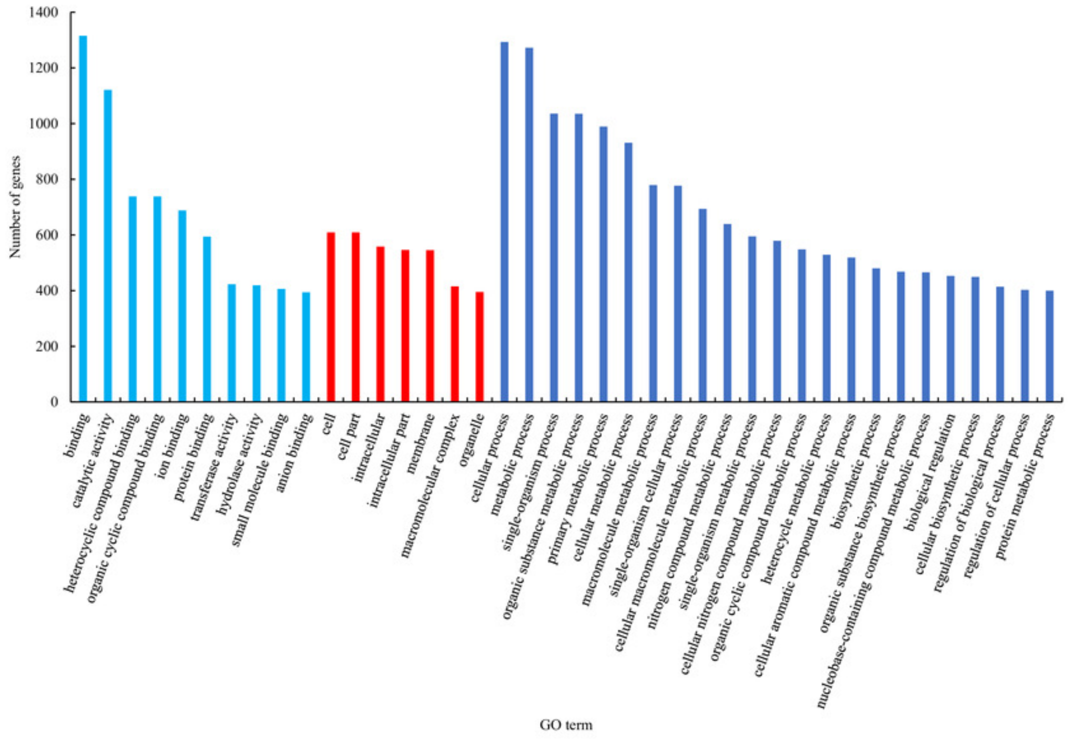


Figure 3

Figure 3 GO significant enrichment analysis of differentially expressed genes

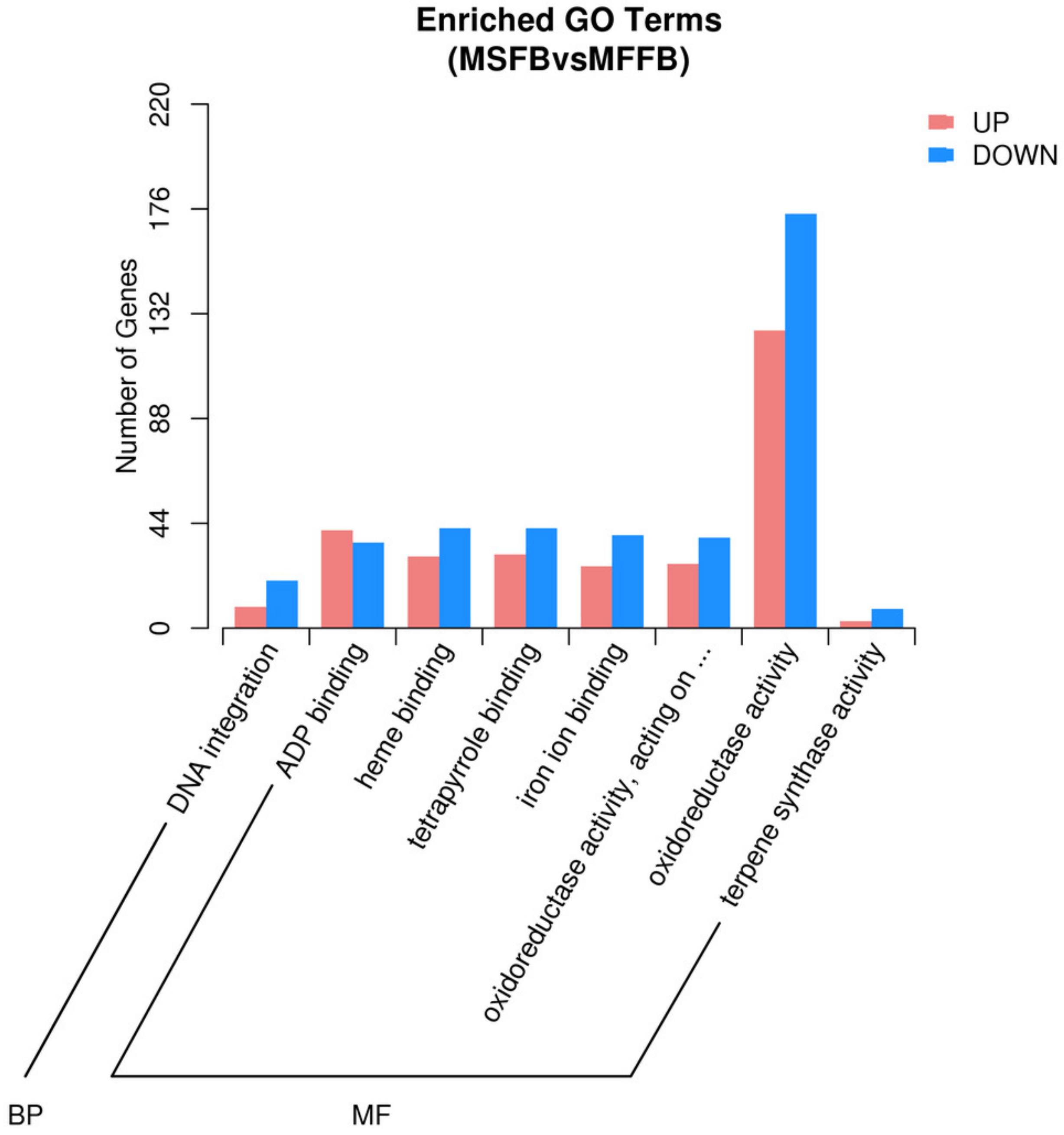


Figure 4

Figure 4 KEGG pathway significant enrichment analysis of differentially expressed genes

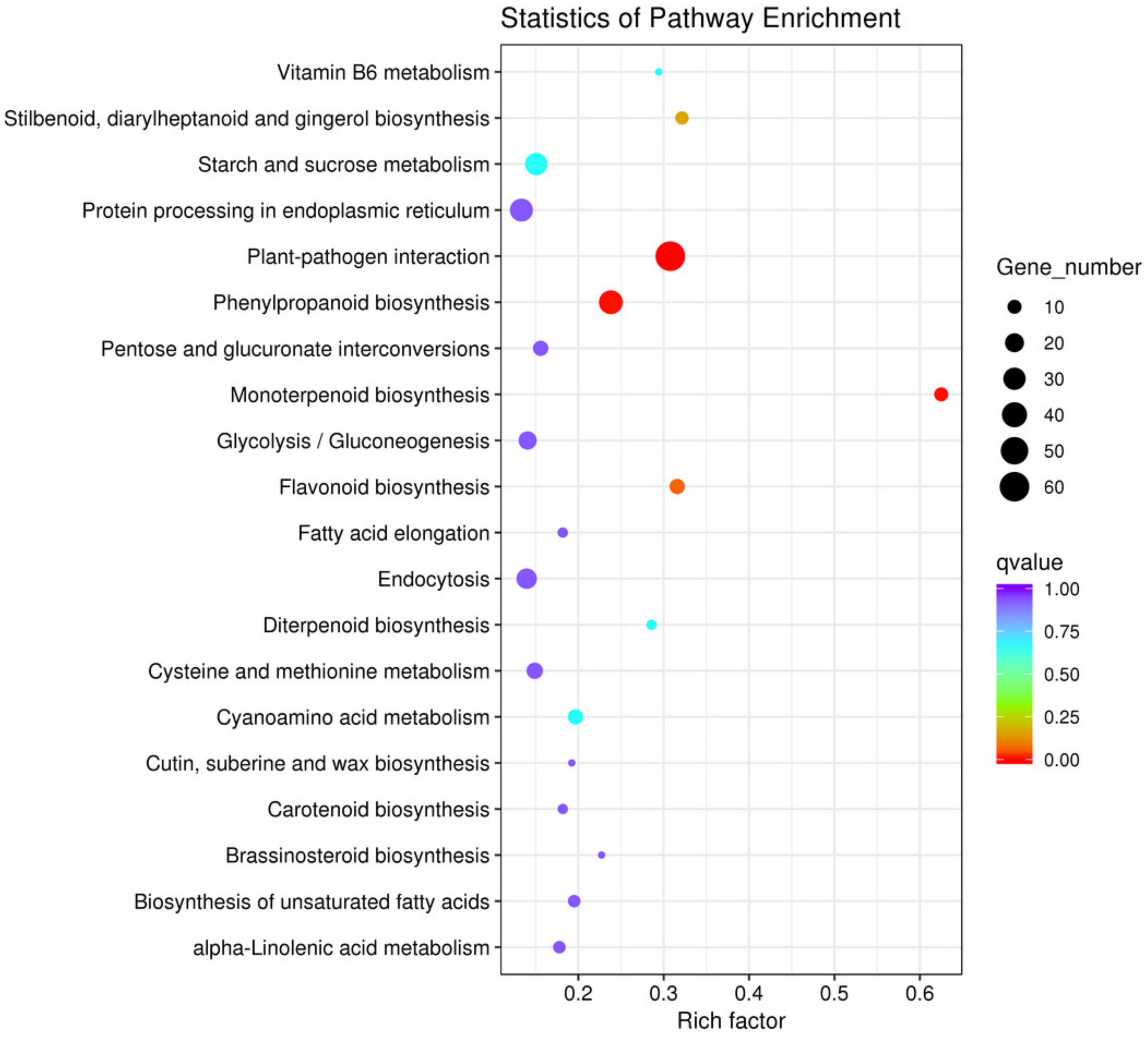


Figure 5

Figure 5 Analysis of related physiological indexes between MSFBs and MFFBs in *Prunus sibirica*

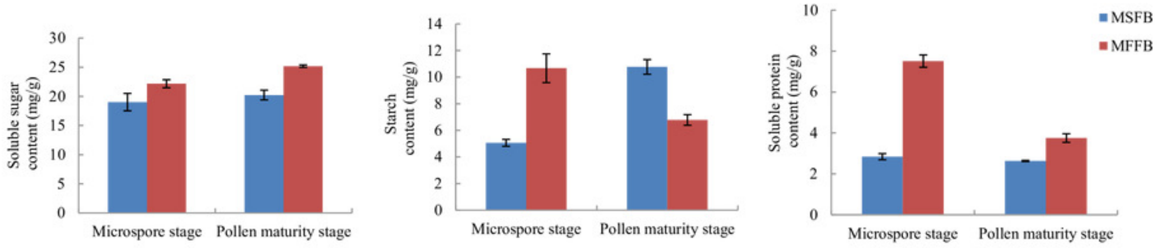
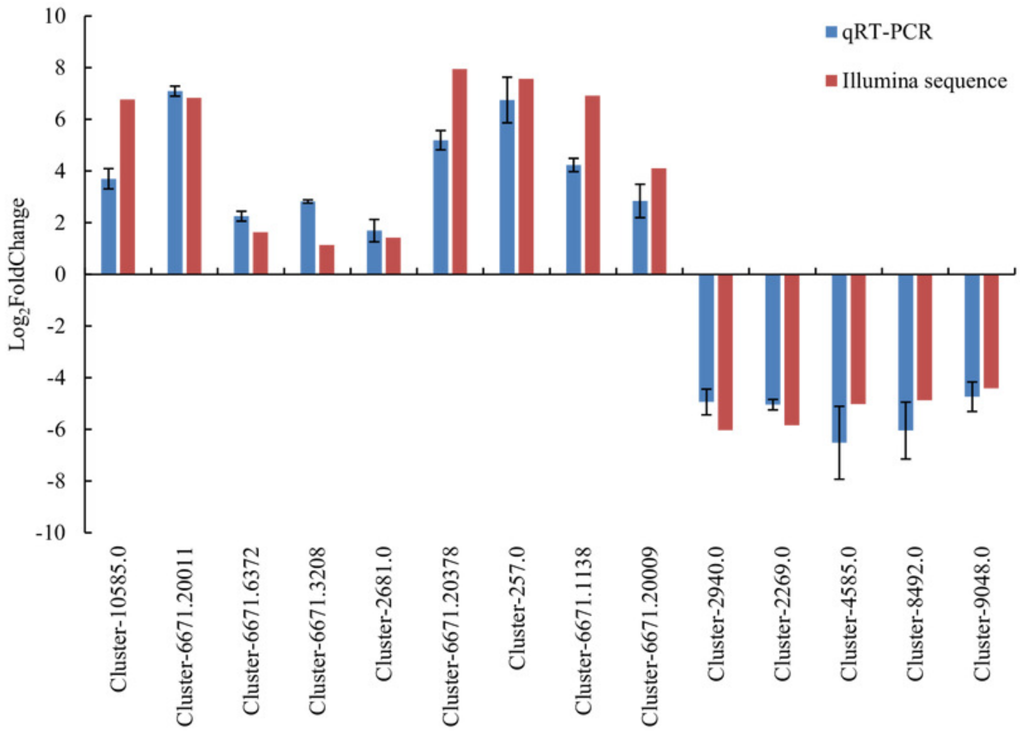


Figure 6

Figure 6 qRT-PCR verification for transcriptome sequencing results


10
8
6
4
2
0
-2
-4
-6
-8
-10
$Log_2FoldChange$
qRT-PCR
Illumina sequence
Cluster-10585.0
Cluster-6671.20011
Cluster-6671.6372
Cluster-6671.3208
Cluster-2681.0
Cluster-6671.20378
Cluster-257.0
Cluster-6671.1138
Cluster-6671.20009
Cluster-2940.0
Cluster-2269.0
Cluster-4585.0
Cluster-8492.0
Cluster-9048.0