



Article

# Using Transcriptome Analysis to Screen for Key Genes and Pathways Related to Cytoplasmic Male Sterility in Cotton (*Gossypium hirsutum* L.)

Yuqing Li <sup>1,2</sup>, Tengfei Qin <sup>2</sup>, Chunyan Wei <sup>2</sup>, Jialiang Sun <sup>2</sup>, Tao Dong <sup>2</sup>, Ruiyang Zhou <sup>3</sup>,  
Quanxia Chen <sup>1,\*</sup> and Qinglian Wang <sup>2,\*</sup>

<sup>1</sup> College of Agriculture, Xinjiang Agricultural University, Urumqi 830000, China; lyq120327@163.com

<sup>2</sup> School of Life Science and Technology, Henan Institute of Science and Technology/Collaborative Innovation Center of Modern Biological Breeding of Henan Province/Henan Key Laboratory Molecular Ecology and Germplasm Innovation of Cotton and Wheat, Xinxiang 453003, China; qintengfeisam@163.com (T.Q.); weichunyan688@163.com (C.W.); sjl5845@126.com (J.S.); dongtao985@163.com (T.D.)

<sup>3</sup> Key Laboratory of Plant Genetics and Breeding, College of Agriculture, Guangxi University, Nanning 530006, China; ruiyangzh@gmail.com

\* Correspondence: chqjia@126.com (Q.C.); wangql985@163.com (Q.W.); Tel.: +0373-3040-001 (Q.W.)

Received: 9 September 2019; Accepted: 12 October 2019; Published: 16 October 2019



**Abstract:** Cotton (*Gossypium hirsutum* L.) is one of the most important cash crops worldwide. Cytoplasmic male sterility (CMS) is an excellent breeding system for exploitation of heterosis, which has great potential to increase crop yields. To understand the molecular mechanism of CMS in cotton, we compared transcriptome, cytomorphological, physiological and bioinformatics data between the CMS line C2P5A and its maintainer line C2P5B. By using high-throughput sequencing technology, 178,166 transcripts were assembled and 2013 differentially expression genes (DEGs) were identified at three different stages of C2P5A anther development. In this study, we identified DEGs associated with reactive oxygen species (ROS), peroxisomes, aldehyde dehydrogenases (ALDH), cytochrome oxidase subunit VI, and cytochrome P450, and DEGs associated with tapetum development, Jojoba acyl-CoA reductase-related male sterility protein, basic helix-loop-helix (bHLH) and MYB transcription factors. The abnormal expression of one of these genes may be responsible for the CMS C2P5A line. In gene ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment, DEGs were mainly related to carbohydrate metabolism, amino acid metabolism, transport and catabolism, and signal transduction. Carbohydrate metabolism provides energy for anther development, starch and sucrose metabolism, fatty acid biosynthesis and metabolism and ascorbate and aldarate metabolism. These results showed that numerous genes and multiple complex metabolic pathways regulate cotton anther development. Weighted correlation network analysis (WGCNA) indicated that three modules, ‘turquoise,’ ‘blue,’ and ‘green,’ were specific for the CMS C2P5A line. The ‘turquoise’ and ‘blue’ modules were mainly related to carbohydrate metabolism, amino acid metabolism, energy metabolism, peroxisomes, pyruvate metabolism as well as fatty acid degradation. The ‘green’ module was mainly related to energy metabolism, carbon metabolism, translation, and lipid metabolism. RNA-sequencing and WGCNA polymerization modules were screened for key genes and pathways related to CMS in cotton. This study presents a new perspective for further research into the metabolic pathways of pollen abortion in the CMS C2P5A line and also provides a theoretical basis for its breeding and production.

**Keywords:** cotton; cytoplasmic male sterility; transcriptome; DEGs

## 1. Introduction

Cotton (*Gossypium hirsutum* L.) is one of the most important cash crops in the world and a renewable source of natural fiber [1]. At present, there are four cultivated varieties of cotton, diploid Asian cotton (*G. arboreum*), African cotton (*G. raimondii*), upland cotton (*G. hirsutum*), and island cotton (*G. barbadense*). Upland cotton accounts for more than 90% of commercial cotton production worldwide because of its excellent characteristics such as high fiber yields, strong adaptability, high resilience, and good quality [2–4].

At present, heterosis is the most effective way of high yield breeding. It makes a great contribution to the increase of agricultural yields because hybrid F1 progenies possess merit traits such as increased adaptability, uniformity, and stronger stress tolerance [5,6]. Heterosis systems have been used to increase yields in several crop species such as wheat [7,8], rice [9], cereals [7,10], soybeans [11], barley [7,12], cotton [13], and rapeseed [14].

In male sterility (MS), the pistil of the bisexual flower plant develops normally, but the stamens fail to produce a dynamic male gamete, therefore, plants can be fertilized by external pollination. MS is widespread in flowering plants as a tool to create new hybrids for breeders, which can greatly reduce labor, costs, and improve production efficiency [15]. According to gene composition, MS includes genic male sterility (GMS) and cytoplasmic male sterility (CMS) [16].

GMS is a two-line breeding system, according to morphological characteristics, the fertile plants and fake hybrids have to be removed, which results in increased labor intensity and operation difficulty. CMS is a three-line breeding system, which can eliminate the need for hand emasculation. CMS has been found in more than 150 species [17], including 7% of the gynodioecy angiosperms [18]. CMS confers an excellent system for hybrid breeding system, especially for monoecious crops, which exploits crop heterosis and produces large-scale commercial F1 hybrid seeds [9,11]. Bohr et al. believed that hybridization or heterosis is the most promising method to solve the challenge of increasing crop yields [11].

Anther development is a complex process involving numerous genomes and multiple metabolic pathways in cells. The carbohydrate metabolism pathway provides energy and carbohydrates for anther development [19–22]. Lipid transport and fatty acid metabolism pathway play crucial roles in signal transduction and sporopollenin synthesis [23–25]. Ascorbate and glutathione metabolism are involved in anther cell development and meiosis [19,26–28]. The flavonoid biosynthesis pathway is involved in the process of forming the pollen coat [19,29,30].

Numerous genes related to anther development have been identified in model plants like *Arabidopsis* (*Arabidopsis thaliana*) and rice (*Oryza sativa*). *SPL/NZZ*, *EMS1/EXS*, *TPD1*, *SERK1*, and *SERK2* regulate the fate of somatic, germ cells and the differentiation of the anther wall during early anther development [31–35]. *BAM1/BAM2* are important regulatory factors for early anther development [36,37]. *MPK3/MPK6* regulate anther chamber formation [38], *DYT1* and *22DYT1* are downstream transcription factors [39,40]. *RPK2* [41], and *TDR* [42–44] regulate and control the degradation of tapetum cells and microspore development. *Defective tapetum Cell Death 1 (DTC1)* controls tapetum degeneration by regulating ROS dynamics in rice [44], *MIL1* and *MIL2* are required for anther and microspore development [45,46]. *UDT1*, *TDR*, and *TIP2* encoding bHLH transcription factors regulate stamen differentiation and anther development [44,47,48]. *PTC1* [49], *EAT1* [50], *API5* [51], and *TIP2/bHLH142* [44] regulate programmed tapetal cell death and pollen development. *WA352* gene interacts with *COX11* gene to cause tapetal programmed cell death (PCD) [52].

The *GhACS1* gene encodes an acyl-CoA synthetase which is essential for normal microsporogenesis in the early anther development of cotton [53]. *GhMYB24* encodes the MYB-like transcription factor that regulates the development of tapetum [54].

With regards to the relationship between male sterility and the homeostasis of reactive oxygen species (ROS), previous studies have shown that the accumulation of ROS is higher in male sterile plants, for instance, in rice [52,55], wheat [56], maize [57] and cotton [27,58], indicating that ROS accumulation may be closely related to the development and functional defects of tapetum [48,52,59].

Previous reports have shown that the formation of CMS mainly involves the mutation or molecular rearrangement of genes in the mitochondria or chloroplast genome [60,61]. Chen et al. believed that CMS is a complex character [62]. Mitochondrial genes that determine male sterility have been found in many plants, most of these genes are chimeric [63–65]. In total, 28 CMS genes have so far been identified in 13 crops [9].

Hua and his colleagues compared mitochondrial genome differences in cotton of the male sterile lines CMS-2074A, CMS-2074S, and maintainer and restorer lines, and identified four chimeric gene ORFs, *Aorf4*, *Aorf9*, *Aorf4-2*, and *Aorf28*, in the male sterile lines (CMS-2074A) and restorer lines with different transcription levels. The study revealed the cause for the formation of CMS in cotton [66].

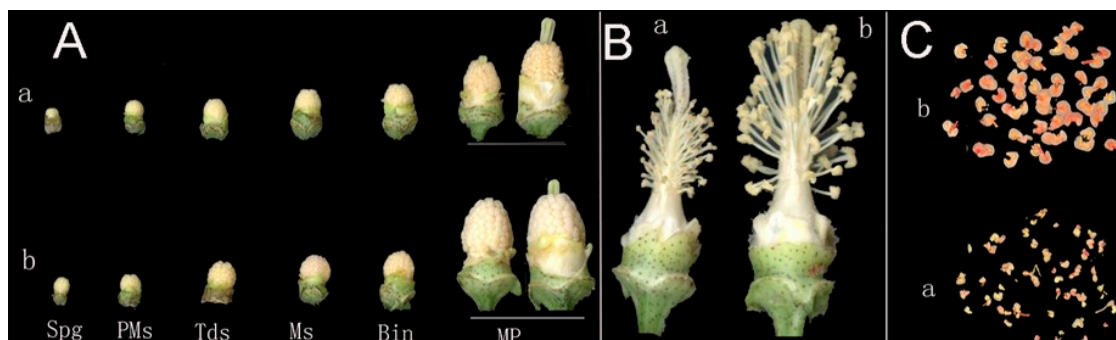
To explore the CMS mechanism of abortion in cotton, the morphological, the physiological, and the molecular characteristics of the CMS line C2P5A and the maintainer line C2P5B were studied. Morphological changes and changes in the tapetum at different stages of anther development were analyzed, pollen viability was detected by triphenyltetrazolium chloride (TTC) staining and observed by scanning electron microscopy (SEM).

In this article, we used high-throughput sequencing technology and WGCNA polymerization modules to screen for key genes and pathways related to CMS in cotton by comparative transcriptome analysis. Besides, this study used quantitative Reverse-Transcription PCR (qRT-PCR) technology to verify the expression of genes related to fertility selected from RNA-seq. The results of this study will provide new data for further research of the metabolic pathway of cotton abortion in the CMS line C2P5A, as well as a theoretical basis for its use in breeding and seed production.

## 2. Results

### 2.1. Morphological Characteristics of the CMS Line C2P5A and the Maintainer Line C2P5B

Comparing the morphology of CMS line C2P5A and maintainer line C2P5B revealed the abnormal development of flower buds and the mature pollen (MP) grain stage in the CMS line C2P5A, and the flower phenotype of C2P5A was small and wizened (Figure 1A).



**Figure 1.** The morphological characteristics of flower buds at different development stages. a, cytoplasmic male sterile line C2P5A; b, maintainer line C2P5B. (A) Flower buds after removing sepals and petals. (B) Flower with sepals and petals removed. (C) Determination of pollen grain activity by TTC. Spg, sporogenous stage; Pms, pollen mother cell stage; Tds, tetrad stage; Ms, mononuclear stage; Bin, binucleate stage; Mps, mature pollen grain stage.

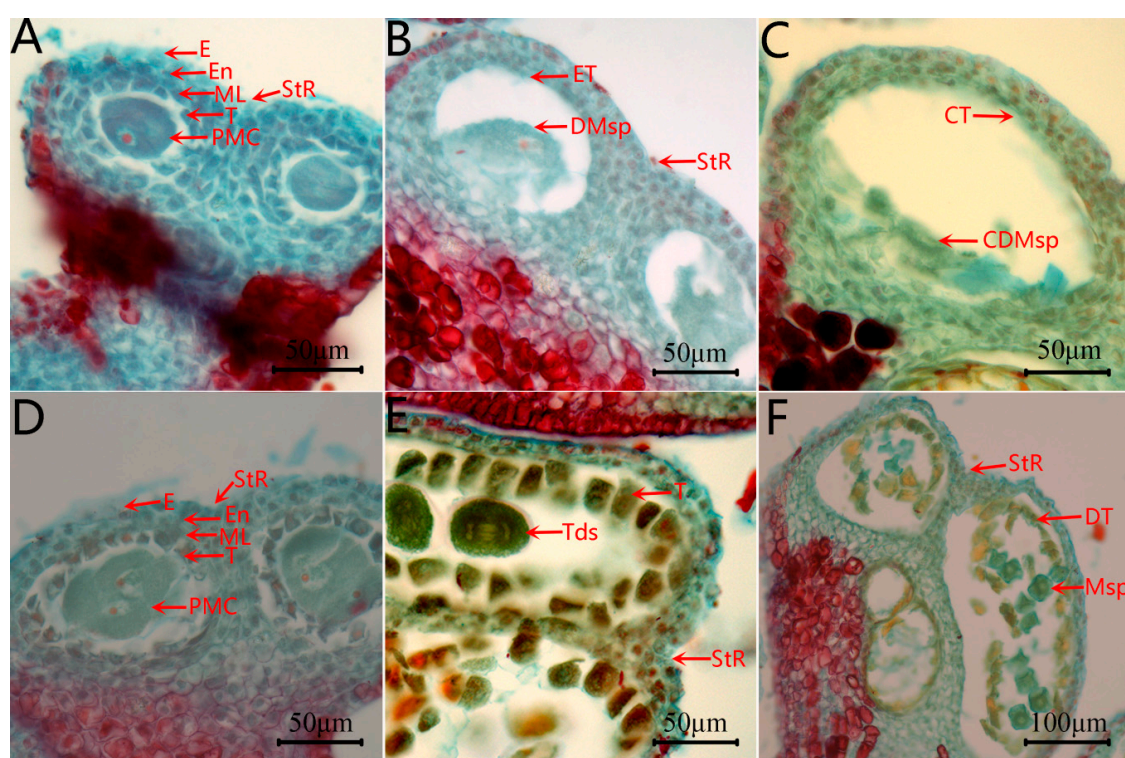
The CMS line with exposed styles and without pollen have significantly shorter filaments than maintainer line (Figure 1B). Pollen grains of CMS line stained by 2% TTC were dry, shape and size were nonuniformity, mostly not colored or lightly colored, and activity was significantly weaker than that of the maintainer line (Figure 1C).

In this study, flower buds were selected from CMS line and maintainer line of anther development at three different stages, and 1.5% (*w/v*) acetocarmine staining and cross-section were used to confirm the developmental stage, pollen mother cell stage (3–4 mm), tetrad stage (4.1–5.0 mm), mononuclear

(5.1–6.0 mm) development at three different stages, and flower bud length (length from nectary to bud apex, mm).

### 2.2. Microstructure of Anther at Different Developmental Stages of the CMS Line C2P5A

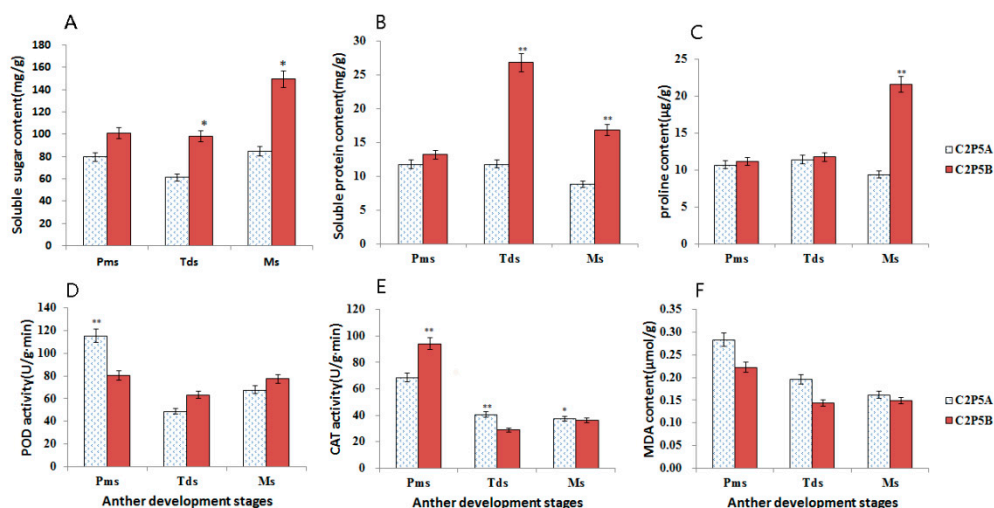
When comparing the anther development of C2P5A with C2P5B by observing cross-sections, there was little difference between the two lines in pollen mother cell stages (Pms). Pollen mother cells were surrounded by four layers of cells, the innermost layer being the tapetum, followed by the middle layer, the endothecium, and the epidermal layer (Figure 2A,D). However, during the tetrad and mononuclear stages, there was a significant difference between the CMS line C2P5A and the maintainer line C2P5B (Figure 2B,C,E,F). The pollen mother cells of the CMS line C2P5A were degraded, the tapetum cells were atrophic, dense, and became a very thin layer (Figure 2B), and during the cells monocyte stage, no monocyte microspores were produced (Figure 2C).



**Figure 2.** Phenotypic comparison between C2P5A and C2P5B at different development stages. (A–C) CMS line C2P5A; (D–F) maintainer line C2P5B. (A,D) Pollen mother cell stage (Pms); (B,E) Tetrad stage (Tds); (C,F) Mononuclear stage (Ms). PMC, pollen mother cell; E, epidermis; En, endothecium; ML, middle layer; Msp, microspores; T, tapetum; DT, degenerated tapetum; Tds, tetrads; Mn, mononuclear; DMsp, degenerated microspores; CDMsp, complete degenerated microspores; CT, compact tapetum; ET, entire tapetum; StR, stomium region.

### 2.3. Measurements of Physiological Indices in the CMS Line C2P5A

Soluble sugar provides energy for anther development and synthesis. Soluble proteins include a variety of amino acids that protect cell development. At different stages of anther development, the soluble sugar and soluble protein content of the maintainer anthers was higher than that of the male sterile anthers. During anther development, significant differences could be observed in Tds and Ms between C2P5A and C2P5B (Figure 3A,B). Proline (Pro) is involved in the stress response in plants, but there was little difference in proline content of Pms and Tds during anther development in C2P5A and C2P5B. However, Pro in C2P5A was lower than in C2P5B at the Ms stage of anther development (Figure 3C).



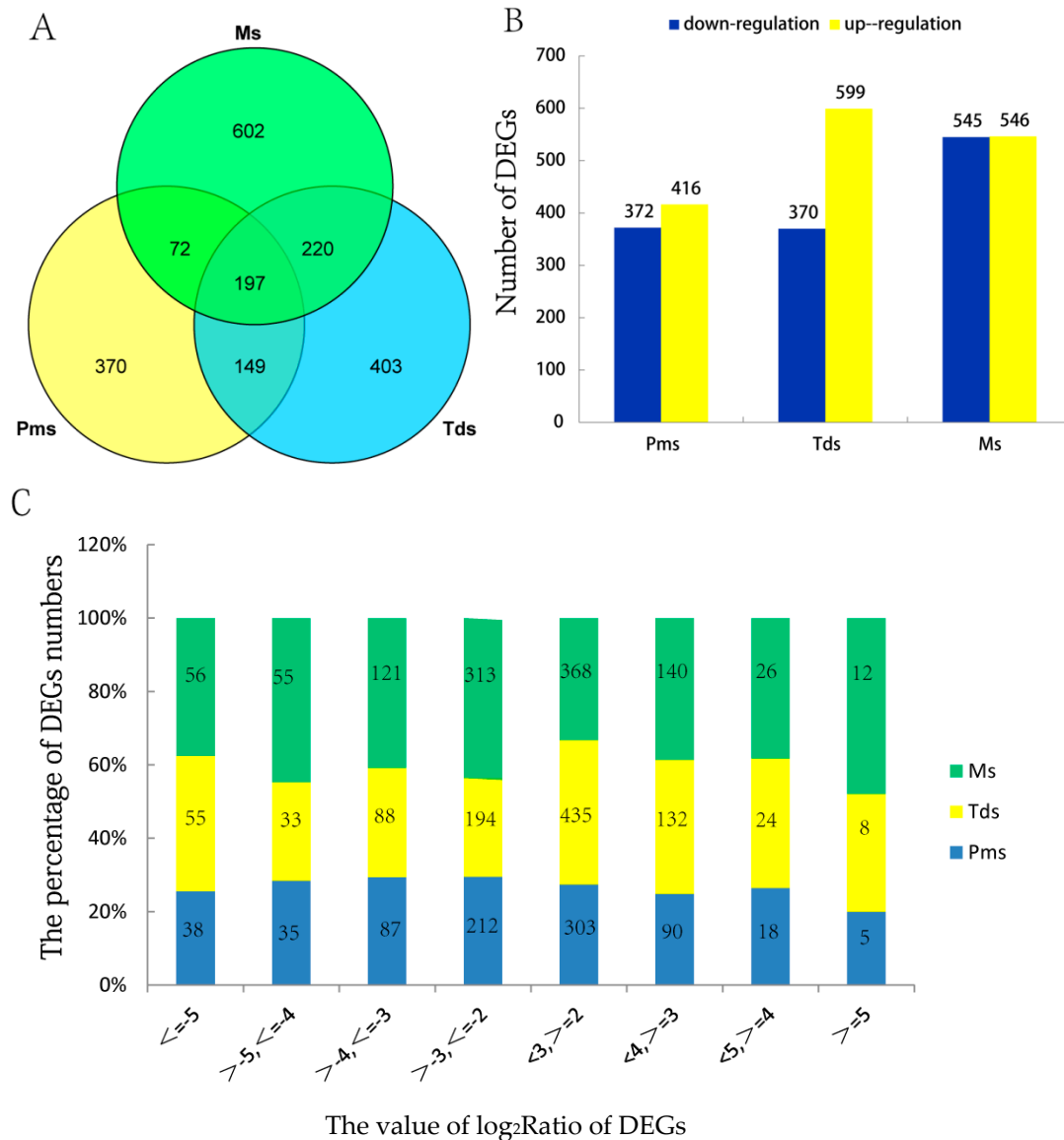
**Figure 3.** Determination of physiological indices of anthers at different developmental stages. (A) Determination of soluble sugar content. (B) Determination of soluble protein content. (C) Determination of proline content. (D) Determination of POD activity. (E) Determination of CAT activity. (F) Determination of MDA activity. Pms, pollen mother cell stage; Tds, tetrad stage; Ms, mononuclear stage; C2P5A, CMS line; C2P5B, maintainer line. The white columns indicate male sterile line C2P5A; red columns indicate the maintainer line C2P5B. The error bars represent SD (\*\*  $p < 0.01$ , \*  $p < 0.05$ ), three biological replicates were performed.

Antioxidase peroxidase (POD) can eliminate the toxicity of hydrogen peroxide, phenols, amines, aldehydes, and benzene, whereas catalase (CAT) can remove hydrogen peroxide ( $H_2O_2$ ) in the organism. At Pms stage in anther development, before anther abortion, POD and CAT activity increased correspondingly and reached significant levels at different stages in both lines. In the Pms stage, POD activity was significantly higher in the male sterile line than in the maintainer line (Figure 3D). However, CAT activity was significantly lower in C2P5A than C2P5B (Figure 3E). At Tds and Ms stage of anther development, the POD activity of male sterile line was lower than that of maintainer line (Figure 3D). However, the CAT activity was higher and reached a significant different level in the male sterile line (Figure 3E). Malondialdehyde (MDA) is an indicator of ROS in organisms, and the results showed that the MDA content in the male sterile line was higher than that in the maintainer line during all three anthers stages (Figure 3F).

#### 2.4. Identification of Differentially Expressed Genes (DEGs)

In this research, a total of 18 *Gossypium hirsutum* samples (three biological replications, three anther stages, two lines C2P5A and C2P5B) were sequenced using the Illumina HiSeq platform. A total of 178,166 transcripts were assembled, including 74,292 novel transcripts and 103,874 known transcripts. The number of total genes and new transcripts per sample were presented in Table 1. Clean reads (77743) were aligned to the reference sequence with Bowtie2 [67], and then the expression levels of genes and transcripts were calculated by RNA-Seq by Expectation Maximization (RSEM) [68] (Supplementary Table S1). Fragments per kilobase of transcripts per million mapped reads (FPKM) is the standard method to calculate gene expression levels by RSEM. Transcripts that were not entirely annotated to cotton were removed [69]. In this study, analysis of the difference between groups was conducted using the software DEGseq [70]. According to the threshold values:  $|\log_2\text{Ratio}| \geq 2$ ,  $\text{readnum} > 3$ ,  $p \leq 0.001$  and the criteria of  $\text{FDR} < 0.01$  (Supplementary Table S2a). Overall, 788, 969, and 1091 differentially expressed genes (DEGs) were identified at the three stages (Pms, Tds, Ms), respectively, and 197 genes were shared among the three stages. A total of 2013 DEGs were identified (Figure 4A,B, Supplementary Table S2a–e). Analysis by  $\log_2\text{Ratio}$  found that the percentage of DEGs were mainly distributed from 2 to 3 and from  $-2$  to  $-3$  during Pms and Tds, and during Ms the percentage of DEGs was distributed

widely (Figure 4C). The results showed that the male sterility had an effect on transcriptional changes in the anthers of C2P5A. At different stages of anther development, DEGs were up or down-regulated compared to C2P5B, indicating that a minimum number of DEGs occurred at the Pms stage of anther development rather than at the Tds and Ms stage of anther development. At Pms and Tds stage of anther development, more DEGs were found to be down-regulated than up-regulated. At Ms stage of anther development, almost an equal number of DEGs were up-regulated and down-regulated (Figure 4B). This may suggest that at Pms and Tds stage of anther development, the DEGs contain critical information related to male sterility.



**Figure 4.** Analysis of DEGs in the male sterile line C2P5A compared to those of the maintainer line C2P5B from Pms to Ms stage of anther development. (A) Venn diagrams showing all DEGs shared among the three stages. (B) Number of DEGs that were up or down-regulated in the three development stages. (C) The distribution of the log<sub>2</sub>Ratio of DEGs in the C2P5A compared to the C2P5B, Pms, Pollen mother cell stage; Tds, Tetrad stage; Ms, Mononuclear stage.

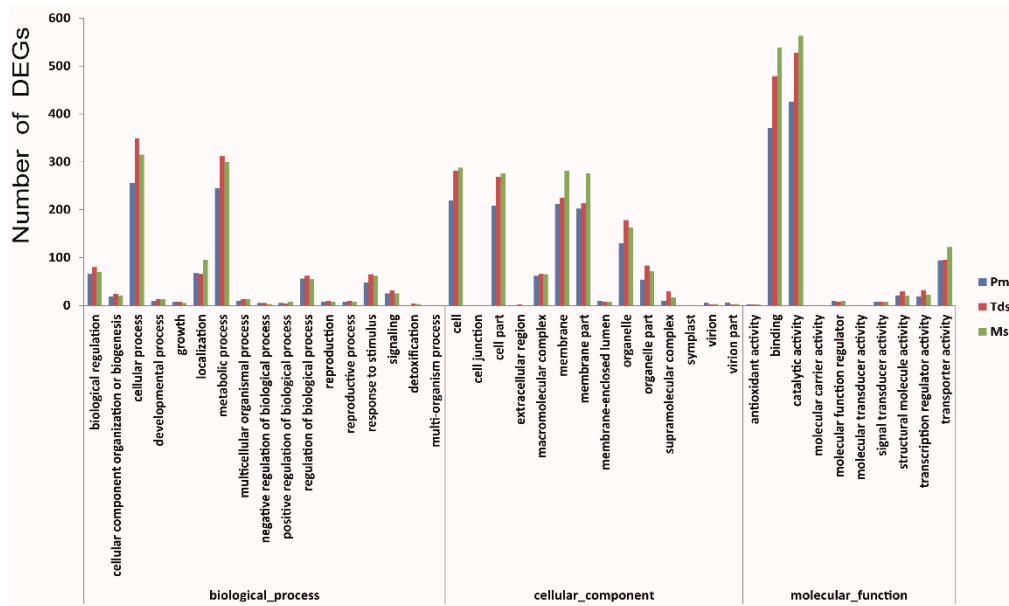
**Table 1.** Summary of tag numbers.

Sample	Raw Data	Clean Data	GC (%)	Q30(%)	(%)of Total Mapped	Total Gene Number	Total Isoforms	Novel Isoforms
MsPms1	96339898	84112026	44.33	97.65	77.42	85012	126455	49311
MsPms2	96339060	84728466	44.43	97.84	75.51	83758	122693	48159
MsPms3	96340692	84822750	44.52	97.81	75.44	85170	125779	49293
MsTds1	66418638	60016542	44.22	98.33	76.38	83808	122636	47699
MsTds2	96341884	84562392	43.57	97.68	77.36	86647	129005	50208
MsTds3	96340878	84830410	44	97.77	74.54	86626	126787	50096
MsMs1	70226770	62112146	45.21	98.35	78.11	82246	117948	45803
MsMs2	96340984	84269862	44.5	97.81	77.01	85808	124373	49192
MsMs3	75363530	68729552	44.43	98.45	79.47	85289	123996	48513
MPms1	78357630	71140510	43.98	98.42	77.1	84009	123651	49187
MPms2	93592588	83300884	43.81	97.97	76.95	84628	124284	49724
MPms3	94122576	81326894	44.24	97.83	73.56	84969	124580	49711
MTds1	93071774	84922682	43.79	98.38	78.87	84527	124186	49781
MTds2	79502370	72818484	43.85	98.23	80.11	83170	121809	48480
MTds3	71210292	64700084	43.83	98.33	79.25	82748	120584	48031
MMs1	96340778	84992800	44.36	98.03	75.37	84192	124176	48798
MMs2	96340324	84734142	43.65	97.83	77.69	84433	124979	49064
MMs3	74569750	67998900	43.6	98.23	75.57	82250	120841	47466

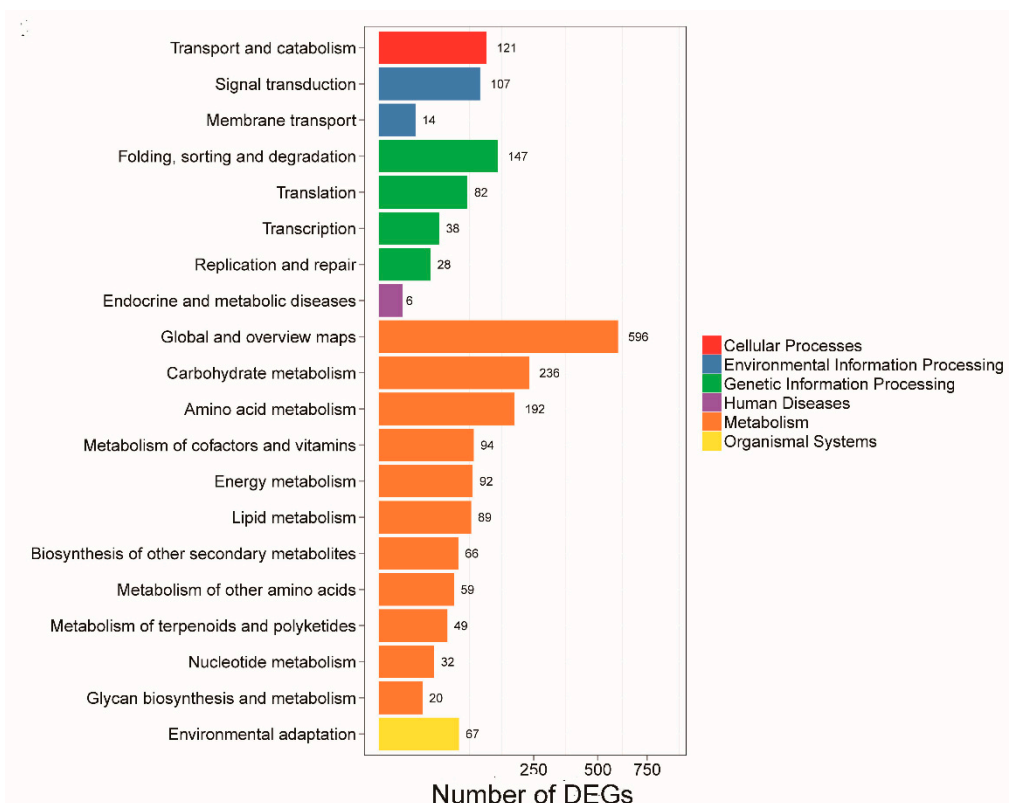
MS, male sterile line C2P5A; M, maintainer line C2P5B; 1, 2, 3 represent the biological duplications; Pms, pollen mother cell stage; Tds, tetrad stage; Ms, mononuclear stage.

### 2.5. Gene Ontology Annotation and Pathway Enrichment Analysis of DEGs

The results of gene ontology (GO) annotation and function enrichment of 2013 DEGs were distributed over 41 GO terms (Figure 5, Supplementary Table S3a). Molecular function, binding, and catalytic activity were related to the highest DEG number, followed by transporter activity. These results suggest that anther development in CMS may be associated with different genes. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment was used to identify major biochemical and signal transduction pathways where the DEGs participated. In this experiment, pathway enrichment analysis showed mainly enrichment in 20 pathway categories including 125 KEGG pathways (Figure 6, Supplementary Table S3b). Global and overview maps, carbohydrate metabolism, folding, amino acid metabolism, sorting and degradation were the four most significant pathways, followed by transport and catabolism, signal transduction, metabolism of cofactors and vitamins, and energy metabolism. These results indicated that multiple complex metabolic pathways were involved in the anther development of cotton.



**Figure 5.** GO function enrichment of DEGs between the C2P5A and the C2P5B from the Pms to Ms stage of anther development. Pms, pollen mother cell stage; Tds, tetrad stage; Ms, mononuclear stage.

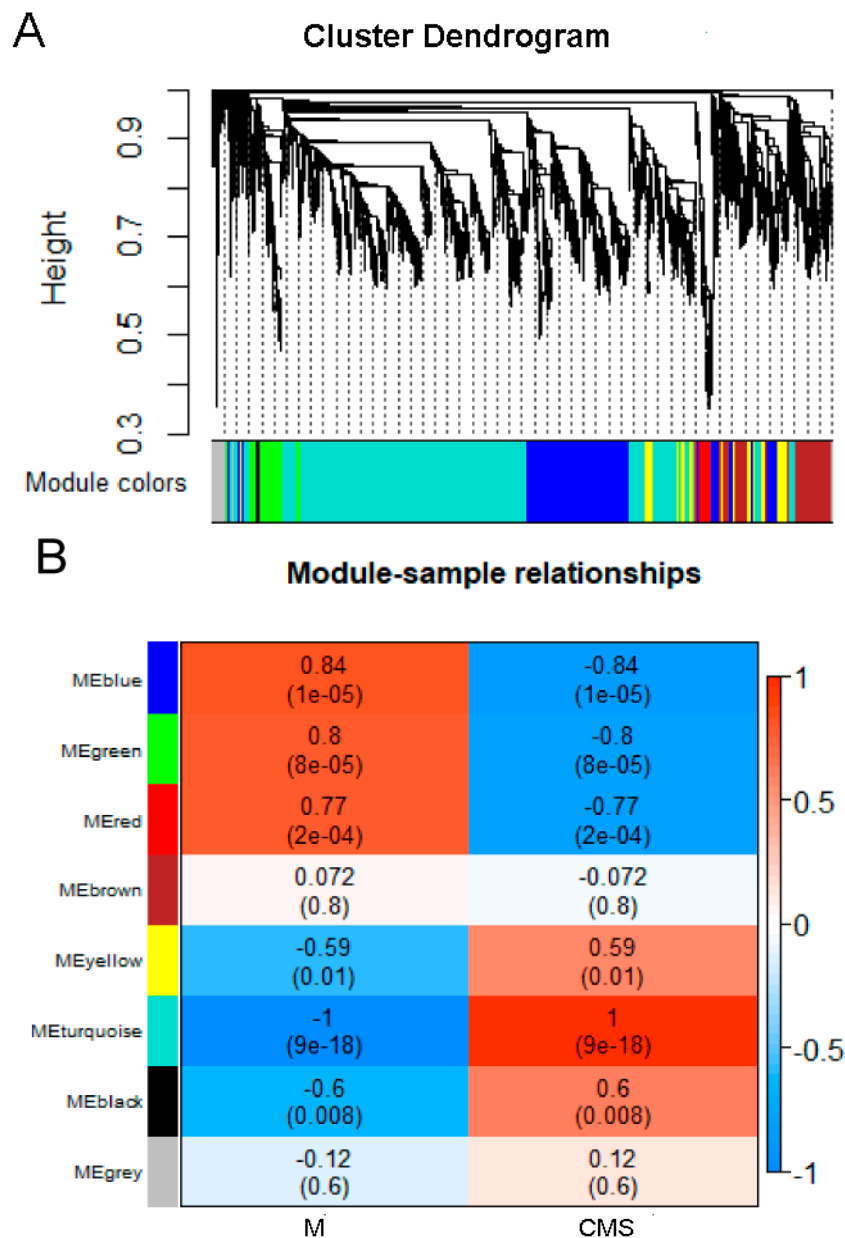


**Figure 6.** KEGG pathway enrichment analysis of 2,013 DEGs.

### 2.6. Correlation Network Analysis with WGCNA

The weighted correlation network (WGCNA) is an alternative tool to explore target genes at a network-level, instead of individual genes [15,71,72]. In WGCNA, a module refers to a highly interconnected gene cluster, and within the same cluster the correlation coefficient is higher in the same tissue. To identify genes associated with CMS, seven modules were identified from the transcriptome data (Figure 7A).





**Figure 7.** WGCNA analysis of CMS related genes. (A) Hierarchical clustering tree showing co-expression modules. In the tree, each leaf represents one gene. The major tree branches constitute seven modules, labeled by different colors. (B) Module-trait relationship. High correlation between a specific module and the sample is shown in dark red or dark blue. M, maintainer line C2P5B; CMS, cytoplasmic male sterile line C2P5A.

Different samples were correlated with seven distinct modules, module 'turquoise' ( $r = 1$ ,  $p = 9 \times 10^{-18}$ ), module 'blue' ( $r = 0.84$ ,  $p = 1 \times 10^{-5}$ ), and module 'green' ( $r = 0.8$ ,  $p = 8 \times 10^{-5}$ ) were highly correlated with the CMS line (Figure 7B).

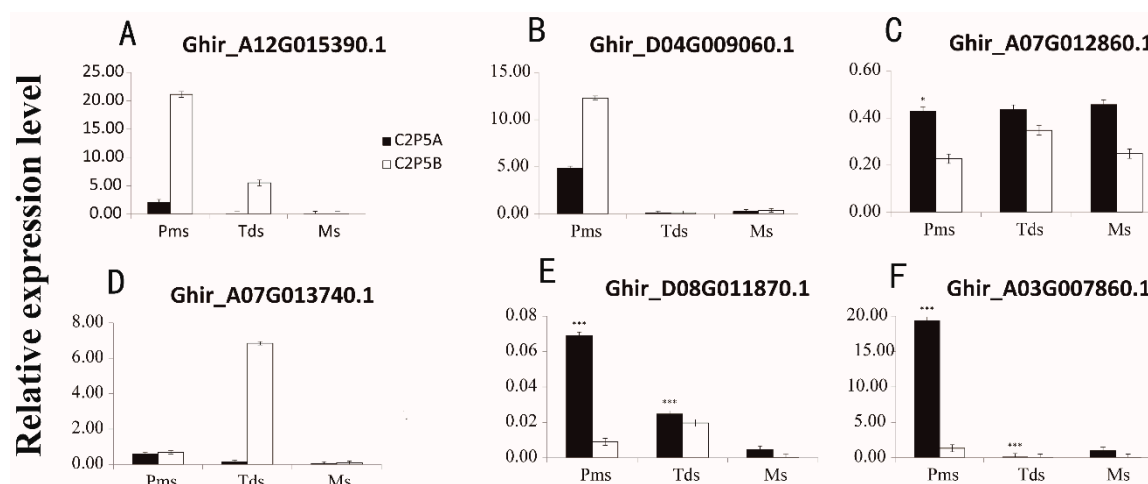
The 'turquoise' module contained 3670 genes that were mainly related to genetic information processing, carbohydrate metabolism, amino acid metabolism, and energy metabolism (Supplementary Table S4a). The 'blue' module contained 1415 genes that were mainly involved in protein folding, sorting and degradation, amino acid metabolism, glycolysis/gluconeogenesis, peroxisomes, pyruvate metabolism, and fatty acid degradation (Supplementary Table S4b). The 359 genes in the 'green' module were connected with metabolism, energy metabolism, and oxidative phosphorylation (Supplementary Table S4c).

To further study the molecular mechanism of CMS, the unigenes of the three modules were analyzed by GO enrichment and KEGG pathway. In GO analysis, the unigenes were mainly enriched in metabolic process (GO: 0030163), cellular process (GO: 0009765), cell (GO: 0005737), catalytic activity (GO: 0008171), and binding (GO: 0008270) (Supplementary Figure S1). In KEGG pathway analysis, the 'turquoise' and 'blue' modules were similar and mainly involved in translation (ko03010), carbohydrate metabolism (ko01200), folding, sorting, and degradation (ko04141), signal transduction (ko04075), and amino acid metabolism (ko01230) (Supplementary Figures S2 and S3). Genes in the 'green' module were mainly related to energy metabolism (ko00910), carbon metabolism (ko01200), translation (ko03010), lipid metabolism (ko00071), and plant hormone signal transduction (ko040750) (Supplementary Figure S4).

### 2.7. Major Pathways and Genes Associated with CMS

It was found that numerous DEGs were involved in the metabolic pathways of male sterility in cotton C2P5A. Carbohydrate metabolism provides energy for anther development, four DEGs were involved in pentose and glucuronate interconversions (ko00040); 67 DEGs in starch and sucrose metabolism (ko00500); 42 DEGs in galactose metabolism (ko00052); 32 DEGs in fatty acid biosynthesis and metabolism (ko00061, ko01212); 17 DEGs in ascorbate and aldarate metabolism (ko00053); and 28 DEGs in glutathione metabolism (ko00480) (Supplementary Table S5a–f).

A number of genes were associated with ROS detoxification. We identified 18 peroxisome (ko04146) genes encoding POD (peroxidase), CAT (catalase), and XOT (alternative oxidase) (Supplement Table S6a). In addition, nine aldehyde dehydrogenase family (ALDH) DEGs were identified (Supplement Table S6b). Of these genes, *Ghir\_A12G015390.1* was down-regulated in three developmental stages (Figure 8A). Cytochrome c oxidase subunit VI mutations are associated with the male sterility of certain plants, we identified three associated DEGs (Supplement Table S6c). *Ghir\_D04G009060.1* encoding cytochrome c oxidase subunit VI was down-regulated at Pms and Tds stage of anther development, then up-regulated at Ms stage (Figure 8B). We identified 13 DEGs which encoded cytochrome P450 (Supplement Table S6d). The expression of *Ghir\_A07G012860.1* was up-regulated at three different stages of the C2P5A anther development (Figure 8C). We identified seven DEGs that encoded acyl-CoA synthetase (ACS1) (Supplement Table S6e) and the expression of *Ghir\_A07G013740.1* was down-regulated at three different stages of the C2P5A anther development (Figure 8D).



**Figure 8.** Expression patterns of DEGs in anther development. (A), (B), (C) and (D) The ROS detoxification related genes. (E) and (F) The tapetum development genes. Pms, pollen mother cell stage; Tds, tetrad stage; Ms, mononuclear stage. C2P5A, cytoplasmic male sterile line; C2P5B, maintainer line. The black columns indicate the male sterile line C2P5A, and the white columns indicate the maintainer line C2P5B. The error bars represent SD (Student's *t*-test), \*\*\*  $p < 0.001$ , \*  $p < 0.05$ , and absolute value of  $\log_2\text{Ratio} \geq 2$ ,  $p = 0$  in (C),  $p = 0$  in (E),  $p = 8.57921\text{E}-05$  in (F), three biological replicates were performed.

Several genes associated with tapetum development were identified in the CMS line C2P5A. Two genes encoded the Jojoba acyl-CoA reductase-related male sterility protein (*Ghir\_A08G000910.1*, *Ghir\_D09G013360.1*). We identified nine up-regulated genes that encoded basic helix-loop-helix (bHLH) DNA-binding family proteins (Supplement Table S6f). *Ghir\_D08G011870.1* was up-regulated at three stages of C2P5A anther development (Figure 8E), the abnormal expression of these genes may be responsible for the male sterile line C2P5A. We identified 55 genes that were MYB domain proteins that regulate tapetum development (supplement Table S6g). *Ghir\_A03G007860.1* was up-regulated at three stages of C2P5A anther development (Figure 8F).

### 2.8. Validation of DEGs by Quantitative Reverse-Transcription PCR (qRT-PCR)

qRT-PCR was used to verify the expression levels of DEGs identified. We identified seven genes associated with fertility at three different stages of anther development, choline/ethanolamine kinase (*Ghir\_A10G019450.1*), ALDH involved in the glyoxylate cycle (*Ghir\_A12G015390.1*), cytochrome c oxidase subunit VI (*Ghir\_D04G009060.1*), cytochrome P450 (*Ghir\_A07G012860.1*), ACS1(*Ghir\_A07G013740.1*), and two transcription factors involved in tapetal development (*Ghir\_D08G011870.1*, *Ghir\_A03G007860.1*) (Table 2). Each gene underwent three biological replicates and each biological replicates was three times. Pearson's correlation was used to evaluate the qRT-PCR and RNA-sequencing data. Overall, the expression levels of the seven DEGs were consistent with RNA-sequencing data (Supplementary Figure S5A; correlation coefficient = 0.68). The coefficients for Pms, Tds and Ms stage of anther development were 0.74, 0.83, 0.54, respectively (Supplementary Figure S5B–D). The results showed that our RNA-seq data were reliable and conducive to the identification of DEGs in anther development.

**Table 2.** qRT-PCR confirmation of DEG expression profiles.

Gene ID	Protein Identity	Pms(Fold Change)		Tds(Fold Change)		Ms(Fold Change)	
		qRT-PCR	RNA-seq	qRT-PCR	RNA-seq	qRT-PCR	RNA-seq
<i>Ghir_A10G019450.1</i>	Cell wall/membrane/envelope biogenesis	−5.41	−2.68	4.51	3.15	3.47	2.37
<i>Ghir_A07G013740.1</i>	Long chain acyl-CoA synthetase	−1.59	−2.56	−3.52	−2.64	−0.16	−2.88
<i>Ghir_A12G015390.1</i>	Aldehyde dehydrogenase family	−3.88	−6.73	−6.10	−9.37	−1.05	−6.86
<i>Ghir_D08G011870.1</i>	bHLH-MYC transcription factors	2.87	2.13	2.00	2.16	4.67	2.52
<i>Ghir_A07G012860.1</i>	Cytochrome P450	0.98	2.12	0.23	2.50	0.98	2.44
<i>Ghir_D04G009060.1</i>	Cytochrome oxidase c subunit VIb	−4.53	−3.49	−0.29	−2.37	1.87	3.11
<i>Ghir_A03G007860.1</i>	Myb-like DNA-binding domain	1.61	2.56	3.70	2.23	5.73	2.19

Fold change refers to the expression in the male sterile line C2P5A when compared to the maintainer line C2P5B. A negative and positive value refers to down- and up-regulation, respectively in the male sterile line C2P5A.

## 3. Discussion

The sequencing of several polyploid cotton plants has been completed, promoting the development of the cotton genome [69,73,74]. With these reference sequences, it was more convenient to use high-throughput sequencing technology to explore the molecular mechanism of CMS. RNA-sequencing techniques were used in cotton CMS-D8, CMS line H276A, CMS line zhong41A, and GMS line 1355A [25,27,61,75]. Except for the CMS line zhong41A, the selected samples are a mixture of anthers, ovaries and stigmas [27]. We selected samples similar to those in the CMS line zhong41A; flower buds of coincident length were selected, sepals and petals were peeled off, pure anthers were collected, and stigmas and ovaries were discarded. To obtain accurate and repeatable RNA-sequencing data, three biological replicates for each sample were used, and a high correlation coefficient (0.99–1) between replicates per sample were obtained, indicating sequencing results were credible and that the consistency of anthers collection was maintained.

### 3.1. The Crucial Period of Abortion in the CMS Line C2P5A

Normal development of anthers is required for the breeding and reproduction of flowering plants. Male sterility is a common phenomenon and often used to produce seeds by heterosis. It has been reported that flowering plants abort at different periods. In the photosensitive male sterility in cotton,

there is no difference between the mutant and the wild type of anthers in pollen mother cells and tetrad stages. The tapetum of the mutant plant is degraded at the microspore stage, but the released microspore is empty and shriveled, so the abortion period of the photosensitive male sterile mutant occurs at the microspore stage [76]. In the GMS 1355A line, anther abortion occurs during the release of microspores (stage 8), microspores are shrunken and the exine lacks spines [25]. In CMS H267A, zhong41A male sterile lines, the pollen mother cell is gradually dissolved at the tetrad stage [27,75]. In their study, Rao and colleagues studied 36 male sterile lines from 14 cytoplasmic sources, but six without cytological material on male sterility. The results showed that in dicotyledonous plants, 27% of abortions occurred at the early stage of meiosis, 58% at the tetrad stage, and 15% at the microspore development stage [77].

In this research, C2P5A anthers do not form tetrads in the tetrad period, indicating that abortion may occur at the tetrad stage. According to previous reports, this period of abortion is called non-pollen male sterility [78], which is appropriate for studying the molecular mechanism of CMS.

### 3.2. Major Genes Regulate Anther Development in Cotton

Anther development is regulated by numerous genes, approximately 3500 genes are explicitly expressed in *Arabidopsis* anthers [79]. In the cotton (*G. hirsutum*) 21A GMS line and its maintainer line, 1742 genes were found to have differential expression [22]. 2446 DEGs were identified in the anthers of 1355AB line [25]. And a comparison of H276A and H276B anthers yielded 3603 DEGs [75]. In this study, we analyzed the transcript profiles of the cotton CMS line C2P5A and the maintainer line C2P5B. A total of 2013 genes were found to have differential expression by DEGseq. GO annotation and KEGG enriched pathway analysis showed that male sterility resulted in the differential expression of many genes.

CMS is associated with recombinant proteins encoded by mitochondria [58,80–82]. Aerobic respiration takes place in the mitochondrion and provides energy for various activities within the cell. The mitochondrion is the main source of ROS and thereby triggers or inhibits apoptosis [27,48,83].

ROS homeostasis is crucial for normal anther development [27]. A subtle ROS concentration is necessary for cell metabolism of various plant growth processes, such as root cell expansion and abiotic stress [27,84,85]. However, when there is an over-accumulation of ROS, which cannot be rapidly removed from cells, a ROS burst occurs and induces cell apoptosis and anther abortion [58,86]. MDA is an indicator of ROS in organisms and results from the enhancement of membrane lipid peroxidation, which damages the cell membrane and disturbs physiological metabolism inside the cell [87]. In the present study, MDA levels in C2P5A male sterile line were higher than those of the C2P5B maintainer line (Figure 3F). This proved that the C2P5A male sterile line accumulate excessive ROS during anther development. The results in cotton are consistent with U87B1-706A (wheat) [88] and 9704A (pepper) [87].

To better resist oxidative stress, plants form an efficient antioxidant defense system in the cell and reduce the impact of ROS on a variety of biological molecules, thus maintaining homeostasis [89]. SOD, POD, and CAT are necessary antioxidant enzymes that scavenge excess ROS [90]. In our study, at Pms stage of anther development and before anther abortion in C2P5A, a maximum increase of MDA content indicated excessive ROS accumulation, as the antioxidant defense system is activated, antioxidant POD and CAT levels increased (Figure 3D,E). Studies have shown that excessive accumulation of ROS may induce an antagonism oxidation system in biological organisms [58]. According to the enzymatic mechanism of ROS scavenging, we identified 18 peroxisome genes encoding POD, CAT, and XOT, at Pms stage of anther development. Seven genes are up-regulated and five down-regulated when ROS levels are highest.

The ALDH family plays an important role in mitochondria, which down-regulates and causes the cytotoxicity of excess acetaldehyde and ethanol in cells, contributes to ROS bursts and inhibits progress of the tricarboxylic acid cycle [91,92]. Nine genes encoding ALDHs were identified (Supplementary Table S6b). In previous study, cytochrome c oxidase subunit VI mutations were associated with male

sterility in certain plants [75], for instance, in Beet G-CMS, *cox2* expression decreased cytochrome c oxidase activity [93], and in the pepper CMS line, the *cox2* gene was inserted into the *orf456* chimeric gene [94].

Cytochrome c release and excessive ROS function are mostly retrograde signals that trigger plant PCD and lead to male sterility [95]. COX11 is the nuclear coding assembly factor of cytochrome c oxidase and is highly conserved in eukaryotes [96], for example, yeast COX11 (ScCOX11) [97] and rice COX11 (OsCOX11) [52]. In our research, three genes encoding cytochrome oxidase have been identified and C2P5A male sterility may be associated with them.

Cytochrome P450 is an oxidase involved in many biosynthetic pathways. In tapetum, fatty acid-CoA enters the endoplasmic reticulum and under the catalysis of cytochrome P450 family proteins performed hydroxylation. This process is essential for the formation of both cuticle and exine [98]. In this study, we identified 13 DEGs that encoded cytochrome P450. The expression of these genes in male sterile line was abnormal compared with the maintainer line, which may be related to pollen abortion in C2P5A.

Transcription factors are associated with anther development. The tapetum is an important structure for anther development. The tapetum and its function are regulated by numerous genes and transcription factors [99]. Basic helix-loop-helix (bHLH) domain transcription factors are mainly involved in the regulation of plant growth and development [100]. Most reported bHLH transcription factors related to tapetum cell and microspore development, abnormal function of bHLH transcription factors cause male sterility mutations, which often have similar characteristics and show precocious or delayed degradation of tapetum cells, resulting in blocked pollen formation and abortion. In *Arabidopsis*, the mutation that causes aborted microspores, *DYT1*, leads to the abnormal development of tapetum and premature degradation of microspores [40,101]. *AtbHLH10*, *AtbHLH89*, and *AtbHLH91* interact with each other, and double or triple mutations show male sterility [47]. *JAM1*, *JAM2*, and *JAM3*, three bHLH transcription factors involved in jasmonic acid signal transduction, play a negative regulatory role of JA-mediated male sterility [102]. In rice, *UDT1*, *TDR*, *ETA1*, and *bHLH142/TIP2* encode bHLH transcription factors. The *UDT1* mutant has highly vacuolated tapetum cells during meiosis, the pollen mother cells do not divide into microspores, and there is no gradual degradation [103]. *TDR*, *ETA1*, and *bHLH142/TIP2* regulate delayed tapetal degradation and aborted pollen formation [42,44,50,104]. *OsbHLH138* regulates thermosensitive GMS in rice [105]. In cotton, Yang and colleagues identified four bHLH DEGs and two MYB130 DEGs that may be involved in abnormal tapetum development in *zhong41A* [27]. In this research, we identified nine up-regulated genes that coded for bHLH DNA-binding family proteins, the up-regulated expression of these genes may be responsible for male sterility of C2P5A.

MYB is an important transcription factor family that regulates anther development. *AtMYB103* regulates the development of the tapetum and its abnormal expression causes male sterility [106]. *AtMYB33* and *AtMYB65* regulate the development of tapetum, and the double mutant has excessive vacuolation, swelling, and hypertrophy in tapetum cells [107]. *TDF1* encodes the R2R3 MYB transcription factor mostly involved in transcriptional regulation of tapetum cells and pollen mother cells [108]. *AtMYB21* mutant or overexpression of *AtMYB24* exhibit shorter anther filaments, delayed or no anther dehiscence, and fewer viable pollen grains [109]. Similarly, the *GhMYB24* transcription factor is specifically expressed in pollen and involved in the regulation of late anther/pollen development [54]. In this study, 55 MYB transcription factors were identified. Their abnormal expression in male sterile line compared with maintainer line may be the cause of male sterility.

*ACS1* encodes acetyl-coenzyme A. In previous studies, it has been reported that ACS proteins may be localized in the outer membrane of pea chloroplasts and maturing safflower seeds, and its activity is associated with fatty acids and microsomes [110,111]. It was found that ACS proteins are essential for normal cuticle development and are involved in peroxisomal and fatty acid  $\beta$ -oxidation in *Arabidopsis* [22,112–114]. However, *GHACS1* gene plays an important role in the formation of microspores in the early anther development of cotton and is highly expressed in sporogenous cells,

pollen mother cells, microspores, and tapetum cells. Inhibition of *GHACS1* seriously affected the development of tapetum cells and the formation of normal microspores [53]. In cotton anthers, the genes were involved in fatty acid metabolism and peroxisome. Abnormal expression was identified in two genes in the male sterile line, which encoded Jojoba acyl-CoA reductase-related male sterility protein.

### 3.3. Major Metabolic Pathways Regulating Anther Development in Cotton

Carbohydrate metabolism, which provides carbon sources and energy for plant growth and development, is the most basic pathway of plant metabolism [115,116]. Carbohydrates provide energy and nutrients for anther development. The period of anther and pollen development are intensively energy-demanding [81,117]. Perturbed carbohydrate metabolism can significantly damage pollen development and cause male sterility [118,119].

It was recently reported that glycolysis was activated in the anthers of 1355A, which caused a decrease of soluble sugars, such as fructose and glucose, and accumulation of acetyl-CoA, which led to significant increases of c14:0 and c18:1 free fatty acids. The genes involved in fatty acid synthesis are crucial to regulate normal pollen hydration and plant fertility. High rates of glucose metabolism may promote fatty acid synthesis to enable anther growth [120].

Previous studies have shown that 1165 genes associated with flavonoid and ascorbate-glutathione cycle, starch and sucrose metabolism are important during anther development in cotton [19]. Using cotton anthers from wild type and GMS lines analyzed by digital gene expression (DGE), identified many of the essential genes required for cotton anther development. The genes were mainly associated with sucrose and starch metabolism, the pentose phosphate pathway, glycolysis, and flavonoid metabolism [20].

Some researchers identified 1742 significant DEGs between anthers of GMS line (B) and its maintainer line (K) in cotton plants by the DGE approach showed that sugar metabolism such as the interconversion of pentose and glucuronate, starch and sucrose metabolism, galactose metabolism were required for the development of cotton anthers [22].

In the present study, we identified 2013 genes that were differentially expressed between the male sterile line C2P5A and the maintainer line C2P5B using a DEG seq approach. These DEGs were analyzed by GO enrichment and KEGG pathways and were found to be associated with starch and sucrose metabolism, galactose metabolism, ascorbate and aldarate metabolism, glutathione metabolism, and pyruvate metabolism, pentose and glucuronate interconversions, and fatty acid biosynthesis and metabolism.

Proline interacts with carbohydrates to provide nutrients, promoting pollen development and pollen tube elongation. Proline also accompanies the development of microspores [121].

Soluble proteins contain enzymes involved in plant growth and development which play crucial roles in the development of anthers and microspores. However, soluble protein in the anther of male sterile line were deficient, resulting in abnormal protein synthesis in the anther [122].

In our research, biochemical determination of soluble sugar and soluble protein content of the CMS line C2P5A was lower than that of the maintainer line C2P5B at the three stages of anther development (Figure 3A–C), which as well as, lack of energy and nutrition may lead to pollen abortion in C2P5A.

## 4. Materials and Methods

### 4.1. Plant Materials

The cotton (*Gossypium hirsutum* L.) plants C2P5A (CMS line) and C2P5B (maintainer line) used in these experiments were cultivated in the experimental field in Henan Institute of Science and Technology, Henan, China. 2018–2019, April 27 sowing, line spacing 1 meter, plant spacing 40 centimeter, field watering was irrigated at the seedling stage, flowering stage and flower bud stage respectively. Organic fertilizer and compound fertilizers were used as field base fertilizers. During anthesis, both lines were identified according to flower apparatus morphology, the anther vigor with 2% TTC staining,

and anther development stage with 1.5% (*w/v*) acetocarmine staining. Plant lines were photographed under a microscope and SEM (Epson Expression 12000XL, Nagoya, Japan).

Flower buds had lengths of 3–4 mm (pollen mother cell stage), 4.1–5.0 mm (tetrad stage), 5.1–6.0 mm (mononuclear stage), and flower bud length (length nectary to bud apex, mm). Anthers were collected from three different periods of the CMS line and maintainer line. Sepals and petals were peeled off, pure anthers were collected, and stigmas and ovaries were discarded. The anthers were carefully and quickly isolated and frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until use.

#### 4.2. Histological Analysis

To identify the anther abortion period of the male sterile line C2P5A, flower buds of different periods were collected and fixed in FAA [10% formalin, 5% acetic acid, and 50% ethanol (*v/v*)].

Samples were sectioned in paraffin and fixed, washed, dyed-saffron, dehydrated, hyalinized, infiltrated, embedded, sectioned, exhibited, re-dyed-solid green, and sealed according to a previous report [75].

#### 4.3. Measurement of Physiological Indices

Buds were collected from three different periods of the CMS line C2P5A and the maintainer line C2P5B, and sepals and petals were removed. Soluble sugar was measured by anthrone [123], pure glucose was used as standard; Soluble protein was measured according to Lowry et al. [124], bovine serum albumin (BSA) was used as standard. Determination of MDA content was by thiobarbituric acid-reactive-substances (TBARS) assay [125], antioxidative enzymes (POD, CAT) were measured using quantificationally by ELISA instrument (TECAN, Tech Spark 10M, Zurich, Switzerland) [126]. There were three biological replicates per sample.

#### 4.4. Screening of DEGs and Co-expression Network Analysis of Unigenes

Total RNA was isolated from harvested anthers using an RNAPrep pure PlantKit (Tiangen, Beijing, China); with three biological replicates per sample. The RNA concentration was assessed by NanoDrop (Waltham, MA, USA). The RNA quality was tested using an Agilent 2100 instrument (Santa Clara, CA, USA) with 28S/18S rRNA ratio  $> 1.6$  (Agilent 2100) per sample for library construction on the BGISEQ-500 platform (BGI, Wuhan, China). SOAP [127], the short oligonucleotide alignment tool, matched raw reads to the ribosome database, a maximum of five mismatches were allowed. Reads with low-quality rRNA, containing adapters, joint contamination, and high levels of unknown bases ("N") were filtered to produce clean reads. Clean reads were matched to the reference genome (<http://cotton.hzau.edu.cn/EN/download.php>) and Cotton Gene database (<https://www.cottongen.org/>) by HISAT [69,128]. The transcriptional assembly was performed by StringTie [129] and Bowtie2 [67]. FPKM was used to calculate gene expression levels by RSEM [68]. DEGs were identified using DEGseq [70] in accordance with the threshold values:  $|\log_2\text{Ratio}| \geq 2$ ,  $\text{readnum} > 3$ ,  $p \leq 0.001$  and the criteria of  $\text{FDR} < 0.01$ . GO annotation and KEGG pathway analyses were performed for the DEGs.

The correlation network of the WGCNA R package is often used for correlation analysis between highly correlated DEG modules and samples to find modules of high specific expression in samples [130]. Module Eigengenes (ME) values were used to estimate associations between modules and the CMS phenotype.

#### 4.5. RT-PCR Analysis for Gene Expression

Total RNA was extracted using an RNA prep pure PlantKit (Aidlab Biotech, Beijing, China), according to the manufacturer's instructions. RNA was reverse transcribed using a TransScript<sup>®</sup> One-Step gDNA Removal kit and a cDNA Synthesis SuperMix kit (TransGen, Beijing, China). Quantification PCR (qRT-PCR) was used to confirm the RNA-seq data. Genes related to C2P5A male sterility of the transcriptome were selected. The specific primers for RT-PCR were designed using Primer Premier 6.0 software (<http://www.premierbiosoft.com/crm/jsp/com/pbi/crm/clientside/ProductList.jsp>)

and synthesized by Sangon Biotech (Shanghai, China) (Supplementary Table S7a,b). The reactions for RT-PCR were performed using cDNA as a template, using a qPCR SuperMix Kit (TransGen, Beijing, China). Each reaction was performed in three biological and three technical replicates on a QuantStudio 6 Flex instrument (Applied Biosystems, Foster city, CA, United States). RT-PCR analysis was performed according to the protocol of the TransStart<sup>®</sup> Top Green qPCR SuperMix with Two-Step kit (TransGen, Beijing, China).

The PCR circulation conditions included denaturation at 95 °C for 30 s, followed by 45 cycles of denaturation at 94 °C for 5 s, annealing and extension at 60 °C for 30 s. The melting curve was determined for each sample. Relative expression was calculated for every sample using the cycle threshold (Ct)2<sup>-ΔΔCt</sup> method [131].

## 5. Conclusions

In this study, we compared the cytology, the physiological characteristics and the transcriptome of the cotton CMS line C2P5A and its maintainer line C2P5B. Cytology and physiological characteristics results indicated that pollen abortion in C2P5A occurred at the tetrad stage. The transcriptome results revealed 2013 DEGs between C2P5A and C2P5B. Bioinformatic analyses showed that DEGs were mainly related to encoding ROS detoxification enzymes, tapetum proteins or transcription factors, vital metabolic pathways including starch and sucrose metabolism, galactose metabolism, ascorbate and aldarate metabolism, glutathione metabolism, and pyruvate metabolism. This research results provide a basic theory and molecular mechanism for CMS research and will accelerate research on CMS in cotton.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/20/5120/s1>.

**Author Contributions:** Q.W. and Q.C. conceived and designed the research; C.W., J.S., T.D. participated in material collection; R.Z. participated in morphological experiments, T.Q. revised the manuscript, and Y.L. performed the experiments and wrote the manuscript.

**Funding:** This work was supported by National Key Research and Development Program of Thirteenth (Project: 2016YFD0101413), Major Science and Technology Projects in Henan Province (Project: 161100510100) and Scientific and Technological Project of Henan Province (182102110003).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

CMS	Cytoplasmic male sterility
GMS	Genic male sterility
MS	Male sterility
DEG	Differentially expressed genes
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
RSEM	RNA-Seq by Expectation Maximization
RNA-seq	RNA-sequencing
qRT-PCR	Quantitative Reverse-Transcription PCR
DGE	Digital gene expression
WGCNA	Weighted correlation network
TTC	Triphenyltetrazolium chloride
SEM	Scanning electron microscopy
PCD	Programmed cell death

## References

1. Zhang, Z.; Li, J.; Muhammad, J.; Cai, J.; Jia, F.; Shi, Y.; Gong, J.; Shang, H.; Liu, A.; Chen, T.; et al. High Resolution Consensus Mapping of Quantitative Trait Loci for Fiber Strength, Length and Micronaire



- on Chromosome 25 of the Upland Cotton (*Gossypium hirsutum* L.). *PLoS ONE* **2015**, *10*, e0135430. [[CrossRef](#)] [[PubMed](#)]
2. Chen, Z.J.; Scheffler, B.E.; Dennis, E.; Triplett, B.A.; Zhang, T.; Guo, W.; Chen, X.; Stelly, D.M.; Rabinowicz, P.D.; Town, C.D.; et al. Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiol.* **2007**, *145*, 1303–1310. [[CrossRef](#)] [[PubMed](#)]
  3. Saha, S.; Raska, D.A.; Stelly, D.M. Breeding and Genetics Upland Cotton (*Gossypium hirsutum* L.) x Hawaiian Cotton (*G. tomentosum* Nutt. ex Seem.) F1 Hybrid Hypoaneuploid Chromosome Substitution Series. *J. Cotton Sci.* **2006**, *10*, 263–272.
  4. Samuel Yang, S.; Cheung, F.; Lee, J.J.; Ha, M.; Wei, N.E.; Sze, S.H.; Stelly, D.M.; Thaxton, P.; Triplett, B.; Town, C.D.; et al. Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. *Plant J.* **2006**, *47*, 761–775. [[CrossRef](#)]
  5. Saxena, K.B.; Kumar, R.V.; Tickle, A.N.; Saxena, M.K.; Gautam, V.S.; Rao, S.K.; Khare, D.K.; Chauhan, Y.S.; Saxena, R.K.; Reddy, B.V.S. ICPH 2671—The world's first commercial food legume hybrid. *Plant Breed.* **2013**, *132*, 479–485. [[CrossRef](#)]
  6. Bvs, R.; Ramesh, S.; Reddy, P.S. Sorghum breeding research at ICRISAT—goals, strategies, methods and accomplishments. *Int. Sorghum Millets Newsl.* **2004**, *2000*, 131–140.
  7. Longin, C.F.H.; Mühleisen, J.; Maurer, H.P.; Zhang, H.; Gowda, M.; Reif, J.C. Hybrid breeding in autogamous cereals. *Appl. Genet.* **2012**, *125*, 1087–1096. [[CrossRef](#)]
  8. Whitford, R.; Fleury, D.; Reif, J.C.; Garcia, M.; Okada, T.; Korzun, V.; Langridge, P. Hybrid breeding in wheat: Technologies to improve hybrid wheat seed production. *J. Exp. Bot.* **2013**, *64*, 5411–5428. [[CrossRef](#)]
  9. Chen, L.; Liu, Y.G. Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.* **2014**, *65*, 579–606. [[CrossRef](#)]
  10. Tester, M.; Langridge, P. Breeding technologies to increase crop production in a changing world. *Science* **2010**, *327*, 818–822. [[CrossRef](#)]
  11. Bohra, A.; Jha, U.C.; Adhimalam, P.; Bisht, D.; Singh, N.P. Cytoplasmic male sterility (CMS) in hybrid breeding in field crops. *Plant Cell. Rep.* **2016**, *35*, 967–993. [[CrossRef](#)] [[PubMed](#)]
  12. Mühleisen, J.; Maurer, H.P.; Stiewe, G.; Bury, P.; Reif, J.C. Hybrid Breeding in Barley. *Crop. Sci.* **2013**, *53*, 819–824. [[CrossRef](#)]
  13. Yu, S.; Fan, S.; Wang, H.; Wei, H.; Pang, C. Progresses in Research on Cotton High Yield Breeding in China. *Chin. Agric. Sci.* **2016**, *49*, 3465–3476.
  14. Yamagishi, H.; Bhat, S.R. Cytoplasmic male sterility in Brassicaceae crops. *Breed. Sci.* **2014**, *64*, 38–47. [[CrossRef](#)] [[PubMed](#)]
  15. Yang, Y.; Bao, S.; Zhou, X.; 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*, 209. [[CrossRef](#)]
  16. Kaul, M.L.H. *Male Sterility in Higher Plants*; Springer: Berlin/Heidelberg, Germany, 1988.
  17. Carlsson, J.; Leino, M.; Sohlberg, J.; Sundström, J.F.; Glimelius, K. Mitochondrial regulation of flower development. *Mitochondrion* **2008**, *8*, 74–86. [[CrossRef](#)]
  18. McCauley, D.E.; Bailey, M.F. Recent advances in the study of gynodioecy: The interface of theory and empiricism. *Ann. Bot.* **2009**, *104*, 611–620. [[CrossRef](#)]
  19. Ma, J.; Wei, H.; Song, M.; Pang, C.; Liu, J.; Wang, L.; Zhang, J.; Fan, S.; Yu, S. Transcriptome profiling analysis reveals that flavonoid and ascorbate-glutathione cycle are important during anther development in Upland cotton. *PLoS ONE* **2012**, *7*, e49244. [[CrossRef](#)]
  20. Wei, M.; Song, M.; Fan, S.; Yu, S. Transcriptomic analysis of differentially expressed genes during anther development in genetic male sterile and wild type cotton by digital gene-expression profiling. *BMC Genom.* **2013**, *14*, 97. [[CrossRef](#)]
  21. Engelke, T.; Hirsche, J.; Roitsch, T. Anther-specific carbohydrate supply and restoration of metabolically engineered male sterility. *J. Exp. Bot.* **2010**, *61*, 2693–2706. [[CrossRef](#)]
  22. Zhang, Y.; Chen, J.; Liu, J.; Xia, M.; Wang, W.; Shen, F. Transcriptome Analysis of Early Anther Development of Cotton Revealed Male Sterility Genes for Major Metabolic Pathways. *J. Plant Growth Regul.* **2015**, *34*, 223–232. [[CrossRef](#)]
  23. Edstam, M.M.; Edqvist, J. Involvement of GPI-anchored lipid transfer proteins in the development of seed coats and pollen in *Arabidopsis thaliana*. *Physiol. Plant* **2014**, *152*, 32–42. [[CrossRef](#)] [[PubMed](#)]

24. Salminen, T.A.; Blomqvist, K.; Edqvist, J. Lipid transfer proteins: Classification, nomenclature, structure, and function. *Planta* **2016**, *244*, 971–997. [[CrossRef](#)] [[PubMed](#)]
25. Wu, Y.; Min, L.; Wu, Z.; Yang, L.; Zhu, L.; Yang, X.; Yuan, D.; Guo, X.; Zhang, X. Defective pollen wall contributes to male sterility in the male sterile line 1355A of cotton. *Sci. Rep.* **2015**, *5*, 9608–9615. [[CrossRef](#)] [[PubMed](#)]
26. Noctor, G.; Foyer, C.H. Ascorbate and Glutathione: Keeping Active Oxygen Under Control. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998**, *49*, 249–279. [[CrossRef](#)] [[PubMed](#)]
27. Yang, L.; Wu, Y.; Zhang, M.; Zhang, J.; Stewart, J.M.; Xing, C.; Wu, J.; Jin, S. Transcriptome, cytological and biochemical analysis of cytoplasmic male sterility and maintainer line in CMS-D8 cotton. *Plant Mol. Biol.* **2018**, *97*, 537–551. [[CrossRef](#)] [[PubMed](#)]
28. Smirnoff, N.; Wheeler, G.L. Ascorbic acid in plants: Biosynthesis and function. *Crit. Rev. Biochem. Mol. Biol.* **2000**, *35*, 291–314. [[CrossRef](#)]
29. Filkowski, J.; Kovalchuk, O.; Kovalchuk, I. Genome stability of vtc1, tt4, and tt5 Arabidopsis thaliana mutants impaired in protection against oxidative stress. *Plant J.* **2004**, *38*, 60–69. [[CrossRef](#)]
30. Hsieh, K.; Huang, A.H. Tapetosomes in Brassica tapetum accumulate endoplasmic reticulum-derived flavonoids and alkanes for delivery to the pollen surface. *Plant Cell.* **2007**, *19*, 582–596. [[CrossRef](#)]
31. Toshiro, I.; Frank, W.; Hao, Y.; Pradeep, D.; Natsuko, I.; Márcio, A.F.; José Luis, R.; Meyerowitz, E.M. The homeotic protein AGAMOUS controls microsporogenesis by regulation of SPOROCTELESS. *Nature* **2004**, *430*, 356–360.
32. Shu-Lan, Y.; Li-Fen, X.; Hui-Zhu, M.; Ching San, P.; Wei-Cai, Y.; Lixi, J.; Venkatesan, S.; De, Y. Tapetum determinant1 is required for cell specialization in the Arabidopsis anther. *Plant Cell.* **2003**, *15*, 2792–2804.
33. Shu-Lan, Y.; Lixi, J.; Ching San, P.; Li-Fen, X.; Xue-Qin, Z.; Li-Qun, C.; Wei-Cai, Y.; De, Y. Overexpression of TAPETUM DETERMINANT1 alters the cell fates in the Arabidopsis carpel and tapetum via genetic interaction with excess microsporocytes1/extra sporogenous cells. *Plant Physiol.* **2005**, *139*, 186–191.
34. Da-Zhong, Z.; Guan-Fang, W.; Brooke, S.; Hong, M. The excess microsporocytes1 gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the Arabidopsis anther. *Genes Dev.* **2002**, *16*, 2021–2031.
35. Catherine, A.; Eugenia, R.; Valerie, H.; Erik, B.; Sacco, D.V. The Arabidopsis thaliana SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 control male sporogenesis. *Plant Cell.* **2005**, *17*, 3337–3349.
36. Hord, C.; Chen, C.; DeYoung, B.; Clark, S.; Ma, H. The BAM1/BAM2 receptor-like kinases are important regulators of Arabidopsis early anther development. *Plant Cell.* **2006**, *18*, 1667–1680. [[CrossRef](#)]
37. Liu, X.; Huang, J.; Parameswaran, S.; Ito, T.; Seubert, B.; Auer, M.; Rymaszewski, A.; Jia, G.; Owen, H.A.; Zhao, D. The SPOROCTELESS/NOZZLE Gene Is Involved in Controlling Stamen Identity in Arabidopsis. *Plant Physiol.* **2009**, *151*, 1401–1411. [[CrossRef](#)]
38. Hord, C.L.; Sun, Y.J.; Pillitteri, L.J.; Torii, K.U.; Wang, H.; Zhang, S.; Ma, H. Regulation of Arabidopsis Early Anther Development by the Mitogen-Activated Protein Kinases, MPK3 and MPK6, and the ERECTA and Related Receptor-Like Kinases. *Mol. Plant* **2008**, *1*, 645–658. [[CrossRef](#)]
39. Wei, Z.; Yujin, S.; Ljudmilla, T.; Changbin, C.; Ueli, G.; Hong, M. Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor. *Development* **2006**, *133*, 3085–3095.
40. Baomin, F.; Dihong, L.; Xuan, M.; Yiben, P.; Yujin, S.; Gang, N.; Hong, M. Regulation of the Arabidopsis anther transcriptome by DYT1 for pollen development. *Plant J. Cell. Mol. Biol.* **2012**, *72*, 612–624.
41. Mizuno, S.; Osakabe, Y.; Maruyama, K.; Ito, T.; Osakabe, K.; Sato, T.; Shinozaki, K.; Yamaguchi-Shinozaki, K. Receptor-like protein kinase 2 (RPK 2) is a novel factor controlling anther development in Arabidopsis thaliana. *Plant J.* **2010**, *50*, 751–766. [[CrossRef](#)]
42. Na, L.; Da-Sheng, Z.; Hai-Sheng, L.; Chang-Song, Y.; Xiao-Xing, L.; Wan-Qi, L.; Zheng, Y.; Ben, X.; Huang-Wei, C.; Jia, W. The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. *Plant Cell.* **2006**, *18*, 2999–3014.
43. Da-Sheng, Z.; Wan-Qi, L.; Zheng, Y.; Na, L.; Jing, S.; Jue, W.; Yu-Min, L.; Wen-Juan, Y.; Da-Bing, Z. Tapetum degeneration retardation is critical for aliphatic metabolism and gene regulation during rice pollen development. *Mol. Plant* **2008**, *1*, 599–610.

44. Fu, Z.; Yu, J.; Cheng, X.; Zong, X.; Xu, J.; Chen, M.; Li, Z.; Zhang, D.; Liang, W. The Rice Basic Helix-Loop-Helix Transcription Factor TDR INTERACTING PROTEIN2 Is a Central Switch in Early Anther Development. *Plant Cell*. **2014**, *26*, 1512–1524. [[CrossRef](#)] [[PubMed](#)]
45. Lilan, H.; Ding, T.; Yi, S.; Qing, H.; Kejian, W.; Ming, L.; Tiegang, L.; Zhukuan, C. MIL2 (MICROSPORELESS2) regulates early cell differentiation in the rice anther. *New Phytol.* **2012**, *196*, 402–413.
46. Lilan, H.; Ding, T.; Keming, Z.; Kejian, W.; Ming, L.; Zhukuan, C. Somatic and reproductive cell development in rice anther is regulated by a putative glutaredoxin. *Plant Cell*. **2012**, *24*, 577–588.
47. Zhu, E.; You, C.; Wang, S.; Cui, J.; Niu, B.; Wang, Y.; Qi, J.; Ma, H.; Chang, F. The DYT1-interacting proteins bHLH010, bHLH089 and bHLH091 are redundantly required for Arabidopsis anther development and transcriptome. *Plant J. Cell. Mol. Biol.* **2015**, *83*, 976–990. [[CrossRef](#)]
48. Lifang, H.; Wanqi, L.; Changsong, Y.; Xiao, C.; Jie, Z.; Xing, W.; Jianping, H.; Dabing, Z. Rice MADS3 regulates ROS homeostasis during late anther development. *Plant Cell*. **2011**, *23*, 515–533.
49. Li, H.; Yuan, Z.; Vizcay-Barrena, G.; Yang, C.; Liang, W.; Zong, J.; Wilson, Z.A.; Zhang, D. PERSISTENT TAPETAL CELL1 encodes a PHD-finger protein that is required for tapetal cell death and pollen development in rice. *Plant Physiol.* **2011**, *156*, 615–630. [[CrossRef](#)]
50. Niu, N.; Liang, W.; Yang, X.; Jin, W.; Wilson, Z.A.; Hu, J.; Zhang, D. EAT1 promotes tapetal cell death by regulating aspartic proteases during male reproductive development in rice. *Nat. Commun.* **2013**, *4*, 1445. [[CrossRef](#)]
51. Li, X.; Gao, X.; Wei, Y.; Deng, L.; Ouyang, Y.; Chen, G.; Li, X.; Zhang, Q.; Wu, C. Rice APOPTOSIS INHIBITOR5 coupled with two DEAD-box adenosine 5'-triphosphate-dependent RNA helicases regulates tapetum degeneration. *Plant Cell*. **2011**, *23*, 1416–1434. [[CrossRef](#)]
52. Luo, D.; Xu, H.; Liu, Z.; Guo, J.; Li, H.; Chen, L.; Fang, C.; Zhang, Q.; Bai, M.; Yao, N.; et al. A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. *Nat. Genet.* **2013**, *45*, 573–577. [[CrossRef](#)] [[PubMed](#)]
53. Wang, X.L.; Li, X.B. The GhACS1 gene encodes an acyl-CoA synthetase which is essential for normal microsporogenesis in early anther development of cotton. *Plant J.* **2009**, *57*, 473–486. [[CrossRef](#)] [[PubMed](#)]
54. Li, Y.; Jiang, J.; Du, M.L.; Li, L.; Wang, X.L.; Li, X.B. A cotton gene encoding MYB-like transcription factor is specifically expressed in pollen and is involved in regulation of late anther/pollen development. *Plant Cell. Physiol.* **2013**, *54*, 893–906. [[CrossRef](#)] [[PubMed](#)]
55. He, L.; Hannon, G.J. MicroRNAs: Small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **2004**, *5*, 522–531. [[CrossRef](#)]
56. Semwal, V.K.; Singh, B.; Khanna-Chopra, R. Delayed expression of SAGs correlates with longevity in CMS wheat plants compared to its fertile plants. *Physiol. Mol. Biol. Plants* **2014**, *20*, 191–199. [[CrossRef](#)]
57. Duan, J.; Liang, C.Y.; Zhang, M.Y. The relationship between membrane lipid peroxidation and cytoplasmic male sterility in maize. *Plant Physiol. Commun.* **1996**, *32*, 331–334.
58. Jiang, P.; Zhang, X.; Zhu, Y.; Zhu, W.; Xie, H.; Wang, X. Metabolism of reactive oxygen species in cotton cytoplasmic male sterility and its restoration. *Plant Cell. Rep.* **2007**, *26*, 1627–1634. [[CrossRef](#)]
59. Yi, J.; Moon, S.; Lee, Y.S.; Zhu, L.; Liang, W.; Zhang, D.; Jung, K.H.; An, G. Defective Tapetum Cell Death 1 (DTC1) Regulates ROS Levels by Binding to Metallothionein during Tapetum Degeneration. *Plant Physiol.* **2016**, *170*, 1611–1623. [[CrossRef](#)]
60. Chase, C.D. Cytoplasmic male sterility: A window to the world of plant mitochondrial-nuclear interactions. *Trends Genet.* **2007**, *23*, 81–90. [[CrossRef](#)]
61. Suzuki, H.; Rodriguez-Urbe, L.; Xu, J.; Zhang, J. Transcriptome analysis of cytoplasmic male sterility and restoration in CMS-D8 cotton. *Plant Cell. Rep.* **2013**, *32*, 1531–1542. [[CrossRef](#)]
62. Chen, Z.; Nan, Z.; Li, S.; Grover, C.E.; Nie, H.; Wendel, J.F.; Hua, J. Plant Mitochondrial Genome Evolution and Cytoplasmic Male Sterility. *Crit. Rev. Plant Sci.* **2017**, *36*, 1–15. [[CrossRef](#)]
63. He, S.; Lyznik, A.; Mackenzie, S. Pollen fertility restoration by nuclear gene Fr in CMS bean: Nuclear-directed alteration of a mitochondrial population. *Genetics* **1995**, *139*, 955–962. [[PubMed](#)]
64. Yi, P.; Wang, L.; Sun, Q.; Zhu, Y. Discovery of mitochondrial chimeric-gene associated with cytoplasmic male sterility of HL-rice. *Chin. Sci. Bull.* **2002**, *47*, 744–747. [[CrossRef](#)]
65. Dewey, R.E.; Timothy, D.H.; Iii, C.S.L. Chimeric mitochondrial genes expressed in the C male-sterile cytoplasm of maize. *Curr. Genet.* **1991**, *20*, 475–482. [[CrossRef](#)] [[PubMed](#)]

66. Li, S.; Chen, Z.; Zhao, N.; Wang, Y.; Nie, H.; Hua, J. The comparison of four mitochondrial genomes reveals cytoplasmic male sterility candidate genes in cotton. *BMC Genom.* **2018**, *19*, 775. [[CrossRef](#)]
67. Finn, R.D.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [[CrossRef](#)]
68. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* **2011**, *12*, 323. [[CrossRef](#)]
69. Wang, M.; Tu, L.; Yuan, D.; Zhu, Shen, C.; Li, J.; Liu, F.; Pei, L.; Wang, P.; Zhao, G.; et al. Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense*. *Nat. Genet.* **2019**, *51*, 224–229. [[CrossRef](#)]
70. Wang, L.; Feng, Z.; Wang, X.; Wang, X.; Zhang, X. DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **2010**, *26*, 136–138. [[CrossRef](#)]
71. Zhang, B.; Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **2005**, *4*, 17.
72. Hollender, C.A.; Kang, C.; Darwish, O.; Geretz, A.; Matthews, B.F.; Slovin, J.; Alkharouf, N.; Liu, Z. Floral Transcriptomes in Woodland Strawberry Uncover Developing Receptacle and Anther Gene Networks. *Plant Physiol.* **2014**, *165*, 1062–1075. [[CrossRef](#)] [[PubMed](#)]
73. Kunbo, W.; Zhiwen, W.; Fuguang, L.; Wuwei, Y.; Junyi, W.; Guoli, S.; Zhen, Y.; Lin, C.; Haihong, S.; Shilin, Z. The draft genome of a diploid cotton *Gossypium raimondii*. *Nat. Genet.* **2012**, *44*, 1098–1103.
74. Fuguang, L.; Guangyi, F.; Kunbo, W.; Fengming, S.; Youlu, Y.; Guoli, S.; Qin, L.; Zhiying, M.; Cairui, L.; Changsong, Z. Genome sequence of the cultivated cotton *Gossypium arboreum*. *Nat. Genet.* **2014**, *3*, 567–572.
75. Kong, X.; Liu, D.; Liao, X.; Zheng, J.; Diao, Y.; Liu, Y.; Zhou, R. Comparative Analysis of the Cytology and Transcriptomes of the Cytoplasmic Male Sterility Line H276A and Its Maintainer Line H276B of Cotton (*Gossypium barbadense* L.). *Int. J. Mol. Sci.* **2017**, *18*, 2240. [[CrossRef](#)]
76. Liu, J.; Pang, C.; Wei, H.; Song, M.; Meng, Y.; Ma, J.; Fan, S.; Yu, S. iTRAQ-facilitated proteomic profiling of anthers from a photosensitive male sterile mutant and wild-type cotton (*Gossypium hirsutum* L.). *J. Proteom.* **2015**, *126*, 68–81. [[CrossRef](#)]
77. Rao, Y. Cytology of cytoplasmic male sterile lines in hybrid rice. Hybrid rice Proceedings of an international symposium, Changsha, China, October 1986; pp. 115–128.
78. Wang, X.; Pan, J. Genetic basis of fertility restoration to cytoplasmic male sterile lines available in China in upland cotton II. Interactive effects between restorer genes and fertility enhancer gene. *Acta. Genet. Sin.* **1997**, *24*, 276–277.
79. Sanders, P.M.; Bui, A.Q.; Weterings, K.; Mcintire, K.N.; Hsu, Y.C.; Pei, Y.L.; Mai, T.T.; Beals, T.P.; Goldberg, R.B. Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. *Sex. Plant Reprod.* **1999**, *11*, 297–322. [[CrossRef](#)]
80. Budar, F.; Pelletier, G. Male sterility in plants: Occurrence, determinism, significance and use. *C R Acad. Sci. Iii.* **2001**, *324*, 543–550. [[CrossRef](#)]
81. Linke, B.; Borner, T. Mitochondrial effects on flower and pollen development. *Mitochondrion* **2005**, *5*, 389–402. [[CrossRef](#)]
82. Touzet, P.; Budar, F. Unveiling the molecular arms race between two conflicting genomes in cytoplasmic male sterility? *Trends Plant Sci.* **2004**, *9*, 568–570. [[CrossRef](#)]
83. Green, D.R.; Reed, J.C. Mitochondria and apoptosis. *Science* **1998**, *281*, 1309–1312. [[CrossRef](#)] [[PubMed](#)]
84. Foreman, J.; Demidchik, V.; Bothwell, J.H.; Mylona, P.; Miedema, H.; Torres, M.A.; Linstead, P.; Costa, S.; Brownlee, C.; Jones, J.D.; et al. Reactive oxygen species produced by NADPH oxidase regulate plant cell growth. *Nature* **2003**, *422*, 442–446. [[CrossRef](#)] [[PubMed](#)]
85. Choudhury, S.; Panda, P.; Sahoo, L.; Panda, S.K. Reactive oxygen species signaling in plants under abiotic stress. *Plant Signal. Behav.* **2013**, *8*, e23681. [[CrossRef](#)] [[PubMed](#)]
86. Li, J.; Dai, X.; Li, L.; Jiao, Z.; Huang, Q. Metabolism of reactive oxygen species in cytoplasmic male sterility of rice by marking upmost pulvinus interval. *Appl. Biochem. Biotechnol.* **2015**, *175*, 1263–1269. [[CrossRef](#)]
87. Ming-Hua, D.; Jin-Fen, W.; Jin-Long, H.; Hai-Shan, Z.; Xiong-Ze, D.; Zhu-Qing, Z.; Xue-Xiao, Z. Relationship of metabolism of reactive oxygen species with cytoplasmic male sterility in pepper (*Capsicum annuum* L.). *Sci. Hortic.* **2012**, *134*, 232–236.

88. Liu, Z.; Shi, X.; Li, S.; Hu, G.; Zhang, L.; Song, X. Tapetal-Delayed Programmed Cell Death (PCD) and Oxidative Stress-Induced Male Sterility of *Aegilops uniaristata* Cytoplasm in Wheat. *Int. J. Mol. Sci.* **2018**, *19*, 1708. [[CrossRef](#)]
89. Sofo, A.; Scopa, A.; Nuzzaci, M.; Vitti, A. Ascorbate Peroxidase and Catalase Activities and Their Genetic Regulation in Plants Subjected to Drought and Salinity Stresses. *Int. J. Mol. Sci.* **2015**, *16*, 13561–13578. [[CrossRef](#)]
90. Wan, C.; Li, S.; Wen, L.; Kong, J.; Wang, K.; Zhu, Y. Damage of oxidative stress on mitochondria during microspores development in Honglian CMS line of rice. *Plant. Cell. Rep.* **2007**, *26*, 373–382. [[CrossRef](#)]
91. Guo, J.; Wang, P.; Cheng, Q.; Sun, L.; Wang, H.; Wang, Y.; Kao, L.; Li, Y.; Qiu, T.; Yang, W.; et al. Proteomic analysis reveals strong mitochondrial involvement in cytoplasmic male sterility of pepper (*Capsicum annuum* L.). *J. Proteom.* **2017**, *168*, 15–27. [[CrossRef](#)]
92. op den Camp, R.G.; Kuhlemeier, C. Aldehyde dehydrogenase in tobacco pollen. *Plant. Mol. Biol.* **1997**, *35*, 355–365. [[CrossRef](#)]
93. Ducos, E.; Touzet, P.; Boutry, M. The male sterile G cytoplasm of wild beet displays modified mitochondrial respiratory complexes. *Plant J.* **2001**, *26*, 171–180. [[CrossRef](#)] [[PubMed](#)]
94. Kim, D.H.; Kang, J.G.; Kim, B.D. Isolation and characterization of the cytoplasmic male sterility-associated orf456 gene of chili pepper (*Capsicum annuum* L.). *Plant Mol. Biol.* **2007**, *63*, 519–532. [[CrossRef](#)] [[PubMed](#)]
95. Greenberg, J.T.; Yao, N. The role and regulation of programmed cell death in plant-pathogen interactions. *Cell. Microbiol.* **2004**, *6*, 201–211. [[CrossRef](#)] [[PubMed](#)]
96. Banting, G.S.; Glerum, D.M. Mutational analysis of the *Saccharomyces cerevisiae* cytochrome c oxidase assembly protein Cox11p. *Eukaryot. Cell.* **2006**, *5*, 568–578. [[CrossRef](#)] [[PubMed](#)]
97. Veniamin, S.; Sawatzky, L.G.; Banting, G.S.; Glerum, D.M. Characterization of the peroxide sensitivity of COX-deficient yeast strains reveals unexpected relationships between COX assembly proteins. *Free Radic. Biol. Med.* **2011**, *51*, 1589–1600. [[CrossRef](#)] [[PubMed](#)]
98. Li, H.; Pinot, F.; Sauveplane, V.; Werck-Reichhart, D.; Diehl, P.; Schreiber, L.; Franke, R.; Zhang, P.; Chen, L.; Gao, Y.; et al. Cytochrome P450 family member CYP704B2 catalyzes the -hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. *Plant Cell.* **2010**, *22*, 173–190. [[CrossRef](#)] [[PubMed](#)]
99. Gu, J.N.; Zhu, J.; Yu, Y.; Teng, X.D.; Lou, Y.; Xu, X.F.; Liu, J.L.; Yang, Z.N. DYT1 directly regulates the expression of TDF1 for tapetum development and pollen wall formation in Arabidopsis. *Plant J.* **2014**, *80*, 1005–1013. [[CrossRef](#)]
100. Andriankaja, M.E.; Danisman, S.; Mignolet-Spruyt, L.F.; Claeys, H.; Kochanek, I.; Vermeersch, M.; De Milde, L.; De Bodt, S.; Storme, V.; Skirycz, A.; et al. Transcriptional coordination between leaf cell differentiation and chloroplast development established by TCP20 and the subgroup Ib bHLH transcription factors. *Plant Mol. Biol.* **2014**, *85*, 233–245. [[CrossRef](#)]
101. Sorensen, A.M.; Krober, S.; Unte, U.S.; Huijser, P.; Dekker, K.; Saedler, H. The Arabidopsis ABORTED MICROSPORES (AMS) gene encodes a MYC class transcription factor. *Plant J.* **2003**, *33*, 413–423. [[CrossRef](#)]
102. Nakata, M.; Ohme-Takagi, M. Two bHLH-type transcription factors, JA-ASSOCIATED MYC2-LIKE2 and JAM3, are transcriptional repressors and affect male fertility. *Plant Signal. Behav.* **2013**, *8*, e26473. [[CrossRef](#)]
103. Jung, K.H.; Han, M.J.; Lee, Y.S.; Kim, Y.W.; Hwang, I.; Kim, M.J.; Kim, Y.K.; Nahm, B.H.; An, G. Rice Undeveloped Tapetum1 is a major regulator of early tapetum development. *Plant Cell.* **2005**, *17*, 2705–2722. [[CrossRef](#)] [[PubMed](#)]
104. Ko, S.S.; Chan, M.T. The bHLH142 Transcription Factor Coordinates with TDR1 to Modulate the Expression of EAT1 and Regulate Pollen Development in Rice. *Plant Cell.* **2014**, *26*, 2486–2504. [[CrossRef](#)] [[PubMed](#)]
105. Wen, J.; Wang, L.; Wang, J.; Zeng, Y.; Xu, Y.; Li, S. The transcription factor OsbHLH138 regulates thermosensitive genic male sterility in rice via activation of TMS5. *Appl. Genet.* **2019**, *132*, 1721–1723. [[CrossRef](#)] [[PubMed](#)]
106. Zhang, Z.B.; Zhu, J.; Gao, J.F.; Wang, C.; Li, H.; Li, H.; Zhang, H.Q.; Zhang, S.; Wang, D.M.; Wang, Q.X.; et al. Transcription factor AtMYB103 is required for anther development by regulating tapetum development, callose dissolution and exine formation in Arabidopsis. *Plant J.* **2007**, *52*, 528–538. [[CrossRef](#)] [[PubMed](#)]
107. Millar, A.A.; Gubler, F. The Arabidopsis GAMBYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell.* **2005**, *17*, 705–721. [[CrossRef](#)] [[PubMed](#)]

108. Zhu, J.; Chen, H.; Li, H.; Gao, J.F.; Jiang, H.; Wang, C.; Guan, Y.F.; Yang, Z.N. Defective in Tapetal development and function 1 is essential for anther development and tapetal function for microspore maturation in Arabidopsis. *Plant J.* **2008**, *55*, 266–277. [[CrossRef](#)] [[PubMed](#)]
109. Mandaokar, A.; Thines, B.; Shin, B.; Lange, B.M.; Choi, G.; Koo, Y.J.; Yoo, Y.J.; Choi, Y.D.; Choi, G.; Browse, J. Transcriptional regulators of stamen development in Arabidopsis identified by transcriptional profiling. *Plant J.* **2006**, *46*, 984–1008. [[CrossRef](#)]
110. Andrews, J.; Keegstra, K. Acyl-CoA Synthetase Is Located in the Outer Membrane and Acyl-CoA Thioesterase in the Inner Membrane of Pea Chloroplast Envelopes. *Plant Physiol.* **1983**, *72*, 735–740. [[CrossRef](#)]
111. Ichihara, K.I.; Nakagawa, M.; Tanaka, K. Acyl-CoA Synthetase in Maturing Safflower Seeds. *Plant Cell. Physiol.* **1993**, *34*, 557–566.
112. Judy, S.; Jay, S.; John, B. The acyl-CoA synthetase encoded by LACS2 is essential for normal cuticle development in Arabidopsis. *Plant Cell.* **2004**, *16*, 629–642.
113. Fulda, M.; Shockey, J.; Werber, M.; Wolter, F.P.; Heinz, E. Two long-chain acyl-CoA synthetases from Arabidopsis thaliana involved in peroxisomal fatty acid beta-oxidation. *Plant J.* **2002**, *32*, 93–103. [[CrossRef](#)] [[PubMed](#)]
114. Katja, S.; Lucie, K.; Elmon, S.; Thomas, C.; Michael, B.; Otto, M.; Claus, W.; Erich, K.; Hans-Peter, S. A new type of peroxisomal acyl-coenzyme A synthetase from Arabidopsis thaliana has the catalytic capacity to activate biosynthetic precursors of jasmonic acid. *J. Biol. Chem.* **2005**, *280*, 13962–13972.
115. Wu, Z.; Cheng, J.; Qin, C.; Hu, Z.; Yin, C.; Hu, K. Differential proteomic analysis of anthers between cytoplasmic male sterile and maintainer lines in *Capsicum annuum* L. *Int. J. Mol. Sci.* **2013**, *14*, 22982–22996. [[CrossRef](#)] [[PubMed](#)]
116. Kretschmer, M.; Croll, D.; Kronstad, J.W. Maize susceptibility to *Ustilago maydis* is influenced by genetic and chemical perturbation of carbohydrate allocation. *Mol. Plant Pathol.* **2017**, *18*, 1222–1237. [[CrossRef](#)] [[PubMed](#)]
117. Tang, H.; Song, Y.; Guo, J.; Wang, J.; Zhang, L.; Niu, N.; Ma, S.; Zhang, G.; Zhao, H. Physiological and metabolome changes during anther development in wheat (*Triticum aestivum* L.). *Plant Physiol. Biochem.* **2018**, *132*, 18–32. [[CrossRef](#)]
118. Datta, R.; Chamusco, K.C.; Chourey, P.S. Starch biosynthesis during pollen maturation is associated with altered patterns of gene expression in maize. *Plant Physiol.* **2002**, *130*, 1645–1656. [[CrossRef](#)]
119. Mamun, E.A.; Alfred, S.; Cantrill, L.C.; Overall, R.L.; Sutton, B.G. Effects of chilling on male gametophyte development in rice. *Cell. Biol. Int.* **2006**, *30*, 583–591. [[CrossRef](#)]
120. Wu, Y.; Li, Y.; Li, Y.; Ma, Y.; Zhao, Y.; Wang, C.; Chi, H.; Chen, M.; Ding, Y.; Guo, X.; et al. Proteomic analysis reveals that sugar and fatty acid metabolisms play a central role in sterility of the male-sterile line 1355A of cotton. *J. Biol. Chem.* **2019**, *294*, 7057–7067. [[CrossRef](#)]
121. Goss, J.A. Development, physiology, and biochemistry of corn and wheat pollen. *Bot. Rev.* **1968**, *34*, 333–358. [[CrossRef](#)]
122. Mascarenhas, J.P. The Biochemistry of Angiosperm Pollen Development. *Bot. Rev.* **1975**, *41*, 259–314. [[CrossRef](#)]
123. Yemm, E.W.; Willis, A.J. The estimation of carbohydrates in plant extracts by anthrone. *Biochem. J.* **1954**, *57*, 508–514. [[CrossRef](#)] [[PubMed](#)]
124. Lowry, O.H.; Rosebrough, N.J.; Farr, A.L.; Randall, R.J. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **1951**, *193*, 265–275. [[PubMed](#)]
125. Landi, M. Commentary to: “Improving the thiobarbituric acid-reactive-substances assay for estimating lipid peroxidation in plant tissues containing anthocyanin and other interfering compounds” by Hodges et al., *Planta*. (1999) 207:604–611. *Planta* **2017**, *245*, 1067. [[CrossRef](#)] [[PubMed](#)]
126. Bibi, N.; Yuan, S.; Zhu, Y.; Wang, X. Improvements of Fertility Restoration in Cytoplasmic Male Sterile Cotton by Enhanced Expression of Glutathione S-Transferase (GST) Gene. *J. Plant Growth Regul.* **2014**, *33*, 420–429. [[CrossRef](#)]
127. Li, R.; Li, Y.; Kristiansen, K.; Wang, J. SOAP: Short oligonucleotide alignment program. *Bioinformatics* **2008**, *24*, 713–714. [[CrossRef](#)] [[PubMed](#)]
128. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [[CrossRef](#)]

129. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [[CrossRef](#)]
130. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinf.* **2008**, *9*, 559. [[CrossRef](#)]
131. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_t}$  Method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).