



Article

A Rapid Pipeline for Pollen- and Anther-Specific Gene Discovery Based on Transcriptome Profiling Analysis of Maize Tissues

Yannan Shi ¹, Yao Li ¹, Yongchao Guo ¹, Eli James Borrego ² , Zhengyi Wei ³ , Hong Ren ⁴, Zhengqiang Ma ^{1,5} and Yuanxin Yan ^{1,5,*}

- ¹ State Key Laboratory for Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China; 2017201046@njau.edu.cn (Y.S.); 2018101126@njau.edu.cn (Y.L.); 2017101142@njau.edu.cn (Y.G.); zqm2@njau.edu.cn (Z.M.)
- ² Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, NY 14623, USA; ejbsbi@rit.edu
- ³ Institute of Agricultural Biotechnology, Jilin Academy of Agricultural Sciences, Changchun 130033, China; weizy80@163.com
- ⁴ Guizhou Institute of Upland Food Crops, Guizhou Academy of Agricultural Sciences, Guiyang 550001, China; rhong666@163.com
- ⁵ Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing Agricultural University, Nanjing 210095, China
- * Correspondence: yuanxin.yan@njau.edu.cn; Tel.: +86-25-84399517



Citation: Shi, Y.; Li, Y.; Guo, Y.; Borrego, E.J.; Wei, Z.; Ren, H.; Ma, Z.; Yan, Y. A Rapid Pipeline for Pollen- and Anther-Specific Gene Discovery Based on Transcriptome Profiling Analysis of Maize Tissues. *Int. J. Mol. Sci.* **2021**, *22*, 6877. <https://doi.org/10.3390/ijms22136877>

Academic Editors: Stefano Del Duca, Giampiero Cai and Iris Aloisi

Received: 17 April 2021

Accepted: 16 June 2021

Published: 26 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Recently, crop breeders have widely adopted a new biotechnology-based process, termed Seed Production Technology (SPT), to produce hybrid varieties. The SPT does not produce nuclear male-sterile lines, and instead utilizes transgenic SPT maintainer lines to pollinate male-sterile plants for propagation of nuclear-recessive male-sterile lines. A late-stage pollen-specific promoter is an essential component of the pollen-inactivating cassette used by the SPT maintainers. While a number of plant pollen-specific promoters have been reported so far, their usefulness in SPT has remained limited. To increase the repertoire of pollen-specific promoters for the maize community, we conducted a comprehensive comparative analysis of transcriptome profiles of mature pollen and mature anthers against other tissue types. We found that maize pollen has much less expressed genes (>1 FPKM) than other tissue types, but the pollen grain has a large set of distinct genes, called pollen-specific genes, which are exclusively or much higher (100 folds) expressed in pollen than other tissue types. Utilizing transcript abundance and correlation coefficient analysis, 1215 mature pollen-specific (MPS) genes and 1009 mature anther-specific (MAS) genes were identified in B73 transcriptome. These two gene sets had similar GO term and KEGG pathway enrichment patterns, indicating that their members share similar functions in the maize reproductive process. Of the genes, 623 were shared between the two sets, called mature anther- and pollen-specific (MAPS) genes, which represent the late-stage pollen-specific genes of the maize genome. Functional annotation analysis of MAPS showed that 447 MAPS genes (71.7% of MAPS) belonged to genes encoding pollen allergen protein. Their 2-kb promoters were analyzed for *cis*-element enrichment and six well-known pollen-specific *cis*-elements (AGAAA, TCCACCA, TGTGGTT, [TA]AAAG, AAATGA, and TTTCT) were found highly enriched in the promoters of MAPS. Interestingly, JA-responsive *cis*-element GCC box (GCCGCC) and ABA-responsive *cis*-element-coupling element1 (ABRE-CE1, CCACC) were also found enriched in the MAPS promoters, indicating that JA and ABA signaling likely regulate pollen-specific MAPS expression. This study describes a robust and straightforward pipeline to discover pollen-specific promoters from publicly available data while providing maize breeders and the maize industry a number of late-stage (mature) pollen-specific promoters for use in SPT for hybrid breeding and seed production.

Keywords: Seed Production Technology; crop breeding; hybrids; male-sterile; maize

1. Introduction

In flowering plants, pollen arises from the sporogenous cells in the loculi of the anther. The sporogenous cells are surrounded by the tapetum, which provides nutrients for sporogenous cells and secretes enzymes for the release of microspores from tetrads at the tetrad stage [1–3]. Based on the anther morphology and tissue structure and cellular events, the anther/pollen development process of *Arabidopsis* has been divided into two phases [1]. During phase I, the anther primordia undergo a series of prescribed cell divisions to form a young anther. This phase includes three early stages of microspores development: archesporial cells (AR), meiocytes (ME), and the quartet stage of microspores (Q). By the end of this phase, tetrads formed and young microspores released from the tetrads. During phase II, the young anthers expand to full-sized anthers which contain mature pollens to be released. Phase II includes three stages of pollen development: uninucleate microspore (UM), binucleate microspore (BM), and mature pollen (MP) [1]. These two phases have been further divided into 14 successive stages using morphological and cellular features viewed under light microscopy [4]. Stages 1–8 make up phase I and stages 9–14 constitute phase II of anther development [4]. For cereals such as maize and rice, anther development has been shown to undergo 14 successive stages, similar to *Arabidopsis*, to produce mature pollen [5–7]. These stages cover four key developmental events, i.e., AR cell specification (phase I, including stages 1–2), anther somatic cell division (phase II, including stages 3–5), tapetum (TA) development and meiosis of pollen mother cell (PMC) (phase III, including stages 6–8), and mature pollen formation and anther dehiscence (phase IV, including stages 9–14) [7]. By the end of phase IV, i.e., at the anthesis time, the mature pollen grains (trinucleate microspore, TM) are released from the stomia of maize anther.

The differentiation and development of the male gametophyte of flowering plants depend on the coordination of thousands of genes in different cells/tissues of the anther as well as in microspores after tetrad formation [8]. Gene expression in the anther has been studied intensively in plant species by methods such as conventional cDNA cloning, in situ hybridization, microarrays, and RNA-seq [8–21]. Microarray-based transcriptome profiling of four stages of rice anther development revealed that ~22,000 genes were expressed in at least one of the anther developmental stages and that mature pollen (MP) has the lowest number of expressed genes (15,465) [8]. In maize, 44k microarrays transcriptome profiling showed that more than 24,400 genes were expressed in at least one of the six anther developmental stages or mature pollen [16]. Maize anthers possess a varied number of detectable transcripts at different developmental stages. At each of the three early stages of anther development (anther size of A1.0, and A1.5, prior to meiosis at A2.0) [16], anthers expressed more than 20,000 transcripts, followed by a 10% decrease by the end of meiosis [16]. Post-meiotic anthers again expressed about 20,000 transcripts and mature pollen expressed about half that number [16]. Regarding pollen formation, the microspore developmental process can be divided into nine consecutive stages: archesporial cell-forming (ACF), pre-meiosis of microspore mother cells (PMe), meiotic leptotene of PMC (microspore mother cells) (Me1), meiotic zygotene-pachytene of PMC (Me2), meiotic diplotene-tetrad of PMC (Me3), uni-nucleate microspore (UM), bi-nucleate microspore (BM), tri-nucleate microspore (TM), and mature pollen grain (MP) [22,23]. The genes expressed specifically in microspores after the uni-nucleate microspore stage (from BM to MP) are known as late-stage pollen-specific genes [22–25].

Male sterility is an important tool for plant breeding to create hybrid varieties. Male-sterile mutations are largely due to abnormal development of either the sporophytic or gametophytic tissues of the anther [26]. Tapetum, a key sporophytic tissue of anther wall, provides nutrients for microspore development, and its dysfunction or delayed degeneration causes microspore abortion [3,26]. A number of studies have shown that anther-specific or tapetum-specific genes are highly linked to male sterility in plants [27–34] and screening for anther-specific genes is an effective strategy to isolate new genes relevant to male sterility. Additionally, anther-specific promoters are widely applied in genetic engineering to create male-sterile lines by driving a lethal gene such as *barnase*, a ribonuclease-

coding gene from *Bacillus amyloliquefaciens* [35]. In rice, microarray analysis of anther tissues revealed that 156 genes were specifically or predominantly expressed in anther tissues [32]. Using Rice Expression Profile Database (RiceXPro, <http://ricexpro.dna.affrc.go.jp/> accessed on 18 June 2021), Akasaka et al. (2018) identified 38 anther-specific genes in the rice genome [35]. In this publication, anther-specific genes were assigned as extremely high/or high expression levels in anthers but no or very low expression in other tissues such as leaf blade, leaf sheath, root, stem, panicle, ovary, pistil, embryo, and endosperm. Seven of these 38 promoters have been tested in their effectiveness for driving *barnase* to create dominant male sterile lines [35]. In maize, *Zm401p10*, was shown to be anther-specific and essential for tapetum degeneration during anther development [36]. Microarray analysis of six anther stages plus mature pollen found 1952 genes specifically expressed in the early stages of anther development (anther 1.0 mm prior to meiosis) and 1921 genes specifically expressed in the latter stages (meiosis to mature pollen) [16].

Pollen grains were found to contain a substantial abundance of transcripts [37–39], which are expected to have important functions during pollen development [40]. These transcripts are produced from more than 20,000 different genes [8,16,41]. Although most of the genes expressed in pollen may function in its development, only a portion (10–20%) have displayed pollen-specifically expressed, i.e., intensively expressed in pollen but marginal expression in the sporophyte tissues [16,41,42].

A number of pollen-specific genes have been isolated from various plant species such as petunia *chiA* [43], tobacco *TA29* [9], *Brassica napus BP4* [44], tomato *LAT52* [45], tobacco *NTP303* [46], maize *Zm13* [47], rice *RA8* [13], and *Arabidopsis MSP1/2/3* [48]. Expression of pollen-specific genes are driven by pollen-specific promoters, many of which have been extensively studied and a number of *cis*-regulatory elements of these promoters have been characterized [13,47,49–52]. The utilization of pollen-specific promoters is an effective tool for generating transgenic male sterile lines or maintainers of natural male sterile lines [53,54], and thus identification of novel pollen-specific promoters across agriculturally-relevant species is of high significance for plant breeding. High-throughput screening of pollen-specific genes is expected to be quickly implemented in plant breeding and genetic engineering for crop improvement. In the last decade, the large-scale microarray or RNA-sequencing technique has been widely applied for profiling transcripts of various plant tissues. The large-scale or genome-wide transcript profiling of reproductive tissues has been reported in *Arabidopsis* [39,55,56], tobacco [57–59], rice [8,23,25,60,61], maize [16,62–66], and other species [42,67,68]. The newly developed technology for crop hybrid seed production, called Seed Production Technology (SPT), creates “maintainers” to propagate the seeds of recessive genic male sterile (*ms*) lines [54]. These maintainers contain a transgenic pollen-inactivation cassette which kills the pollen with genotype of dominant *MS* transgene but not of recessive *ms* gene in the late stages (at or after the binucleate and trinucleate stages) of pollen development [54]. Pollen-inactivation occurs from a late-stage pollen-specific promoter driving a “pollen-killer” gene such as *zm-aa1*, α -amylase gene [54]. Therefore, the late-stage pollen-specific promoters in plants are important resources for SPT application in plant breeding. Transcriptome-profiling is an effective approach to discover a large number of late-stage pollen-specific promoters. To the best of our knowledge, the transcriptome features of the matured anther and/or pollen of maize in comparison with other sporophytic tissues have not yet been clarified. In this study, we conducted genome-wide transcriptome analysis of maize mature anther and pollen tissue from publicly-accessible RNA-seq datasets [69,70] and screened hundreds of genes specifically expressed in the mature anther or pollen tissues. Candidate genes were analyzed for predicted functional roles and their promoters were analyzed for *cis*-regulatory elements to understand their expression specificity. This study provides useful knowledge for fast-tracking plant breeding and seed production via SPT technology and a rapid and robust pollen and anther specific-candidate gene discovery pipeline.

2. Results

2.1. Maize Pollen Possesses a Specialized Transcriptome

To characterize the transcriptome features of mature pollen and identify mature pollen-specific genes (MPS) in maize, we assessed a publicly available RNA-seq-based developmental atlas of maize tissue for its applicability. This dataset contained RNA-seq profiling of 23 tissues including mature pollen of B73 inbred plants [69]. A Pearson correlation coefficient analysis of transcript abundance in the 23 tissues showed that transcripts of mature pollen have an insignificant correlation with all tissue types analyzed (Figure S1), indicating that mature pollen possesses a unique transcript profile compared with other maize tissues. On the other hand, some tissues displayed strong correlations, e.g., 6–7th internode and 7–8th internode, have very similar *r* values with other tissues (Figure S1), indicating that 6–7th internode and 7–8th internode have substantially overlapped in their transcript profiles. Therefore, only 12 tissue types representative of transcript profiles for the major maize organs were included in downstream analysis.

To understand the expression diversity and abundance in each of the selected tissue a criterion of a 1.00 FPKM cut-off was established which roughly corresponds to 1 mRNA molecule per cell for expressed genes (EG) [71]. The abundances of EG in pollen and other tissues were divided into four transcript level intervals: very highly expressed gene (VHEG, ≥ 400 FPKM), highly expressed gene (HEG, 40–400 FPKM), medium expressed gene (MEG, 10–40 FPKM), and low expressed gene (LEG, 1–10 FPKM). The genes expressed below 1.00 FPKM were termed trace-expressed genes or non-expressed genes. Of the over 32,00 genes reported in the maize B73 reference genome [72], our analysis showed that pollen grains expressed less than 30% of known maize genes with only 9242 EG meeting the 1.00 FPKM cut-off (Table 1). In contrast, the vegetative tissues (roots, meristem, internodes, and leaves) expressed $\sim 80\%$ of maize genes with 24,728 to 27,167 EG, and reproductive tissues (except pollen) (ear primordia, female spikelets, silks, embryo, endosperm, and pericarp/aleurone) had 24,175 to 27,659 EG (Table 1). These results indicated that the diversity of genes expressed in mature maize pollen was largely reduced compared to other tissues.

Table 1. Expressed Genes (EG) and average transcriptional levels of top 300 EG and top 1000 EG in 12 selected maize tissue types.

Tissues ^a	Number of EG ^b	Average Transcription of EG (FPKM) ^c	Average Transcription of Top 300 EG (FPKM) ^d	Average Transcription of Top 1000 EG (FPKM) ^e	Percentage of Top 300 EG in Transcriptome ^f	Percentage of Top 1000 EG in Transcriptome ^g
Mature Pollen	9242	99.1	2305.0	839.7	75.5%	91.7%
Primary Root_5 Days	24,799	37.9	1093.4	531.2	34.9%	56.5%
Secondary Root_7–8 Days	24,728	37.3	1039.9	511.4	33.8%	55.5%
Vegetative Meristem_16–19 Days	26,880	39.4	1153.3	542.8	32.7%	51.3%
7–8th Internode_28–30 Days	25,148	37.7	1002.5	490.4	31.7%	51.7%
Mature Leaf_8_43 Days	27,167	39.1	1408.5	598.0	39.8%	56.3%
Ear Primordium_6–8 mm	27,659	33.4	868.7	400.7	28.2%	43.4%
Female Spikelets_silking	25,281	42.5	1390.3	624.8	38.8%	58.1%
Unpollinated Silks	24,825	47.5	1854.3	758.7	47.2%	64.4%
Embryo_20 DAP	27,415	38.1	1199.9	534.9	34.4%	51.2%
Endosperm_12 DAP	24,175	42.5	1455.4	626.2	42.5%	61.0%
Pericarp/Aleurone_27 DAP	26,044	39.1	2000.9	704.3	59.0%	69.2%

^a 12 selected tissues samples and their collection time: mature pollen, primary root_5 days, secondary root_7–8 days, vegetative meristem_16–19 days, 7–8th internode_28–30 days, and mature leaf_8 at 43 days old, female spikelet collected at silking day, unpollinated_silks, ear primordium_6–8 mm, embryo_20 DAP, endosperm_12 DAP, pericarp/aleurone_27 DAP. DAP, days after pollination. ^b Expressed genes (EG) whose transcriptional level is higher than 1 FPKM. The transcript levels of transcribed genes were calculated in FPKM (Fragments per kilobase pair of exon model per million fragments mapped) using transcriptome profiling dataset of B73 tissues [69]. ^c The average FPKM value of all EG of a tissue. ^d or ^e Average transcription of top 300 EG or 1000 EG is the average FPKM value of top abundant 300 EG or 1000 EG. ^f or ^g Percentage of top 300 EG or 1000 EG in transcriptome is the percentage that the transcripts of top abundant 300 or 1000 EG occupied the whole transcriptome.

In terms of transcript abundance, pollen grain possessed a substantially higher average FPKM per EG than other tissues especially with the top 300 and 1000 EG (Table 1). The transcripts of the top 300 EG in pollen occupied 75.5% of the pollen transcriptome while transcripts of the top 300 EG in the other 11 tissues accounted for only 28.2–47.2% of their

whole transcriptomes. For the top 1000 EG, this percentage increased to 91.7% in pollen compared to 43.4–69.2% in the other 11 tissues. These results indicated that maize pollen invested transcription on a small number of genes, implying that highly expressed genes in maize pollen are extremely specialized and function in processes crucial for pollen maturation and other activities associated with fertilization.

As tissue types displayed dramatically different levels of highly expressed genes compared, to understand the differences in expressional levels across their transcriptomes, transcript abundances were enumerated for the selected tissue types at four transcript level intervals. The results showed that pollen has 380 VHEG while other tissues averaged 356 across the other 11 tissues (Figure 1), suggesting that each tissue type, including pollen, possesses comparable numbers of VHEG. However, in sharp contrast, pollen posed substantially reduced numbers of other gene levels with only 1061 HEG, 1535 MEG, and 6266 LEG, compared to the average of the other tissue types which had 3628 HEG, 8071 MEG, and 13,956 LEG (Figure 1), which indicate that the transcriptome complexity of maize pollen is strongly reduced compared to other tissues.

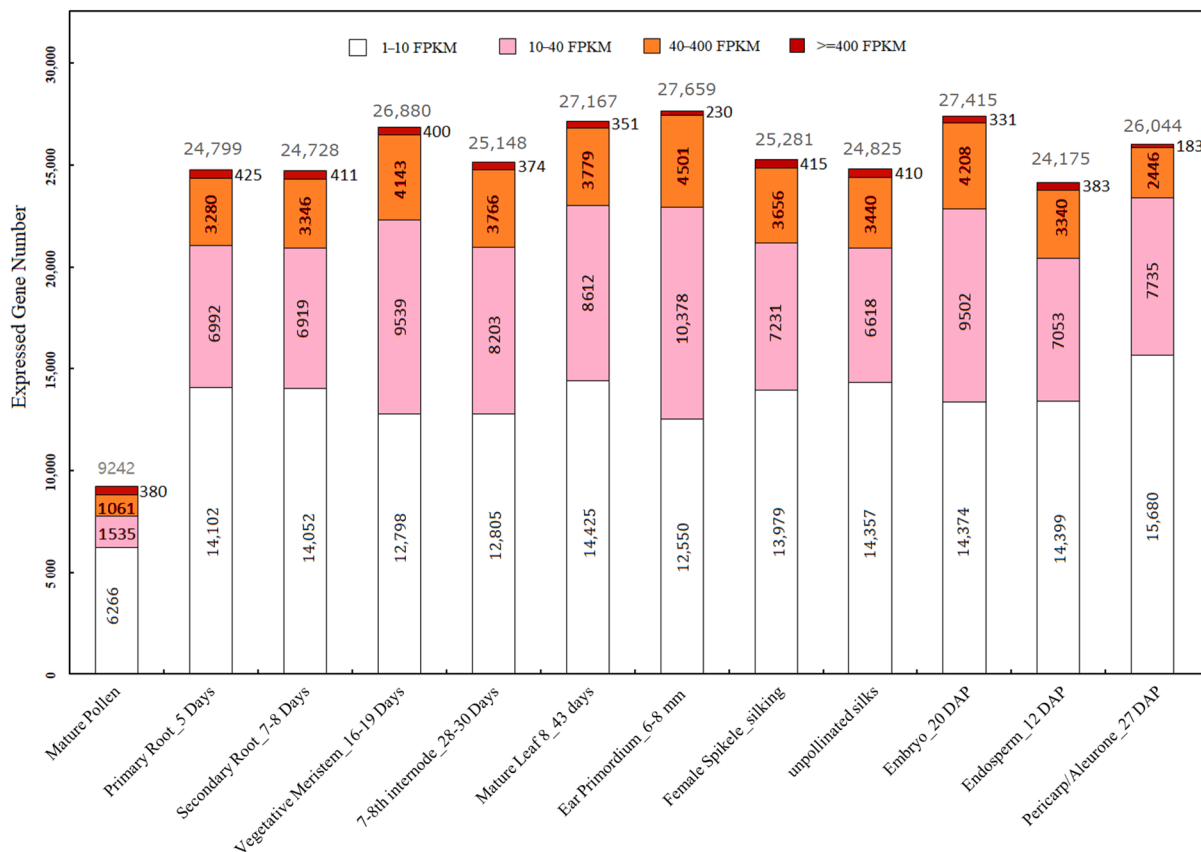


Figure 1. Unique transcripts identified in the 12 selected maize tissue types of this study grouped by levels according to their abundances. Very highly expressed genes (VHEG), highly expressed genes (HEG), medium expressed genes (MEG), and low expressed genes (LEG) are indicated in dark red, orange, pink, and white color, respectively. VHEG, HEG, MEG, and LEG possess the transcript levels of ≥ 400 , 40–400, 10–40, and 1–10 FPKM, respectively. The total EG number of each tissue was indicated on the top of the columns in grey color. The transcript profiles of 12 tissues were extracted from the dataset of the B73 developmental atlas [69].

Pollen of diverse plant species contains a large number of genes, termed pollen-specifically expressed genes and defined as displaying detectable only in pollen compared to any other tissue type analyzed [25,39]. This definition however useful for traditional methods such Northern blot, RT-PCR, q-PCR, and microarrays, is impractical for next-generation sequencing technologies whose increased sensitivity allows the detection of transcripts with very low expression. Here, we screened the maize transcriptome for

mature pollen-specific genes (MPS) with the following criteria: (1) an FPKM of at least 1 in mature pollen, and (2) expressed at least 100 folds higher in pollen compared to other sporophytic or reproductive tissues. With these criteria, we identified 1215 MPS in mature maize mature pollen based on their transcript levels in the 12 tissues (Table S1). Of the 9242 EG in maize pollen grains, 13.1% had pollen-specific expression and of these, 952 genes were predicted to encode proteins while the others were predicted to be pseudogenes or transposable elements. Of the protein-coding genes, 631 were annotated with a putative functional description while the remaining encoded uncharacterized proteins.

To understand tissue-dependent gene expression in the 12 selected tissues, we performed a scatter-plot analysis of mature pollen against each of the other tissues (Figure S2). This analysis found a low correlation between pollen and other tissues for gene expression. As the number of pollen-preferentially expressed genes are different when compared across the tissues, we analyzed the mature pollen against the average of other 11 tissues (Figure 2) and found 1307 gene models distributed above the 100-fold threshold (orange line) (Figure 2) and of those 1215 displayed expression above 1 FPK (Table S1). Additionally, expression of several genes was observed over 400 FPKM (red line) and classified as mature pollen abundant genes (MPA). In total, 380 MPA were identified with 293 or ~77% of those also being MPS (Figure 2). Together these results find that maize pollen expresses a highly unique and specialized transcriptome.

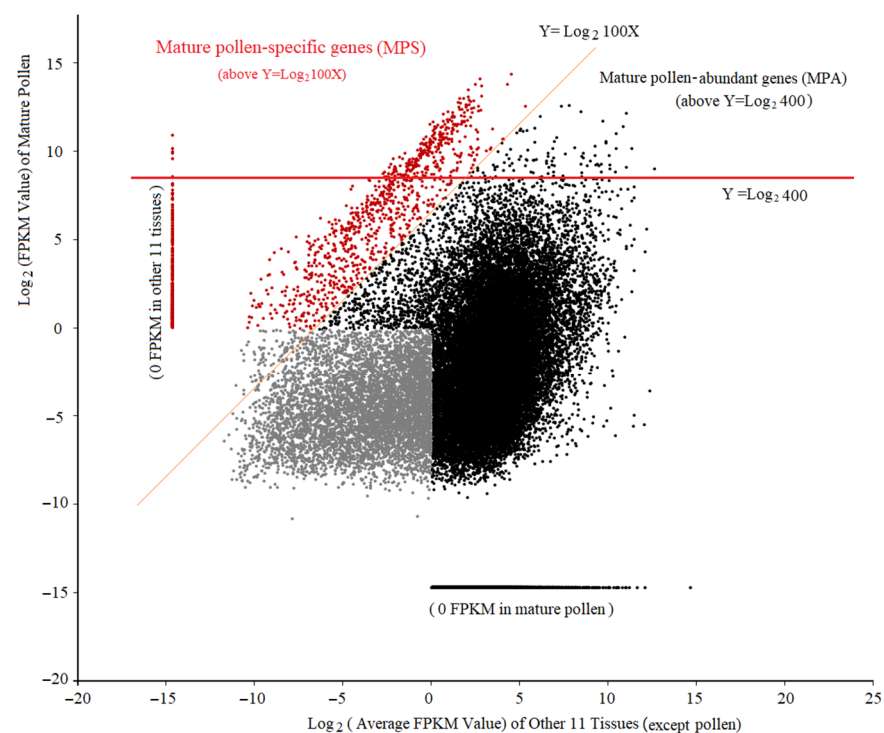


Figure 2. Two-dimension scatter plot of transcript abundance in mature pollen vs. the mean of other 11 tissues. The *x*-axis indicates the average of transcript abundances in the other 11 tissues (except pollen) and the *y*-axis indicates the transcript abundance detected in mature pollen [69] with each point representing a unique gene model. Mature pollen-specific genes (MPS) are distributed above the orange line ($y = \log_2 100x$) and mature pollen-abundant genes (MPA) are located above the red line ($y = \log_2 400$). Red points represent all MPS and the grey points represent the trace-expressed genes in all tissues.

2.2. Maize Anthers Express a Unique Transcriptome

In maize, pollen grains develop in the loculi of anthers and these anthers have to undergo 14 successive stages from archesporial cells to mature pollen grain [5,7] while reprogramming their transcriptome at each developmental stage to meet the needs of male

microspores [16]. Mature anthers contain transcripts expressed in the microspores and in the anther walls which include the epidermis, endothecium, middle wall layers, and additional connective tissues [16]. To reveal the transcriptomic features of mature anthers (anther at R1 stage) in comparison with other tissues, a RNA-seq dataset of 79 maize tissues of the B73 inbred line [70,73] were downloaded from the maize genetics and genomics database (maizeGDB, <https://maizegdb.org/> accessed on 18 June 2021). This dataset provided an expanded maize gene expression atlas covering 79 tissue types including anthers. It is necessary to note that mature pollen was not included in the profiled tissue of this dataset. From the 79 tissues, 14 unique tissues were selected for the anther transcriptomic analysis (Figure 3) and these tissues were similar to those from the pollen transcriptome analysis (Table 1, Figure 1). EG were filtered with a cut-off of 1 FPKM and categorized into four levels of expression: very highly expressed gene (VHEG) which transcripts 2000 FPKM or greater, highly expressed gene (HEG) with 200–2000 FPKM, medium expressed gene (MEG) with 50–200 FPKM, and low expressed gene (LEG) with 1–50 FPKM. The limits for these categorical groups were assigned based on the higher overall FPKM values in this dataset [70,73] compared with the previous [69] due to different experimental and sequencing strategies (Figure S3). The results showed that maize mature anthers have EG comparable to the other tissue types analyzed, with 23,519 EG found in anthers (Table S2, Figure 3) while the other 13 tissues have an average of 22,845 EG. A similar comparable number EGs for the other categorical groups was found between anthers and the other tissues.

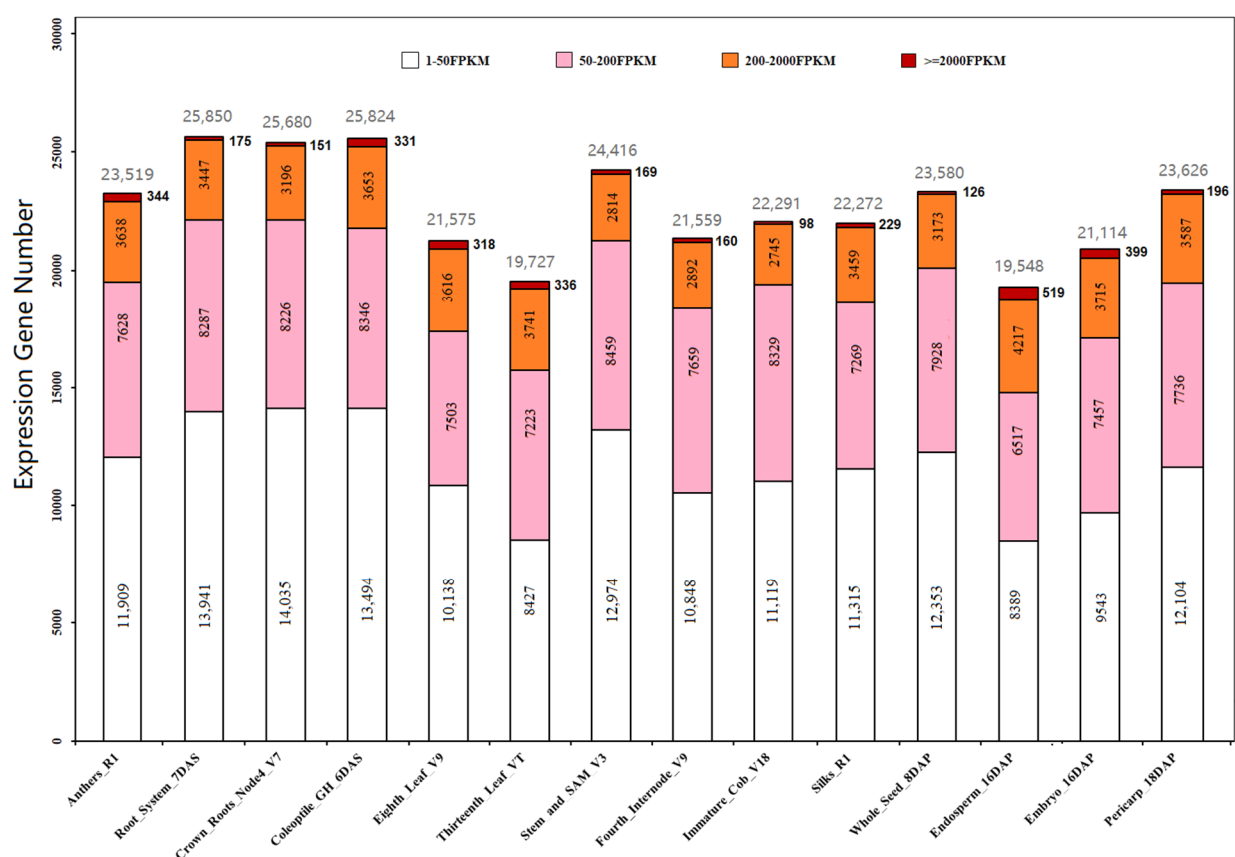


Figure 3. Unique transcripts identified in 14 maize tissue types and grouped according to their abundance. Very highly expressed genes (VHEG), highly expressed genes (HEG), medium expressed genes (MEG), and low expressed genes (LEG) are indicated in dark red, orange, pink and white color, respectively, in the columns. VHEG, HEG, MEG, and LEG possess the transcript levels of ≥ 2000 , 200–2000, 50–200, and 1–50 FPKM, respectively. The total EG number of each tissue was indicated on the top of the columns in grey color. The transcriptome profiling data of 14 selected tissues were extracted from the transcriptome profiling dataset [70,73].

To understand the similarity between tissue transcriptome profiles, a correlation analysis was performed and it found a very poor correlation (r of 0.05–0.18) of the levels transcript expression in anthers compared to other tissue types, indicating that mature anther has a distinct transcriptome pattern as compared to other tissues (Figure 4).

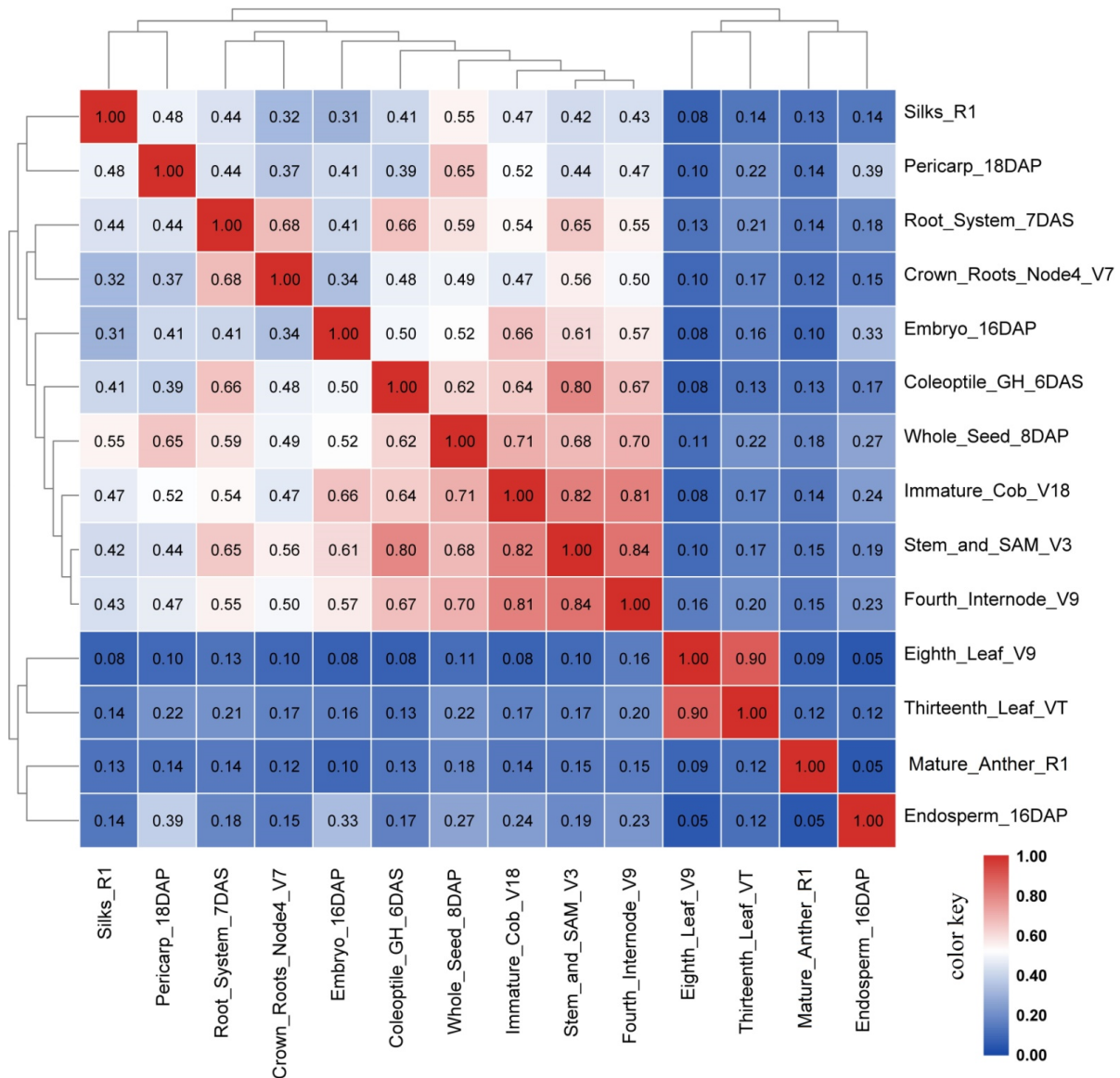


Figure 4. Pearson's correlation coefficient r matrix of 14 selected tissues including mature anther. FPKM values for the transcripts of 39,456 genes in 14 selected tissues [70] were used to perform this correlation analysis. Pearson's correlation matrix was drawn by software TBtools [74]. The numbers in this matrix are the Pearson's correlation coefficients (r) between two tissues of 14 tissues selected.

We further generated a two-dimension scatterplot of transcripts in mature anthers against the average of the other 13 tissues (Figure 5). The result showed that 1009 unique gene models, termed mature anther-specific genes (MAS) displaying at least 100-fold increased expression in anthers compared to other tissue types (points above the orange line) (Table S3). Additionally, 344 unique gene models were identified that displayed high expression, termed mature anther abundant gene (MAA) with an FPKM of at least 2000 (red line) (Figures 3 and 5). Of the MAA genes, 222 or ~65% are also MAS, suggesting that a high portion of the anther transcriptome is specific to this tissue. Of the 1009 MAS,

nearly all (1002 gene models) were predicted to encode proteins while the remainings were predicted as pseudogenes or transposable elements (Table S3).

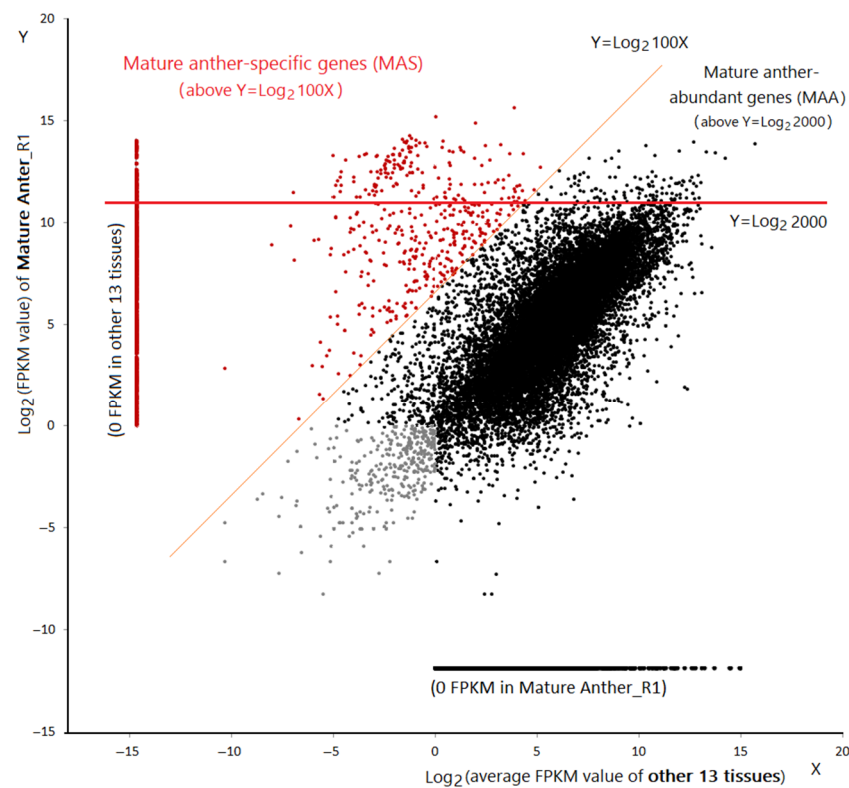


Figure 5. Two-dimension scatter plot of transcript abundance in mature anther vs. other 13 tissues. The x -axis indicates the average of transcript abundances in the other 13 tissues and the y -axis indicates the transcript abundance detected in mature anthers by RNA-seq [70]. Each point represents a unique gene model. Mature anther-specific genes (MAS) are distributed above the orange line ($y = \log_2 100x$) and mature anther-abundant genes (MAA) are located above the red line ($y = \log_2 400$). MAS are indicated in red dots and the grey dots represent the trace-expressed genes.

2.3. MPS and MAS Are Predicted to Function in Similar Roles for the Molecular Regulation of Maize Reproduction

As MPS and MAS are specifically expressed genes in mature pollen and mature anther, respectively, it is reasonable to expect that they play important roles in the reproductive processes of maize e.g., anthesis, anther dehiscence, and double fertilization. To understand their function, gene ontology (GO) enrichment analysis of MPS and MAS was performed through the Blast2GO software (biobam, <https://www.blast2go.com/>, accessed on 18 June 2021). A total of 1215 MPS genes were assigned to 40 GO terms (Figure 6A), which are classified into three major categories (biological process, cellular component, and molecular function), while 1009 MAS genes were assigned to 42 GO terms (Figure 6B). For GO terms of biological process, MPS were significantly overrepresented in 19 biological processes including cellular, metabolic, and single-organism processes (Figure 6A). Similar to MPS, MAS were also significantly overrepresented in these 19 processes (Figure 6B) and MAS, interestingly, were also enriched in locomotion. For GO terms of molecular function, MPS and MAS displayed similar patterns with substantial overrepresented in binding and catalytic activity. Both tissues were also significantly overrepresented in seven other activities (molecular function regulator, transporter activity, electron carrier activity, antioxidant activity, nucleic acid binding transcription factor activity, structural molecule activity, and signal transducer activity) (Figure 6). For GO terms of cellular component, MPS and MAS were significantly over-enriched in 12 and 13 cellular components, respectively, and they are super highly over-represented in cell, cell part, and membrane (Figure 6).

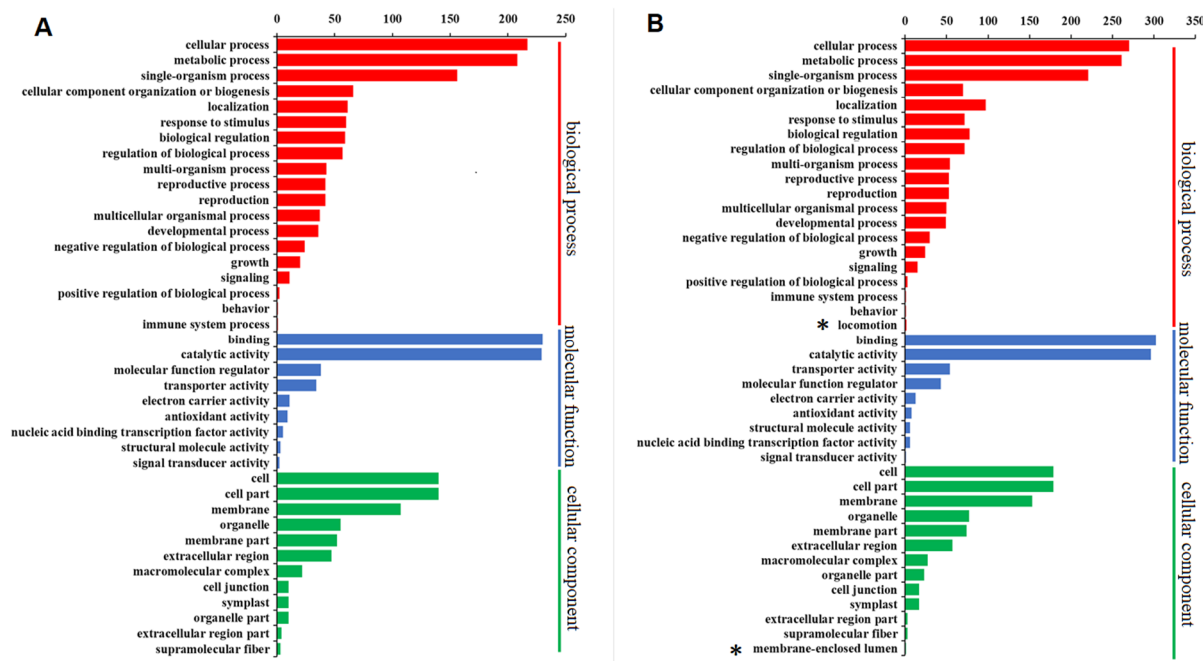


Figure 6. Gene ontology (GO) enrichment analysis and KEGG pathway assignments of MPS and MAS in B73 inbred using Blast2GO. **(A)** GO term enrichment analysis of MPS; **(B)** GO term enrichment analysis of MAS; In **(A,B)** the x -axis indicates the gene number assigned to a GO term and the y -axis is the list of significantly enriched GO terms. These significantly enriched GO terms were selected based on a p -value < 0.05 and FDR < 0.05 . GO terms of the categories of biological processes, molecular functions, and cellular components are depicted in red, blue, and green, respectively. The asterisk indicates the two additional GO terms assigned to MAS genes.

To understand the pathways impacted by MPS and MAS, a KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis was performed, however, only a limited number of genes were able to be assigned to known pathways (Figure S4), suggesting that the vast majority of MPS or MAS genes are not involved in known metabolism or signaling pathways. The results found 89 MPS genes and 95 MAS genes could be assigned to 29 and 34 KEGG pathways, respectively (Tables S4 and S5), and for the top 24 most enriched pathways, MPS and MAS have similar patterns (Figure S4). MPS and MAS were dramatically overrepresented in metabolic pathways, including pentose-glucuronate interconversions, plant-pathogen interaction, biosynthesis of secondary metabolites, ascorbate and aldarate metabolism, starch and sucrose metabolism, nitrogen metabolism, and phenylpropanoid biosynthesis. Together, the similarity of GO terms and KEGG pathway enrichment suggests that the specialized transcriptomes represented in MPS and MAS impact similar functions in the reproductive processes of maize.

2.4. Promoters of Mature Anther- and Pollen-Specific Genes (MAPS) Are Enriched for Pollen-Specific and Hormone-Responsive cis-Elements

Both MPS and MAS represented the exclusively expressed genes in the final stage of anther development and as both groups have similar GO term and KEGG pathway enrichment patterns (Figure 7), it suggests they play similar functions during pollen development. Using the online software Venny 2.10 (Venny 2.10, <https://bioinfoq.cnb.csic.es/tools/venny/index.html> accessed on 18 June 2021), we observed that 1215 MPS and 1009 MAS genes identified in our study share 623 overlapping genes (Figure S5), designated as mature anther- and pollen-specific genes (MAPS) (Table S6). A comprehensive literature analysis found reports for 20 pollen-specific genes of maize (Table S7), 18 of which were found in the collection of 623 MAPS genes (Table S6), providing support of this pipeline for pollen-specific gene discovery.

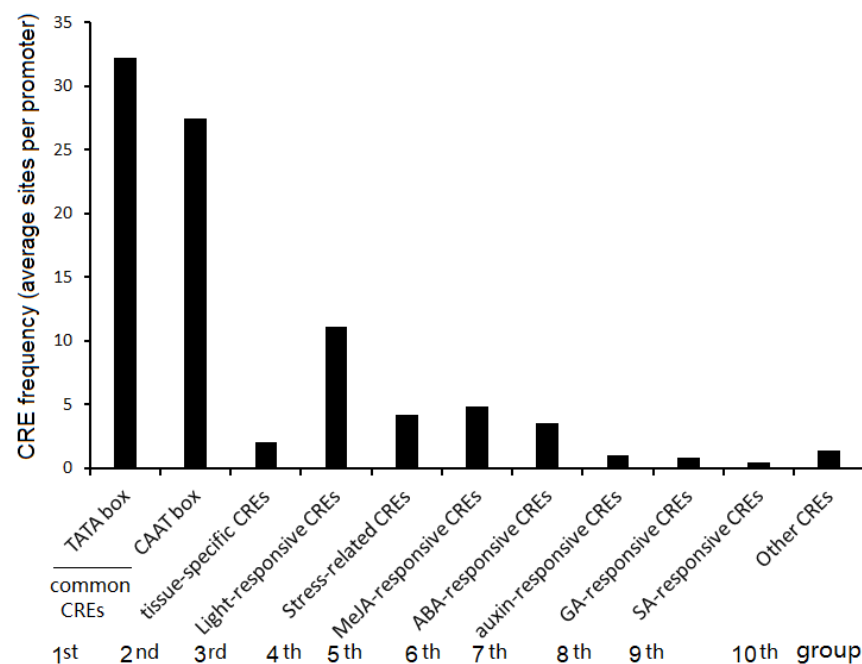


Figure 7. Classification of CREs identified in the promoters of MAPS genes; 2-kb promoter regions upstream of the start codon (ATG) of 618 MAPS genes were scanned to obtain the CRE sites by “Search for CARE” from PlantCARE web. The search retrieved 95,631 CRE sites, which belong to 123 CREs. Those CREs were classified manually into 10 groups according to their functional annotation.

To further validate the pollen-specificity of MAPS, we performed a q-PCR analysis of 12 randomly selected MAPS genes in 12 tissues of the B73 inbred line. The results showed that all 12 genes displayed substantially increased expression in mature pollen grain compared to all other tissues (Figure S6); however, one gene, GRMZM2G703173, also showed moderate expressions in the tassel spikelet and third leaf blade/sheathe (Figure S6). This result provides strong evidence that the majority of MAPS genes identified in this study are pollen-specific.

To reveal the potential functions of MAPS genes involving in pollen development and pollination, we searched maizeGDB for information of gene models for functional descriptions of MAPS genes (Table S6) and manually classified MAPS genes into nine categories (Figure S7) based on this information. Those groups contained 447 pollen allergen proteins, 4 cell wall metabolism, 5 transposable elements, 6 transcriptional factors, 1 JA signaling protein, 21 calcium signal related proteins, 5 kinases, 92 unknown proteins, and 42 other functional proteins. This result found that a large portion (71.7%) of MAPS encode pollen-allergen proteins which are typical pollen-specific proteins of the plant kingdom [75]. Pollen allergies have long been the respiratory illnesses like asthma for humans with significant economic repercussions, however, the biological function of these proteins remains to be characterized [75].

To further reveal the molecular function of MAPS genes, we performed functional classification of MAPS genes using the MapMan tool kit [76]. The results showed that in the metabolism overview, cell wall organization and modification processes employed 80 MAPS genes including 41 genes for cell wall pectin esterase, seven for cell wall protein, four for cellulose/hemicelluloses synthesis, two for cell wall precursor synthesis, 25 for cell wall dehydration, and one for inositol 3-phosphate synthase (Figure S8). Seven genes were also involved in lipid metabolism (Figure S8). This result indicated that MAPS genes are primarily involved in cell wall organization and composition of pollen grains important for processes such as germination, pollen tube growth, and pollen-sigma interaction. For molecular functions, MapMan predicted MAPS proteins include 24 transcriptional factors, 25 proteins for protein-modification processes, and 18 proteins for protein-degradation

processes, suggesting a portion of MAPS govern the pollen proteome (Figure S8). There is also evidence that MAPS genes may function in signaling with receptor kinases, G-proteins, and calcium signaling proteins (Figure S8).

To understand the molecular basis of the tissue-specific expression of MAPS, we analyzed 2 kb promoter regions upstream of the start codon (ATG) of MAPS genes for *cis*-regulatory elements (CREs) via PlantCARE [77]. The analysis found that 123 *cis*-regulatory elements (CREs) (Table S8) were identified in 618 MAPS promoters with the top 10 frequent CREs being the TATA-box, CAAT-box, MYB-binding site, G-box, MYC (bHLH)-binding site, STRE (stress-responsive element), ABRE (ABA-responsive element), as-1, CGTCA-motif, and TGACG-motif. We further divided the 123 CREs into 10 groups (1st–10th group) according to their function annotations (Figure 7; Table S8). The first group contained the common *cis*-elements, TATA-box and CAAT-box, which act as the universal CREs for transcription initiation of eukaryotic genes [78] and were the most enriched CREs in the promoter of MAPS. On average, one promoter of MAPS had 32 copies of TATA-boxes and 27 copies of CAAT-boxes (Figure 7). Usually, a TATA-box with a consensus sequence of TATAATAAT is located around –30 bp before the transcription start site, and a CAAT-box with a consensus sequence of GG(T/C)CAATCT can be found around –75 bp upstream of the transcriptional start site [78]. The second group includes CREs specific or related to some tissues such as endosperm, meristems, roots, and leaf mesophylls. The third group contained more than 10 CREs involving light responsiveness which interestingly, *cis*-elements of light responsiveness are highly enriched in the promoters of MAPS. Typically, one MAPS promoter contained 11 copies of light-responsive CREs, indicating that light is an important factor for pollen-specific gene expression. The fourth group included stress-related CREs, which are involved in responses to diverse stresses such as drought, low temperature, salt stress, anoxia, and pathogens, indicating that environmental stresses can modify the gene expression in maize pollen. The fifth and sixth group included *cis*-elements responsive to MeJA/JA and ABA signaling, respectively. These two groups of CREs are enriched in the promoters of MAPS with the frequency of 3.5 and 4.8 copies in one MAPS promoter for MeJA/JA and for ABA, respectively (Figure 7). This result indicated that jasmonic acid and ABA signals are involved in the regulation of pollen-specific gene expression. The promoters of MAPS also contained auxin-, GA-, and SA-responsive CREs (seventh, eighth, and the ninth group), providing evidence that auxin, GA, and SA may also regulate MAPS gene expression. An assortment of other CREs located in the MAPS promoters e.g., A-box, box III, 3-AF3 binding site are classified into the tenth group (Figure 7).

Unfortunately, the PlantCARE platform does not provide information on pollen-specific *cis*-elements, thus we manually characterized the promoters of MAPS for the enrichment of pollen-specific CREs (Table S9). Our analysis found that six typical pollen-specific *cis*-elements (AGAAA, TCCACCA, TGTGGTT, [TA]AAAG, AAATGA, and TTTCT) have an enrichment index more than 1, indicating that MAPS promoters are highly enriched for these pollen specific *cis*-elements (Figure 8). The four most enriched *cis*-elements were the *LeLAT52* PB motif (TGTGGTT), the *NPT303* novel *cis*-element (AAATGA), the *SBgLR* *cis*-element (TTTCT), and the *LeLAT52* pollen-specific *cis*-element (AGAAA) (Figure 8). MAPS promoters contained 9–49 pollen-specific CRE sites and averaged 27.7 pollen-specific CRE sites per promoter (Figure S9).

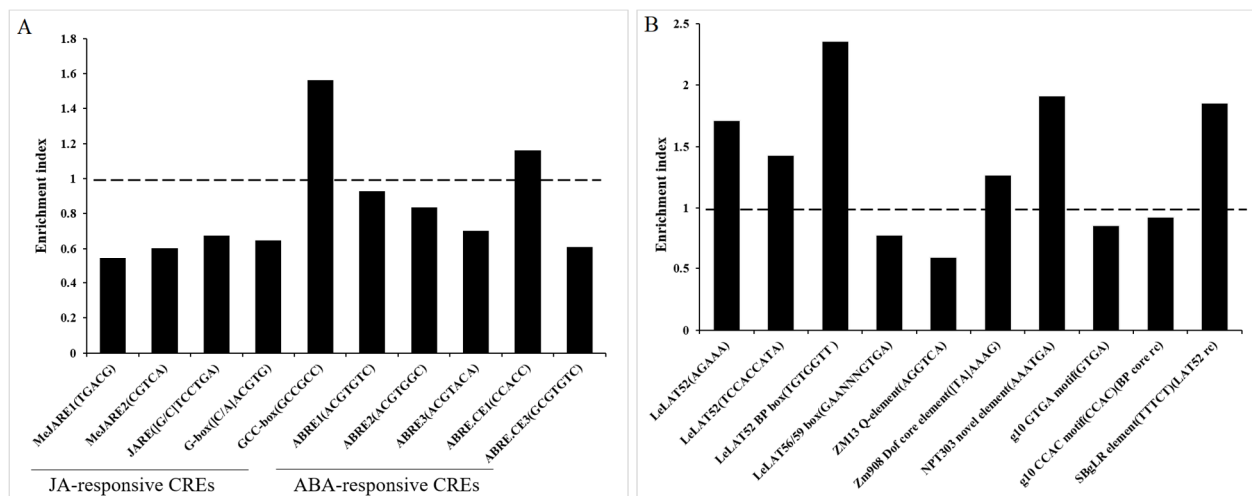


Figure 8. Enrichment analyses of known pollen-specific and ABA-/JA- responsive *cis*-elements in the promoters of MAPS. (A) Enrichment analyses of known pollen-specific *cis*-elements in MAPS promoters. The actual frequency of 10 known pollen-specific *cis*-elements appeared in the 2-kb promoters of the 618 MAPS genes were counted by software Python 3.7.3. The expected frequency of a specific *cis*-elements occurred in a random DNA sequence can be simply calculated based on the sequence length. For example, the expected frequency of a 5-nt *cis*-element such as AGAAA was expected to be one time in a length of 4⁵-bp DNA. TCCACCA, a 7-nt *cis*-element was expected to meet once in 4⁷-bp DNA. The enrichment index is the actual frequency divided by the expected frequency. (B) Enrichment analyses of known JA- and ABA-responsive *cis*-elements in MAPS promoters.

In Figure 7, 123 CREs found in the MAPS promoters included MeJA/JA-responsive CREs and ABA-responsive CREs which encouraged us to perform enrichment analysis of other well-known JA- and ABA-responsive CREs in the MAPS promoters. We found that five MeJA/JA-responsive CREs are present in the promoters of MAPS but only the GCC box (GCCGCC) is highly enriched (enrichment index > 1) (Figure 8). Additionally, five ABA-responsive CREs were located in the MAPS promoters, however, only one ABA-responsive *cis*-element, the ABRE-coupling element1 (ABRE-CE1, CCACC), was highly enriched (Figure 8). This result indicated that JA and ABA signals may involve in the expression regulation of MAPS genes in maize.

Together these results provide insight into the mechanisms underpinning transcriptional regulation of the MAPS gene specificity and suggest that expression will be impacted by environmental or other physiological stresses.

3. Discussion

3.1. MAPS Genes Are Expressed Specifically at the Late Stage of Anther Development

Precise tissue- and cell-specific expression is a major goal in biotechnology, genetic engineering, and synthetic biology. For plant breeding, pollen-specific expression, in particular, is of wide interest because of the potential for pollen-engineering and crop improvement. Genes specifically expressed in pollen are classified into two categories according to the timing: (1) early pollen-specific genes are expressed before and during the division phase of the pollen mother cell meiosis (from archesporial cell-forming stage to uni-cellular gametophyte stage), and (2) late pollen-specific gene are expressed during pollen maturation (from bi-cellular gametophyte stage to germinating pollen) [1,22–25]. Many pollen-specific genes have been isolated from diverse plant species, such as *LAT52/LAT56/LAT59* in tomato [79], *RA8* in rice [13], *TUA1* & *NTM19* in *Arabidopsis thaliana* [80,81], *g10* & *NPT303* [46,82] in tobacco, *Bp4* in *Brassica napus* [44], and *CHI* in petunia [43]. In maize, several pollen-specific genes such as *PG47* [83], *Zm58.1 (mgs2)* [84], *Pex1* [85], *ZmABP1/2* [86], *Zm13 (mgs1)* [47,87], *ZmC5* [88], *MATRILINEAL* or *NOT LIKE DAD* [89,90], *ZmSTK2_USP* [91] have been reported. A summary of pollen-specific genes in maize with their expression across tissues is provided in Table S7.

Using a microarray profiling approach, Ma et al. (2008) examined the maize anthers at seven stages: two pre-meiotic stages (A1.0 and A1.5), two meiotic stages (A2.0 and quartet), and three microspore maturation stages (uninucleate haploid microspore, binucleate microspore, and mature pollen). This study detected 1952 transcripts only in the pre-meiotic stages, 974 only in the uninucleate or binucleate microspores, and 245 only in mature pollen (trinucleate microspore) [16]. However, this study could not discern expression from pollen from—sporophytic or other reproductive tissues. In this study, we used a maize gene developmental expression atlas of 23 tissues of the B73 inbred line [69] and identified 1215 mature pollen-specific genes (MPS) (Table S1) whose transcripts were at least 100-fold higher in mature pollen compared to other sporophytic or reproductive tissues. Different strategies have been employed to identify pollen-specific and -preferential genes. Jiang et al. (2011) identified 82 pollen-specific genes in the rice genome using MPSS (Massively Parallel Signature Sequencing) which included 70 sequencing libraries of rice tissues. Using the definition of pollen-specific as genes whose expression in the pollen libraries were higher than their sum of expression in the remaining libraries, the study identified 82 pollen-specific genes with transcript levels over 1000 TPM (transcript per million) and a 200-fold increase in pollen compared to other tissues [92]. In another microarray approach, Oo et al. (2014) identified 261 late pollen-specific rice genes (LPS) with expression at least 500-fold higher in late-stage pollen compared to other tissue [24].

In a next-generation sequencing-based study, Wang et al. (2020) identified 269 late-stage pollen-specific (LSP) genes from expression levels comparison across anthers, roots, shoots, seeds, and leaves with at least five-fold increase in mature anther FPKM values compared to other tissue types [25]. Moon et al. (2018) identified 627 late pollen-preferred genes that are preferentially expressed (the ratio of pollen versus other tissues > 2X) in the late stages of pollen grains in rice by comparing the microarrays-based transcriptome profiling of sporophytes and male gametes over anther developmental stages [23]. Honys et al. (2004) profiled the transcriptome of Arabidopsis tissues using Affymetrix ATH1 microarrays and 1355 transcripts were found to be transcribed specifically at the stages of male gametophyte development [93]. Of 1355 transcripts, 625 were expressed at the mature stage of the pollen [93]. In our study, a more stringent 100-fold criterion was used to select pollen-specific genes which identified 1215 MPS (Table S1) expressed specifically at the trinucleate stage of maize mature pollen. The R1 stage is the last stage of anther development [70] and their mature pollen (trinucleate microspore) will shed all the mature pollen in one–two days. In this study, we analyzed transcriptomic data of mature anther using a maize gene developmental expression atlas of 79 tissues of the B73 inbred line [70,73]. Using similar criteria for the selection of MPS, we identified 1009 mature anther-specific genes (MAS) (Table S3) and found similar GO terms and KEGG pathway assignments of their functional annotations (Figure 7). We identified 623 genes that overlapped in the two gene sets which we termed mature anther and pollen-specific genes (MAPS) and the reason that MAPS represents the late-stage pollen-specific genes of the maize genome. Function annotation analysis showed that MAPS genes included 447 pollen allergen protein genes (Table S6) and expressional analysis provided strong evidence that MAPS are specifically expressed in the mature pollen (Figure S6). Our results are consistent with the studies in Arabidopsis and rice, in which hundreds of genes were identified as pollen-specifically expressed genes at the late stages of male gametophyte development [23,25,93]. In this study we combined the RNA-seq-based transcriptome analysis of mature pollen and mature anther and identified 623 MAPS in maize. The MAPS genes of maize have 449 or 550 orthologues in Arabidopsis or rice. Most of the orthologues are pollen-specifically/preferentially expressed in Arabidopsis and rice genomes [23,93], indicating that MAPS may serve as conservative genes in monocotyledonous and dicotyledonous plants.

3.2. MAPS Gene Promoters Are Enriched with Pollen-Specific cis-Elements

Organ- and tissue-specific gene expression largely depends on presences of *cis*-elements (CEs) in their promoter and their respective transcription factors. A number of conserved *cis*-elements of pollen-specific promoters have been intensively investigated in plant species. For example, the 52/56 box, also called the PB core motifs (TGTGGTT) and the 56/59 box (TGTGA) were identified as crucial *cis*-elements of the *LAT52*, *LAT56*, and *LAT59* promoters in tomato [94]. Expression and activity of the *LAT52* gene was strictly dependent upon the PB core motif and a downstream pollen-specific activator (AGAAATAATAGCTCCACCATA) which contains two *cis*-regulatory elements AGAAA and TCCACCATA [49]. Furthermore, a 30-bp-region (TTTAGTTTCAAAACAAGTGACTGTGCGCA) downstream of the 56/59 box was also shown to be essential for *LAT56/59* promoter activity [95]. In *Arabidopsis*, a *cis*-element, GAATATTCCT, located in the 88-bp-*cis*-regulatory module (CRM) determines the specificity of *ACA7* promoter [96]. The GTGA motifs of the *g10* promoter [50] and the AAATGA motif of the *NTP303* promoter [46] in tobacco were showed to be the *cis*-elements for their pollen-specificity. In potato, two palindrome motifs (TTTCTATTATAATAGAAA and AGAATTGGAAATTCT) play an important role in pollen-specific expression of *SBgLR* [97,98]. In maize, a number of these previously described *cis*-elements are also frequently found in maize pollen-specific genes, such as *PG47* [83], *Zm13* [47,99], *Zm908* [52], and *ZmSTK2_USP* [91]. In addition, expression of the *Zm13* gene is attributed to both the presence of the TTTCT motif, a reverse-complementary sequence of the AGAAA pollen-specific motif AGAAA and the AGGTCA motif which acts as a quantitative element (called the Q-element) [47]. In this study, we calculated the enrichment analysis of pollen-specific *cis*-elements in the promoters of MAPS genes, and found that the *LeLA52* elements (AGAAA) and its reverse complement element (TTTCT), the *LeLat52* second *cis*-element (TCCACCA), the *LeLat52/56* box (TGTGGTT), the Dof core elements ([T/A]AAAG), and the tobacco *NPT303* *cis*-element (AAATGA) have higher enrichment index than other *cis*-elements (Figure 8). These results suggest that these six *cis*-elements are the major players for expression specificity of MAPS and are consistent with Wang et al., (2020), who showed that “TTTCT”, and “AGAAA” were the most frequent pollen-specific *cis*-motifs present in the LSP promoters of rice [25]. We also found a high frequency (appearance per kb) of these *cis*-elements in the promoters (2 kb sequence upstream of the start codon) of MAPS genes. The PB core motif TGTGG or TGTGGA and *cis*-elements AGAAA of tomato pollen-specific promoters *LAT52/56/59* were especially abundant in the promoters of MAPS (Figure 8). The enhancer element Box core-R (TTTCT) and Q-element (AGGTCA) of *Zm13* also appeared frequency. Other pollen-specific *cis*-elements such as the *NPT 303* enhancer element (AAATGA), the *MGSA* element (GAAACG), and the *cis*-elements of *ACA7* gene (GAATAT), were found also over-represented in the promoters of MAPS (Figure 8).

Surprisingly, we also found JA- and abscisic acid-responsive elements in promoters of MAPS (Figure 8). Other hormone responsive *cis*-elements found enriched in these promoters included those for MeJA-, ABA-, auxin-, and GA- and SA-regulatory elements. Together this suggests environmental and developmental conditions impact global expression of pollen-specific genes and additional efforts are needed to comprehensively understand pollen-specific expression for advancements in tissue- and organ-specific expression of agriculturally important crop species.

4. Materials and Methods

4.1. RT-qPCR to Validate the Tissue-Specific Expression Pattern of MAPS

The B73 seeds were sowed in the plots of net-houses located in Pailou experimental garden of Nanjing Agricultural University, Nanjing city, China. The plants grew in the net-houses during the spring-summer time (April to August) of 2019 under local climate conditions with a regular management procedure for green houses including water supply, pest/weed control, and fertilization. The plant tissues/organs were collected as RNA-extraction samples at different stages indicated below. The samples included seedling

shoots and root system at V1 stage, stem and SAM (stem apical meristem) at V3 stage, 3rd leaf blades and sheathes at V3, 7th leaf (middle section of the leaf) at V7, young tassel at V8, tassel spikelets at VT, silk at R1, female spikelets at R1, mature pollen at R1, embryo at 20 DAP (days after pollination). The plant growth stages are recorded as described by the webpage of PennState extension (Corn Growth Stages, <https://extension.psu.edu/corn-growth-stages/> accessed on 18 June 2021). The tissue samples were frozen in liquid nitrogen immediately after sampling and stored at $-80\text{ }^{\circ}\text{C}$ for further use. Three biological replicates were taken for each tissue sample.

The total RNA was extracted using Total RNA Extractor (Trizol) (Sangon Biotech, Shanghai, China) following the manufacturer's instructions. The cDNA was synthesized in 20 μL reaction including 1 μg of total RNA, 500 ng of oligod(T)₁₈ primer, and 10 μL 2 \times EasyScript First-Strand cDNA Synthesis SuperMix (TransGen Biotech, Beijing, China) according to the manufacturer's instructions. RT-qPCR was conducted in 96-well plates and performed on the Bio-Rad CFX96 real-time PCR System (Bio-Rad, Hercules, CA, USA) under cycling conditions (95 $^{\circ}\text{C}$ for 1 min; 40 thermo-cycles of 95 $^{\circ}\text{C}$ for 10 s and 58 $^{\circ}\text{C}$ for 30 s). Each reaction mix contained 1 μL of 1/10 diluted cDNA mix (above), 10 μL 2 \times ChamQ Universal SYBR qPCR Master Mix (Vazyme), 8 μL RNase free water, and 0.5 μL (10 mM) of each primer, for a final volume of 20 μL . A no-template control was also included in each run for each gene. Each sample was conducted in technical triplicates with at least two biological replicates. In addition, melting curves were generated at 65–95 $^{\circ}\text{C}$ after 40 cycles to check for primer specificity. The primer sets listed in Table S10. Maize house-keeping gene β -tubulin4 (GRMZM066191) was used as the reference gene for RT-qPCR [100].

4.2. Identification of MAS, MPS, and MAPS in Maize Genome Using RNA-seq Datasets

Walley et al. (2016) published a developmental atlas of maize which contained transcriptome and proteome profiles of 23 tissues of B73 inbred including mature pollen [69]. The dataset of this publication was adopted by maizeGDB as transcript and proteomic profiling data for gene models of B73_RefGen_v4. In this study, we downloaded this dataset for transcriptome profiling analysis of mature pollen and rest tissues. Stelpflug et al. (2016) published an expanded maize gene expression atlas of maize which contained transcriptome data of 76 tissues of B73 inbred including mature anther [70]. This dataset was adopted by maizeGDB as the expression resource of B73_RefGen_v3 for gene models. Here, we downloaded this dataset for transcriptome profiling analysis of mature anther (anther_R1) against other tissues. The two datasets that we downloaded are presented in the files of excel format. Using Microsoft software Excel, we calculate the transcript abundance of each gene ID of the selected tissues. To identify mature pollen specific genes (MPS) we selected 12 tissues of the dataset of Walley et al. (2016) and for mature anther-specific genes (MAS) we selected 14 tissues of the dataset of Stelpflug et al. (2016) [69,70]. The 12 or 14 tissues that we selected represent the major organs of maize plant covering the major sporophytic and reproductive organs/tissues of maize plant. The MPS and MAS were selected by ratio of transcript abundance in the mature pollen or mature anther versus the average transcript abundance in other selected tissues. We tried different thresholds of ratio (10-, 50-, and 100-fold) to selected MPS and MAS, and we found 100-fold is the best for our selections. The mature pollen- and anther-specific genes (MAPS) are the shared genes of the MPS and MAS. The function categories of MAPS were carried out manually based on the molecular functional description of the genes.

4.3. GO Term and KEGG Pathway Enrichment Analysis of MAS and MPS

The gene IDs (B73_RefGen_v3) of MPS and MAS were submitted to an online tool of cloud tool platform, Omicshare (GO Enrichment Analysis Advanced <https://www.omicshare.com/tools/Home/Soft/gogseasenor> accessed on 18 June 2021) to perform the GO term enrichment analysis. A false discovery rate (FDR) ≤ 0.05 and p -value ≤ 0.05 were used as the criteria to obtain significantly enriched GO terms. KEGG (Kyoto Ency-

lopedia of Genes and Genomes) pathway enrichment analysis was performed using a KEGG enrichment tool of the platform OmicShare (Pathway Enrichment Analysis Advanced, <https://www.omicshare.com/tools/Home/Soft/doi enrichment> accessed on 18 June 2021).

4.4. MapMan Analysis of Functional Categories of MAPS Genes

MapMan is a software tool that supports the visualization of profiling data sets through various diagrams in the context of existing knowledge [76]. To classify the molecular function annotation of MAPS, we uploaded gene IDs of MAPS (B73_RefGen_v3) to the MapMan tool kit. We searched the overview of metabolism, carbohydrate, nucleic acid and protein synthesis, and protein modification and degradation pathway, hormone signaling pathways, calcium signaling pathways, and transcriptional factors for MAPS genes. The gene enrichments in overall metabolism and pathway were showed in Figure S8.

4.5. cis-Elements Analysis of MAPS Genes

The 2-kb upstream sequences before the translation start codon ATG of maize MAPS genes were taken as the gene's promoters. The *cis*-regulatory elements (CREs) of each promoter sequence were predicted by searching the PlantCARE database (PlantCARE, <http://bioinformatics.psb.ugent.be/webtools/plantcare/html/> accessed on 18 June 2021) with an FDR < 0.1% [77]. Predicted CREs existing in the promoters of MAPS were classified into ten groups based on the functional annotation of the CREs. Furthermore, 11 known pollen specific *cis*-elements of reported pollen-specific genes of plant species were searched out from published papers and the frequency of known pollen specific *cis*-elements appeared in the 2-kb promoters of MAPS were counted by the software Python 3.7.3.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ijms22136877/s1>, Figure S1: Pearson's correlation coefficient r matrix of 23 tissues of B73 inbred according to the RNA-seq dataset of Walley et al. (2016) [69], Figure S2: Two-dimension scatter plot of transcript abundance in mature pollen vs. each single tissue, Figure S3: FPKM distribution plots of 12 tissues (pollen-grouping tissues) and 14 tissues (anther-grouping tissues), Figure S4: Top 24 most represented categories of KEGG pathways of MPS and MAS, Figure S5: 632 MAPS genes, Figure S6: Q-PCR validation of the expression pattern of 12 MAPS genes, Figure S7: Manual classification of MAPS genes by genes' functional description, Figure S8: MapMan analysis of MAPS genes involving in metabolism or biological processes, Figure S9: Distribution of pollen-specific *cis*-elements and JA-/ABA-responsive *cis*-elements in the promoters of 618 MAPS genes, Table S1: The gene ID and transcript abundance (FPKM) of 1215 mature pollen-specific (MPS) genes in 12 tissues (including mature pollen) of B73 (.exls), Table S2: The number Expressed genes (EG) and average transcriptional levels of EG, top 300 EG, and top 1000 EG in 14 selected tissues of B73 inbred (.doc), Table S3: The gene ID and transcript abundance (FPKM) of 1009 mature anther-specific (MAS) genes in 14 tissues (including mature anther) of B73 (.exls), Table S4: GO term enrichment analysis of MPS and MAS genes (.exls), Table S5: KEGG pathway analysis of MPS and MAS genes (.exls), Table S6: The gene ID, description, mRNA abundance (FPKM) of 623 MAPS in 23 tissue and 79 tissue (.exls), Table S8: *cis*-element searched by PlantCare in 2-kb promoters of 618 MAPS (.exls), Table S9: Known pollen-specific *cis*-elements and JA- and ABA-responsive *cis*-elements of plant gene promoters (.exls), Table S10: The primer used in this study (.doc).

Author Contributions: The work presented here was carried out in collaboration among all the authors. Y.S., Y.L. and Y.G. performed the RT-qPCR experiments, analyzed data and contributed to produce figures and tables of the manuscript. E.J.B. contributed to manuscript writing. Z.W. and H.R. provided important helps to the research work of the team. Z.M. contributed to manuscript writing and funding. Y.Y. designed the project, wrote the drafts of the manuscript and obtained funds for the project. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (Grant No. 31571580), and the Fundamental Research Funds for the Key Universities Subordinated to the Ministry of Education of China (KYZ202101 and KYRC201404).

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

References

1. Goldberg, R.B.; Beals, T.P.; Sanders, P.M. Anther development: Basic principles and practical applications. *Plant Cell* **1993**, *5*, 1217–1229. [[PubMed](#)]
2. Varnier, A.-L.; Mazeyrat-Gourbeyre, F.; Sangwan, R.S.; Clément, C. Programmed cell death progressively models the development of anther sporophytic tissues from the tapetum and is triggered in pollen grains during maturation. *J. Struct. Biol.* **2005**, *152*, 118–128. [[CrossRef](#)]
3. Li, N.; Zhang, D.S.; Liu, H.S.; Yin, C.S.; Li, X.X.; Liang, W.Q.; Yuan, Z.; Xu, B.; Chu, H.W.; Wang, J.; et al. The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. *Plant Cell* **2006**, *18*, 2999–3014. [[CrossRef](#)]
4. Sanders, P.M.; Bui, A.Q.; Weterings, K.; McIntire, K.N.; Hsu, Y.-C.; Lee, P.Y.; Truong, M.T.; Beals, T.P.; Goldberg, R.B. Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. *Sex. Plant Reprod.* **1999**, *11*, 297–322. [[CrossRef](#)]
5. Zhang, D.; Luo, X.; Zhu, L. Cytological analysis and genetic control of rice anther development. *J. Genet. Genom.* **2011**, *38*, 379–390. [[CrossRef](#)]
6. An, X.; Dong, Z.; Tian, Y.; Xie, K.; Wu, S.; Zhu, T.; Zhang, D.; Zhou, Y.; Niu, C.; Ma, B.; et al. ZmMs30 Encoding a Novel GDSL Lipase Is Essential for Male Fertility and Valuable for Hybrid Breeding in Maize. *Mol. Plant* **2019**, *12*, 343–359. [[CrossRef](#)] [[PubMed](#)]
7. Wan, X.; Wu, S.; Li, Z.; Dong, Z.; An, X.; Ma, B.; Tian, Y.; Li, J. Maize Genic Male-Sterility Genes and Their Applications in Hybrid Breeding: Progress and Perspectives. *Mol. Plant* **2019**, *12*, 321–342. [[CrossRef](#)] [[PubMed](#)]
8. Deveshwar, P.; Bovill, W.D.; Sharma, R.; Able, J.A.; Kapoor, S. Analysis of anther transcriptomes to identify genes contributing to meiosis and male gametophyte development in rice. *BMC Plant Biol.* **2011**, *11*, 78. [[CrossRef](#)] [[PubMed](#)]
9. Koltunow, A.M.; Truettner, J.; Cox, K.H.; Wallroth, M.; Goldberg, R.B. Different temporal and spatial gene expression patterns occur during anther development. *Plant Cell* **1990**, *2*, 1201–1224. [[CrossRef](#)]
10. Tsuchiya, T.; Toriyama, K.; Ejiri, S.; Hinata, K. Molecular characterization of rice genes specifically expressed in the anther tapetum. *Plant Mol. Biol.* **1994**, *26*, 1737–1746. [[CrossRef](#)]
11. Hihara, Y.; Hara, C.; Uchimiyama, H. Isolation and characterization of two cDNA clones for mRNAs that are abundantly expressed in immature anthers of rice (*Oryza sativa* L.). *Plant Mol. Biol.* **1996**, *30*, 1181–1193. [[CrossRef](#)]
12. Rubinelli, P.; Hu, Y.; Ma, H. Identification, sequence analysis and expression studies of novel anther-specific genes of *Arabidopsis thaliana*. *Plant Mol. Biol.* **1998**, *37*, 607–619. [[CrossRef](#)]
13. Jeon, J.-S.; Chung, Y.-Y.; Lee, S.; Yi, G.-H.; Oh, B.-G.; An, G. Isolation and characterization of an anther-specific gene, RA8, from rice (*Oryza sativa* L.). *Plant Mol. Biol.* **1999**, *39*, 35–44. [[CrossRef](#)] [[PubMed](#)]
14. Amagai, M.; Ariizumi, T.; Endo, M.; Hatakeyama, K.; Kuwata, C.; Shibata, D.; Toriyama, K.; Watanabe, M. Identification of anther-specific genes in a cruciferous model plant, *Arabidopsis thaliana*, by using a combination of *Arabidopsis* macroarray and mRNA derived from *Brassica oleracea*. *Sex. Plant Reprod.* **2003**, *15*, 213–220. [[CrossRef](#)]
15. Masuko, H.; Endo, M.; Saito, H.; Hakozaiki, H.; Park, J.-I.; Kawagishi-Kobayashi, M.; Takada, Y.; Okabe, T.; Kamada, M.; Takahashi, H.; et al. Anther-specific genes, which expressed through microsporogenesis, are temporally and spatially regulated in model legume, *Lotus japonicus*. *Genes Genet. Syst.* **2006**, *81*, 57–62. [[CrossRef](#)]
16. Ma, J.; Skibbe, D.S.; Fernandes, J.; Walbot, V. Male reproductive development: Gene expression profiling of maize anther and pollen ontogeny. *Genome Biol.* **2008**, *9*, R181. [[CrossRef](#)] [[PubMed](#)]
17. Suwabe, K.; Suzuki, G.; Takahashi, H.; Shiono, K.; Endo, M.; Yano, K.; Fujita, M.; Masuko, H.; Saito, H.; Fujioka, T.; et al. Separated Transcriptomes of Male Gametophyte and Tapetum in Rice: Validity of a Laser Microdissection (LM) Microarray. *Plant Cell Physiol.* **2008**, *49*, 1407–1416. [[CrossRef](#)]
18. Chen, C.; Farmer, A.D.; Langley, R.J.; Mudge, J.; Crow, J.A.; May, G.D.; Huntley, J.; Smith, A.G.; Retzel, E.F. Meiosis-specific gene discovery in plants: RNA-Seq applied to isolated *Arabidopsis* male meiocytes. *BMC Plant Biol.* **2010**, *10*, 280. [[CrossRef](#)]
19. Aya, K.; Suzuki, G.; Suwabe, K.; Hobo, T.; Takahashi, H.; Shiono, K.; Yano, K.; Tsutsumi, N.; Nakazono, M.; Nagamura, Y.; et al. Comprehensive network analysis of anther-expressed genes in rice by the combination of 33 laser microdissection and 143 spatiotemporal microarrays. *PLoS ONE* **2011**, *6*, e26162. [[CrossRef](#)] [[PubMed](#)]
20. Klepikova, A.V.; Kasianov, A.S.; Gerasimov, E.S.; Logacheva, M.D.; Penin, A.A. A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *Plant J.* **2016**, *88*, 1058–1070. [[CrossRef](#)]
21. Lin, Y.; Zhang, L.; Zhao, Y.; Wang, Z.; Liu, H.; Zhang, L.; Zhang, Y.; Fu, Y.; Wu, J.; Ge, Y.; et al. Comparative analysis and functional identification of temperature-sensitive miRNA in *Arabidopsis* anthers. *Biochem. Biophys. Res. Commun.* **2020**, *532*, 1–10. [[CrossRef](#)] [[PubMed](#)]
22. Nguyen, T.D.; Moon, S.; Nguyen, V.N.; Gho, Y.; Chandran, A.K.; Soh, M.S.; Song, J.T.; An, G.; Oh, S.A.; Park, S.K.; et al. Genome-wide identification and analysis of rice genes preferentially expressed in pollen at an early developmental stage. *Plant Mol. Biol.* **2016**, *92*, 71–88. [[CrossRef](#)] [[PubMed](#)]
23. Moon, S.; Oo, M.M.; Kim, B.; Koh, H.J.; Oh, S.A.; Yi, G.; An, G.; Park, S.K.; Jung, K.H. Genome-wide analyses of late pollen-preferred genes conserved in various rice cultivars and functional identification of a gene involved in the key processes of late pollen development. *Rice* **2018**, *11*, 28. [[CrossRef](#)]

24. Oo, M.M.; Bae, H.K.; Nguyen, T.D.; Moon, S.; Oh, S.A.; Kim, J.H.; Soh, M.S.; Song, J.T.; Jung, K.H.; Park, S.K. Evaluation of rice promoters conferring pollen-specific expression in a heterologous system, *Arabidopsis*. *Plant Reprod.* **2014**, *27*, 47–58. [[CrossRef](#)] [[PubMed](#)]
25. Wang, M.; Yan, W.; Peng, X.; Chen, Z.; Xu, C.; Wu, J.; Deng, X.W.; Tang, X. Identification of late-stage pollen-specific promoters for construction of pollen-inactivation system in rice. *J. Integr. Plant Biol.* **2020**, *62*, 1246–1263. [[CrossRef](#)] [[PubMed](#)]
26. Zheng, X.; He, L.; Liu, Y.; Mao, Y.; Wang, C.; Zhao, B.; Li, Y.; He, H.; Guo, S.; Zhang, L.; et al. A study of male fertility control in *Medicago truncatula* uncovers an evolutionarily conserved recruitment of two tapetal bHLH subfamilies in plant sexual reproduction. *New Phytol.* **2020**, *228*, 1115–1133. [[CrossRef](#)] [[PubMed](#)]
27. Wilson, Z.A.; Morroll, S.M.; Dawson, J.; Swarup, R.; Tighe, P.J. The *Arabidopsis* MALE STE-RILITY1 (MS1) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors. *Plant J.* **2001**, *28*, 27–39. [[CrossRef](#)]
28. Sorensen, A.-M.; Kröber, 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](#)]
29. Ariizumi, T.; Hatakeyama, K.; Hinata, K.; Inatsugi, R.; Nishida, I.; Sato, S.; Kato, T.; Tabata, S.; Toriyama, K. Disruption of the novel plant protein NEF1 affects lipid accumulation in the plastids of the tapetum and exine formation of pollen, resulting in male sterility in *Arabidopsis thaliana*. *Plant J.* **2004**, *39*, 170–181. [[CrossRef](#)]
30. Yang, C.; Vizcay-Barrena, G.; Conner, K.; Wilson, Z.A. MALE STERILITY1 is required for ta-petal development and pollen wall biosynthesis. *Plant Cell* **2007**, *19*, 3530–3548. [[CrossRef](#)]
31. Luo, H.; Lee, J.-Y.; Hu, Q.; Nelson-Vasilchik, K.; Eitas, T.K.; Lickwar, C.; Kausch, A.P.; Chandlee, J.M.; Hodges, T.K. RTS, a rice anther-specific gene is required for male fertility and its promoter sequence directs tissue-specific gene expression in different plant species. *Plant Mol. Biol.* **2006**, *62*, 397–408. [[CrossRef](#)] [[PubMed](#)]
32. Endo, M.; Tsuchiya, T.; Saito, H.; Matsubara, H.; Hakozaki, H.; Masuko, H.; Kamada, M.; Higashitani, A.; Takahashi, H.; Fukuda, H.; et al. Identification and molecular characterization of novel anther-specific genes in *Oryza sativa* L. by using cDNA microarray. *Genes Genet. Syst.* **2004**, *79*, 213–226. [[CrossRef](#)] [[PubMed](#)]
33. Ma, Y.; Kang, J.; Wu, J.; Zhu, Y.; Wang, X. Identification of tapetum-specific genes by comparing global gene expression of four different male sterile lines in *Brassica oleracea*. *Plant Mol. Biol.* **2015**, *87*, 541–554. [[CrossRef](#)] [[PubMed](#)]
34. Qi, J.; Ni, F.; Wang, X.; Sun, M.; Cui, Y.; Wu, J.; Caplan, A.; Fu, D. The anther-specific CYP704B is potentially responsible for MSG26 male sterility in barley. *Theor. Appl. Genet.* **2019**, *132*, 2413–2423. [[CrossRef](#)] [[PubMed](#)]
35. Akasaka, M.; Taniguchi, Y.; Oshima, M.; Abe, K.; Tabei, Y.; Tanaka, J. Development of trans-genic male-sterile rice by using anther-specific promoters identified by comprehensive screening of the gene expression profile database ‘RiceXPro’. *Breed. Sci.* **2018**, *68*, 420–431. [[CrossRef](#)]
36. Wang, D.; Li, C.; Zhao, Q.; Zhao, L.; Wang, M.; Zhu, D.; Ao, G.; Yu, J. Zm401p10, encoded by an anther-specific gene with short open reading frames, is essential for tapetum degeneration and anther development in maize. *Funct. Plant Biol.* **2009**, *36*, 73–85. [[CrossRef](#)]
37. Frankis, R.; Mascarenhas, J.P. Messenger RNA in the Ungerminated Pollen Grain: A Direct Demonstration of its Presence. *Ann. Bot.* **1980**, *45*, 595–599. [[CrossRef](#)]
38. Tupý, J. Alterations in polyadenylated RNA during pollen maturation and germination. *Biol. Plant* **1982**, *24*, 331–340. [[CrossRef](#)]
39. Becker, J.D.; Boavida, L.C.; Carneiro, J.; Haury, M.; Feijó, J.A. Transcriptional profiling of *Arabidopsis* tissues reveals the unique characteristics of the pollen transcriptome. *Plant Physiol.* **2003**, *133*, 713–725. [[CrossRef](#)]
40. Mascarenhas, J.P. Gene activity during pollen development. *Annu. Rev. Plant Bio.* **1990**, *41*, 317–338. [[CrossRef](#)]
41. Willing, R.P.; Bashe, D.; Mascarenhas, J.P. An analysis of the quantity and diversity of messenger RNAs from pollen and shoots of *Zea mays*. *Theor. Appl. Genet.* **1988**, *75*, 751–753. [[CrossRef](#)]
42. Haerizadeh, F.; Wong, C.E.; Bhalla, P.L.; Gresshoff, P.M.; Singh, M.B. Genomic expression profiling of mature soybean (*Glycine max*) pollen. *BMC Plant Biol.* **2009**, *9*, 25. [[CrossRef](#)]
43. Van Tunen, A.J.; Mur, L.A.; Brouns, G.S.; Rienstra, J.D.; Koes, R.E.; Mol, J.N. Pollen- and anther-specific chi promoters from petunia: Tandem promoter regulation of the chiA gene. *Plant Cell* **1990**, *2*, 393–401. [[PubMed](#)]
44. Albani, D.; Robert, L.S.; Donaldson, P.A.; Altosaar, I.; Arnison, P.G.; Fabijanski, S.F. Characterization of a pollen-specific gene family from *Brassica napus* which is activated during early microspore development. *Plant Mol. Biol.* **1990**, *15*, 605–622. [[CrossRef](#)]
45. Twell, D.; Yamaguchi, J.; McCormick, S. Pollen-specific gene expression in transgenic plants: Coordinate regulation of two different tomato gene promoters during microsporogenesis. *Development* **1990**, *109*, 705–713. [[CrossRef](#)] [[PubMed](#)]
46. Weterings, K.; Schrauwen, J.; Wullems, G.; Twell, D. Functional dissection of the promoter of the pollen-specific gene NTP303 reveals a novel pollen-specific, and conserved cis-regulatory element. *Plant J.* **1995**, *8*, 55–63. [[CrossRef](#)]
47. Hamilton, D.A.; Schwarz, Y.H.; Mascarenhas, J.P. A monocot pollen-specific promoter contains separable pollen-specific and quantitative elements. *Plant Mol. Biol.* **1998**, *38*, 663–669. [[CrossRef](#)]
48. Honys, D.; Oh, S.A.; Renák, D.; Donders, M.; Solcová, B.; Johnson, J.A.; Boudová, R.; Twell, D. Identification of microspore-active promoters that allow targeted manipulation of gene expression at early stages of microgametogenesis in *Arabidopsis*. *BMC Plant Biol.* **2006**, *6*, 31. [[CrossRef](#)]
49. Bate, N.; Twell, D. Functional architecture of a late pollen promoter: Pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. *Plant Mol. Biol.* **1998**, *37*, 859–869. [[CrossRef](#)] [[PubMed](#)]

50. Rogers, H.; Bate, N.; Combe, J.; Sullivan, J.; Sweetman, J.; Swan, C.; Lonsdale, D.; Twell, D. Functional analysis of *cis*-regulatory elements within the promoter of the tobacco late pollen gene *g10*. *Plant Mol. Biol.* **2001**, *45*, 577–585. [[CrossRef](#)] [[PubMed](#)]
51. Li, J.; Yuan, J.; Li, M. Characterization of putative *cis*-regulatory elements in genes preferentially expressed in Arabidopsis male meiocytes. *Biomed. Res. Int.* **2014**, *2014*, 708364. [[CrossRef](#)] [[PubMed](#)]
52. Peng, J.; Qi, X.; Chen, X.; Li, N.; Yu, J. ZmDof30 Negatively Regulates the Promoter Activity of the Pollen-Specific Gene Zm908. *Front. Plant Sci.* **2017**, *8*, 685. [[CrossRef](#)]
53. Zhan, X.-Y.; Wu, H.-M.; Cheung, A. Nuclear male sterility induced by pollen-specific expression of a ribonuclease. *Sex. Plant Reprod.* **1996**, *9*, 35–43. [[CrossRef](#)]
54. Wu, Y.; Fox, T.W.; Trimmell, M.R.; Wang, L.; Xu, R.; Cigan, A.M.; Huffman, G.A.; Garnaat, C.W.; Hershey, H.P.; Albertsen, M.C. Development of a novel recessive genetic male sterility system for hybrid seed production in maize and other cross-pollinating crops. *Plant Biotechnol. J.* **2015**, *14*, 1046–1054. [[CrossRef](#)]
55. Honys, D.; Twell, D. Comparative Analysis of the Arabidopsis Pollen Transcriptome. *Plant Physiol.* **2003**, *132*, 640–652. [[CrossRef](#)]
56. Pina, C.; Pinto, F.; Feijó, J.A.; Becker, J. Gene Family Analysis of the Arabidopsis Pollen Transcriptome Reveals Biological Implications for Cell Growth, Division Control, and Gene Expression Regulation. *Plant Physiol.* **2005**, *138*, 744–756. [[CrossRef](#)] [[PubMed](#)]
57. Hafidh, S.; Breznenová, K.; Růžička, P.; Feciková, J.; Capková, V.; Honys, D. Comprehensive analysis of tobacco pollen transcriptome unveils common pathways in polar cell expansion and underlying heterochronic shift during spermatogenesis. *BMC Plant Biol.* **2012**, *16*, 24. [[CrossRef](#)] [[PubMed](#)]
58. Bokvaj, P.; Hafidh, S.; Honys, D. Transcriptome profiling of male gametophyte development in *Nicotiana tabacum*. *Genom. Data* **2015**, *3*, 106–111. [[CrossRef](#)]
59. Conze, L.L.; Berlin, S.; Le Bail, A.; Kost, B. Transcriptome profiling of tobacco (*Nicotiana tabacum*) pollen and pollen tubes. *BMC Genom.* **2017**, *18*, 581. [[CrossRef](#)]
60. Huang, M.-D.; Wei, F.-J.; Wu, C.-C.; Hsing, Y.-I.C.; Huang, A.H. Analyses of Advanced Rice Anther Transcriptomes Reveal Global Tapetum Secretory Functions and Potential Proteins for Lipid Exine Formation. *Plant Physiol.* **2009**, *149*, 694–707. [[CrossRef](#)]
61. Yan, S.; Wang, Z.; Liu, Y.; Li, W.; Wu, F.; Lin, X.; Meng, Z. Functional architecture of two exclusively late stage pollen-specific promoters in rice (*Oryza sativa* L.). *Plant Mol. Biol.* **2015**, *88*, 415–428. [[CrossRef](#)]
62. Huang, M.-D.; Hsing, Y.-I.C.; Huang, A.H.C. Transcriptomes of the Anther Sporophyte: Availability and Uses. *Plant Cell Physiol.* **2011**, *52*, 1459–1466. [[CrossRef](#)] [[PubMed](#)]
63. Xu, X.H.; Chen, H.; Sang, Y.L.; Wang, F.; Ma, J.P.; Gao, X.Q.; Zhang, X.S. Identification of genes specifically or preferentially expressed in maize silk reveals similarity and diversity in transcript abundance of different dry stigmas. *BMC Genom.* **2012**, *13*, 294. [[CrossRef](#)]
64. Zhang, H.; Egger, R.L.; Kelliher, T.; Morrow, D.; Fernandes, J.; Nan, G.L.; Walbot, V. Transcriptomes and proteomes define gene expression progression in pre-meiotic maize anthers. *G3 Genes Genom. Genet.* **2014**, *4*, 993–1010.
65. Chettoor, A.M.; Phillips, A.R.; Coker, C.T.; Dilkes, B.; Evans, M.M.S. Maternal Gametophyte Effects on Seed Development in Maize. *Genetics* **2016**, *204*, 233–248. [[CrossRef](#)] [[PubMed](#)]
66. Yuan, T.-L.; Huang, W.-J.; He, J.; Zhang, D.; Tang, W.-H. Stage-Specific Gene Profiling of Germinal Cells Helps Delineate the Mitosis/Meiosis Transition. *Plant Physiol.* **2018**, *176*, 1610–1626. [[CrossRef](#)]
67. Li, Z.; Zhang, P.; Lv, J.; Cheng, Y.; Cui, J.; Zhao, H.; Hu, S. Global Dynamic Transcriptome Programming of Rapeseed (*Brassica napus* L.) Anther at Different Development Stages. *PLoS ONE* **2016**, *11*, e0154039. [[CrossRef](#)]
68. Feng, N.; Song, G.; Guan, J.; Chen, K.; Jia, M.; Huang, D.; Wu, J.; Zhang, L.; Kong, X.; Geng, S.; et al. Transcriptome Profiling of Wheat Inflorescence Development from Spikelet Initiation to Floral Patterning Identified Stage-Specific Regulatory Genes. *Plant Physiol.* **2017**, *174*, 1779–1794. [[CrossRef](#)]
69. Walley, J.; Sartor, R.C.; Shen, Z.; Schmitz, R.J.; Wu, K.J.; Urich, M.A.; Nery, J.R.; Smith, L.G.; Schnable, J.C.; Ecker, J.R.; et al. Integration of omic networks in a developmental atlas of maize. *Science* **2016**, *353*, 814–818. [[CrossRef](#)]
70. Stelplflug, S.C.; Sekhon, R.S.; Vaillancourt, B.; Hirsch, C.N.; Buell, C.R.; De Leon, N.; Kaeppler, S.M. An Expanded Maize Gene Expression Atlas based on RNA Sequencing and its Use to Explore Root Development. *Plant Genome* **2016**, *9*, 1–16. [[CrossRef](#)]
71. Lindskog, C.; Linné, J.; Fagerberg, L.; Hallström, B.M.; Sundberg, C.J.; Lindholm, M.; Huss, M.; Kampf, C.; Choi, H.; Liem, D.A.; et al. The human cardiac and skeletal muscle proteomes defined by transcriptomics and antibody-based profiling. *BMC Genom.* **2015**, *16*, 475. [[CrossRef](#)] [[PubMed](#)]
72. Schnable, P.S.; Ware, D.; Fulton, R.S.; Stein, J.C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T.A.; et al. The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science* **2009**, *326*, 1112–1115. [[CrossRef](#)]
73. Hoopes, G.M.; Hamilton, J.P.; Wood, J.; Esteban, E.; Pasha, A.; Vaillancourt, B.; Provart, N.J.; Buell, C.R. An updated gene atlas for maize reveals organ-specific and stress-induced genes. *Plant J.* **2018**, *97*, 1154–1167. [[CrossRef](#)]
74. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)] [[PubMed](#)]
75. Chen, M.; Xu, J.; Devis, D.; Shi, J.; Ren, K.; Searle, I.; Zhang, D. Origin and functional prediction of pollen allergens in plants. *Plant Physiol.* **2016**, *172*, 341–357. [[CrossRef](#)]
76. Usadel, B.; Poree, F.; Nagel, A.; Lohse, M.; Czedik-Eysenberg, A.; Stitt, M. A guide to using MapMan to visualize and compare Omics data in plants: A case study in the crop species, Maize. *Plant Cell Environ.* **2009**, *32*, 1211–1229. [[CrossRef](#)]

77. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van De Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327. [[CrossRef](#)]
78. Nussinov, R. Compilation of eukaryotic sequences around transcription initiation sites. *J. Theor. Biol.* **1986**, *120*, 479–487. [[CrossRef](#)]
79. Twell, D.; Klein, T.M.; Fromm, M.E.; McCormick, S. Transient Expression of Chimeric Genes Delivered into Pollen by Microprojectile Bombardment. *Plant Physiol.* **1989**, *91*, 1270–1274. [[CrossRef](#)]
80. Carpenter, J.L.; Ploense, S.E.; Snustad, D.P.; Silflow, C.D. Preferential expression of an α -tubulin gene of Arabidopsis in pollen. *Plant Cell* **1992**, *4*, 557–571.
81. Oldenhof, M.T.; De Groot, P.F.M.; Visser, J.H.; Schrauwen, J.A.M.; Wullems, G.J. Isolation and characterization of a microspore-specific gene from tobacco. *Plant Mol. Biol.* **1996**, *31*, 213–225. [[CrossRef](#)]
82. Rogers, H.J.; Harvey, A.; Lonsdale, D.M. Isolation and characterization of a tobacco gene with homology to pectate lyase which is specifically expressed during microsporogenesis. *Plant Mol. Biol.* **1992**, *20*, 493–502. [[CrossRef](#)]
83. Allen, R.L.; Lonsdale, D.M. Molecular characterization of one of the maize polygalacturonase gene family members which are expressed during late pollen development. *Plant J.* **1993**, *3*, 261–271. [[CrossRef](#)]
84. Turcich, M.P.; Hamilton, D.A.; Mascarenhas, J.P. Isolation and characterization of pollen-specific maize genes with sequence homology to ragweed allergens and pectate lyases. *Plant Mol. Biol.* **1993**, *23*, 1061–1065. [[CrossRef](#)]
85. Rubinstein, A.L.; Broadwater, A.H.; Lowrey, K.B.; Bedinger, P.A. Pex1, a pollen-specific gene with an extensin-like domain. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3086–3090. [[CrossRef](#)]
86. Lopez, I.; Anthony, R.G.; Maciver, S.K.; Jiang, C.-J.; Khan, S.; Weeds, A.G.; Hussey, P.J. Pollen specific expression of maize genes encoding actin depolymerizing factor-like proteins. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7415–7420. [[CrossRef](#)]
87. Hanson, D.D.; Hamilton, D.A.; Travis, J.L.; Bashe, D.M.; Mascarenhas, J.P. Characterization of a pollen-specific cDNA clone from *Zea mays* and its expression. *Plant Cell* **1989**, *1*, 173–179.
88. Wakeley, P.R.; Rogers, H.J.; Rozycka, M.; Greenland, A.J.; Hussey, P.J. A maize pectin methylesterase-like gene, ZmC5, specifically expressed in pollen. *Plant Mol. Biol.* **1998**, *37*, 187–192. [[CrossRef](#)]
89. Kelliher, T.; Starr, D.; Richbourg, L.; Chintamanani, S.; Delzer, B.; Nuccio, M.L.; Green, J.; Chen, Z.; McCuiston, J.; Wang, W.; et al. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. *Nat. Cell Biol.* **2017**, *542*, 105–109. [[CrossRef](#)]
90. Gilles, L.M.; Khaled, A.; Laffaire, J.; Chaignon, S.; Gendrot, G.; Laplaige, J.; Bergès, H.; Beydon, G.; Bayle, V.; Barret, P.; et al. Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. *EMBO J.* **2017**, *36*, 707–717. [[CrossRef](#)]
91. Wang, H.; Fan, M.; Wang, G.; Zhang, C.; Shi, L.; Wei, Z.; Ma, W.; Chang, J.; Huang, S.; Lin, F. Isolation and characterization of a novel pollen-specific promoter in maize (*Zea mays* L.). *Genome* **2017**, *60*, 485–495. [[CrossRef](#)]
92. Jiang, S.-Y.; Ramachandran, S. Functional Genomics of Rice Pollen and Seed Development by Genome-wide Transcript Profiling and Ds Insertion Mutagenesis. *Int. J. Biol. Sci.* **2011**, *7*, 28–40. [[CrossRef](#)]
93. Honys, D.; Twell, D. Transcriptome analysis of haploid male gametophyte development in Arabidopsis. *Genome Biol.* **2004**, *5*, R85. [[CrossRef](#)]
94. Twell, D.; Yamaguchi, J.; Wing, R.A.; Ushiba, J.; McCormick, S. Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. *Genes Dev.* **1991**, *5*, 496–507. [[CrossRef](#)]
95. Eyal, Y.; Curie, C.; McCormick, S. Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes. *Plant Cell* **1995**, *7*, 373–384.
96. Hoffmann, R.D.; Olsen, L.I.; Husum, J.O.; Nicolet, J.S.; Thøfner, J.F.; Wätjen, A.P.; Ezike, C.V.; Palmgren, M. A *cis*-Regulatory Sequence Acts as a Repressor in the Arabidopsis thaliana Sporophyte but as an Activator in Pollen. *Mol. Plant* **2017**, *10*, 775–778. [[CrossRef](#)]
97. Chang, Y.; Yu, J.; Zhu, D.; Zhao, Q.; Yan, M. The 5' untranslated region of potato SBgLR gene contributes to pollen-specific expression. *Planta* **2017**, *157*, 1–403. [[CrossRef](#)]
98. Lang, Z.; Zhou, P.; Yu, J.; Ao, G.; Zhao, Q. Functional characterization of the pollen-specific SBgLR promoter from potato (*Solanum tuberosum* L.). *Planta* **2007**, *227*, 387–396. [[CrossRef](#)]
99. Hamilton, D.A.; Roy, M.; Rueda, J.; Sindhu, R.K.; Sanford, J.; Mascarenhas, J.P. Dissection of a pollen-specific promoter from maize by transient transformation assays. *Plant Mol. Biol.* **1992**, *18*, 211–218. [[CrossRef](#)]
100. Lin, Y.; Zhang, C.; Lan, H.; Gao, S.; Liu, H.; Liu, J.; Cao, M.; Pan, G.; Rong, T.; Zhang, S. Validation of Potential Reference Genes for qPCR in Maize across Abiotic Stresses, Hormone Treatments, and Tissue Types. *PLoS ONE* **2014**, *9*, e95445. [[CrossRef](#)]