Xu *et al. BMC Plant Biology* (2022) 22:199
https://doi.org/10.1186/s12870-022-03581-1

BMC Plant Biology

RESEARCH Open Access

Transcriptome profiling of flower buds of male-sterile lines provides new insights into male sterility mechanism in alfalfa

Bo Xu, Rina Wu, Fengling Shi*, Cuiping Gao and Jia Wang

Abstract

Background: The use of heterosis to produce hybrid seeds is a challenge to breeding for improved crop yield. In previous studies, we isolated a male sterile alfalfa hybrid and successfully obtained a genetically stable alfalfa male sterile line through backcrossing, henceforth named MS-4. In this study, we used RNA-seq technology to analyze the transcriptome profiles of the male sterile line (MS-4) and the male fertile line (MF) of alfalfa to elucidate the mechanism of male sterility.

Results: We screened a total of 11,812 differentially expressed genes (DEGs) from both MS-4 and MF lines at three different stages of anther development. Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses revealed that these DEGs are mainly involved in processes such as energy metabolism, lipid and amino acid metabolism, carbohydrate metabolism, in addition to cell synthesis and aging. The results from protein–protein interaction (PPI) network analysis showed that the ribosomal protein (*MS.Gene25178*) was the core gene in the network. We also found that transcriptional regulation was an influential factor in the development of anthers.

Conclusions: Our findings provide new insights into understanding of the fertility changes in the male sterile (MS-4) of alfalfa.

Keywords: Alfalfa, Ribosomal protein, Male sterility, Transcriptome, DEGs

Introduction

Heterosis or hybrid vigor is a natural phenomenon whereby the hybrid offspring of a genetically diverse cross outperform the parents in multiple traits, including yield, adaptability, and resistance to biotic and abiotic stresses. It is one of the most effective approaches for improving both yield and quality of crops [1, 2]. Using male sterile lines to achieve heterosis can effectively minimize self-degeneration, save the laborious steps of manual emasculation, and facilitate the harvesting of hybrid seeds. The use of male sterile lines to produce hybrid seeds has been widely tested and successfully applied to numerous crops [3, 4]. At the same time, male sterile lines are valuable materials in studying changes occurring during the development of stamens and anthers due to the interaction of cytoplasmic and nuclear genomes [5, 6].

Various conditions may lead to male sterility. Most of the male sterile lines were produced by impairment of anther formation or abnormal development. The development and formation of anthers is a complex process involving many metabolic pathways [7]. Metabolic pathways related to the development of anthers (that produce pollen) mainly include processes such as amino acid transport and metabolism [8], carbohydrate metabolism [9], lipid metabolism [10], energy conduction and metabolism [11]. Fang [12] reported reduced expression of glutamine synthetase in the amino acid biosynthesis

*Correspondence: nmczysfl@126.com
Key Laboratory of Grassland Resources of the Ministry of Education, College of Grassland Resources and Environment, Inner Mongolia Agricultural University, Hohhot, China

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

pathway in cytoplasmic male sterile pepper, while Yang [7], detected significant differences in the expression of glycosyltransferase and glycosylhydrolase in the carbohydrate metabolism pathway between male fertile and male sterile lines of eggplant. Moreover, genes of carbon, starch, and sucrose metabolism were most abundant in the KEGG enrichment analysis. The *CaMF2* gene, which encodes a lipid transfer protein (LTP), is also considered to have an important function, where its silencing influenced pollen development in peppers [13]. For energy production and transformation, some enzymes and carrier proteins related to the mitochondrial respiratory chain are possibly related to male sterility in plants, and these have been explored by using the MS lines [14, 15]. Some genes involved in protein processing may also cause male sterility. The hub genes UbL40s (ubiquitin-60S ribosomal protein L40) and HSPs (heat shock protein) were found to be highly expressed at restrictive temperature (30 °C) and were suggested to be involved in fertility alteration in male sterile lines of rice [16].

Furthermore, the occurrence of male sterility is also related to abnormal tapetum development. The tapetum is the closest connection to the pollen mother cell (PMC) in the anther wall, providing nutrients for pollen development, which degrades later to release the microspores [17]. Therefore, the premature or delayed degradation of the tapetum can induce abnormal pollen development and could lead to male sterility [18]. Previous studies have connected specific genes to the tapetal programmed cell death (PCD) [19]. The mutant gene PTC2 (PERSISTENT TAPETAL CELL2) encodes the AT hook nucleus localization (AHL) protein in rice (*Oryza sativa*). This mutant exhibited defects in the pollen wall such as collapsed bacula and disordered tectum [20]. Yang [21] used the CRISPR/Cas9 system to obtain the male sterile mutant *osms1*. *OsMS1* encodes a plant homeodomain (PHD) finger protein, which acts as a transcriptional activator and regulates chorionic PCD in rice by interacting with tapered regulatory factors. Some transcription factors are also involved in the PCD of the tapetum and could play a vital role in the development of anthers [22, 23]. In Arabidopsis, the gene cascade DYT1-TDF1-AMS-MYB103-MS1 has been proposed to be involved in the development and timely degradation of tapetum and play an important role in regulating the complex changes associated with male fertility [24].

In previous studies, we isolated a male sterile alfalfa hybrid and successfully obtained a genetically stable alfalfa male sterile line through backcrossing, henceforth named MS-4 [25, 26]. Our previous research revealed that the pollen abortion of MS-4 mainly occurred during the mononuclear pollen formation period (flower bud length was 2 ~ 3 mm), resulting from the tapetal premature disintegration and creating shriveled and distorted mature pollen grains [27]. Since the mechanism of male sterility in the MS-4 line is not yet fully understood, we have not yet initiated commercial hybridization, as is common in other countries. Thus, there is an urgent need to conduct in-depth research to elucidate the mechanism of male sterility in the MS-4 line. This study used RNA-seq technology to perform comparative transcriptome analysis using alfalfa male sterile and male fertile lines. Our approach included the use of GO, KEGG pathways, PPI network analysis, and transcription factor family analysis to identify the gene networks and metabolic pathways that are related to male sterility in alfalfa. The results from this study will hopefully contribute to elucidating the mechanism of male sterility in alfalfa.

Results

An overview of the transcriptome data

A total of 18 libraries (2 genotypes × 3 developmental stages × 3 biological duplications) were sequenced, and a total of 293.01 Gb of clean data was obtained. A sum of 1,978,058,230 original reads was generated where each sample reached at least 14.43 Gb. The basic percentage values observed were: Q20 > 96.57%, Q30 > 90.37%, and the GC content > 42.2% (Supplementary Table 1). After filtering low-quality reads, a total of 1,956,318,172 clean reads was finally obtained. Clean reads from each sample were compared to the designated reference genome, and the comparison rate ranged from 86.14% to 90.38%. A total of 145,544 expressed genes was detected in this analysis, including 127,953 known genes and 17,591 new genes.

Identification of differentially expressed genes

At three different stages of another development, a total of 11,812 genes was found to be differentially expressed between the flower buds of MS and MF plants. Among these differentially expressed genes (DEGs), 10,672 (52.72%) were down-regulated, and 9,572 (47.28%) were up-regulated (Fig. 1A). The number of DEGs at the first developmental stage was the largest (8,273), of which 4,181 were up-regulated, and 4,092 were down-regulated. At the second stage, there were more down-regulated genes than the up-regulated ones where 2,803 (46.82%) were up-regulated and 3,184 (53.18%) were down-regulated. At the third stage, the proportion of the down-regulated genes was higher (56.75%) and the number (3,396) was 31.22% higher than the number of up-regulated genes (2,588). Up to 2,834 DEGs were shared among the three developmental stages, while there were 3,177 most specifically expressed DEGs at the first stage (Fig. 1B). In summary, the 2,834 DEGs shared across the three

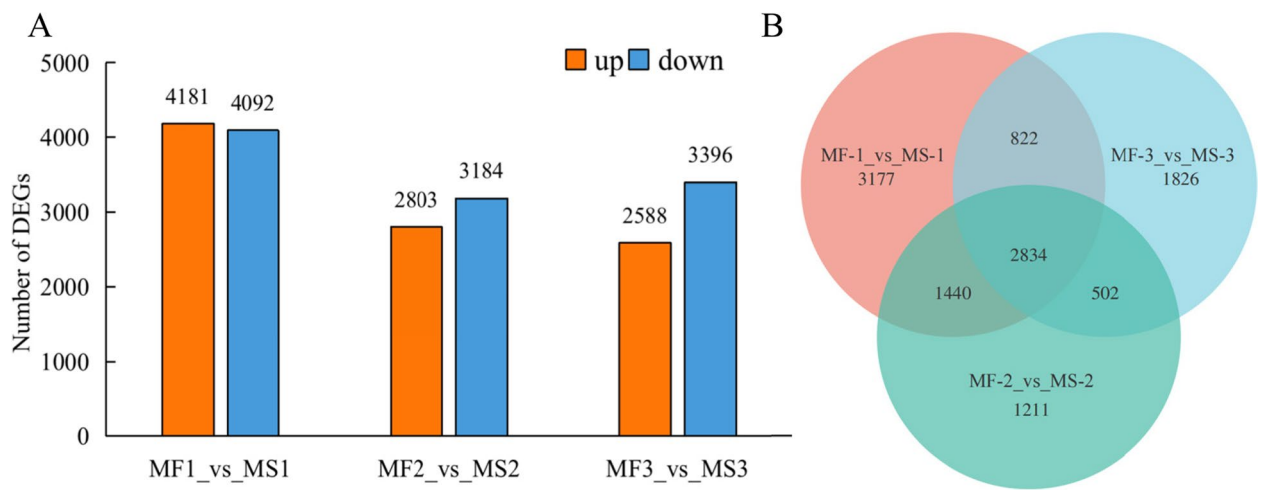

Fig. 1 Statistics of DEGs. **A** The number of up- and down-regulated DEGs in three developmental stages. **B** Veen diagram analyses of differentially and stage-specific expression genes at three developmental stages

developmental stages between MS-4 and MF lines may play a key role in another development.

Hierarchical clustering analysis

Based on the trend of gene expression under the same conditions, hierarchical clustering of 2,834 DEGs shared across the three developmental stages was conducted (Fig. 2A). The hierarchical clustering of gene expression profiles of MS-4 and MF lines at different developmental stages revealed that the gene expression patterns of these samples were different. The DEGs were subsequently divided into six clusters (Fig. 2B). Among them, cluster 1 and cluster 3 exhibited higher transcription levels in the male sterile line MS-4, with a total of 1,220 up-regulated DEGs. Cluster 2 and cluster 4 showed the opposite trend with a total of 1,608 DEGs suppressed in the male sterile line, MS-4. The number of DEGs showing down-regulated expression was more than those with up-regulated expression in the MS-4.

Gene ontology and KEGG pathway analyses of DEGs

To understand the functional annotation of male fertility changes, GO annotation and enrichment analyses was performed on all DEGs at the three developmental stages (Fig. 3). The highly enriched terms of "Molecular Function" (MF) at the three stages of anther were binding (GO: 0,005,488) and catalytic activity (GO: 0,003,824). In the "cell component" (CC), these DEGs mainly participated in GO: 0,044,464 (cell part), GO: 0,044,425 (membrane part), GO: 0,044,422 (organelle part) and GO: 0,032,991 (protein-containing complex). In addition, "biological process" (BP) in terms of the cellular process (GO: 0,009,987), metabolic process (GO: 0,008,152), biological regulation (GO: 0,065,007), localization (GO: 0,051,179), response to stimulus (GO: 0,050,896) and cellular component organization or biogenesis (GO: 0,071,840) were highly enriched at all stages. The 2,834 DEGs shared at the three stages enriched the cellular process, cell part and catalytic activity (Supplementary Fig. 1).

In order to determine the main biochemical and signal transduction pathways involved in DEGs, an enrichment analysis of the KEGG pathway was performed (Supplementary Fig. 2). A total of 2,834 DEGs was significantly enriched in ribosomal protein, fatty acid biosynthesis and degradation, pyrimidine and purine metabolism, cellular senescence, necroptosis, lipid and amino acid metabolism, and other pathways (Fig. 4). These annotations provided valuable information to determine the specific structure, function, process and pathway of the mechanism of pollen development in alfalfa.

Protein–protein interaction network and topological analysis

Protein–protein interaction (PPI) network and topological attribute analysis can identify important network relationships between proteins from complex biological data. To further understand the functional categories of 2,834 DEGs, a PPI network analysis was conducted for the top 500 ones with a comprehensive score of protein interactions. After removing the effect of free protein pairs (interacting proteins < 3), the PPI networks with 179 nodes and 500 edges were obtained, which mainly formed three networks with 84, 12 and 6 DEGs (Fig. 5, Supplementary Table 2). Most of the genes were assembled into Group 1, mainly with ribosomal protein family members, triose-phosphate transporter family protein

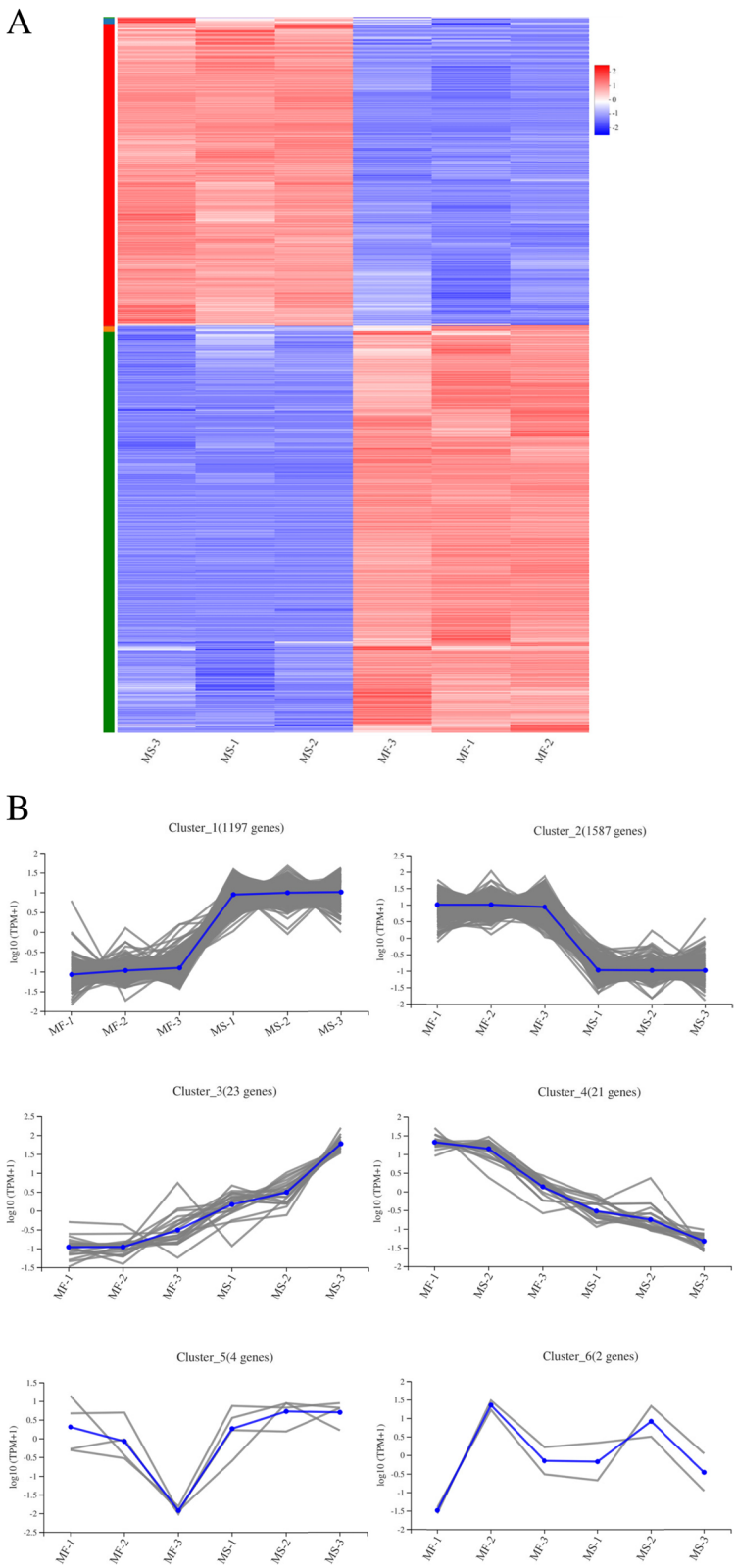

Fig. 2 Clustering analysis of the 2,834 DEGs between the MF line and the MS-4 line. **A** Heat map diagram of the 2,834 DEGs. **B** Expression patterns of the 2,834 DEGs in the six clusters. The number of genes in each cluster was shown in parentheses

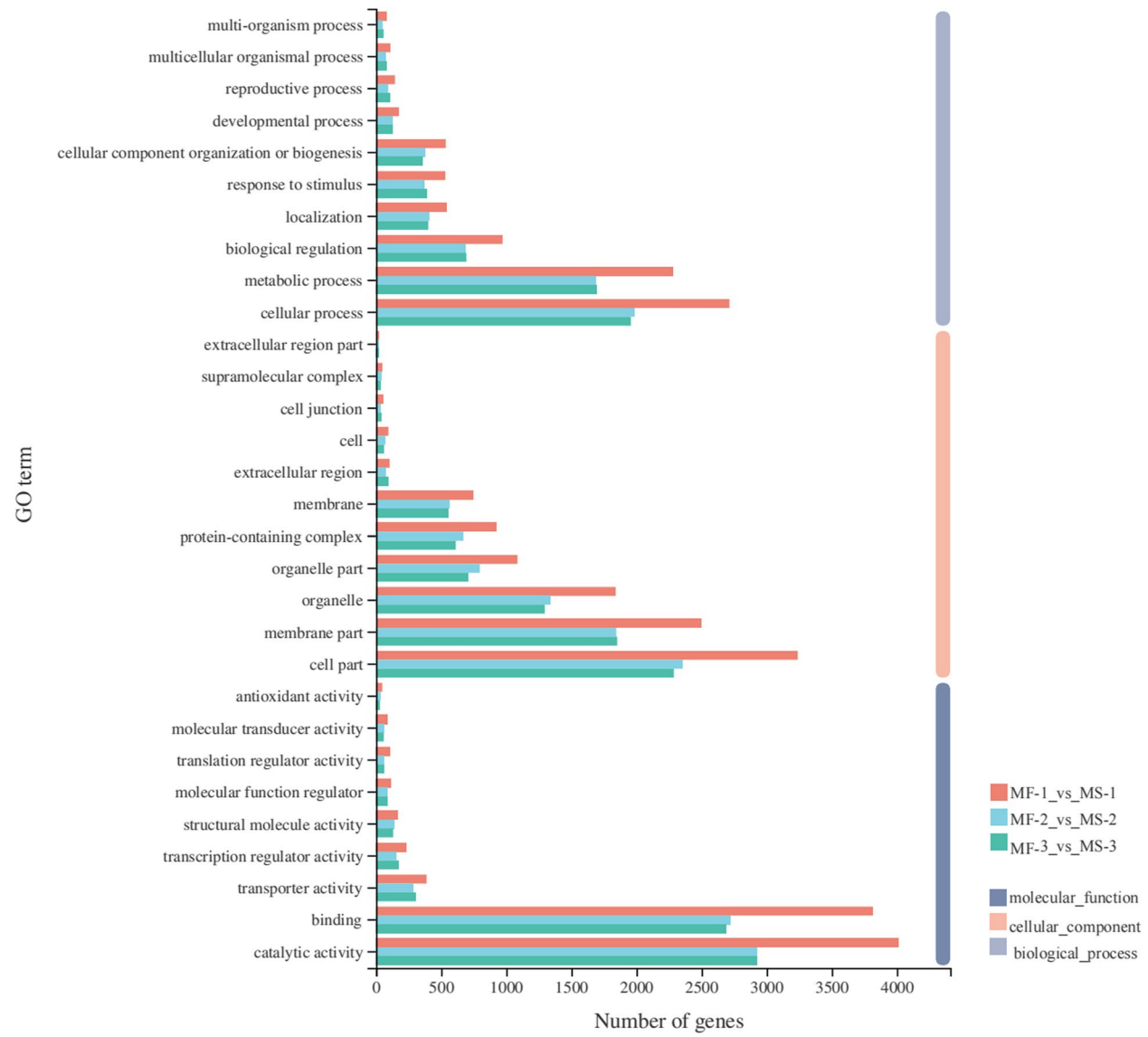

Fig. 3 GO functional enrichment analysis of DEGs at the three stages of bud development

(*MS.gene29581*) and aminoacyl- tRNA synthetase family (*MS.gene006544*, *MS.gene41385*) and others. These genes are involved in processes such as transport, translation and lipid and amino acid metabolism. Among them, ribosomal protein family members mainly included uS family (*MS.gene25178*, *MS.gene39902*, *MS.gene20037*, *MS.gene062731*), eS family (*MS.gene69921*, *MS.gene26343*, *MS.gene95803*, *MS.gene051746*), uL family (*MS.gene08434*, *MS.gene016301*, *MS.gene91152*). The genes in group 2 participate in DNA replication, including four DNA replication licensing factors that belong to the MCM family (*MS.gene36485*, *MS.gene23323*, *MS.gene07048*, *MS.gene061260*). In addition, the genes in group 3 are involved in the catalytic regulation process, including two peptidase T1B family members (*MS.gene040231* and *MS.gene40359*).

Identification of the differentially expressed transcription factors

The data showed that some DEGs belong to the genes encoding transcription factors (TFs). To better understand the molecular mechanism of male sterility, we analyzed the differentially expressed TFs. We identified 639 TFs from all DEGs. These TFs were divided into 40 families (Supplementary Table 3). The most abundant genes belonged to the MYB (12.36%), ERF

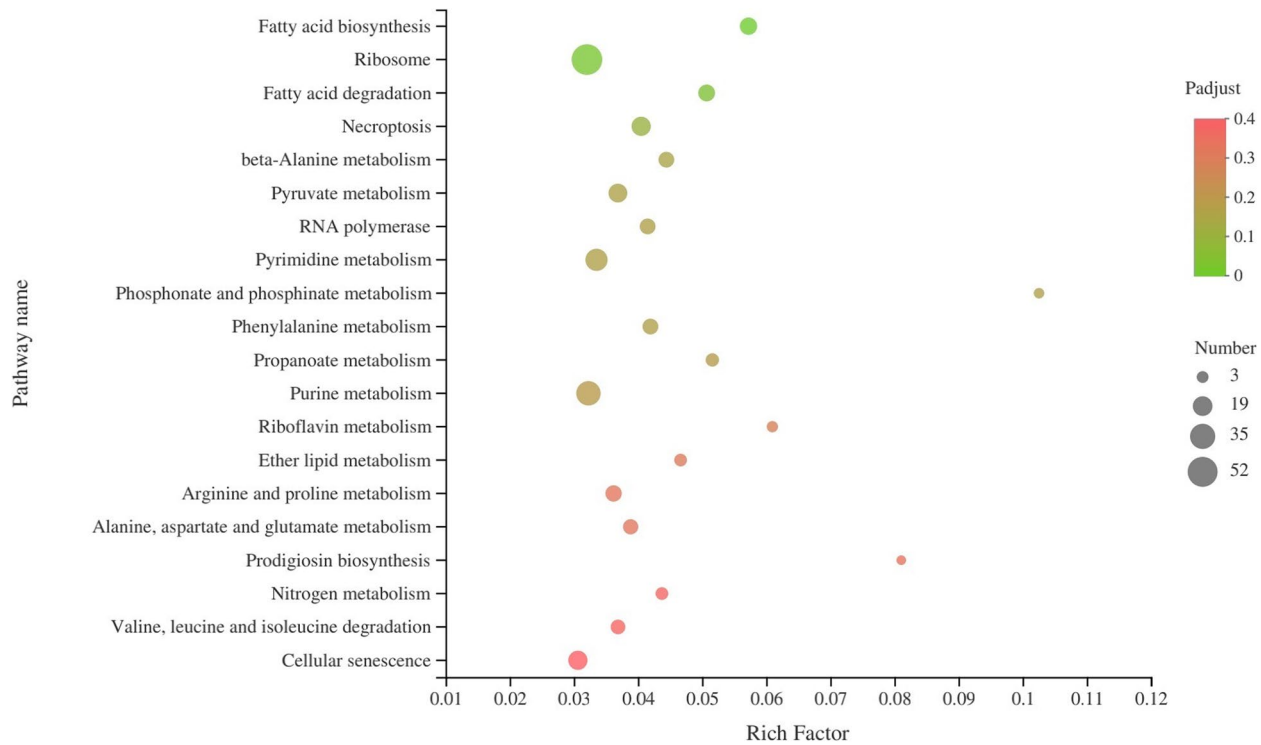

Fig. 4 KEGG pathway enrichment analysis of the 2,834 DEGs between MF and MS-4

(9.08%), WRKY (7.67%), B3 (6.89%), bHLH (6.73%), HB-other (5.79%), bZIP (5.48%) and NAC families (4.69%, Fig. 6). In addition, 132 TFs out of the 2,834 DEGs were shared in the three development stages (Supplementary Fig. 3). We speculate that these TFs form a highly complex transcription network and could play an important role in the mechanism of male sterility.

Analysis of DEGs by qRT-PCR

According to RNA-seq data, the expression of nine selected genes was verified by qRT-PCR (Fig. 7). The genes included three genes related to tapetum and pollen development (*MS.gene35590*, *MS.gene047030*, *MS.gene58163*), ABCG39 (*MS.gene007255*), sugar transport protein (*MS.gene052566*), putative carboxylesterase (*MS.gene042789*), transcription factor MYB4 isoform X2 (*MS.gene024535*), NAC domain-containing protein 2 (*MS.gene013270*), and transcription factor bHLH18 (*MS.gene033528*). These genes exhibited similar expression trends to those obtained from the RNA-seq at the three stages of another development, confirming the reliability of RNA-seq analysis.

Discussion

Genes involved in tapetum and pollen development are related to male sterility

During male gametogenesis in flowering plants, the tapetum in the anther provides nutrients for the development and maturation of microspores and pollen walls [28]. Defective development of the tapetum was found to be responsible for male sterility in watermelon [29]. In this study, we identified 17 genes that were related to tapetum and pollen development (Supplementary Table 4). These genes were differentially expressed between the fertile and sterile lines. These differentially-expressed genes may have contributed to male sterility. The timely programmed cell death (PCD) of the tapetum is crucial where early or delayed start of PCD can cause male sterility [30, 31]. For example, SaNa-1A, a novel cytoplasmic male sterility (CMS) line of *Brassica napus*, was used to validate that abnormal tapetum degradation and anther development were the primary causes for anther abortion [32]. Previously, we observed that the tapetum layer of the sterile line MS-4 was prematurely disintegrated, and in this study, we identified six genes that are related to tapetum PCD. Furthermore, in the male sterility group, a gene labeled as male sterility (*MS.gene78221*)

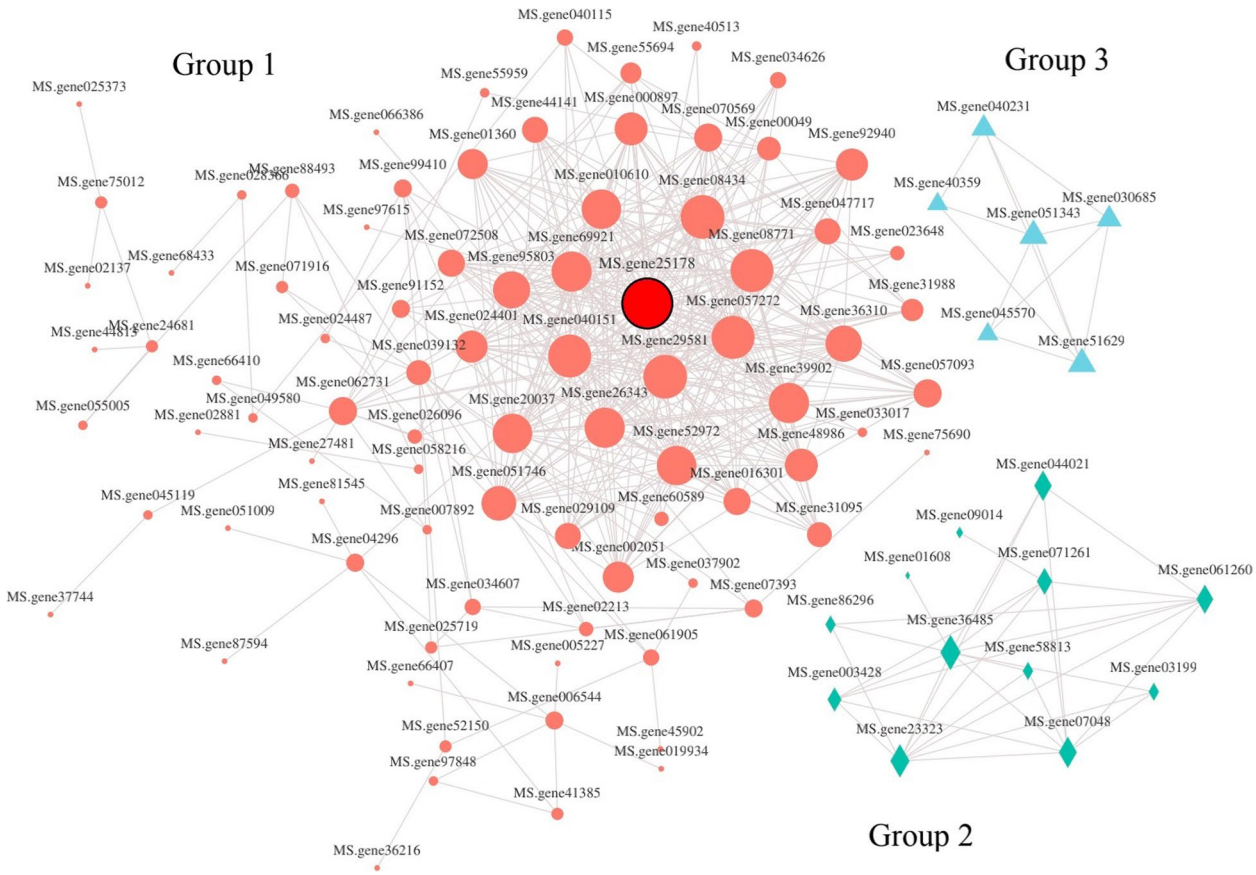

Fig. 5 Co-expression network analysis of the partially predicted PPIs. Different colors and shapes represent different groups; the size of the nodes represents the importance in the network, the larger the node, the greater the importance of the gene in the network; the red-framed circle represents the hub gene

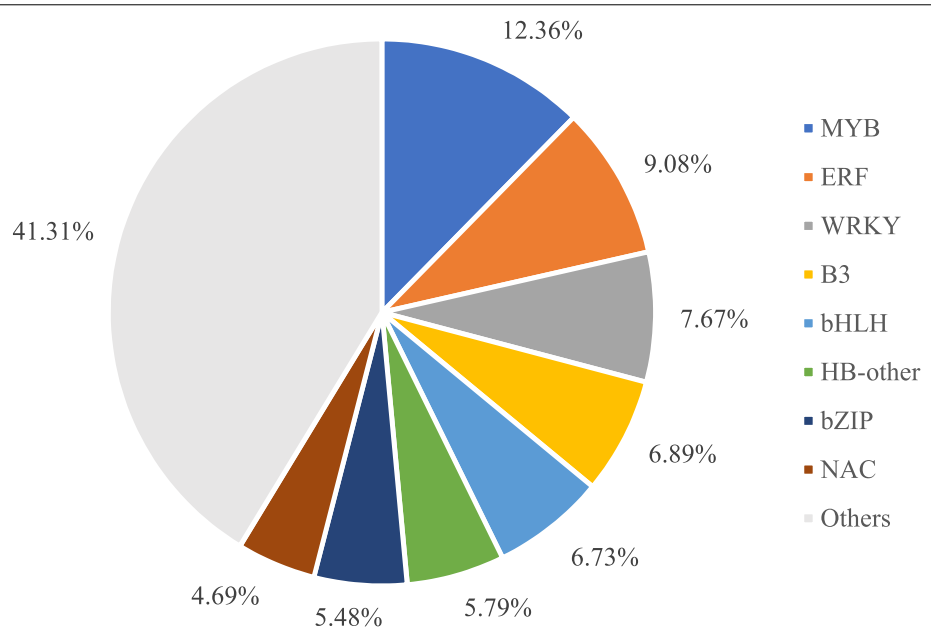

Fig. 6 Distribution of differentially expressed transcription factors

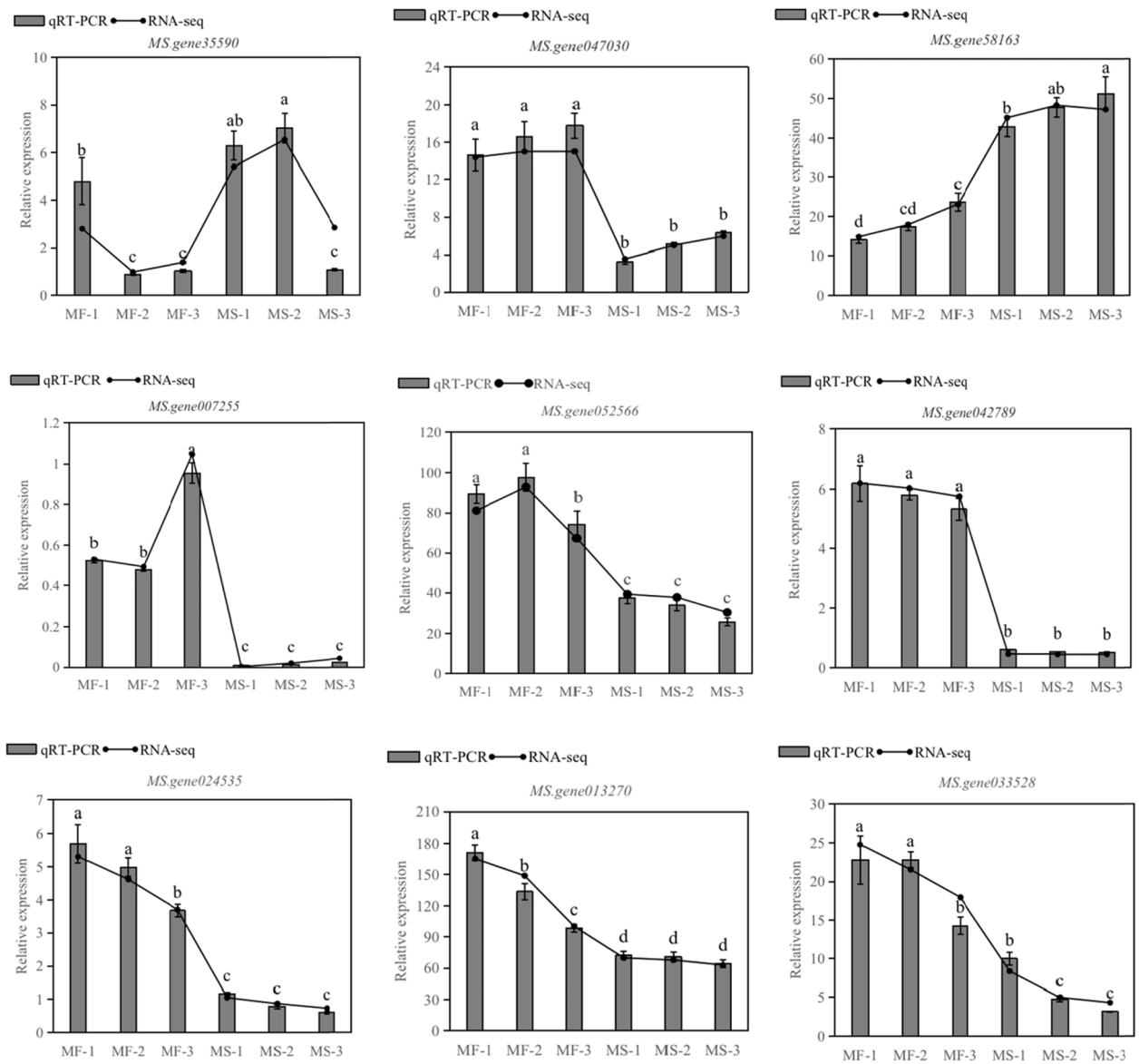

Fig. 7 Verification of the expression of the selected DEGs was carried out via qRT-PCR. Error bars indicate the standard deviation of three biological replicates

was down-regulated (Supplementary Fig. 4). These genes can be used as critical genes to test the effects of tapetum and pollen development on male sterility in alfalfa.

Genes involved in carbohydrate and lipid metabolism are related to male sterility

Functional fertile pollen grains have a complete and complex outer wall structure [33], resulting from the joint action of multiple metabolic pathways and functional genes [34]. However, the mature pollen grains of MS-4 were shriveled and deformed. To explore the potential causes of these features, we performed enrichment analysis of, and annotated the DEGs at the three stages of bud development. The GO enrichment results of alfalfa anther transcriptome showed that DEGs related to biological processes were significantly enriched in "cellular processes", "metabolic processes" and "biological regulation". The KEGG pathway analysis of DEGs revealed that genes of "fatty acid biosynthesis", "lipid metabolism", "amino acid metabolism", "pyrimidine and purine metabolism", and "cellular senescence", were among those significantly enriched. These genes are involved in energy

metabolism, lipid and amino acid metabolism, carbohydrate metabolism, and some cell synthesis and aging processes. Carbon pathways provide energy and carbohydrates for plant growth and development [35], while glutamine plays a central role in pollen development and amino acid metabolism [36]. Fatty acids (FA) regulate the homeostasis of reactive oxygen species (ROS), and dysfunction of *FAX1* can destroy oxygen free radicals, which could contribute to male sterility [34]. Changes in the expression level of genes in these pathways may influence fertility. In our study, one carbohydrate metabolism pathway DEG (*MS.gene052566*) and one lipid metabolism pathway DEG (*MS.gene042789*) were selected for RT-PCR analysis to compare their expression levels in the male sterile and fertile (Fig. 7). The data showed that the expression levels of the three genes were lower in MS-4 than in the MF line. Therefore, we speculated that reducing carbohydrate and lipid metabolism affected the transmission of energy and nutrients, leading to male sterility.

ABCG transporters involved in altering the fertility

As an important component of the process of transport of nutrients and energy across the cell membranes, the ABC transporters play a key role in another development [37]. The ABCG transporters constitute the largest subfamily of ABC transporters [38], which have been reported to be involved in the development of anther or pollen (pollen wall). Zhou et al. [39]. speculated that β-ketoacyl-[acyl carrier protein] synthase I (*SiKASI*) was crucial for fatty acid metabolism and might interact with *ABCG18* causing abnormal pollen development in *Arabidopsis*. The rice male sterile mutant *osabcg26* exhibited defects in tapetal cells, lipid Ubisch bodies and pollen exine, possibly due to a disturbance in the homeostasis of anther lipid metabolism and transport [40]. Moreover, other plant ABCG transporters such as *AtABCG1*, *AtABCG16* [41], *AtABCG9*, *AtABCG31* [42], *OsABCG3* [43], and *OsABCG15* have been reported to be involved in the development of pollen or anther development in *Arabidopsis* and *Oryza sativa* [44]. In this study, according to RT-PCR results, a gene labeled as ABC transporter G family member 39 (ABCG39, *MS.gene007255*) was down-regulated in the male sterile group. This suggests that ABCG39 (*MS.gene007255*) may participate in the regulation of anther development to ultimately cause male sterility. However, this hypothesis needs further experiments to be verified.

Genes related to ribosomal proteins are involved in male sterility

Cytoplasmic male sterility (CMS) in plants is caused by genetic conflicts or communication barriers between nuclear and mitochondrial genes [45]. The genes of CMS are typically new chimeric regions genes, and most of these genes are known to encode ribosome and mitochondrial electron transport chain complexes [46]. Additionally, ribosomal protein content is an indicator of mitochondrial function, where some studies have shown that ribosomal proteins were closely related to fertility changes. Lower content of ribosomal protein during anther abortion was observed in cotton, suggesting that insufficient synthesis of ribosomal protein may have hindered pollen production [47]. However, Zhao et al. [48] found that the 60S ribosomal protein L13a-4-like was up-regulated in the male sterile line and hence was related to the abortion of microspores in cotton. These results indicate that the up- or down-regulation of some key ribosomal proteins may lead to male sterility. Sunok [49] and Ding et al. [50] speculated that ribosomal proteins may participate in tapetum PCD and could play an important regulatory role in tapetum development. The results from the PPI network analysis of this study indicate interactions between ribosomal protein family members, triose-phosphate transporter protein family, and aminoacyl-tRNA synthetase family. Topological analysis (Fig. 5) indicated that the 40S ribosomal protein S2 (*MS.gene25178*) was the core gene in the network and was up-regulated in the male sterile line (MS-4). The up-regulated expression of this gene in the present study (Supplementary Fig. 5), consistently with that of Zhao's on cotton, suggests a critical role for 40S ribosomal proteins in regulating male sterility in alfalfa. These results also indicate that ribosomal protein genes could also induce fertility changes, and future research should focus on in-depth analysis of these genes.

The TF genes are related to male sterility

Many studies on male sterility have shown that TFs play an important role in plant growth and anther development [51, 52]. The helix-loop-helix transcription factor (bHLH) has been reported to regulate the development of the tapetum [53, 54]. *OsbHLH138* was identified as a key gene that regulates thermo-sensitive male sterility in rice [55]. Moreover, *MYB80* (formerly *MYB103*) can delay PCD by activating transcription of the UNDEAD gene that results in male sterility [56]. These results illustrate the importance of bHLH transcription factors. Han et al. [57] obtained several male sterile lines in which the MYB transcription factor *OsMS188* was knocked out in rice, in which *OsMS188* was verified as the key regulator of tapetum development and sporopollenin synthesis. Furthermore, Cheng et al. [28] suggested that NAC transcription factor was a central gene related to another development and may cause male sterility in cotton. In this study, we found that 639 genes belonged to 40 families of TFs in all DEGs, and that some TFs belonging to the same family

were differentially up- or down-regulated. These results are consistent with recent studies on bud dormancy in the grapevine [58], suggesting that members of the same TF family may play different roles in the development of pollen in alfalfa, forming a more complex transcriptional regulatory network. We selected two DEGs annotated as transcription factors (transcription factor MYB4, *MS.gene024535*; and transcription factor bHLH18, *MS.gene033528*) for RT-PCR analysis (Fig. 7). The data showed that the expression of the two genes in MS-4 was reduced (relative to the MF line). These results indicate that TFs play an essential role in regulating male sterility.

Conclusions

Transcriptomic analysis of anthers at three different developmental stages in the male sterile (MS-4) of alfalfa as compared to the male fertile (MF) line revealed that the underlying mechanism of male sterility is complex. Male sterility in alfalfa is a network process involving many biological processes and pathways, including catalytic activity, cell component, cellular processes, ribosomal proteins, fatty acid biosynthesis, and metabolic function. The findings from this study provide several insights to elucidate the mechanism of male sterility based on changes observed from anther development in the male sterile line, MS-4 of alfalfa. In addition, this study also indicated that the 40S ribosomal protein S2 (*MS.gene25178*) plays an important role in inducing male sterility. Further research is needed on the key metabolic pathways and candidate genes to define the mechanism of male sterility in alfalfa.

Material and methods

Plant materials

We isolated a natural sterile mutant of *Medicago varia Martin*. cv. 'Caoyuan No. 1', and we obtained a stable male sterile line MS-4 from it after a few years of cultivation. The tapetum layer of the sterile line MS-4 was prematurely disintegrated, and the mature pollen grains were shriveled and deformed. MS-4 plants were grown under normal conditions in the fields on the Inner Mongolia Agricultural University campus in China. In our previous research, we conducted cytological studies on MS-4 and the male fertile (MF) plant and found that the pollen abortion observed with MS-4 mainly occurred during mononuclear pollen formation [59]. Therefore, flower buds from MS-4 and MF plants were collected at the three stages of anther development: pollen mother cell formation stage (bud length 0.8 ~ 2 mm), mononuclear pollen formation stage (bud length 2 ~ 3 mm), and pollen maturity stage (bud length 4 ~ 5 mm). Flower buds of the same plant at the same developmental stage were peeled off by removing the calyx and were flash frozen before being stored (-80 °C) for subsequent transcriptome sequencing. Each sample at each developmental stage included three biological replicates, which were sequenced separately.

RNA extraction, sequencing, and RNA-seq analysis

Trizol (TaKaRa, Dalian, China) was used to extract total RNA from alfalfa anther tissues, and DNase I (Takara, Japan) was used to remove genomic DNA. The RNA library was established using TruSeq™ RNA sample preparation Kit (Illumina, San Diego, CA). SuperScript double-stranded cDNA synthesis kit (Invitrogen, CA) was used with six-base random primers (Illumina) to synthesize one-strand cDNA using mRNA as template, and then to perform two-strand synthesis to form a stable double-stranded structure. After the cDNA was enriched by PCR, a 200–300 bp band was recovered using 2% agarose gel. After being quantified with TBS380 (Picogreen), the library was sequenced with an Illumina HiSeq Xten/NovaSeq 6000 sequencing platform for high-throughput sequencing, with a sequencing read length of PE 150. Complete transcriptome analysis of 18 libraries was conducted. Reference gene source used was from:

https://figshare.com/projects/whole_genome_sequencing_and_assembly_of_Medicago_sativa/66380 [60].

Identification and bioinformatic analysis of DEGs

DESeq2 software based on negative binomial distribution was used to analyze raw counts, and p-adjust < 0.05 and $|\log2^{FC}| \geq 1$ (the expression difference was more than twice) were used as the criteria for selecting significant differential expression of genes. Bioinformatic analyses of DEGs were carried out on the Meiji Bio-Cloud Platform (https://www.i-sanger.com/). GO (http://www.geneontology.org/) used Fisher method for exact test, and BH was the multiple test correction method. When the corrected p value ($p$adjust) < 0.05 was applied, it was considered that genes of GO function and KEGG pathways (http://www.genome.jp/kegg/) [61, 62] were significantly enriched [63]. The DEGs of protein–protein interaction (PPI) network were predicted by STRING [64] (https://string-db.org/), and the drawn network diagram and network node degree were analyzed by Cytoscape 3.5.1. Transcription factors (TFs) were identified using PlantTFDB 4.0 (http://planttfdb.cbi.pku.edu.cn/).

qRT-PCR verification

UEIris II First-Strand cDNA Synthesis Kit (Code NO. R2028 US EVERBRIGHT INC, Suzhou, China) was used to reverse transcribe RNA. Fast Super EvaGreen qPCR Master Mix kit (Code NO. S2008 US EVERBRIGHT INC, Suzhou, China) was used, which contained a 20 μL solution (50 ng cDNA as a template;

10 × ROX reference dye), for qRT-PCR in a LightCycler 480 (Roche). This was conducted on each sample. The amplification protocol was as follows: 95 °C for 2 min, then 40 cycles at 95 °C for 10 s, 56 °C for 1 min and 72 °C for 10 s, and then melt curve at 95 °C 15 s, 60 °C 1 min, and 95 °C 15 s. Supplementary Table 5 lists the sequence of the specific primers used for qRT-PCR. β-Actin was used as an internal control. All qRT-PCR analyses were performed in three replicates. The relative gene expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method [65]. The relative gene expression levels were analyzed via analysis of variance (ANOVA, $p < 0.05$).

Abbreviations

MS-4: Male sterile line; DEGs: Differentially expressed genes; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: Protein–protein interaction; LTP: Lipid transfer protein; PMC: Pollen mother cell; PCD: Programmed cell death; PTC2: Persistent tapetal cell2; AHL: AT hook nucleus localization; PHD: Plant homeodomain; TFs: Transcription factors; BP: Biological process; MF: Molecular function; CC: Cell component; CMS: Cytoplasmic male sterility; FA: Fatty acids; ROS: Reactive oxygen species; bHLH: Helix-loop-helix transcription factor; UbL40: Ubiquitin-60S ribosomal protein L40; HSP: Heat shock protein.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-022-03581-1.

Additional file1: Table S1. An overview of the transcriptome data.

Additional file2: Fig. S1. GO functional enrichment analysis of the 2,834 DEGs.

Additional file3: Fig. S2. Pathway classification statistics histogram of all DEGs.

Additional file4: Table S2. PPI network information.

Additional file5: Table S3. Statistics of transcription factor families.

Additional file6: Fig. S3. Transcription factors distribution among 2,834 DEGs.

Additional file7: Table S4. 17 genes related to tapetum and pollen development.

Additional file8: Fig. S4. Relative expression of MS.gene78221 via qRT-PCR.

Additional file9: Fig. S5. Relative expression of MS.gene25178 via qRT-PCR.

Additional file10: Table S5. Primer information.

Acknowledgements

Thanks to the Key Laboratory of Grassland Resources, Ministry of Education of P.R. of China for providing facilities. The authors would like to express their gratitude to EditSprings (https://www.editsprings.com/) for their expert linguistic services.

Authors' contributions

FL S designed the experiment, B X, RN W and CP G performed the experiments, B X and J W analyzed the data, B X wrote the manuscript. All authors reviewed the manuscript. The author(s) read and approved the final manuscript.

Funding

This study was supported by the financial project of the Inner Mongolia Autonomous Region (RZ1900008322) and the Science and technology project of the Inner Mongolia Autonomous Region (2019GG244).

Availability of data and materials

Datasets are available at NCBI project PRJNA751239. The transcriptome datasets are available in the NCBI Short Read Archive (SRA) database with the accession number: SAMN20512309, SAMN20512310, SAMN20512311, SAMN20512312, SAMN20512313, SAMN20512314, SAMN20512315, SAMN20512316, SAMN20512317, SAMN20512318, SAMN20512319, SAMN20512320, SAMN20512321, SAMN20512322, SAMN20512323, SAMN20512324, SAMN20512325, SAMN20512326.

Declarations

Ethics approval and consent to participate

The collection of plant seeds conform China's regulatory standards and current laws.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests. Table S1. An overview of the transcriptome data. Fig. S1. GO functional enrichment analysis of the 2,834 DEGs. Fig. S2. Pathway classification statistics histogram of all DEGs. Table S2. PPI network information. Table S3. Statistics of transcription factor families. Fig. S3. Transcription factors distribution among 2,834 DEGs. Table S4. 17 genes related to tapetum and pollen development. Fig. S4. Relative expression of MS.gene78221 via qRT-PCR. Fig. S5. Relative expression of MS.gene25178 via qRT-PCR.

Received: 30 September 2021 Accepted: 31 March 2022
Published online: 15 April 2022

References

1. Om PM, Major SD, Salesh KJ. Heterosis breeding in chilli pepper by using cytoplasmic male sterile lines for high-yield production with special reference to seed and bioactive compound content under temperature stress regimes. Sci Hortic. 2020;262:1–18.
2. Wang ZG, Ma BL, Yu XF, Gao JL, Sun JY, Su ZJ, et al. Physiological basis of heterosis for nitrogen use efficiency of maize. Sci Rep. 2019;9(1):18708.
3. Yan XH, Dong CH, Yu JY, Liu WH, Jiang CH, Liu J, et al. Transcriptome profile analysis of young floral buds of fertile and sterile plants from the self-pollinated offspring of the hybrid between novel restorer line NR1 and Nsa CMS line in Brassica napus. BMC Genomics. 2013;14(1):26.
4. Liu YM, Wei G, Xia YY, Liu XW, Tang J, Lu YL, et al. Comparative transcriptome analysis reveals that tricarboxylic acid cycle-related genes are associated with maize CMS-C fertility restoration. BioMed Central. 2018;18(1):190.
5. Chase CD. Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet. 2007;23:81–90.
6. Guo JX, Liu YG. Molecular control of male reproductive development and pollen fertility in rice. J Integr Plant Biol. 2012;54(12):967–78.
7. Yang Y, Bao SY, Zhou XH, Liu J, Zhuang Y. The key genes and pathways related to male sterility of eggplant revealed by comparative transcriptome analysis. BMC Plant Biol. 2018;18(1):209.
8. Yuan C, Zhang SW, Hu RL, Wei DY, Tang QL, Wang YQ, et al. Comparative transcriptome analysis provides insight into the molecular mechanisms of anther dehiscence in eggplant (Solanum melongena L.). Genomics. 2021;2020(113):497–506.
9. Sangu E, Tibazarwa FI, Nyomora A, Symonds RC. Expression of genes for the biosynthesis of compatible solutes during pollen development under heat stress in tomato (Solanum lycopersicum). J Plant Physiol. 2015;178:10–6.
10. Salminen TA, Blomqvist K, Edqvist J. Lipid transfer proteins: classification, nomenclature, structure, and function. Planta. 2016;244(5):971–97.

11. Su AG, Song W, Shi Z, Zhao YX, Xing JF, Zhang RY, et al. Exploring differentially expressed genes associated with fertility instability of S-type cytoplasmic male-sterility in maize by RNA-seq. J Integr Agric. 2017;16(8):1689–99.
12. Fang X, Fu H, Gong Z, Chai W. Involvement of a universal amino acid synthesis impediment in cytoplasmic male sterility in pepper. Sci Rep. 2016;6:23357.
13. Chen CM, Chen GJ, Hao XF, Cao BH, Chen QH, Liu SQ, et al. CaMF2, an antherspecific lipid transfer protein (LTP) gene, affects pollen development in capsicum annuum L. Plant Sci. 2011;181(4):439–48.
14. Nie Z, Zhao T, Yang S, Gai J. Development of a cytoplasmic male-sterile line NJCMS4A for hybrid soybean production. Plant Breed. 2017;136(4):516–25.
15. Liu Q, Lan Y, Wen C, Zhao H, Wang J, Wang Y. Transcriptome sequencing analyses between the cytoplasmic male sterile line and its maintainer line in welsh onion (Allium fistulosum L.). Int J Mol Sci. 2016;17(7):1058.
16. Li C, Tao RF, Li Y, Duan MH, Xu JH. Transcriptome analysis of the thermosensitive genic male-sterile line provides new insights into fertility alteration in rice (Oryza sativa). Genomics. 2020;112(3):2119–29.
17. Zhang W, Sun YJ, Timofejeva LM, Chen CB, Grossniklaus U, Ma H. Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor. Development. 2006;133(16):3085–95.
18. Balk J, Leaver CJ. The PET1-CMS mitochondrial mutation in sunflower is associated with premature programmed cell death and cytochrome c release. Plant Cell. 2001;13(8):1803–18.
19. Xie HT, Wan ZY, Li S, Zhang Y. Spatiotemporal production of reactive oxygen species by NADPH oxidase is critical for tapetal programmed cell death and pollen development in Arabidopsis. Plant Cell. 2014;26(5):2007–23.
20. Uzair M, Xu DW, Schreiber L, Shi JX, Liang WQ, Jung KH, et al. PERSISTENT TAPETAL CELL2 is required for normal tapetal programmed cell death and pollen wall patterning. Plant Physiol. 2020;182(2):962–76.
21. Yang ZF, Liu L, Sun LP, Yu P, Zhang PP, Abbas A, et al. OsMS1 functions as a transcriptional activator to regulate programmed tapetum development and pollen exine formation in rice. Plant Mol Biol. 2019;99(1–2):175–91.
22. Zou T, He ZY, Qu LY, Liu MX, Zeng J, Liang YL, et al. Knockout of OsACOS12 caused male sterility in rice. Mol Breeding. 2017;37(10):126.
23. Distelfeld A, Pearce SP, Avni R, Scherer B, Uauy C, et al. Divergent functions of orthologous NAC transcription factors in wheat and rice. Plant Mol Biol. 2012;78(4–5):515–24.
24. Ariizumi T, Toriyama K. Genetic regulation of sporopollenin synthesis and pollen exine development. Annu Rev Plant Biol. 2012;62(1):437–60.
25. Wu YF, Wei L, Wu RY, Duan QX. A study on flower bud differentiation, the microspore development and the reason of male sterility in alfalafa. Grassl China. 1986;6:23–7.
26. Wu YF, Temue B, Xu ZY, Lou HN. Use of male sterile line in alfalfa hybrid seed production. Grassl China. 1998;5:9–12.
27. Gao CP, Shi FL, Li H. Cyto-morphologyical studies on the stamen development of male sterile line MS-4 of alfalfa. J Inner Mongolia Univ. 2005;36:288–93.
28. Cheng XQ, Zhang XY, Xue F, Zhu SH, Li YJ, Zhu QH, et al. Characterization and transcriptome analysis of a dominant genic male sterile cotton mutant. BMC Plant Biol. 2020;20(1):312–312.
29. Wei CH, Zhang RM, Yue Z, Yan X, Cheng DH, Li JY, et al. The impaired biosynthetic networks in defective tapetum lead to male sterility in watermelon. J Proteomics. 2021;243:104241.
30. Bai WT, Wang PR, Hong J, Kong WY, Xiao YJ, Yu XW, et al. Earlier degraded tapetum1 (EDT1) encodes an ATP-citrate lyase required for tapetum programmed cell death. Plant Physiol. 2019;181(3):1223–38.
31. Gao YG, Zhang LJ, Zhao SC, Yan YX. Comparative analysis of the male inflorescence transcriptome profiles of an ms22 mutant of maize. PloS One. 2018;13(7):e0199437.
32. Kun D, Xiao YY, Liu QE, Wu XY, Jiang JJ, Wu J, et al. Abnormal tapetum development and energy metabolism associated with sterility in SaNa-1A CMS of Brassica napus L. Plant Cell Rep. 2019;38(5):545–58.
33. Kobayashi K, Akita K, Suzuki M, Ohta D, Nagata N. Fertile Arabidopsis cyp704b1 mutant, defective in sporopollenin biosynthesis, has a normal pollen coat and lipidic organelles in the tapetum. Plant Biotechnol. 2021;38(1):109–16.
34. Zhu L, He SY, Liu YY, Shi JX, Xu J. Arabidopsis FAX1 mediated fatty acid export is required for the transcriptional regulation of anther development and pollen wall formation. Plant Mol Biol. 2020;104(1–2):187–201.
35. Kretschmer M, Croll D, Kronstad JW. Maize susceptibility to Ustilago maydis is influenced by genetic and chemical perturbation of carbohydrate allocation. Mol Plant Pathol. 2017;18:1222–37.
36. Ribarits A, Mamun ANK, Li S, Resch T, Fiers M, Heberle-Bors E, et al. Combination of reversible male sterility and doubled haploid production by targeted inactivation of cytoplasmic glutamine synthetase in developing anthers and pollen. Plant Biotechnol J. 2007;5(4):483–94.
37. Do THT, Martinoia E, Lee Y. Functions of ABC transporters in plant growth and development. Curr Opin Plant Biol. 2018;41:32–8.
38. Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, et al. Plant ABC proteins - a unified nomenclature and updated inventory[J]. Trends Plant Sci. 2008;13(4):151–9.
39. Zhou T, Yang YX, Li TY, Liu HY, Zhou F, Zhao YZ. Sesame β-Ketoacyl-ACP synthase I regulates pollen development by interacting with an ABC transporter in transgenic Arabidopsis. Physiol Plant. 2021;173(3):1048–62.
40. Chang ZY, Chen ZF, Yan W, Xie G, Lu JW, Wang N, et al. An ABC transporter, OsABCG26, is required for anther cuticle and pollen exine formation and pollen-pistil interactions in rice. Plant Sci. 2016;253:21–30.
41. Yadav V, Molina I, Ranathunge K, Castillo IQ, Rothstein SJ, Reed JW. ABCG transporters are required for suberin and pollen wall extracellular barriers in Arabidopsis. Plant Cell. 2014;26:3569–88.
42. Choi H, Ohyama K, Kim YY, Jin JY, Lee SB, Yamaoka Y, et al. The role of Arabidopsis ABCG9 and ABCG31 ATP binding cassette transporters in pollen fitness and the deposition of steryl glycosides on the pollen coat. Plant Cell. 2014;26:310–24.
43. Luo T, Zou T, Yuan GQ, He ZY, Li WJ, Tao Y, et al. Less and shrunken pollen 1(LSP1) encodes a member of the ABC transporter family required for pollen wall development in rice(Oryza sativa L.). The Crop J. 2020;8(3):492–504.
44. Qin P, Tu B, Wang Y, Deng L, Quilichini TD, Li T, et al. ABCG15 encodes an ABC transporter protein, and is essential for post-meiotic anther and pollen exine development in rice. Plant Cell Physiol. 2013;54:138–54.
45. Ferreira JJ, Ramos A, Chambergo F, Stambuk B, Muschellack L, Schumacher R, et al. Functional expression of the maize mitochondrial URF13 down-regulates galactose-induced GAL1 gene expression in Saccharomyces cerevisiae. Biochem Biophys Res Commun. 2006;1:30–6.
46. Li M, Chen L, Tang D, Liao X, Kong X, Li B, et al. Discovery of four novel ORFs responsible for cytoplasmic male sterility (CMS) in cotton (Gossypium hirsutum L.) through comparative analysis of the mitochondrial genomes of four isoplasmic lines. Agronomy. 2020;10:765.
47. Li M, Chen L, Khan A, Kong XJ, Khan MR, Rao MJ, et al. Transcriptome and MiRNAomics analyses identify genes associated with cytoplasmic male sterility in cotton (Gossypium hirsutum L.). Int J Mol Sci. 2021;22(9):4684.
48. Zhao HY, Wang JS, Qu YF, Peng RH, Magwanga RO, Liu F, et al. Transcriptomic and proteomic analyses of a new cytoplasmic male sterile line with a wild Gossypium bickii genetic background[J]. BMC Genomics. 2020;21(1):1–17.
49. Moon S, Hong WJ, Kim YJ, NChandran AK, Gho YS, Yoo YH, et al. Comparative transcriptome analysis reveals gene regulatory mechanism of UDT1 on another development. J Plant Biol. 2020;63:289–96.
50. Ding ZJ, Wu XH, Wang T. The rice tapetum-specifc gene RA39 encodes a type I ribosome-inactivating protein. Sex Plant Reprod. 2002;15:205–12.
51. Zhang W, Sun YJ, Timofejeva L, Chen CB, Grossniklaus U, Ma H. Regulation of Arabidopsis tapetum development and function by Dysfynctional Tapetum1 (DYT1) encoding a putative bHLH transcription factor. Development. 2006;133:3085–95.
52. Gu JN, Zhu J, Yu Y, Teng XD, Lou Y, Xu XF, et al. DYT1 directly regulates the expression of TDF1 for tapetum development and pollen wall formation in Arabidopsis. Plant J. 2014;80(6):1005–13.
53. Jung KH, Han MJ, Lee YS, Kim YW, Hwang I, Kim MJ, et al. Rice undeveloped tapetum1 is a major regulator of early tapetum development. Plant Cell. 2005;17:2705–22.
54. Zahn LM, Ma X, Altman NS, Zhang Q, Wall PK, Tian DL, et al. Comparative transcriptomics among floral organs of the basal eudicot eschscholzia Californica as reference for floral evolutionary developmental studies. Genome Biol. 2010;11:R101.

55. Wen JY, Wang LT, Wang J, Zeng YF, Xu YH, Li SQ. The transcription factor OsbHLH138 regulates thermosensitive genic male sterility in rice via activation of TMS5. Theor Appl Genet. 2019;132(6):1721–32.
56. Huy AP, Sylvana I, Song FL, Roger WP. The MYB80 transcription factor is required for pollen development and the regulation of tapetal programmed cell death in Arabidopsis thaliana. Plant Cell. 2011;23(6):2209–24.
57. Han Y, Zhou SD, Fan JJ, Zhou L, Shi QS, Zhang YF, et al. OsMS188 is a key regulator of tapetum development and sporopollenin synthesis in rice. Rice. 2021;14(1):4.
58. Min Z, Zhao X, Li R, Yang B, Liu M, Fang Y. Comparative transcriptome analysis provides insight into differentially expressed genes related to bud dormancy in grapevine (Vitis vinifera). Sci Horti. 2018;225:213–20.
59. Gao CP. Studies on crossing improvement of male sterile lines Ms-4 of alfalfa. Hohhot: Inner Mongolia University; 2007.
60. Chen HT, Zeng Y, Yang YZ, Huang LL, Tang BL, Zhang H, Hao F, Liu W, Li YH, Liu YB, Zhang XS, Zhang R, Zhang YS, Li YX, Wang K, He H, Wang ZK, Fan GY, Yang H, Bao AK, Shang ZH, Chen JH, Wang W, Qiu Q. Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. Nat Commun. 2020;11:2494.
61. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000;28:27–30.
62. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. Protein Sci. 2019;28:1947–51.
63. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. Nucleic Acids Res. 2021;49:545–51.
64. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45:362–8.
65. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using realtime quantitative PCR and the 2(T)(-Delta Delta C) method. Methods. 2001;2001(25):402–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

BMC